

La communication d'attitudes sociales par la musique :
une étude de décodage

Hugo Trad
IRCAM STMS UMR9912

Juin 2014

« Music has all the attributes of a communicative system that is highly adapted to facilitate the management of the uncertainties of social interaction » (Ian Cross, The Evolutionary Nature of Musical Meaning, p. 190)

« La musique improvisée entretient de nombreux rapports avec ce qui se vit de (plus) fort à deux et/ou à plusieurs dans le monder tel qu'il est. [...] A chaque concert, c'est comme autant de réinvention de l'entraide (la boxe aussi en est une). Les improvisations sont autant de schémas possibles d'une généalogie des conflits et des alliances entre les hommes, elles sont comme la remise en jeu perpétuelle de ce qu'a pu être la construction progressive de l'altruisme, du jeu, ou des raisons de l'encadrement des luttes, mais aussi comme le tracé à vif (et toujours en-musique) de ce que peut être l'histoire d'un couple, l'histoire d'un amour (les duos), d'une rupture » (JL Guionney, Filigrane, 8, 2008, p. 134).

1. Introduction

S'il est possible d'estimer l'ancienneté de la musique (découverte de flûtes en ivoire datant de 40 000 av. J.C; Conard, Malina & Münzel, 2009), les origines évolutives de cette activité proprement humaine semblent à jamais perdues dans l'histoire de nos ancêtres, et ceci pour la simple raison qu'un trait cognitif ne peut être fossilisé (Lewontin, 1998; Honing 2012). Néanmoins, il est évident que la musique accompagne de nombreuses activités humaines, et ceci est particulièrement visible dans les sociétés préhistoriques (Becker, 2001; Gregory, 1997); de ce constat, de nombreuses théories en psychologie évolutionniste suggèrent que la musique aurait pu jouer un rôle secondaire ou indirect dans l'évolution en tant qu'elle permettrait d'organiser et renforcer la cohésion du groupe social (Cross, 2007; Merker 2007; Honing 2012).

Jusqu'à présent, le pouvoir de la musique à instaurer et réguler des comportements sociaux a été surtout mis en évidence par des effets indirects. Par exemple, le fait de faire de la musique ensemble peut amener des gens à coopérer plus volontiers avec le groupe (Wiltermuch & Heath, 2009), à être plus serviables et enclins à aider les autres (Kirschner & Tomasello, 2010) ou à éprouver une plus grande confiance vis à vis d'autrui (Anshel & Kipper, 1988). Dans des contextes cliniques, les activités musicales de groupe amènent certains patients à mieux coopérer aux actes thérapeutiques, par exemple dans le cas de réhabilitation après un accident vasculaire cérébral (Nayak et al., 2000) ou lors d'une prise en charge pour la maladie d'Alzheimer (Thomas et al., 1997). Cependant ces effets étant indirects, il n'est pas clair si la musique arrive à ces effets par son évocation émotionnelle, un effet relaxant, ou une connotation sémantique etc.; en bref, la question reste entière de savoir si la musique est capable de communiquer directement (i.e. encoder, et décoder) des indices sociaux.

Il existe une grande littérature sur la fonction expressive de la musique, notamment sur sa capacité à communiquer des émotions (Justin & Laukka, 2003) : il a été montré qu'elle était capable de communiquer des émotions de base (Ekman, 1992) dont la colère, la peur, la joie, et la tristesse. De la même manière, il a également été montré que des participants pouvaient discriminer précisément des musiques de types émotionnels différents (par exemple des musiques festives, guerrières, funéraires etc...) provenant de cultures différentes (Eggebrecht, 1993; Balkwill & Thompson, 1999), et qu'il existait des caractéristiques acoustiques communes (i.e. tempo, niveau sonore, timbre, mode) entre des morceaux d'un même type émotionnel (Gagnon & Peretz, 2003). Seulement, si ces études peuvent contribuer à confirmer la thèse selon laquelle la musique participerait à la cohésion sociale, elles n'en restent pas moins incomplètes : de fait, en étudiant des effets unilatéraux (i.e. un seul musicien qui encode et un participant qui décode) elles ne considèrent que des émotions intrapersonnelles. Dès lors, la question des attitudes sociales dyadiques (e.g. autoritaire, dédaigneux, conciliant - Wichmann, 2000) - pourtant indispensables à l'organisation du groupe social - est mise de côté car peu naturelle dans ces conditions expérimentales.

Si le fait que la musique communique des émotions intrapersonnelles permette déjà de médier des relations interpersonnelles (Keltner & Haidt, 1999), il serait d'autant plus intéressant de montrer que la musique peut communiquer des intentions sociales au même titre que le langage (Fitch, 2005). Ainsi, ce travail s'inscrit à la fois dans la lignée des études sur les origines évolutionnistes de la musique (Honing, 2012), et dans un certain nombre de théories récentes en musicologie qui considèrent qu'une lecture « agentielle » de certaines musiques est possible, la musique étant capable d'incarner tour à tour différents agents abstraits ou personae. (Maus, 1988).

Afin de poser cette question, des travaux préliminaires à notre expérience ont été réalisés dans une situation de duo improvisé entre deux musiciens (un encodeur et un décodeur-interlocuteur qui accompagne simplement le premier). Ceux-ci ont montré que les instrumentistes interlocuteurs étaient capables de décoder les attitudes encodées avec des performances similaires à l'expression des émotions de base dans la musique et la voix (cf section 2.1 pour le protocole, section 3.1 pour les résultats).

A la suite de ces travaux préliminaires, le travail de ce stage consistait à :

- (a) Tester si ces bonnes performances de décodage par des musiciens sont également observées dans la population générale
- (b) Tester l'hypothèse selon laquelle ce décodage d'attitudes sociales est favorisé par l'information dyadique - c'est-à-dire mettant en jeu une interaction entre deux instruments : un encodeur qui joue une attitude et un deuxième qui réagit

librement à cette proposition - et pénalisé par la seule écoute de l'instrument encodeur.

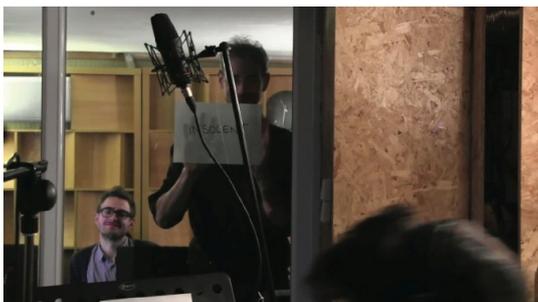
Pour cela, nous avons réalisé une étude de performance sur jugement catégoriel (i.e. de décodage) au casque et en cabine insonorisée auprès d'un ensemble de participants.

2. Méthodes

2.1. Stimuli

Les stimuli utilisés dans l'expérience ont été sélectionnés à partir du corpus de duos enregistrés en studio au mois de mars 2014. Afin de réaliser ces enregistrements, nous avons fait appel à des instrumentistes de la classe d'improvisation générative du CNSMDP (Conservatoire National Supérieur de Musique et de Danse de Paris). Pour chaque duo, un des musiciens (l'encodeur) devait exprimer une attitude tandis que l'autre (le décodeur) était chargé d'improviser avec lui et, à la suite de l'improvisation, de reconnaître l'attitude qui lui était communiquée par l'encodeur (voir figure 1).

Figure 1 :
Protocole de
l'enregistrement
du corpus de
sons utilisés lors
de l'expérience.



Avant chaque duo, un des deux musiciens, l'«encodeur», est informé d'une attitude à exprimer, à l'insu de son compagnon de jeu qui sera, lui, chargé de la «décoder».



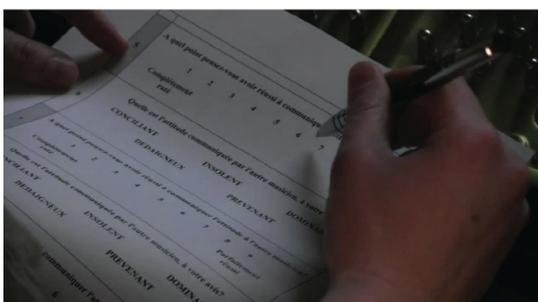
Les deux musiciens, «encodeur» et «décodeur», sont placés dans des cabines d'enregistrement distinctes. Ils jouent ensemble, à l'aveugle, en s'écoutant au casque.



Au cours du duo, l'encodeur essaie de faire comprendre au décodeur, par sa seule interaction musicale avec lui, quelle attitude il est en train d'exprimer.



De son côté, le musicien décodeur a pour consigne d'interagir musicalement avec l'encodeur, dans un duo improvisé librement, pendant environ 2 minutes.



A la fin de chaque duo, l'encodeur évalue à quel point il pense avoir réussi à communiquer l'attitude qu'il était tâché d'exprimer...



... tandis que le décodeur essaie de deviner quelle était justement cette attitude. Puis les rôles sont inversés, une nouvelle attitude est tirée au sort, et on recommence.

Dix duos ont été sollicités, chaque instrumentiste jouant les 5 attitudes sélectionnées (cf 2.2) : nous avons donc réalisé au total 100 enregistrements, soit 20 enregistrements par attitudes chacun joué par un instrumentiste différent.

Nous avons finalement sélectionné 25 stimuli - soit 5 par attitude - sur les 100 enregistrés. Cette sélection a été faite selon plusieurs critères : (a) nous n'avons considéré que les duos qui avaient été correctement reconnus (c'est-à-dire décodés) par les musiciens décodeurs (64/100). (b) Parmi ceux-ci, nous n'avons considéré que ceux qui avaient été bien notés par les musiciens encodeurs (entre 7/10 et 9/10). (c) Ensuite, nous avons enlevé ceux qui étaient trop longs, de façon à ne retenir que des improvisations durant entre 1:30 et 2:00 (d) ainsi que ceux qui faisaient des références culturelles (e.g. « Joyeux anniversaire »), (e) et dans le

reste, nous en avons sélectionné 25 de façon à maximiser la diversité des modes de jeu et des instruments.

Deux conditions ont été réalisées afin de tester notre deuxième hypothèse (i.e. que l'attitude est mieux décodée lorsque l'information est dyadique) :

- (1) Condition dyadique : un montage audio des 25 duos a été réalisé afin de mixer l'instrument « acteur » dans le canal de gauche et l' «interlocuteur » dans le canal de droite (effet de panoramique polyphonique, ou "panning"). Les participants doivent alors reconnaître l'attitude de l'instrument acteur (à gauche) envers l'instrument interlocuteur (à droite).
- (2) Condition monadique : seul l'instrument acteur, mixé au centre en stéréo, est gardé. Le participant doit essayer de reconnaître l'attitude de cet instrument.

2.2 Attitudes

Nous avons retenu 5 attitudes différentes: autoritaire, conciliant, dédaigneux, effronté, et prévenant. Ces attitudes ont été expliquées aux instrumentistes et aux participants sous formes de définitions (voir table 1) et de « moodboards », c'est-à-dire des plaquettes d'images représentatives des attitudes sélectionnées (voir figures 2).

Table 1: Définition des attitudes utilisées dans l'expérience.

Autoritaire, dominant, agressif, directif (AUT)	Quelqu'un dans cette attitude est sûr de sa supériorité ; il fait sentir à l'autre qu'il doit le suivre ou lui obéir, et qu'il n'est pas en mesure de le contester. Il n'hésite pas à le lui rappeler fermement s'il ne lui témoigne pas le respect qui lui est dû.
Conciliant, faire un pas vers l'autre, accommodant (CON)	Bien qu'à priori très différent de l'autre, voire en désaccord avec lui, quelqu'un dans cette attitude fait l'effort de faire un pas vers l'autre dans l'espoir que lui aussi fasse un pas vers lui, et que le dialogue puisse ainsi s'établir.
Dédaigneux, indifférent, sans égard (DED)	Quelqu'un dans cette attitude veut marquer sa distance avec l'autre, et bien lui faire sentir qu'il n'est pas vraiment du même monde que lui. Il fait en sorte que l'autre comprenne qu'il ne se sent aucune obligation vis-à-vis de lui, voire qu'il l'ignore.
Effronté, insolent, impudent (EFF)	Quelqu'un dans cette attitude veut que l'autre comprenne qu'il ne le respecte pas, qu'il pense que sa position est illégitime. Il tente donc de miner ce que l'autre tente de faire par un comportement irrévérencieux, provocateur ou ironique, sans se soucier des conséquences éventuelles que pourra entraîner son attitude.
Prévenant, être aux petits soins (PRE)	Quelqu'un dans cette attitude est à l'écoute des moindres souhaits et désirs de l'autre, il tente même de les devancer. Il fait en sorte de se faire oublier tout en donnant à l'autre l'impression que tout est fait pour lui faciliter la tâche.

Ces attitudes ont été choisies selon différents critères : (a) il devait s'agir de « speaker behavior » (Wichmann, 2000), i.e. des attitudes qui émergent forcément dans une interaction sociale et qui ne peuvent pas être endossées par un individu isolé (e.g. on est forcément "effronté" envers quelqu'un. A l'inverse, on peut être joyeux ou jaloux tout en étant isolé). (b) De plus, nous avons sélectionné des attitudes proches (e.g « autoritaire » et « effronté », des attitudes jugées toutes deux négatives) afin d'éviter que la tâche ne puisse être ramenée par les participants à une simple reconnaissance de valence émotionnelle (par exemple juger un stimulus comme « autoritaire » simplement parce qu'il s'agit d'une attitude négative, et « prévenant » simplement parce qu'il s'agit d'une attitude positive).

2.3 Protocole

On demande aux participants d'écouter 25 stimuli, présentés aléatoirement l'un après l'autre. Le participant doit reconnaître :

(a) En condition dyadique, l'attitude de l'instrument de gauche (encodeur) envers l'instrument de droite (décodeur)

(b) En condition monadique, l'attitude de l'unique instrument entendu.

La réponse du participant est un choix forcé entre les 5 attitudes possibles (voir figure 2).

Quel est le comportement de l'instrument de gauche (Clarinette) vis-a-vis de l'instrument de droite (Guitare)?

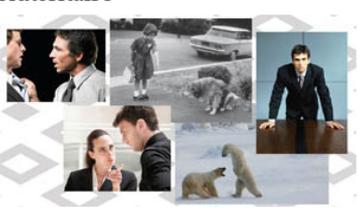



Votre réponse:
Prévenant

Prévenant



Autoritaire



Insolent



Conciliant



Dédaigneux



NEXT

2.4 Participants

N=20 adultes (M= 21.9, SD = 1.95, 10 hommes) ont participé à la condition dyadique; et N=20 adultes (M = 24,4, SD= 5.0, 10 hommes) ont participé à la condition monadique.

Les participants ont été recrutés sur le site du RISC (Relai d'information sur les sciences de la cognition - <http://www.risc.cnrs.fr>) sous la condition de ne pas être musiciens, i.e. ne pas avoir pas reçu de formation musicale; être incapables de déchiffrer une partition; ne pratiquer aucun instrument de manière assidue.

3. Résultats

3.1. Mesures de performances

Les tables 2.a, 2.b, 2.c présentent les matrices de confusion dénombrant le nombre de stimuli encodés et décodés dans chaque attitude; la table 2.a présente les résultats pour les musiciens en session d'enregistrement; la table 2.b présente les résultats pour les non-musiciens, en cabine, dans la condition dyadique; la table 2.c pour les non-musiciens, en condition monadique. Seules les tables 2.b et 2.c correspondent aux résultats du travail mené en stage, la table 2.a étant donnée pour comparaison.

Table 2: Matrices de confusion dénombrant le nombre de stimuli encodés et décodés dans chacune des 5 attitudes (AUT: autoritaire ; CON: conciliant ; DED: dédaigneux ; EFF: effronté ; PRE: prévenant), au cours de la session d'enregistrement en studio (2a) et dans les deux conditions dyadique (2b) et monadique (2c) réalisées au cours du stage. Les performances sont exprimées en *hit rate* H (rapport du nombre de stimuli correctement décodés par rapport au nombre de stimuli encodés dans chaque attitude) et en *proportion index* $\pi = H(K - 1)/(1 + H(K - 2))$, avec $K = 5$ (Rosenthal & Rubin,1989).

(a) En studio d'enregistrement

Encodés	Décodés					Total
	AUT	CON	DED	EFF	PRE	
AUT	14	0	3	2	1	20
CON	0	12	1	2	5	20
DED	3	3	13	0	1	20
EFF	5	2	4	9	0	20
PRE	0	3	0	1	16	20
Total	22	20	21	14	23	100
H	63%	60%	62%	45%	70%	
π	0.90	0.86	0.88	0.77	0.94	

(b) En cabine, condition dyadique

Encodés	Décodés					Total
	AUT	CON	DED	EFF	PRE	
AUT	32	20	7	18	23	100
CON	23	33	14	14	16	100
DED	3	24	44	13	16	100
EFF	45	18	10	24	3	100
PRE	1	36	12	14	37	100
Total	104	131	87	83	95	500
H	32%	33%	44%	24%	37%	
π	0.65	0.66	0.76	0.56	0.70	

(c) En cabine, condition monadique

Encodés	Décodés					Total
	AUT	CON	DED	EFF	PRE	
AUT	33	20	12	21	14	100
CON	19	24	32	16	9	100
DED	6	18	39	12	25	100
EFF	28	12	12	41	7	100
PRE	8	34	23	7	28	100
Total	94	108	118	97	83	500
H	33%	24%	39%	41%	28%	
π	0.66	0.56	0.72	0.74	0.61	

A partir de ces tables de confusion, plusieurs indices de performances peuvent être calculés:

- (a) Le hit rate (H), qui correspond à la probabilité qu'un stimulus d'une classe donnée soit bien identifié pourvu qu'il ait été présenté. A titre d'exemple, dans le cas d'effronté pour les musiciens en studio, 9 des 20 stimuli « effrontés » ont été reconnus comme « effronté », ce qui correspond à un hit rate = $9/20=45\%$.
- (b) La « differential accuracy » (D), qui correspond à la probabilité qu'une catégorie soit correctement appliquée à un stimulus pourvu qu'elle soit appliquée. Toujours dans le cas d'effronté, pour les musiciens en studio, la catégorie « effronté » n'a été utilisée que 14 fois seulement (pour 20 stimuli « effrontés » présentés), i.e. certaines personnes n'ont jamais répondu « effronté ». La differential accuracy est donc $D = 9/14 = 65\%$, c'est-à-dire que 65% des fois où quelqu'un répond « effronté », c'est effectivement sur un stimulus « effronté ».

- (c) Pour tenir compte de ces deux mesures, Wagner (1993) propose le unbiased hit rate H_u , qui est le produit de hit rate par la diff. acc. ($H \times D$). Cette mesure est la meilleure pour faire des comparaisons entre groupes, car elle tient compte non seulement de la précision des réponses, mais aussi du biais des participants à utiliser un type de réponse plutôt qu'un autre. C'est le unbiased hit rate que nous utiliserons dans la section suivante.
- (d) Enfin, Rosenthal & Rubin (1989) proposent un hit rate normalisé, le proportion index (PI), qui permet de comparer les hit rates entre expériences indépendamment du nombre de classes. Le PI ramène toute tâche de décodage à une tâche de discrimination entre 2 classes, où $PI = 0.50$ correspond donc au hasard. C'est une mesure de « taille d'effet » qui est utile pour donner une interprétation intuitive des résultats. $PI = \frac{H \cdot (K-1)}{1 + H \cdot (K-2)}$ ou K est le nombre de classes ($K=5$), et H le hit rate.

Comme vu dans la table 2.a, les hit rates des musiciens en condition studio vont de 70% (PRE) à 45% (EFF). Ces hit rates, mesurés pour une tâche à 5 classes, correspondent à des proportion index allant de 0.94 (PRE) à 0.77 (EFF). Ces scores sont comparables aux mesures rapportées par la meta-étude de Juslin & Laukka (2003), qui donnent des PI entre 80% et 95% pour le décodage des émotions de base dans la musique et l'expression vocale.

Les hit rates en cabine, pour les échantillons de non-musiciens, sont sévèrement dégradés par rapport aux musiciens en studio, quoique supérieur au hasard ($H=20\%$ entre 5 attitudes). En condition dyadique, les hit rates vont de 44% (DED) à 24% (EFF), ce qui correspond à des PI de 0.76 à 0.56. En condition monadique, les hit rates vont de 41% (EFF) à 24% (CON), ce qui correspond à des PI de 0.74 à 0.56. Ces mesures sont dans la limite inférieure des mesures de performances rapportée par Juslin & Laukka (2003).

3.2. Tests de significativité statistique

Notre première hypothèse expérimentale est que les capacités cognitives permettant de décoder les attitudes sociales communiquées par la musique sont présentes dans la population générale. Pour tester cette hypothèse, nous posons l'hypothèse statistique "nulle" (H_0) que la distribution des indices de performances sur la population de non-musiciens, évaluée sur l'échantillon de participants, ne diffère pas significativement (au critère $p < 0.05$) de la distribution des indices de performance qui serait observée dans le cas de réponses aléatoires à la même tâche. Notre hypothèse expérimentale sera vérifiée si les données observées permettent de rejeter cette hypothèse nulle.

Pour ce faire, comme suggéré dans Wagner (1993), nous calculons pour chaque participant son unbiased hit rate H_u pour chacune des 5 attitudes (cf ci-dessus). Pour comparer cet indice de performance à celui qui serait mesuré dans le cas de réponses au hasard, nous estimons, pour chaque participant, la probabilité conjointe p_c d'observer à la fois, par hasard et indépendamment, un stimulus et une réponse de la même catégorie. Cette probabilité est obtenue en multipliant la probabilité d'occurrence d'un stimulus d'une attitude ($p=5/25$) avec la probabilité d'occurrence d'une réponse de cette attitude ($p = N_{att}/25$ où N_{att} est le nombre de fois où le participant a utilisé cette attitude comme réponse). Nous comparons ensuite les distributions de H_u et de p_c , appariées pour chacun des $N=20$ participants dans chaque condition, avec un test t de Student pour échantillon apparié.

Dans la condition dyadique, les performances des participants diffèrent du hasard pour chacune des 5 attitudes: AUT: $t(19)=3.13$, $p=0.005$; CON: $t(19)=2.55$, $p=0.019$; DED: $t(19)=5.95$, $p=0.00001$; EFF: $t(19)=3.00$, $p=0.007$; PRE: $t(19)=3.51$, $p=0.002$.

Dans la condition monadique, les performances des participants diffèrent elles-aussi du hasard pour 4 des attitudes (AUT: $t(19)=3.44$, $p=0.002$; DED: $t(19)=3.81$, $p=0.001$; EFF: $t(19)=4.30$, $p=0.0004$; PRE: $t(19)=2.77$, $p=0.012$), mais pas pour l'attitude CON ($t(19)=1.73$, $p=0.09$).

Notre seconde hypothèse expérimentale est que le décodage est favorisé par l'informatique dyadique. Nous testons donc la différence statistique entre la distribution d'indices de performances du groupe dyadique et celle du groupe monadique. Nous procédons comme précédemment, en comparant les unbiased hit rates pour chaque attitude, entre groupes, avec un test t de Student pour échantillons indépendants. Les performances entre les 2 groupes ne diffèrent que pour deux des 5 attitudes: l'attitude DED est mieux décodée en condition dyadique ($t(38)=2.14$, $p=0.03$) mais l'attitude EFF est mieux décodée en condition monadique ($t(38)=-2.93$, $p=0.005$). Il n'y a pas de différence statistiquement significative entre les performances dyadiques et monadiques pour les attitudes AUT, CON et PRE.

4. Eléments de discussion

Ce travail de stage complète une expérience plus globale sur la capacité de la musique à communiquer des attitudes sociales. Le résultat global est que cette capacité existe : les performances de reconnaissance d'attitudes sociales dans la musique observées en conditions optimales (musiciens en studio) sont comparables à celles d'émotions intrapersonnelle, que ce soit dans la musique ou par le langage. De façon globale, ce travail établit donc la possibilité d'utiliser la musique

comme moyen de communication pour médier des comportements sociaux liés à l'autorité, la hiérarchie, la coopération, etc. Ces capacités ont été principalement prêtées au langage jusqu'à présent (Dunbar, 2003), et ce travail est le premier à établir cette possibilité pour la musique.

Le travail précis de ce stage consistait d'une part à tester les conditions d'application de ce résultat, établi sur des musiciens en condition studio, dans la population générale. Notre étude montre que les performances y sont dégradées (-20% de hit rate, ce qui correspond à une détection de moins sur 5 attitudes présentées en moyenne), mais ces performances restent supérieures au hasard, et ce de façon statistiquement significative. Ce résultat exclut donc que la capacité à décoder des indices sociaux dans la musique soit une capacité purement culturelle spécifique aux musiciens; au contraire, elle constitue vraisemblablement un trait partagé dans la population générale, qui est renforcé par l'expertise dans le cas des musiciens.

Le travail de ce stage consistait d'autre part à tester l'hypothèse que cette communication - étant une interaction sociale - repose sur l'information dyadique (i.e.. l'interaction entre les deux agents) et non le simple une simple information monadique (i.e. écoute d'un seul des deux agents). Les résultats sur ce point sont contrastés. Même si le hit rate de 4 attitudes sur 5 est plus important en condition dyadique qu'en condition monadique, la plupart de ces différences ne sont pas statistiquement significatives. L'attitude DED est significativement mieux reconnue en condition dyadique : ceci semble être cohérent dans la mesure où être dédaigneux consiste à éviter l'interlocuteur : il semble donc logiquement difficile de reconnaître cette attitude si le comportement de l'interlocuteur n'est pas observé conjointement. Cependant, l'attitude EFF est significativement moins bien reconnue en condition dyadique qu'en monadique: les données indiquent que la non-présentation du comportement de l'interlocuteur est corrélée à une baisse de la confusion de l'attitude EFF avec l'attitude voisine AUT. Cet effet est surprenant, et mériterait d'être examiné plus avant dans de futurs travaux.

Cette différence de performances entre les deux conditions peut être due : (a) à la différence d'expertise entre non-musiciens et musiciens, ou (b) à la différence de situation entre écoute passive en cabine et écoute participative en studio, ou (c) à la différence de difficulté de la tâche (5 stimuli pour 5 attitudes en studio, contre 25 stimuli pour 5 attitudes en cabine). En l'état, nos résultats ne nous permettent pas de choisir avec certitude une de ces hypothèses. Pour discriminer ces interprétations, il sera utile par exemple de faire passer en condition cabine des musiciens de niveau comparable à ceux enregistrés en studio. Si ce groupe de musiciens présente des performances dégradées de façon similaire au groupe non-musicien testé ici, cela attestera qu'il ne s'agit pas d'un effet d'expertise, mais plutôt d'une différence de tâche.

5. Remerciements

Ce travail a été effectué au cours du mois de Juin 2014, dans le laboratoire IRCAM STMS UMR 9912 à Paris, au sein de l'équipe « Perception et Design Sonore » dirigée par Patrick Susini. Le stage a été encadré par Jean-Julien Aucouturier (CNRS), Clément Canonne (Université de Bourgogne) et Guillaume Lemaitre (IRCAM).

L'Ircam (Institut de Recherche et Coordination Acoustique/Musique) est une institution dédiée à la recherche et la création musicale contemporaine, associée au Centre Pompidou.

6. Références

Anshel, A., & Kipper, D. (1988). The influence of group singing on trust and cooperation. *Journal of Music Therapy*, 25, 145-155.

Balkwill, LL and Thompson, W.F. (1999) A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music perception*: 43-64.

Becker, J. (2001). Anthropological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 135-160). New York: Oxford University Press.

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, 460(7256), 737-740.

Cross, I. (2007). Music and cognitive evolution. In R. I. M. Dunbar & L. Barrett (Eds.), *Oxford handbook of evolutionary psychology* (pp. 649-667). Oxford, England: Oxford University Press.

Dunbar, R. I. M. (2003). The Social Brain : Mind, Language, and Society in Evolutionary Perspective. *Annual Review of Anthropology*, Vol. 32 (2003), pp. 163-181

Eggebrecht, R. (1983). Sprachmelodie und musikalische Forschungen im Kulturvergleich [Speech melody and music research in cross-cultural comparison]. Doctoral dissertation, University of Munich, Munich, Germany.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6, 169-200.

Fitch, W. T. (2005). The Evolution of Language: A Comparative Review. *Biology and Philosophy*, 20(2-3):193-203.

Gagnon, L. and Peretz, I. (2003) "Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies." *Cognition & Emotion* 17.1, 25-40

Gregory, A. H. (1997). The roles of music in society: The ethnomusico- logical perspective. In D. J. Hargreaves & A. C. North (Eds.), *The social psychology of music* (pp. 123-140). Oxford, England: Oxford University Press.

Honing, H., & Ploeger, A. (2012) *Cognition and the Evolution of Music: Pitfalls and Prospects*. *Topics in Cognitive Science*, 1-12.

Juslin, P. N., & Laukka, P. (2003). Communication of Emotions in Vocal Expression and Music Performance : Different Channels, Same Code ? *Psychological Bulletin*, Vol. 129, No. 5, 770-814.

Keltner, D., & Haidt, J. (1999). Social Functions of Emotions at Four Levels of Analysis. *Cognition and Emotion*, 1999, 13 (5), 505-521.

Kirschner, S. & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior* 31 (2010) 354-364.

Lewontin, R. C. (1998). The evolution of cognition: Questions we will never answer. In D. Scarborough & S. Sternberg (Eds.), *Methods, models, and conceptual issues: An invitation to cognitive science*, Vol. 4(pp. 107-132). Cambridge, MA: MIT Press.

Maus, F., E. (1988). Music as Drama. *Music Theory Spectrum*, Vol. 10, 10th Anniversary Issue (Spring, 1988), pp. 56-73.

Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315-327). Cambridge, MA: MIT Press.

Nayak, S., Wheeler, B. L., Shiflett, S. C., & Agostinelli, S. (2000). Effect of Music Therapy on Mood and Social Interaction among individuals with acute traumatic brain injury and stroke. *Rehabilitation psychology*, Vol. 45, No. 3, 274 - 283.

Rosenthal, R., & Rubin, D. B. (1989). Effect size estimation for one-sample multiple-choice-type data: Design, analysis, and meta-analysis. *Psychological Bulletin*, 106(2), 332.

Thomas, D. W., Heitman, R. J., & Alexander T. (1994). The Effects of Music on Bathing Cooperation for Residents with Dementia. *Journal of Music Therapy*, Vol. 34, No. 4, 246 - 259.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, 17(1), 3-28.

Wichmann, A.(2000): "The attitudinal effects of prosody, and how they relate to emotion", In SpeechEmotion-2000, 143-148.

Wiltermuth, S.S., & Heath, C. (2008). Synchrony and Cooperation. Psychological Science. Vol 20, No. 1.

7. Annexes

Figure 3 : Instructions préliminaires à la passation de l'expérience

Instructions

Vous allez écouter une suite de 25 morceaux de musique, d'une durée d'une ou deux minutes chacun. Ces morceaux sont des duos, mettant en scène deux instruments différents (par exemple piano et saxophone) : vous entendrez l'un de ces instruments dans votre oreille gauche, et l'autre dans votre oreille droite.

Pour chacun de ces duos, nous vous demandons d'indiquer quelle vous semble être **l'attitude de l'instrument que vous entendez à gauche par rapport à l'instrument que vous entendez à droite, la manière dont l'instrument de gauche vous semble se comporter avec l'instrument de droite**. Il est souhaitable d'avoir écouté le morceau jusqu'au bout avant d'entrer votre réponse et vous pouvez écouter chaque morceau autant de fois que vous le souhaitez.

Instructions (2)

Vous aurez à choisir votre réponse parmi cinq attitudes possibles : autoritaire, conciliant, dédaigneux, insolent et prévenant. Il est indispensable de sélectionner une (et une seule) de ces attitudes avant de pouvoir passer au morceau suivant.

Chacune de ces attitudes est définie précisément dans le document que nous vous avons distribué. **Il est impératif que vous lisiez attentivement ces définitions avant de commencer l'expérience** ; vous pourrez toujours vous reporter à ces définitions en cas de doute. Par ailleurs, vous disposez d'une série de cinq tableaux destinés à vous aider à mieux visualiser le sens de chacune de ces attitudes, et les situations qui peuvent lui être associées. Là encore, n'hésitez pas à les consulter lors de l'expérience, par exemple avant d'entrer votre réponse.

Instructions (3)

Une dernière chose : les instruments présents sur chaque morceau seront indiqués à chaque fois. Mais vous entendrez parfois des sons « étranges », que vous n'associez peut-être pas à l'instrument indiqué : le piano pourra parfois sonner comme des percussions, le violon ou l'alto comme de la guitare, la flûte comme de la voix, etc... Pas d'inquiétude ! Tous les sons que vous entendrez du même côté proviennent bien du même instrument, et il s'agit bien à chaque fois de l'instrument indiqué, même si les musiciens utilisent parfois des techniques de jeu peu habituelles. L'essentiel est de bien faire attention au rapport entre les sons que vous entendez à gauche et les sons que vous entendez à droite pour essayer d'en déduire l'attitude adoptée par l'instrument de gauche.