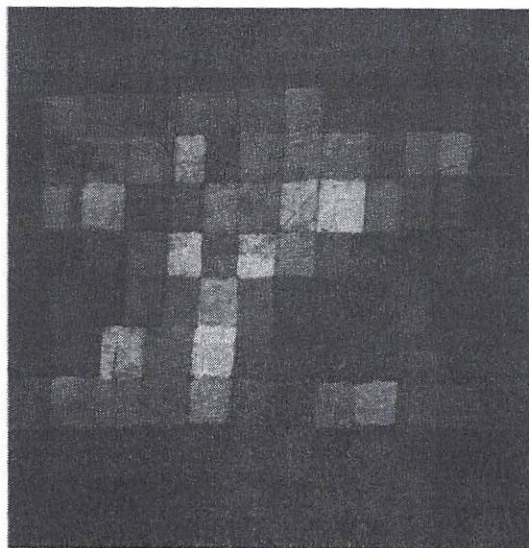
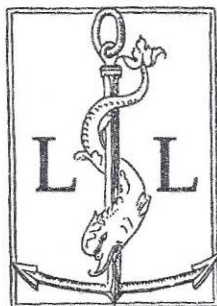


Sciences du langage et neurosciences

Actes du colloque 2015
de l'Association des Sciences du Langage



Textes réunis et présentés par Alain Rabatel,
Malika Temmar et Jean-Marc Leblanc



Lambert-Lucas

Les structures hiérarchiques linguistiques et musicales partagent-elles le même substrat neural ? Études comportementales et d'imagerie cérébrale fonctionnelle en lien avec l'expertise musicale

Élodie Cauvet ^{1,2}, Pierre Pica ³, Stanislas Dehaene ¹, Christophe Pallier ¹

¹ INSERM U992 Unité de Neuroimagerie Cognitive, Neurospin, Gif-sur-Yvette, France

² Center of Neurodevelopmental Disorders at Karolinska Institute, Stockholm, Sweden

³ UMR 7023, CNRS Université Paris VIII, Paris, France

Introduction

Structure syntaxique en langage et en musique

Les langues humaines peuvent être décrites comme des systèmes combinatoires hiérarchiques. À partir d'un ensemble fini d'éléments, les mots, un nombre infini de séquences, les phrases, peuvent être créés en suivant un ensemble de règles syntaxiques. D'autres domaines comme la planification de l'action, les mathématiques, la danse ou encore la musique peuvent également être décrits en utilisant des structures hiérarchiques syntaxiques.

Depuis l'apparition des théories génératives linguistiques dans les années 1960 (Chomsky 1957), certains musiciens ont cherché à décrire la musique et sa structure en utilisant les mêmes outils formels. Les premiers à formaliser une théorie générative de la musique occidentale tonale ont été le compositeur Lerdahl et le linguiste Jackendoff. Leur ouvrage fondateur (Lerdahl & Jackendoff 1983) définit les unités structurelles de la musique et fournit une description des règles gouvernant leurs combinaisons possibles. Les unités de base sont des notes et des accords. D'après Lerdahl et Jackendoff, la structure de surface musicale d'une pièce peut être décrite en utilisant des règles de construction "*well-formedness rules*". Il est possible de décomposer ces règles en quatre sous-ensembles interagissant pour former la structure entendue par l'auditeur. Ces quatre composantes structurales ¹ permettent d'obtenir une structure hiérarchique cons-

1. Ces quatre ensembles de règles sont appelées structure de groupe, structure métrique, réduction sur la durée et réduction *prolongationnelle*. Elles permettent d'incorporer la description des différents aspects musicaux comme le tempo, le rythme, la tonalité, l'harmonie, etc.

truite par l'auditeur que l'on peut représenter sous forme d'arbre. À la différence du langage, ces règles syntaxiques seraient soumises à une pondération grâce à des règles de préférences dont le but est d'obtenir une structure maximale stable et perçue comme correcte. La relation qu'entretiennent les règles les unes par rapport aux autres dépend du morceau et de l'auditeur qui pourra favoriser une analyse syntaxique particulière de la pièce. Même si tous les auditeurs d'un morceau de musique ne l'interprètent pas de la même façon, il existe, le plus souvent, une structure hiérarchique partagée par la majorité des auditeurs, comme en langage. Cependant, musique et langage sont deux domaines cognitifs bien distincts. Une première différence importante avec le langage provient de l'existence de catégories syntaxiques (nom, verbe, adjectif, etc.) qu'on ne retrouve pas en musique. Une seconde différence notable entre les deux domaines provient du fait qu'il existe un référent externe liant le signifiant (le mot) et le signifié (le sens), en langage alors que les éléments musicaux en sont dépourvus. Ainsi, aucune condition de vérité ne peut être associée à une phrase musicale au contraire d'une phrase linguistique.

Malgré ces différences et à mesure que la recherche avance en linguistique, les chercheurs trouvent de plus en plus de similitudes dans l'organisation structurale de ces deux domaines. Récemment, deux linguistes ont même proposé la « thèse de l'identité » (Katz & Pesetsky *submitted*). Selon cette thèse énoncée dans le cadre de la grammaire générative, « il y aurait identité entre les descriptions syntaxiques du langage et de la musique ». Dans leur article, ils appliquent les règles de la syntaxe linguistique aux unités de base de la musique pour en décrire la structure. Ils soutiennent que

Toutes les différences formelles entre le langage et la musique sont la conséquence des différences dans les unités structurales à savoir des paires arbitraires entre sons et sens pour le langage et des paires entre classes de hauteurs de notes et leurs combinaisons pour la musique. Autrement, le langage et la musique sont identiques en tous points. (Katz & Pesetsky *submitted*)

Dans la même optique, Martin Rohrmeier (2011) a développé une grammaire syntagmatique pour décrire la musique tonale.

Les similarités formelles entre musique et langage, comme l'existence de structure syntaxique hiérarchique, posent la question de partage d'un même système computationnel au niveau neuronal. Selon l'hypothèse de partage des ressources syntaxiques – *shared syntactic integration resource hypothesis* développée par Patel (2003 et 2008) – les réseaux neuronaux intégrant la structure linguistique et musicale seraient au moins en partie partagés. L'existence et la localisation de

ce potentiel réseau partagé font actuellement l'objet de débat.

Examiner le traitement cérébral des structures syntaxiques en langage et en musique

Comparer les traitements syntaxiques dans différents domaines comme la musique et le langage nécessite des dessins expérimentaux équivalents. Deux grands paradigmes ont été principalement utilisés à la fois en psychologie comportementale et en neuroimagerie :

1. les violations syntaxiques
2. le contraste entre des phrases ou des morceaux de musique bien formés et des listes de mots ou d'accords déstructurées.

Le paradigme de violation syntaxique compare des phrases contenant des violations syntaxiques comme **Les arbres peuvent grandi*, avec des phrases correctes *Les arbres peuvent grandir* (par exemple Ni 2000, Musso 2015). De façon similaire en musique, les paradigmes de violations syntaxiques comparent des séquences d'accords finissant sur un accord régulier, comme la tonique qui est l'accord le plus régulier et le plus attendu en position finale selon la théorie de l'harmonie (Piston 1987, Schoenberg 1969), avec des séquences d'accords terminant sur un accord inattendu et syntaxiquement illégal comme un accord majeur sur le second degré diminué qui est parfaitement incorrect en position finale. Au niveau cérébral, les violations syntaxiques en langage, mesurées par électroencéphalogramme (EEG), provoquent une négativité antérieure gauche précoce (Friederici 2002), tandis que les violations musicales provoquent une négativité antérieure précoce dans l'hémisphère droit (Koelsch 2003 et 2005). De plus, les violations syntaxiques en musique et en langage provoquent une réponse positive en EEG, dite P600, indistinguishable chez un même sujet (Patel 1998). La combinaison des violations syntaxiques linguistiques et musicales simultanées semble entrer en compétition pour les mêmes ressources neuronales (Koelsch 2005, Slevc 2009, Fedorenko 2009). Cependant, les paradigmes de violations reflètent, de par leur nature, plusieurs processus cérébraux différents difficiles à démêler : une première analyse automatique de la phrase, puis une détection d'erreur et enfin une tentative de réparation de la structure syntaxique. Il est donc difficile de déduire de ce type de paradigme quels processus sont partagés, et en particulier s'il s'agit de la première analyse automatique qui nous intéresse dans l'étude présente.

Le second grand type de paradigme compare des listes de mots, « chose arbre que signaler elles instruments regardant tendu », (Humphries 2005 et 2006, Rogalsky 2011), ou des mélodies déstructurées (Levitin 2003, Abrams 2011) avec des phrases syntaxiquement

correctes, « *je pense que tu devrais accepter la proposition de ton nouvel associé* », ou de mélodies intactes. De tels paradigmes linguistiques ont été utilisés en Topographie par Emission de Positron (TEP) (Mazoyer 1993) et en Imagerie par Résonance Magnétique fonctionnelle (IRMf) (Humphries 2005, 2006, Snijders 2009, Pallier 2011) et adaptés en musique (Levitin 2003, Abrams 2011). Cependant, ces paradigmes ne dissocient pas structure syntaxique et structure sémantique. En effet, les séparer est chose ardue puisque les deux sont étroitement liées : les *items* lexicaux possèdent un sens. Des stratégies ont donc été employées comme l'utilisation de pseudo-mots pour remplacer les mots de contenu (ces stimuli sont souvent appelés du *Jabberwocky* en référence à un poème de Lewis Carroll où ce dernier avait remplacé les mots de contenu par des pseudo-mots aux consonances lugubres). Ces stratégies qui diminuent la charge sémantique ont mis en évidence un réseau de régions cérébrales inféro-frontal et temporal à gauche, répondant spécifiquement aux processus syntaxiques linguistiques (Humphries 2006, Pallier 2011).

Cependant, le calcul de structure syntaxique est un procédé dynamique qui consiste à intégrer successivement chacun des mots jusqu'à obtenir une structure entière, un percept complet. La comparaison de listes de mots et de phrases ne donne donc qu'une information partielle quant à la dynamique de ce processus. Un dessin expérimental paramétrique permet d'accéder en temps réel au processus de construction de la structure syntaxique : chaque nouvelle unité, mot ou note de musique, est intégrée progressivement dans la structure syntaxique préexistante. Par exemple, Pallier *et alii* en 2011, ont fait varier paramétriquement la quantité de structure syntaxique en langage. Des phrases correctes ont été coupées en constituants syntaxiques de taille variable, d'un mot à douze mots, et ce en français mais aussi en *Jabberwocky*. Sept régions latéralisées à gauche (pôle temporal, sillon temporal supérieur antérieur et postérieur, jonction temporo-parietale, gyrus frontal inférieur triangularis et orbitalis et putamen) étaient de plus en plus activées à mesure que la taille des constituants augmentait. Parmi ces sept régions, seules quatre ont répondu de la même façon à la manipulation syntaxique en *Jabberwocky* à savoir le gyrus frontal inférieur orbitalis, triangularis, la partie postérieure du sillon temporal supérieur et le putamen. Ces régions semblent donc sensibles à la structure syntaxique indépendamment de la sémantique et constituent donc des candidats crédibles pour un réseau syntaxique indépendant du domaine. De plus, ces régions coïncident avec les régions rapportées dans d'autres études portant sur la structure musicale (Levitin 2003, Maess 2001). À ce jour, il existe

très peu d'études comparant chez les mêmes sujets la structure linguistique et musicale (Abrams 2011, Rogalsky 2011, Musso 2015, Kunert 2015, Norman-Haignere 2015, Fedorenko 2012). Les résultats varient selon les stimuli et les méthodes d'analyses employées. La moitié montre que musique et langage recrutent des populations neuronales différentes pour le traitement de leur structure notamment structure temporelle fine, notamment dans les régions inféro-frontales et temporales mentionnées ci-dessus (Abrams 2011, Fedorenko 2012, Norman-Haignere 2015). L'autre moitié montre qu'une de ces mêmes régions serait commune au traitement syntaxique linguistique et musical (Kunert 2015, Musso 2015). La disparité des acceptions du terme « structure » dans les expériences mentionnées pourrait expliquer leurs différents résultats. Il est donc crucial de définir le type de structure testée. Dans cette expérience, nous étudions les processus cérébraux conduisant à la construction d'une structure hiérarchique en musique et en langage.

Effet de l'expertise dans le traitement de la structure syntaxique

Aucune étude ne teste l'effet de l'expertise musicale sur les processus syntaxiques neuronaux à la fois en langage et en musique. Cependant, il existe une différence cruciale d'expertise entre ces deux domaines qui pourraient affecter leur traitement neuronal. En effet, les êtres humains sont des experts en langage, auquel ils sont exposés continuellement et ce même avant leur naissance. L'exposition à la musique, elle, varie grandement selon les individus au sein du même socle culturel, mais plus encore si on considère différentes cultures. Le rôle et l'importance de la musique diffèrent à travers le monde. Dans la société occidentale, l'exposition passive à la musique tonale peut être considérée comme importante. Par contre, les connaissances explicites varient beaucoup – de l'exposition passive pour les non-musiciens jusqu'à la connaissance des règles syntaxiques du rythme, de la mélodie, de l'harmonie, etc., pour les musiciens. À l'opposé du même spectre de connaissances et d'exposition à la musique, on trouve les Mundurucus, un peuple indigène vivant de façon isolée dans l'Amazonie brésilienne. Chez les Mundurucus, la musique est rarement utilisée et a pour but principal d'effrayer les esprits et les animaux. La langue Mundurucu a un seul mot correspondant à musique et chant ; le vocabulaire musical est quasi inexistant. Il n'y a pas de radio et le peu de musique occidentale est entendu durant les messes dans les missions religieuses. En conséquence, la comparaison entre musiciens occidentaux experts, non-musiciens occidentaux et Mundurucus permet de couvrir le spectre d'expertise musicale et d'explorer le traitement des structures syntaxiques avec un regard

particulier sur son caractère inné et automatique.

En effet, nous savons que l'expertise façonne le cerveau humain aussi bien en langage qu'en musique. L'acquisition d'une nouvelle langue implique un réseau cérébral différent de celui utilisé pour traiter la langue maternelle. Par exemple, plus une personne devient experte dans une seconde langue, plus les réseaux cérébraux utilisés pour les deux langues vont être semblables (Perani 1998). En musique, l'expertise musicale modifie l'anatomie cérébrale (Bermudez 2009), en particulier dans les régions auditives (Meyer 2012), les régions motrices (Gaser 2003), mais aussi dans son fonctionnement (Ellis 2012, 2013, Bever & Chiarello 1974, Ohnishi 2001). Des études ont montré que des non-musiciens occidentaux percevaient les relations hiérarchiques entre les accords dans la musique, à savoir la tonalité (Krumhansl 1982, Bharucha 1983, Tillmann 2000), et que leurs réponses cérébrales étaient différentes lors de violations (Koelsch 2000). Cependant, la structure musicale ne se résume pas seulement à la tonalité (Lerdahl & Jackendoff 1983). Des non-musiciens sont-ils capables de percevoir la structure et sa déstructuration progressive en absence d'information tonale ? Cette capacité est-elle liée à l'expertise acquise par exposition passive ?

Expériences présentes comparant la perception et le traitement cérébral de la structure syntaxique en langage et en musique en tenant compte de l'expertise des participants

Dans cet article, nous décrivons une série d'expériences testant la perception et le traitement cérébral de la structure syntaxique musicale et linguistique, de façon symétrique avec une attention particulière sur l'effet de l'expertise musicale. Pour refléter le processus dynamique de construction syntaxique, la taille des domaines syntaxiquement cohérents varie paramétriquement dans nos stimuli alors que leur longueur globale reste constante. Le premier ensemble de trois expériences comportementales compare la perception et le traitement de la structure en étudiant l'effet de l'exposition passive à la musique. Deux tâches comportementales comparent Mundurucus et non-musiciens occidentaux quant à leur perception de la structure musicale. La première expérience teste le traitement implicite de la structure musicale en demandant si deux extraits sont différents ou identiques. Dans une deuxième expérience, les participants devaient noter la qualité des extraits musicaux. Cette dernière tâche est finalement répétée dans une troisième expérience incluant stimuli musicaux et linguistiques comparant la perception de la structure dans les deux domaines par des non-musiciens occidentaux. Enfin, une expérience d'imagerie fonctionnelle utilise ces mêmes stimuli linguistiques et

musicaux pour comparer leurs réponses neurales chez des musiciens et des non-musiciens occidentaux.

1. Effet de l'exposition sur le traitement de la structure musicale et comparaison avec le langage : expériences comportementales

Dans le but de tester l'effet de l'exposition passive sur le traitement de la structure musicale, deux groupes d'individus non experts différant selon leur degré d'exposition à la musique tonale occidentale ont été testés : des Mundurucus avec une exposition passive extrêmement minimale et des occidentaux non-musiciens sans éducation musicale formelle.



Figure 1. Schéma des stimuli musicaux. La longueur totale est constante à travers toutes les conditions mais la durée sur laquelle un percept cohérent peut être construit varie.

La structure en musique tonale occidentale repose sur des indices de tempo, de rythme, de mélodie et d'harmonie. Dans cet ensemble d'expériences, tous les indices structuraux musicaux ont été manipulés et en particulier l'effet de la tonalité. Notre manipulation principale de la structure musicale porte sur la taille des fragments intégrables dans une structure cohérente. Les stimuli sont donc de longueurs identiques mais la durée sur laquelle un percept est cohérent varie paramétriquement depuis seize temps, huit temps, quatre temps, deux temps et un

temps, correspondant ainsi à cinq conditions de structure différentes. La figure 1 présente un exemple de stimuli musicaux. De plus, la tonalité est un élément crucial de ce type de musique. Elle correspond à l'échelle de sons sur laquelle les notes sont organisées de façon hiérarchique. Les notes appartenant à la même tonalité, la même gamme, s'intègrent plus facilement les unes avec les autres par rapport aux notes en dehors de celle-ci. On peut ainsi prédire que nos stimuli contenant des notes issues de tonalités différentes seront perçus comme plus dissemblables que des stimuli du même niveau de structure mais contenant uniquement des notes de la même tonalité.

La perception implicite et explicite de la structure est testée en utilisant deux tâches : une de jugement explicite de cohérence et une de discrimination. L'expérience de discrimination examine l'impact de la structure musicale sur la capacité à mémoriser les mélodies. Les participants devaient juger si deux extraits musicaux présentés successivement étaient identiques ou différents. Si notre manipulation de structure est perçue, nous prédisons que plus les stimuli seront déstructurés, plus il sera difficile de les comparer. L'expérience de notation teste l'évaluation explicite de la quantité de structure perçue dans les stimuli. Les sujets étaient d'abord entraînés avec des exemples de phrases dans leur langue maternelle (français ou mundurucu) plus ou moins structurées. On leur expliquait que les phrases normales étaient associées avec une bonne note (représentée par un visage souriant) et que les phrases déstructurées étaient mauvaises (représentées par un visage triste). L'expérience commençait alors et les participants entendaient chaque extrait musical après lequel ils devaient noter à quel point ils étaient corrects. Ils répondaient soit pour les Mundurucus en pointant un visage plus ou moins souriant sur une échelle de cinq, soit pour les sujets français en cliquant sur le visage. Finalement, la dernière expérience permet de comparer la perception de la structure des stimuli musicaux et linguistiques en testant un ensemble indépendant de sujets occidentaux non-musiciens qui devaient noter leur structure. Dans cette dernière expérience, seuls les stimuli musicaux provenant de tonalités différentes sont utilisés dans un souci de maximiser l'effet de déstructuration tout en conservant une durée de test acceptable. Les stimuli linguistiques ont été créés en suivant le même modèle que les stimuli musicaux toujours de longueur constante mais dont la quantité de structure varie de seize mots, huit mots, quatre mots, deux mots et un mot. Le tableau 1 présente un exemple de stimuli linguistiques.

Condition	Taille constituant	Exemples
C16	16 mots	Mon père parle souvent de choses très savantes qui fascinent les copains qui viennent chez moi
C08	8 mots	Elle couvre la marmite qui réchauffe tout doucement / il passe à côté du parc du manoir
C04	4 mots	Il lance la balle / les grimaces des clowns / la biche sort vite / qui sent les fleurs
C02	2 mots	Il manque / est tombé / les fleurs / avec soin / très raide / du Brésil / le blé / les femmes
C01	1 mot	La / à / sortent / tarte / le / un / grimpe / Luc / les / qui / siècle / pains / qui / de / réchauffe / jaunes

Tableau 1. Exemple de stimuli linguistiques pour chaque condition. La longueur totale est toujours constante mais la taille sur laquelle un percept structuré peut être construit varie paramétriquement. Les phrases cohérentes de seize mots ont été prononcées mot par mot avec une pause égale entre chacun des mots afin de permettre le découpage en constituants de tailles variables qui sont recombinaés différemment selon les conditions. Dans toutes les conditions, les mots sont espacés du même temps de pause.

1.1 Résultats de la Tâche de discrimination

L'indice de discriminabilité, mesuré par le d' , a été calculé en fonction de la cohérence des stimuli. La figure 2 montre la discriminabilité moyenne en fonction du niveau de structure en séparant Mundurucus des occidentaux non musiciens. Chaque ligne représente l'indice de discriminabilité d'un participant. L'analyse statistique (Analyse de Variance avec les facteurs : structure, groupe et tonalité) montre que les sujets ont répondu différemment pour les différents niveaux de structure. De plus, les Mundurucus et les non-musiciens occidentaux discriminent les stimuli différemment : les Mundurucus font globalement plus d'erreurs que les non-musiciens occidentaux. Enfin, il est plus facile de discriminer des fragments contenant des tonalités différentes. En particulier, on notera que l'amélioration de la discriminabilité avec la cohérence de stimuli est présente dans les deux groupes de sujets sans différence significative.

Des analyses *post hoc* séparant les deux groupes de sujets montrent que pour chacun des groupes, plus les stimuli sont structurés, plus il est facile de discriminer les deux extraits présentés : leur discriminabilité augmente. L'analyse des temps de réaction, col-

lectés pour les occidentaux non musiciens, montre que les participants ont été de plus en plus rapides à comparer des stimuli de plus en plus structurés. L'absence d'indice de tonalité n'a par contre pas influencé significativement le temps nécessaire à la comparaison des stimuli.

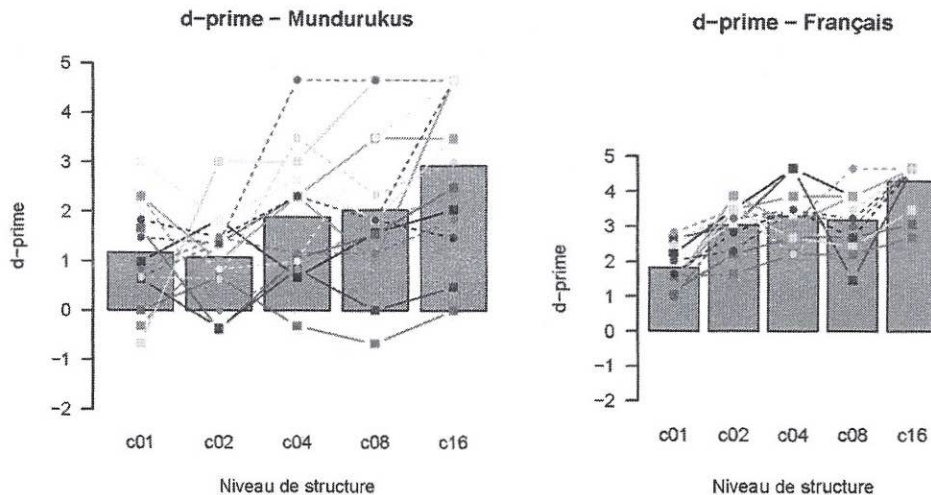


Figure 2. Indice de discriminabilité des deux extraits incluant tous les stimuli présentés. Zéro reflète le niveau du hasard, plus le score est haut, plus les extraits ont été discriminés correctement.

La tâche de discrimination a donc été correctement réalisée y compris par les Mundurucus (seuls 2 participants de ce groupe sont au niveau du hasard). Elle nous permet de mettre en évidence le traitement automatique de la structure musicale de nos stimuli puisque ce paramètre influence une tâche qui ne requiert pas de jugement explicite de la structure. Plus les stimuli sont structurés, plus il est plus facile de les encoder en mémoire en vue de les comparer et ce quelle que soit l'exposition. L'absence d'indices de tonalité aux frontières entre les fragments semble avoir eu un léger impact sur cette tâche, rendant un peu plus difficile la comparaison des niveaux les moins structurés puisque cet effet est marginal dans chacun des groupes.

1.2 Résultats de la Tâche d'Évaluation de la structure

Les notes moyennes données aux stimuli sont représentées figure 3. Chaque ligne représente un individu. L'analyse statistique (identique à l'expérience précédente) révèle que plus les stimuli sont déstructurés, plus les notes données sont basses pour les deux groupes même si l'amplitude de l'effet diffère entre les deux : la relation entre le niveau de structure et le jugement des sujets est plus forte pour les participants français.

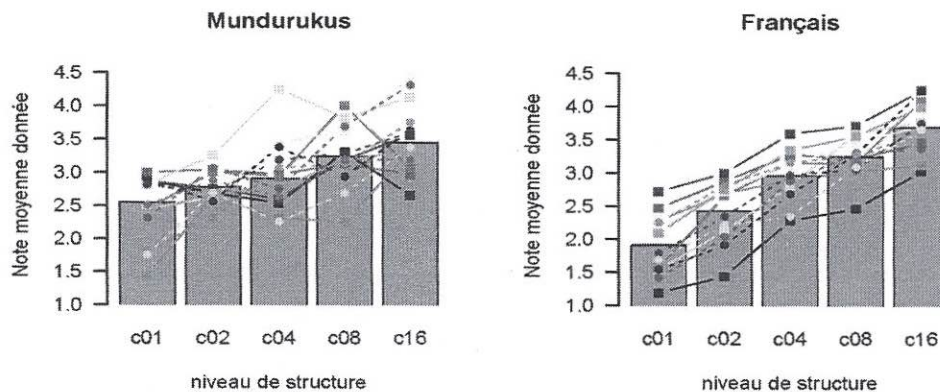


Figure 3. Note moyenne donnée en fonction du niveau de structure.

De plus, les notes données diffèrent selon que les fragments sont tous dans la même tonalité ou qu'ils sont dans des tonalités différentes : les notes données pour les stimuli issus de tonalités différentes sont meilleures que celles toutes issues de la même tonalité. Il semble que l'évaluation ait été rendue plus difficile par l'absence de changement de tonalité. En comparant les notes données aux stimuli issus de la même tonalité aux réponses données aux stimuli avec un changement de tonalité entre les fragments, nous pouvons déduire que le changement de tonalité a aidé à la perception de structure mais qu'elle n'est pas la seule responsable puisque l'effet est toujours visible en ayant ôté ce paramètre. Les Mundurucus, comme les Français, sont capables de noter le degré de structure d'un morceau de musique classique occidentale. Cela est remarquable dans la mesure où leur exposition à la musique tonale est minimale voire nulle.

1.3 Résultats de la Tâche d'évaluation de la structure : Comparaison musique et langage

La figure 4 présente la moyenne des notes données par sujet en fonction du niveau de structure construit.

Chaque ligne représente un sujet. L'analyse statistique (anova avec facteurs domaine et structure) montre que les sujets notent et perçoivent différemment les niveaux de structure en musique et en langage. De plus, la structure est perçue différemment en langage et en musique : la pente de la courbe plus grande pour le langage montre que la déstructuration des phrases en langage est perçue de façon plus prononcée. Cela revient à dire que pour un même niveau de structure, la note donnée va être plus basse pour le langage que la musique. Ce phénomène est d'autant plus prononcé que les stimuli sont déstructurés. Les sujets sont plus « choqués » par la déstructuration en

langage qu'en musique. Les stimuli langagiers portent un sens qui va être perdu assez rapidement avec la déstructuration, la cohérence perçue est influencée par des facteurs sémantiques et pas seulement syntaxiques. En musique, la structure vient non seulement de l'organisation temporelle mais aussi du rapport entre les différentes hauteurs des notes ; on ne peut pas assigner un sens aux notes.

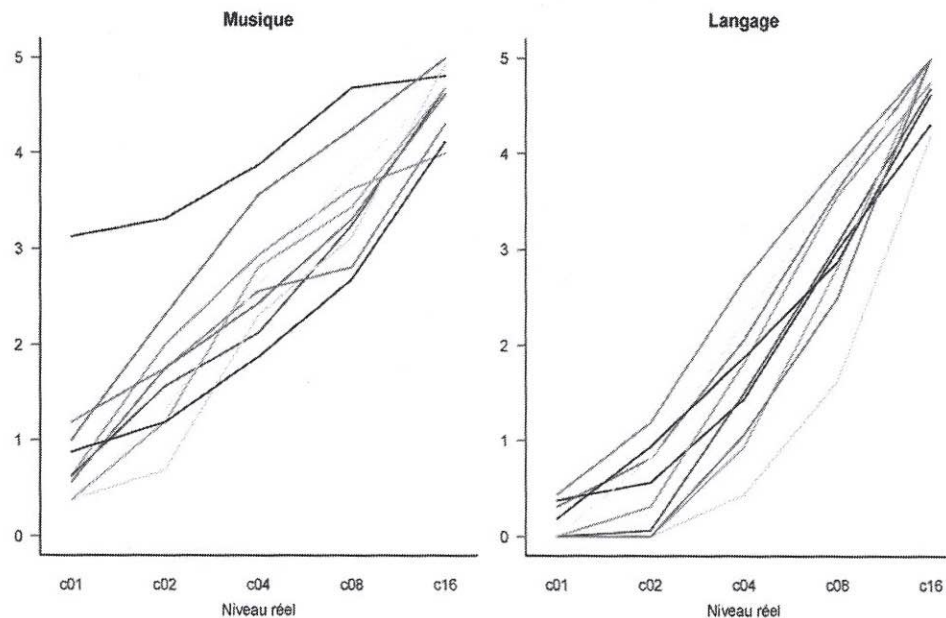


Figure 4. Notes moyennes données en fonction du niveau de structure par des non-musiciens occidentaux. Stimuli musicaux et linguistiques.

2. Traitement cérébral de la structure musicale et comparaison avec le langage : comparaison entre musiciens et non-musiciens

Cette étude de neuroimagerie utilise les mêmes stimuli linguistiques et musicaux que les expériences comportementales précédentes. Cette étude en IRM fonctionnelle cherche à identifier les régions cérébrales impliquées dans le traitement de la structure syntaxique en langage et en musique. Nous avons comparé des musiciens et des non-musiciens, qui ont tous eu une exposition passive importante à la musique tonale occidentale depuis l'enfance.

Afin de s'assurer que les participants prêtaient bien attention aux stimuli, on leur demandait de détecter un changement d'instrument (clavecin à la place de piano) pour les stimuli musicaux et un changement de voix (masculine à la place de féminine) pour les stimuli linguistiques. Cette tâche nous permet de tester le traitement cérébral automatique de la structure syntaxique sans attirer l'attention explicitement sur la structure.

2.1 Réponses globales au langage et à la musique

Les aires cérébrales activées par l'écoute globale de tous les stimuli linguistiques et musicaux sont bilatérales et correspondent aux aires classiques périsylviennes (gyrus supérieur temporal, gyrus frontal inférieur), le thalamus et l'aire supplémentaire motrice. La comparaison entre tous les stimuli linguistiques et musicaux montre des réseaux similaires bien que les réponses pour le langage soient plus importantes que la musique dans les régions temporales bilatérales et dans le gyrus inférieur frontal à gauche. Aucune région n'est plus activée par l'écoute globale de la musique que par le langage.

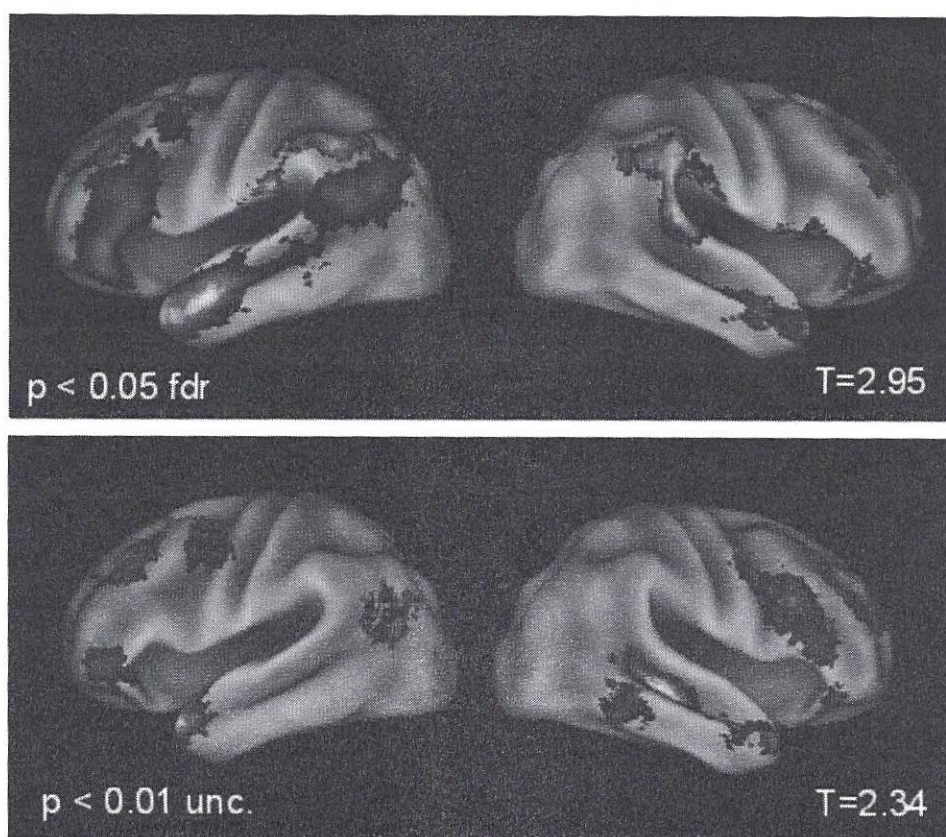


Figure 5. Carte d'activation cérébrale en fonction de la structure en langage à gauche, et de la structure en musique à droite. En grisé les activations et désactivations qui ne sont pas discutées ici.

2.2 Effet de structure

Nous avons cherché les régions dont l'activité cérébrale augmentait avec la taille des constituants syntaxiques en langage, figure 5. Un réseau de régions de l'hémisphère gauche est mis en évidence incluant le pôle temporal, la région antérieure du sillon temporal supérieur, la

Les structures hiérarchiques linguistiques et musicales partagent-elles le même substrat neural ? Études comportementales et d'imagerie cérébrale fonctionnelle en lien avec l'expertise musicale

Élodie Cauvet ^{1,2}, Pierre Pica ³, Stanislas Dehaene ¹, Christophe Pallier ¹

¹ INSERM U992 Unité de Neuroimagerie Cognitive, Neurospin, Gif-sur-Yvette, France

² Center of Neurodevelopmental Disorders at Karolinska Institute, Stockholm, Sweden

³ UMR 7023, CNRS Université Paris VIII, Paris, France

Introduction

Structure syntaxique en langage et en musique

Les langues humaines peuvent être décrites comme des systèmes combinatoires hiérarchiques. À partir d'un ensemble fini d'éléments, les mots, un nombre infini de séquences, les phrases, peuvent être créés en suivant un ensemble de règles syntaxiques. D'autres domaines comme la planification de l'action, les mathématiques, la danse ou encore la musique peuvent également être décrits en utilisant des structures hiérarchiques syntaxiques.

Depuis l'apparition des théories génératives linguistiques dans les années 1960 (Chomsky 1957), certains musiciens ont cherché à décrire la musique et sa structure en utilisant les mêmes outils formels. Les premiers à formaliser une théorie générative de la musique occidentale tonale ont été le compositeur Lerdahl et le linguiste Jackendoff. Leur ouvrage fondateur (Lerdahl & Jackendoff 1983) définit les unités structurelles de la musique et fournit une description des règles gouvernant leurs combinaisons possibles. Les unités de base sont des notes et des accords. D'après Lerdahl et Jackendoff, la structure de surface musicale d'une pièce peut être décrite en utilisant des règles de construction "*well-formedness rules*". Il est possible de décomposer ces règles en quatre sous-ensembles interagissant pour former la structure entendue par l'auditeur. Ces quatre composantes structurales ¹ permettent d'obtenir une structure hiérarchique cons-

1. Ces quatre ensembles de règles sont appelées structure de groupe, structure métrique, réduction sur la durée et réduction *prolongationnelle*. Elles permettent d'incorporer la description des différents aspects musicaux comme le tempo, le rythme, la tonalité, l'harmonie, etc.

truite par l'auditeur que l'on peut représenter sous forme d'arbre. À la différence du langage, ces règles syntaxiques seraient soumises à une pondération grâce à des règles de préférences dont le but est d'obtenir une structure maximale stable et perçue comme correcte. La relation qu'entretiennent les règles les unes par rapport aux autres dépend du morceau et de l'auditeur qui pourra favoriser une analyse syntaxique particulière de la pièce. Même si tous les auditeurs d'un morceau de musique ne l'interprètent pas de la même façon, il existe, le plus souvent, une structure hiérarchique partagée par la majorité des auditeurs, comme en langage. Cependant, musique et langage sont deux domaines cognitifs bien distincts. Une première différence importante avec le langage provient de l'existence de catégories syntaxiques (nom, verbe, adjectif, etc.) qu'on ne retrouve pas en musique. Une seconde différence notable entre les deux domaines provient du fait qu'il existe un référent externe liant le signifiant (le mot) et le signifié (le sens), en langage alors que les éléments musicaux en sont dépourvus. Ainsi, aucune condition de vérité ne peut être associée à une phrase musicale au contraire d'une phrase linguistique.

Malgré ces différences et à mesure que la recherche avance en linguistique, les chercheurs trouvent de plus en plus de similitudes dans l'organisation structurale de ces deux domaines. Récemment, deux linguistes ont même proposé la « thèse de l'identité » (Katz & Pesetsky *submitted*). Selon cette thèse énoncée dans le cadre de la grammaire générative, « il y aurait identité entre les descriptions syntaxiques du langage et de la musique ». Dans leur article, ils appliquent les règles de la syntaxe linguistique aux unités de base de la musique pour en décrire la structure. Ils soutiennent que

Toutes les différences formelles entre le langage et la musique sont la conséquence des différences dans les unités structurales à savoir des paires arbitraires entre sons et sens pour le langage et des paires entre classes de hauteurs de notes et leurs combinaisons pour la musique. Autrement, le langage et la musique sont identiques en tous points. (Katz & Pesetsky *submitted*)

Dans la même optique, Martin Rohrmeier (2011) a développé une grammaire syntagmatique pour décrire la musique tonale.

Les similarités formelles entre musique et langage, comme l'existence de structure syntaxique hiérarchique, posent la question de partage d'un même système computationnel au niveau neuronal. Selon l'hypothèse de partage des ressources syntaxiques – *shared syntactic integration resource hypothesis* développée par Patel (2003 et 2008) – les réseaux neuronaux intégrant la structure linguistique et musicale seraient au moins en partie partagés. L'existence et la localisation de

ce potentiel réseau partagé font actuellement l'objet de débat.

Examiner le traitement cérébral des structures syntaxiques en langage et en musique

Comparer les traitements syntaxiques dans différents domaines comme la musique et le langage nécessite des dessins expérimentaux équivalents. Deux grands paradigmes ont été principalement utilisés à la fois en psychologie comportementale et en neuroimagerie :

1. les violations syntaxiques
2. le contraste entre des phrases ou des morceaux de musique bien formés et des listes de mots ou d'accords déstructurées.

Le paradigme de violation syntaxique compare des phrases contenant des violations syntaxiques comme **Les arbres peuvent grandi*, avec des phrases correctes *Les arbres peuvent grandir* (par exemple Ni 2000, Musso 2015). De façon similaire en musique, les paradigmes de violations syntaxiques comparent des séquences d'accords finissant sur un accord régulier, comme la tonique qui est l'accord le plus régulier et le plus attendu en position finale selon la théorie de l'harmonie (Piston 1987, Schoenberg 1969), avec des séquences d'accords terminant sur un accord inattendu et syntaxiquement illégal comme un accord majeur sur le second degré diminué qui est parfaitement incorrect en position finale. Au niveau cérébral, les violations syntaxiques en langage, mesurées par électroencéphalogramme (EEG), provoquent une négativité antérieure gauche précoce (Friederici 2002), tandis que les violations musicales provoquent une négativité antérieure précoce dans l'hémisphère droit (Koelsch 2003 et 2005). De plus, les violations syntaxiques en musique et en langage provoquent une réponse positive en EEG, dite P600, indistinguishable chez un même sujet (Patel 1998). La combinaison des violations syntaxiques linguistiques et musicales simultanées semble entrer en compétition pour les mêmes ressources neuronales (Koelsch 2005, Slevc 2009, Fedorenko 2009). Cependant, les paradigmes de violations reflètent, de par leur nature, plusieurs processus cérébraux différents difficiles à démêler : une première analyse automatique de la phrase, puis une détection d'erreur et enfin une tentative de réparation de la structure syntaxique. Il est donc difficile de déduire de ce type de paradigme quels processus sont partagés, et en particulier s'il s'agit de la première analyse automatique qui nous intéresse dans l'étude présente.

Le second grand type de paradigme compare des listes de mots, « *chose arbre que signaler elles instruments regardant tendu* », (Humphries 2005 et 2006, Rogalsky 2011), ou des mélodies déstructurées (Levitin 2003, Abrams 2011) avec des phrases syntaxiquement

correctes, « *je pense que tu devrais accepter la proposition de ton nouvel associé* », ou de mélodies intactes. De tels paradigmes linguistiques ont été utilisés en Topographie par Emission de Positron (TEP) (Mazoyer 1993) et en Imagerie par Résonance Magnétique fonctionnelle (IRMf) (Humphries 2005, 2006, Snijders 2009, Pallier 2011) et adaptés en musique (Levitin 2003, Abrams 2011). Cependant, ces paradigmes ne dissocient pas structure syntaxique et structure sémantique. En effet, les séparer est chose ardue puisque les deux sont étroitement liées : les *items* lexicaux possèdent un sens. Des stratégies ont donc été employées comme l'utilisation de pseudo-mots pour remplacer les mots de contenu (ces stimuli sont souvent appelés du *Jabberwocky* en référence à un poème de Lewis Carroll où ce dernier avait remplacé les mots de contenu par des pseudo-mots aux consonances lugubres). Ces stratégies qui diminuent la charge sémantique ont mis en évidence un réseau de régions cérébrales inféro-frontal et temporal à gauche, répondant spécifiquement aux processus syntaxiques linguistiques (Humphries 2006, Pallier 2011).

Cependant, le calcul de structure syntaxique est un procédé dynamique qui consiste à intégrer successivement chacun des mots jusqu'à obtenir une structure entière, un percept complet. La comparaison de listes de mots et de phrases ne donne donc qu'une information partielle quant à la dynamique de ce processus. Un dessin expérimental paramétrique permet d'accéder en temps réel au processus de construction de la structure syntaxique : chaque nouvelle unité, mot ou note de musique, est intégrée progressivement dans la structure syntaxique préexistante. Par exemple, Pallier *et alii* en 2011, ont fait varier paramétriquement la quantité de structure syntaxique en langage. Des phrases correctes ont été coupées en constituants syntaxiques de taille variable, d'un mot à douze mots, et ce en français mais aussi en *Jabberwocky*. Sept régions latéralisées à gauche (pôle temporal, sillon temporal supérieur antérieur et postérieur, jonction temporo-parietale, gyrus frontal inférieur triangularis et orbitalis et putamen) étaient de plus en plus activées à mesure que la taille des constituants augmentait. Parmi ces sept régions, seules quatre ont répondu de la même façon à la manipulation syntaxique en *Jabberwocky* à savoir le gyrus frontal inférieur orbitalis, triangularis, la partie postérieure du sillon temporal supérieur et le putamen. Ces régions semblent donc sensibles à la structure syntaxique indépendamment de la sémantique et constituent donc des candidats crédibles pour un réseau syntaxique indépendant du domaine. De plus, ces régions coïncident avec les régions rapportées dans d'autres études portant sur la structure musicale (Levitin 2003, Maess 2001). À ce jour, il existe

très peu d'études comparant chez les mêmes sujets la structure linguistique et musicale (Abrams 2011, Rogalsky 2011, Musso 2015, Kunert 2015, Norman-Haignere 2015, Fedorenko 2012). Les résultats varient selon les stimuli et les méthodes d'analyses employées. La moitié montre que musique et langage recrutent des populations neuronales différentes pour le traitement de leur structure notamment structure temporelle fine, notamment dans les régions inféro-frontales et temporales mentionnées ci-dessus (Abrams 2011, Fedorenko 2012, Norman-Haignere 2015). L'autre moitié montre qu'une de ces mêmes régions serait commune au traitement syntaxique linguistique et musical (Kunert 2015, Musso 2015). La disparité des acceptions du terme « structure » dans les expériences mentionnées pourrait expliquer leurs différents résultats. Il est donc crucial de définir le type de structure testée. Dans cette expérience, nous étudions les processus cérébraux conduisant à la construction d'une structure hiérarchique en musique et en langage.

Effet de l'expertise dans le traitement de la structure syntaxique

Aucune étude ne teste l'effet de l'expertise musicale sur les processus syntaxiques neuronaux à la fois en langage et en musique. Cependant, il existe une différence cruciale d'expertise entre ces deux domaines qui pourraient affecter leur traitement neuronal. En effet, les êtres humains sont des experts en langage, auquel ils sont exposés continuellement et ce même avant leur naissance. L'exposition à la musique, elle, varie grandement selon les individus au sein du même socle culturel, mais plus encore si on considère différentes cultures. Le rôle et l'importance de la musique diffèrent à travers le monde. Dans la société occidentale, l'exposition passive à la musique tonale peut être considérée comme importante. Par contre, les connaissances explicites varient beaucoup – de l'exposition passive pour les non-musiciens jusqu'à la connaissance des règles syntaxiques du rythme, de la mélodie, de l'harmonie, etc., pour les musiciens. À l'opposé du même spectre de connaissances et d'exposition à la musique, on trouve les Mundurucus, un peuple indigène vivant de façon isolée dans l'Amazonie brésilienne. Chez les Mundurucus, la musique est rarement utilisée et a pour but principal d'effrayer les esprits et les animaux. La langue Mundurucu a un seul mot correspondant à musique et chant ; le vocabulaire musical est quasi inexistant. Il n'y a pas de radio et le peu de musique occidentale est entendu durant les messes dans les missions religieuses. En conséquence, la comparaison entre musiciens occidentaux experts, non-musiciens occidentaux et Mundurucus permet de couvrir le spectre d'expertise musicale et d'explorer le traitement des structures syntaxiques avec un regard

particulier sur son caractère inné et automatique.

En effet, nous savons que l'expertise façonne le cerveau humain aussi bien en langage qu'en musique. L'acquisition d'une nouvelle langue implique un réseau cérébral différent de celui utilisé pour traiter la langue maternelle. Par exemple, plus une personne devient experte dans une seconde langue, plus les réseaux cérébraux utilisés pour les deux langues vont être semblables (Perani 1998). En musique, l'expertise musicale modifie l'anatomie cérébrale (Bermudez 2009), en particulier dans les régions auditives (Meyer 2012), les régions motrices (Gaser 2003), mais aussi dans son fonctionnement (Ellis 2012, 2013, Bever & Chiarello 1974, Ohnishi 2001). Des études ont montré que des non-musiciens occidentaux percevaient les relations hiérarchiques entre les accords dans la musique, à savoir la tonalité (Krumhansl 1982, Bharucha 1983, Tillmann 2000), et que leurs réponses cérébrales étaient différentes lors de violations (Koelsch 2000). Cependant, la structure musicale ne se résume pas seulement à la tonalité (Lerdahl & Jackendoff 1983). Des non-musiciens sont-ils capables de percevoir la structure et sa déstructuration progressive en absence d'information tonale ? Cette capacité est-elle liée à l'expertise acquise par exposition passive ?

Expériences présentes comparant la perception et le traitement cérébral de la structure syntaxique en langage et en musique en tenant compte de l'expertise des participants

Dans cet article, nous décrivons une série d'expériences testant la perception et le traitement cérébral de la structure syntaxique musicale et linguistique, de façon symétrique avec une attention particulière sur l'effet de l'expertise musicale. Pour refléter le processus dynamique de construction syntaxique, la taille des domaines syntaxiquement cohérents varie paramétriquement dans nos stimuli alors que leur longueur globale reste constante. Le premier ensemble de trois expériences comportementales compare la perception et le traitement de la structure en étudiant l'effet de l'exposition passive à la musique. Deux tâches comportementales comparent Mundurucus et non-musiciens occidentaux quant à leur perception de la structure musicale. La première expérience teste le traitement implicite de la structure musicale en demandant si deux extraits sont différents ou identiques. Dans une deuxième expérience, les participants devaient noter la qualité des extraits musicaux. Cette dernière tâche est finalement répétée dans une troisième expérience incluant stimuli musicaux et linguistiques comparant la perception de la structure dans les deux domaines par des non-musiciens occidentaux. Enfin, une expérience d'imagerie fonctionnelle utilise ces mêmes stimuli linguistiques et

musicaux pour comparer leurs réponses neurales chez des musiciens et des non-musiciens occidentaux.

1. Effet de l'exposition sur le traitement de la structure musicale et comparaison avec le langage : expériences comportementales

Dans le but de tester l'effet de l'exposition passive sur le traitement de la structure musicale, deux groupes d'individus non experts différant selon leur degré d'exposition à la musique tonale occidentale ont été testés : des Mundurucus avec une exposition passive extrêmement minimale et des occidentaux non-musiciens sans éducation musicale formelle.



Figure 1. Schéma des stimuli musicaux. La longueur totale est constante à travers toutes les conditions mais la durée sur laquelle un percept cohérent peut être construit varie.

La structure en musique tonale occidentale repose sur des indices de tempo, de rythme, de mélodie et d'harmonie. Dans cet ensemble d'expériences, tous les indices structuraux musicaux ont été manipulés et en particulier l'effet de la tonalité. Notre manipulation principale de la structure musicale porte sur la taille des fragments intégrables dans une structure cohérente. Les stimuli sont donc de longueurs identiques mais la durée sur laquelle un percept est cohérent varie paramétriquement depuis seize temps, huit temps, quatre temps, deux temps et un

temps, correspondant ainsi à cinq conditions de structure différentes. La figure 1 présente un exemple de stimuli musicaux. De plus, la tonalité est un élément crucial de ce type de musique. Elle correspond à l'échelle de sons sur laquelle les notes sont organisées de façon hiérarchique. Les notes appartenant à la même tonalité, la même gamme, s'intègrent plus facilement les unes avec les autres par rapport aux notes en dehors de celle-ci. On peut ainsi prédire que nos stimuli contenant des notes issues de tonalités différentes seront perçus comme plus dissemblables que des stimuli du même niveau de structure mais contenant uniquement des notes de la même tonalité.

La perception implicite et explicite de la structure est testée en utilisant deux tâches : une de jugement explicite de cohérence et une de discrimination. L'expérience de discrimination examine l'impact de la structure musicale sur la capacité à mémoriser les mélodies. Les participants devaient juger si deux extraits musicaux présentés successivement étaient identiques ou différents. Si notre manipulation de structure est perçue, nous prédisons que plus les stimuli seront déstructurés, plus il sera difficile de les comparer. L'expérience de notation teste l'évaluation explicite de la quantité de structure perçue dans les stimuli. Les sujets étaient d'abord entraînés avec des exemples de phrases dans leur langue maternelle (français ou mundurucu) plus ou moins structurées. On leur expliquait que les phrases normales étaient associées avec une bonne note (représentée par un visage souriant) et que les phrases déstructurées étaient mauvaises (représentées par un visage triste). L'expérience commençait alors et les participants entendaient chaque extrait musical après lequel ils devaient noter à quel point ils étaient corrects. Ils répondaient soit pour les Mundurucus en pointant un visage plus ou moins souriant sur une échelle de cinq, soit pour les sujets français en cliquant sur le visage. Finalement, la dernière expérience permet de comparer la perception de la structure des stimuli musicaux et linguistiques en testant un ensemble indépendant de sujets occidentaux non-musiciens qui devaient noter leur structure. Dans cette dernière expérience, seuls les stimuli musicaux provenant de tonalités différentes sont utilisés dans un souci de maximiser l'effet de déstructuration tout en conservant une durée de test acceptable. Les stimuli linguistiques ont été créés en suivant le même modèle que les stimuli musicaux toujours de longueur constante mais dont la quantité de structure varie de seize mots, huit mots, quatre mots, deux mots et un mot. Le tableau 1 présente un exemple de stimuli linguistiques.

Condition	Taille constituant	Exemples
C16	16 mots	Mon père parle souvent de choses très savantes qui fascinent les copains qui viennent chez moi
C08	8 mots	Elle couvre la marmite qui réchauffe tout doucement / il passe à côté du parc du manoir
C04	4 mots	Il lance la balle / les grimaces des clowns / la biche sort vite / qui sent les fleurs
C02	2 mots	Il manque / est tombé / les fleurs / avec soin / très raide / du Brésil / le blé / les femmes
C01	1 mot	La / à / sortent / tarte / le / un / grimpe / Luc / les / qui / siècle / pains / qui / de / réchauffe / jaunes

Tableau 1. Exemple de stimuli linguistiques pour chaque condition. La longueur totale est toujours constante mais la taille sur laquelle un percept structuré peut être construit varie paramétriquement. Les phrases cohérentes de seize mots ont été prononcées mot par mot avec une pause égale entre chacun des mots afin de permettre le découpage en constituants de tailles variables qui sont recombinaisonnés différemment selon les conditions. Dans toutes les conditions, les mots sont espacés du même temps de pause.

1.1 Résultats de la Tâche de discrimination

L'indice de discriminabilité, mesuré par le d' , a été calculé en fonction de la cohérence des stimuli. La figure 2 montre la discriminabilité moyenne en fonction du niveau de structure en séparant Mundurucus des occidentaux non musiciens. Chaque ligne représente l'indice de discriminabilité d'un participant. L'analyse statistique (Analyse de Variance avec les facteurs : structure, groupe et tonalité) montre que les sujets ont répondu différemment pour les différents niveaux de structure. De plus, les Mundurucus et les non-musiciens occidentaux discriminent les stimuli différemment : les Mundurucus font globalement plus d'erreurs que les non-musiciens occidentaux. Enfin, il est plus facile de discriminer des fragments contenant des tonalités différentes. En particulier, on notera que l'amélioration de la discriminabilité avec la cohérence de stimuli est présente dans les deux groupes de sujets sans différence significative.

Des analyses *post hoc* séparant les deux groupes de sujets montrent que pour chacun des groupes, plus les stimuli sont structurés, plus il est facile de discriminer les deux extraits présentés : leur discriminabilité augmente. L'analyse des temps de réaction, col-

lectés pour les occidentaux non musiciens, montre que les participants ont été de plus en plus rapides à comparer des stimuli de plus en plus structurés. L'absence d'indice de tonalité n'a par contre pas influencé significativement le temps nécessaire à la comparaison des stimuli.

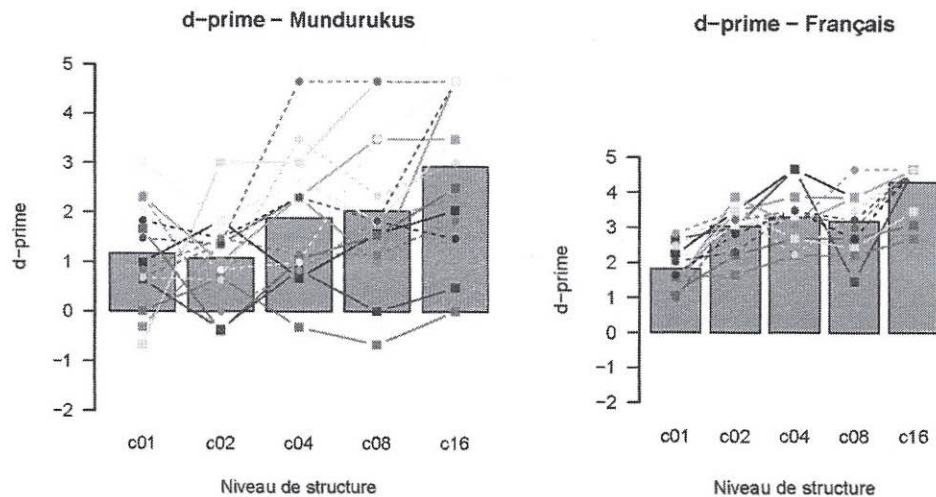


Figure 2. Indice de discriminabilité des deux extraits incluant tous les stimuli présentés. Zéro reflète le niveau du hasard, plus le score est haut, plus les extraits ont été discriminés correctement.

La tâche de discrimination a donc été correctement réalisée y compris par les Mundurucus (seuls 2 participants de ce groupe sont au niveau du hasard). Elle nous permet de mettre en évidence le traitement automatique de la structure musicale de nos stimuli puisque ce paramètre influence une tâche qui ne requiert pas de jugement explicite de la structure. Plus les stimuli sont structurés, plus il est plus facile de les encoder en mémoire en vue de les comparer et ce quelle que soit l'exposition. L'absence d'indices de tonalité aux frontières entre les fragments semble avoir eu un léger impact sur cette tâche, rendant un peu plus difficile la comparaison des niveaux les moins structurés puisque cet effet est marginal dans chacun des groupes.

1.2 Résultats de la Tâche d'Évaluation de la structure

Les notes moyennes données aux stimuli sont représentées figure 3. Chaque ligne représente un individu. L'analyse statistique (identique à l'expérience précédente) révèle que plus les stimuli sont déstructurés, plus les notes données sont basses pour les deux groupes même si l'amplitude de l'effet diffère entre les deux : la relation entre le niveau de structure et le jugement des sujets est plus forte pour les participants français.

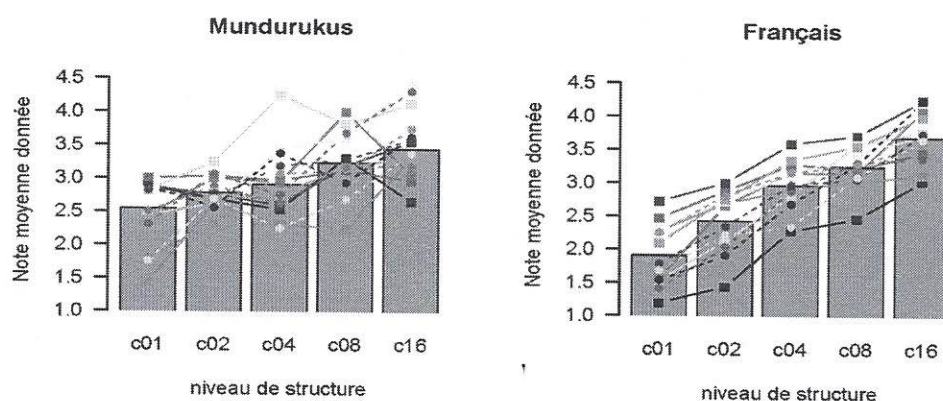


Figure 3. Note moyenne donnée en fonction du niveau de structure.

De plus, les notes données diffèrent selon que les fragments sont tous dans la même tonalité ou qu'ils sont dans des tonalités différentes : les notes données pour les stimuli issus de tonalités différentes sont meilleures que celles toutes issues de la même tonalité. Il semble que l'évaluation ait été rendue plus difficile par l'absence de changement de tonalité. En comparant les notes données aux stimuli issus de la même tonalité aux réponses données aux stimuli avec un changement de tonalité entre les fragments, nous pouvons déduire que le changement de tonalité a aidé à la perception de structure mais qu'elle n'est pas la seule responsable puisque l'effet est toujours visible en ayant ôté ce paramètre. Les Mundurucus, comme les Français, sont capables de noter le degré de structure d'un morceau de musique classique occidentale. Cela est remarquable dans la mesure où leur exposition à la musique tonale est minimale voire nulle.

1.3 Résultats de la Tâche d'évaluation de la structure : Comparaison musique et langage

La figure 4 présente la moyenne des notes données par sujet en fonction du niveau de structure construit.

Chaque ligne représente un sujet. L'analyse statistique (anova avec facteurs domaine et structure) montre que les sujets notent et perçoivent différemment les niveaux de structure en musique et en langage. De plus, la structure est perçue différemment en langage et en musique : la pente de la courbe plus grande pour le langage montre que la déstructuration des phrases en langage est perçue de façon plus prononcée. Cela revient à dire que pour un même niveau de structure, la note donnée va être plus basse pour le langage que la musique. Ce phénomène est d'autant plus prononcé que les stimuli sont déstructurés. Les sujets sont plus « choqués » par la déstructuration en

langage qu'en musique. Les stimuli langagiers portent un sens qui va être perdu assez rapidement avec la déstructuration, la cohérence perçue est influencée par des facteurs sémantiques et pas seulement syntaxiques. En musique, la structure vient non seulement de l'organisation temporelle mais aussi du rapport entre les différentes hauteurs des notes ; on ne peut pas assigner un sens aux notes.

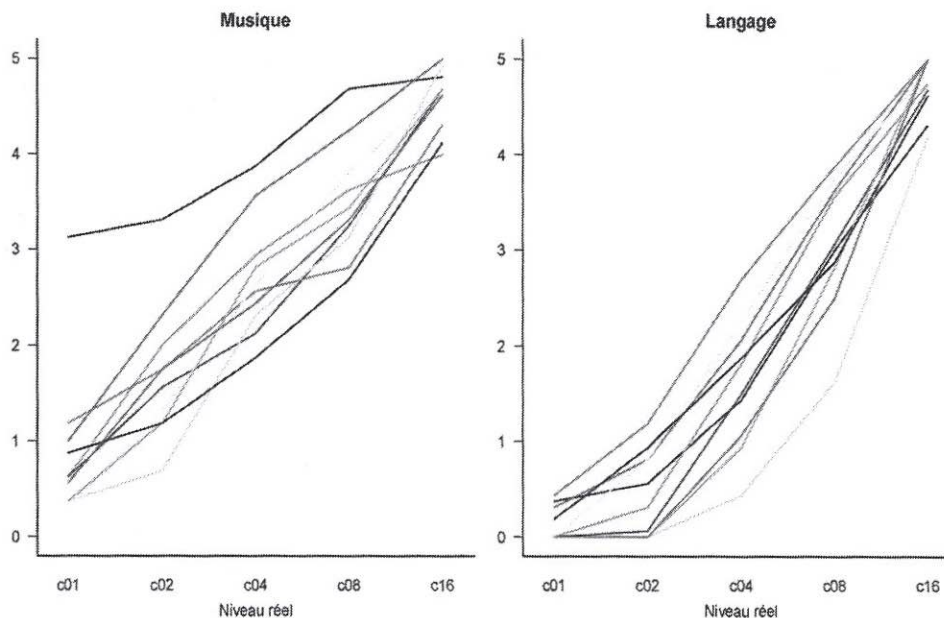


Figure 4. Notes moyennes données en fonction du niveau de structure par des non-musiciens occidentaux. Stimuli musicaux et linguistiques.

2. Traitement cérébral de la structure musicale et comparaison avec le langage : comparaison entre musiciens et non-musiciens

Cette étude de neuroimagerie utilise les mêmes stimuli linguistiques et musicaux que les expériences comportementales précédentes. Cette étude en IRM fonctionnelle cherche à identifier les régions cérébrales impliquées dans le traitement de la structure syntaxique en langage et en musique. Nous avons comparé des musiciens et des non-musiciens, qui ont tous eu une exposition passive importante à la musique tonale occidentale depuis l'enfance.

Afin de s'assurer que les participants prêtaient bien attention aux stimuli, on leur demandait de détecter un changement d'instrument (clavecin à la place de piano) pour les stimuli musicaux et un changement de voix (masculine à la place de féminine) pour les stimuli linguistiques. Cette tâche nous permet de tester le traitement cérébral automatique de la structure syntaxique sans attirer l'attention explicitement sur la structure.

2.1 Réponses globales au langage et à la musique

Les aires cérébrales activées par l'écoute globale de tous les stimuli linguistiques et musicaux sont bilatérales et correspondent aux aires classiques périsylviennes (gyrus supérieur temporal, gyrus frontal inférieur), le thalamus et l'aire supplémentaire motrice. La comparaison entre tous les stimuli linguistiques et musicaux montre des réseaux similaires bien que les réponses pour le langage soient plus importantes que la musique dans les régions temporales bilatérales et dans le gyrus inférieur frontal à gauche. Aucune région n'est plus activée par l'écoute globale de la musique que par le langage.

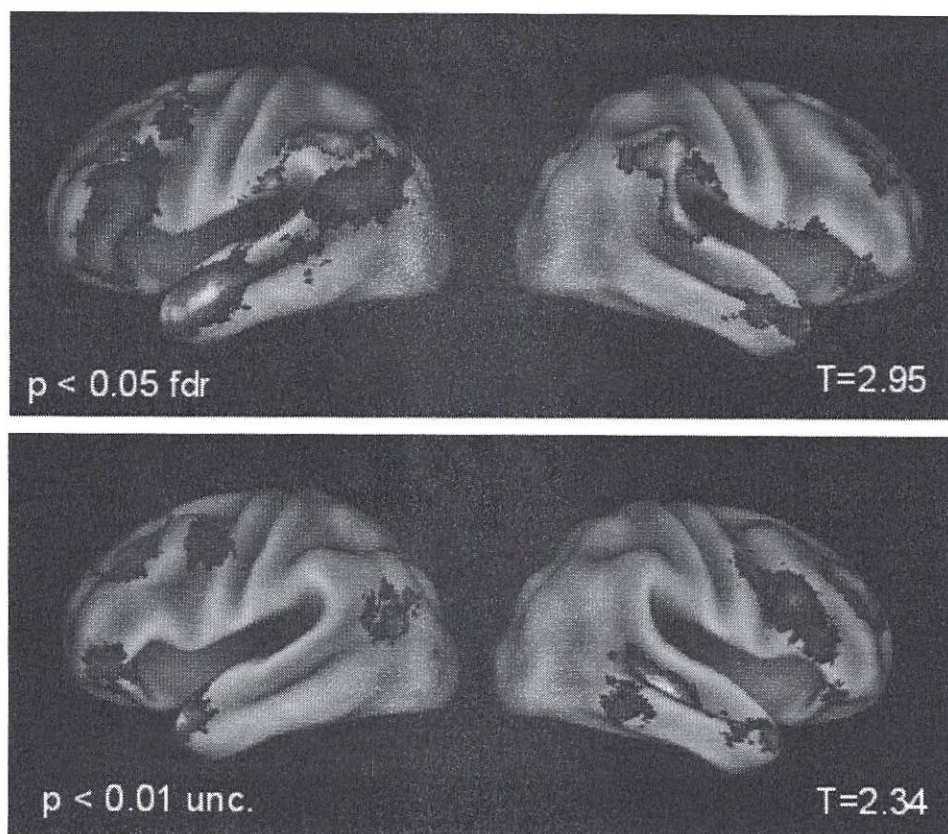


Figure 5. Carte d'activation cérébrale en fonction de la structure en langage à gauche, et de la structure en musique à droite. En grisé les activations et désactivations qui ne sont pas discutées ici.

2.2 Effet de structure

Nous avons cherché les régions dont l'activité cérébrale augmentait avec la taille des constituants syntaxiques en langage, figure 5. Un réseau de régions de l'hémisphère gauche est mis en évidence incluant le pôle temporal, la région antérieure du sillon temporal supérieur, la

2.1 Réponses globales au langage et à la musique

Les aires cérébrales activées par l'écoute globale de tous les stimuli linguistiques et musicaux sont bilatérales et correspondent aux aires classiques périsylviennes (gyrus supérieur temporal, gyrus frontal inférieur), le thalamus et l'aire supplémentaire motrice. La comparaison entre tous les stimuli linguistiques et musicaux montre des réseaux similaires bien que les réponses pour le langage soient plus importantes que la musique dans les régions temporales bilatérales et dans le gyrus inférieur frontal à gauche. Aucune région n'est plus activée par l'écoute globale de la musique que par le langage.

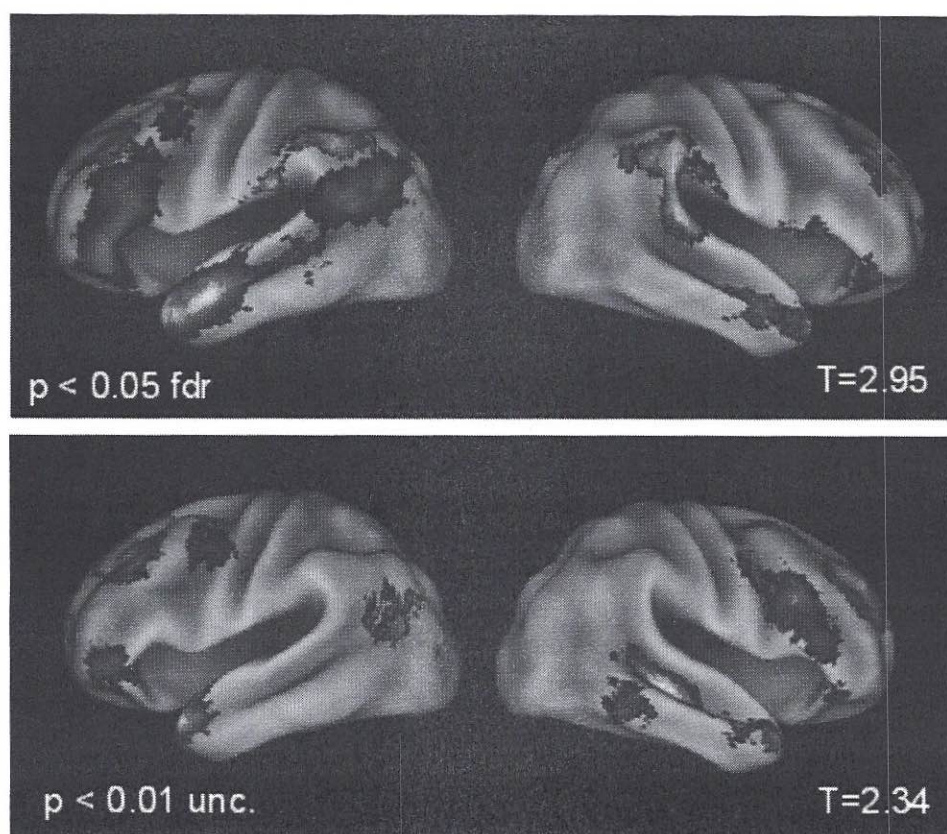


Figure 5. Carte d'activation cérébrale en fonction de la structure en langage à gauche, et de la structure en musique à droite. En grisé les activations et désactivations qui ne sont pas discutées ici.

2.2 Effet de structure

Nous avons cherché les régions dont l'activité cérébrale augmentait avec la taille des constituants syntaxiques en langage, figure 5. Un réseau de régions de l'hémisphère gauche est mis en évidence incluant le pôle temporal, la région antérieure du sillon temporal supérieur, la

jonction temporo-pariétale et les aires inféro-frontale. Dans l'hémisphère droit, on trouve le pôle temporal et une portion du gyrus frontal inférieur.

Pour les stimuli musicaux, le même seuil statistique ne révèle pas de région répondant de façon linéaire à l'accroissement de la taille des fragments. Cependant, en abaissant le seuil de significativité, l'analyse révèle un réseau symétrique en termes d'hémisphère à celui rapporté en langage. Particulièrement, les régions inféro-frontales, la partie médiane du sillon temporal supérieur, le pôle temporal dans l'hémisphère droit et le pôle temporal et une partie de la région inféro-frontale dans l'hémisphère gauche sont de plus en plus actives à mesure que la taille des fragments cohérents augmente. Le réseau ressemble à celui pour le langage mais basé cette fois-ci dans l'hémisphère droit.

Dans le but de savoir si les mêmes zones étaient sensibles à la structure en langage et en musique, nous avons calculé l'interaction entre le degré de structure et le domaine (langage vs musique). Cette comparaison des aires sensibles à la structure en langage avec celles sensibles à la structure en musique révèle des interactions dans un seul sens à savoir des régions qui répondent plus fortement à la structure en langage qu'en musique. Les régions de l'hémisphère gauche sont plus sensibles à la structure en langage qu'en musique : le pôle temporal gauche, la partie antérieure du sillon temporal supérieur et la jonction temporo-pariétale.

Les analyses précédentes reposent sur des statistiques de groupe. Or on sait que les activations cérébrales peuvent être localisées légèrement différemment entre individus. Pour répondre à la question de savoir si les mêmes régions répondent à la structure musicale et linguistique chez un même sujet, il faut donc procéder à une analyse plus poussée. Nous avons regardé la réponse à la structure en musique et en langage dans les mêmes voxels (les plus petites unités, pixel en trois dimensions, observables en IRM). Nous nous sommes basés sur le réseau de sept régions d'intérêt trouvées dans l'expérience indépendante étudiant la structure linguistique (Pallier 2011). Cette analyse prouve que les réponses hémodynamiques dans le gyrus frontal inférieur gauche (orbitalis et triangularis), le pôle temporal, la partie postérieure du gyrus temporal supérieur sont corrélées de façon positive avec l'augmentation de la taille des fragments en musique.

2.3 Comparaison entre musiciens et non-musiciens

Concernant le langage, aucune différence dans le traitement de la structure syntaxique n'a été trouvée au niveau de l'analyse de groupe entre les musiciens et les non-musiciens, ni lors de l'augmentation de

la taille des constituants, ni lors de leur diminution. Concernant le traitement de la structure en musique, les deux analyses de groupe utilisant les statistiques individuelles concordent sur le fait que les musiciens et les non-musiciens ont des profils d'activation différents en localisation et en amplitude.

Notre étude est la première à étudier l'effet de la structure en musique et en langage en comparant musiciens et non-musiciens. Nous montrons que les musiciens et les non-musiciens ne recrutent pas les mêmes aires cérébrales pour calculer la structure temporelle en musique. Les mêmes zones sont engagées par le traitement de la structure linguistique et musicale pour les musiciens, à savoir le gyrus inférieur frontal, le pôle temporal et le sillon temporal supérieur dans l'hémisphère gauche.

Conclusion

Nous avons présenté des stimuli musicaux plus ou moins structurés – construit en concaténant des fragments de plus en plus courts de sonates de Mozart – à trois groupes de personnes : des Amazoniens n'ayant pratiquement aucune exposition préalable à la musique tonale occidentale, des Européens sans éducation musicale formelle, et des musiciens pratiquant depuis un très jeune âge.

Des expériences comportementales utilisant des tâches de discrimination perceptive ou de jugement explicite de degré de cohérence montrent que les sujets, qu'ils soient de culture occidentale ou non et donc exposés ou non à la musique tonale occidentale, sont sensibles à la structure de ces stimuli musicaux. Plus le niveau de structure est élevé, plus il est facile de mettre en mémoire et de comparer les extraits musicaux, et plus la note donnée pour évaluer la structure est élevée. Notre manipulation des stimuli musicaux est effectivement perçue même par les membres de la tribu amazonienne. De plus, la perception et le traitement de la structure ne sont pas seulement dus à la présence d'un changement de tonalité entre les fragments, même si ce paramètre améliore les performances. Bien que l'exposition passive à la musique tonale occidentale aide à percevoir sa structure, des individus pour qui elle est absente sont capables de la percevoir, montrant que certains aspects structurels semblent être perçus de façon innée. Nous ne pouvons pas, grâce à ces expériences, préciser quels aspects de la structure musicale sont responsables du changement de perception et de traitement de nos stimuli. Cependant, il est possible d'imaginer des stimuli musicaux dont la structure ne varierait que sur un seul ou deux paramètres orthogonalement

Dans une expérience en imagerie fonctionnelle (IRMf), nous avons comparé les activations cérébrales en fonction de la structure linguistique et musicale chez des musiciens et non-musiciens occidentaux. Nos résultats sont en accord avec la littérature tant pour le traitement de la structure en langage qu'en musique. Les régions inféro-frontales, le pôle temporal et la partie postérieure du sillon temporal supérieur de l'hémisphère gauche sont engagés dans le traitement de la structure en langage. L'expertise musicale affecte les réseaux recrutés par le traitement automatique de la structure musicale. De façon générale les musiciens tendent à recruter les mêmes zones pour le traitement du langage et de la musique. Pour eux, les zones inféro-frontales et temporales seraient responsables de l'intégration d'éléments pour former une structure syntaxique indépendamment du sens, comme dans l'expérience de (Pallier 2011). Les non-musiciens n'activent pas ces mêmes régions pour traiter la structure musicale. Il est possible de mettre en parallèle l'acquisition de la musique avec l'acquisition d'une seconde langue. À mesure que la maîtrise de celle-ci augmente, les aires cérébrales recrutées deviendraient de plus en plus similaires.

Références

- ABRAMS Daniel, BHATARA Anjali, RYALI Srikanth, BALABAN Evan, LEVITIN Daniel and MENON Vinod, 2011, "Decoding Temporal Structure in Music and Speech Relies on Shared Brain Resources but Elicits Different Fine-scale Spatial Patterns", *Cerebral Cortex*, n° 21, p. 1507-1518.
- BERMUDEZ Patrick, LERCH Jason, EVANS Alan and ZATORRE Robert, 2009, "Neuroanatomical Correlates of Musicianship as Revealed by Cortical Thickness and Voxel-based Morphometry", *Cerebral Cortex*, n° 19, p. 1583-1596.
- BEVER Thomas and CHIARELLO Robert, 1974, "Cerebral Dominance in Musicians and Nonmusicians", *Science*, n° 185, p. 537-539.
- BHARUCHA Jamshed and KRUMHANSL Carol, 1983, "The Representation of Harmonic Structure in Music: Hierarchies of Stability as a Function of Context", *Cognition*, n° 13, p. 63-102.
- CHOMSKY Noam, 1957, *Syntactic Structures*, Mouton, The Hague.
- ELLIS Robert, BRUIJN Bente, NORTON Andrea, WINNER Ellen and SCHLAUG Gottfried, 2013, "Training-mediated Leftward Asymmetries During Music Processing: A Cross-sectional and Longitudinal fMRI Analysis", *Neuroimage*, n° 75, p. 97-107.
- ELLIS Robert, NORTON Andrea, OVERY Katie, WINNER Ellen, ALSOP David and SCHLAUG Gottfried, 2012, "Differentiating Maturational and

- Training Influences on fMRI Activation During Music Processing”, *Neuroimage*, n° 60, p. 1902-1912.
- FEDORENKO Evelina, PATEL Aniruddh, CASASANTO Daniel, WINAWER Jonathan and GIBSON Edward, 2009, “Structural Integration in Language and Music: Evidence for a Shared System”, *Memory & Cognition*, n° 37, p. 1-9.
- FEDORENKO Evelina, McDERMOTT Josh, NORMAN-HAIGNERE Samuel and KANWISHER Nancy, 2012, “Sensitivity to Musical Structure in the Human Brain”, *Journal of Neurophysiology*, n° 108, p. 2389-3300.
- FRIEDERICI Angela, 2002, “Towards a Neural Basis of Auditory Sentence Processing”, *Trends in Cognitive Science*, n° 6, p. 78-84.
- GASER Christian and SCHLAUG Gottfried, 2003, “Brain Structures Differ between Musicians and Non-Musicians”, *The Journal of Neuroscience*, n° 23, p. 9240-9245.
- HUMPHRIES Colin, BINDER Jeffrey, MEDLER David and LIEBENTHAL Einat, 2006, “Syntactic and Semantic Modulation of Neural Activity During Auditory Sentence Comprehension”, *Journal of Cognitive Neuroscience*, n° 18, p. 665-679.
- HUMPHRIES Colin, LOVE Tracy, SWINNEY David and HICKOK Gregory, 2005, “Response of Anterior Temporal Cortex to Syntactic and Prosodic Manipulations During Sentence Processing”, *Human Brain Mapping*, n° 26, p. 128-138.
- KATZ Jonah and PESETSKY David, Submitted, “The Identity Thesis for Language and Music.”
- KOELSCH Stefan and FRIEDERICI Angela, 2003, “Toward the Neural Basis of Processing Structure in Music. Comparative Results of Different Neurophysiological Investigation Methods”, *Annals of the New York Academy of Sciences*, n° 999, p. 15-28.
- KOELSCH Stefan, GUNTER Tomas, FRIEDERICI Angela and SCHRÖGER Erich, 2000, “Brain indices of music processing: ‘nonmusicians’ are musical”, *Journal of Cognitive Neuroscience*, n° 12, p. 520-541.
- KOELSCH Stefan, GUNTER Tomas, WITTFOTH Matthias and SAMMLER Daniela, 2005, “Interaction between Syntax Processing in Language and in Music: An ERP Study”, *Journal of Cognitive Neuroscience*, n° 17, p. 1565-1577.
- KRUMHANSL Carol and KESSLER Edward, 1982, “Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys”, *Psychological Review*, n° 89, p. 334-368.
- KUNERT Richard, WILLEMS Roel, CASASANTO Daniel, PATEL Aniruddh and HAGOORT Peter, 2015, “Music and Language Syntax Interact in Broca’s Area: An fMRI Study”, *PLoS One*, n° 10, e0141069.
- LERDAHL Fred and JACKENDOFF Ray, 1983, *A Generative Theory of Tonal Music*, Cambridge (MA), The MIT Press.

- LEVITIN Daniel and MENON Vinod, 2003, "Musical Structure is Processed in 'Language' Areas of the Brain: A Possible Role for Brodmann Area 47 in Temporal Coherence", *Neuroimage*, n° 20, p. 2142-2152.
- MAESS Burkhard, KOELSCH Stefan, GUNTER Tomas and FRIEDERICI Angela, 2001, "Musical Syntax is Processed in Broca's Area: An MEG Study", *Nature Neuroscience*, n° 4, p. 540-545.
- MAZOYER Bernard, TZOURIO Nathalie, FRAK Victor, SYROTA André, MURAYAMA Noriko, LEVRIER Olivier, SALAMON George, DEHAENE Stanislas, COHEN Laurent and MEHLER Jacques, 1993, "The Cortical Representation of Speech", *Journal of Cognitive Neuroscience*, n° 5, p. 467-479.
- MEYER Martin, ELMER Stefan, JÄNCKE Lutz, 2012, "Musical Expertise Induces Neuroplasticity of the Planum Temporale", *Annals of New York Academy of Sciences*, n° 1252, p. 116-123.
- MUSSO Mariacristina, WEILLER Cornelius, HORN Andreas, GLAUCHE Volkmer, UMAROVA Roza, HENNIG Jürgen, SCHNEIDER Albrecht and RIJNTJES Michel, 2015, "A Single Dual-Stream Framework for Syntactic Computations in Music and Language", *NeuroImage*, n° 117, p. 267-283 .
- NI Weijia, CONSTABLE Todd, MENCL Einar, PUGH Kenneth, FULBRIGHT Robert, SHAYWITZ Sally, SHAYWITZ Bennett, GORE John and SHANKWEILER Donald, 2000, "An Event-related Neuroimaging Study Distinguishing form and Content in Sentence Processing", *Journal of Cognitive Neuroscience*, n° 12, p. 120-133.
- NORMAN-HAIGNERE Sam, KANWISHER Nancy and McDERMOTT Josh, 2015, "Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition", *Neuron*, n° 88, p. 1281-1296.
- OHNISHI Takashi, MATSUDA Hiroshi, ASADA Takashi, ARUGA Makoto, HIRAKATA Makiko, NISHIKAWA Masami, KATOH Asako, IMABAYASHI Etsuko, 2001, "Functional Anatomy of Musical Perception in Musicians", *Cerebral Cortex*, n° 11, p. 754-760.
- PALLIER Christophe, DEVAUCHELLE Anne-Dominique and DEHAENE Stanislas, 2011, "Cortical Representation of the Constituent Structure of Sentences", *Proceedings of National Academy of Science USA*, n° 108, p. 2522-2527.
- PATEL Aniruddh, 2003, "Language, Music, Syntax and the Brain", *Nature Neuroscience*; n° 6, p. 674-681.
- PATEL Aniruddh, 2008, *Music, Language and the Brain*, Oxford, Oxford University Press.
- PATEL Aniruddh, GIBSON Edward, RATNER Jennifer, BESSON Mireille and HOLCOMB Phillip, 1998, "Processing Syntactic Relations in Language

and Music: an Event-related Potential Study”, *Journal of Cognitive Neuroscience*, n° 10, p. 717-733.

- PERANI Daniela, PAULESU Eraldo, GALLES Nuria Sebastian, DUPOUX Emmanuel, DEHAENE Stanislas, BETTINARDI Valentino, CAPPÀ Stefano, FAZIO Ferruccio and MEHLER Jacques, 1998, “The Bilingual Brain. Proficiency and Age of Acquisition of the Second Language”, *Brain*, n° 121, p. 1841-1852.
- PISTON Walter and DeVOTO Mark, 1987, *Harmony*, New York, Norton.
- ROGALSKY Corianne, RONG Feng, SABERI Kourosh and, HICKOK Gregory, 2011, “Functional Anatomy of Language and Music Perception: Temporal and Structural Factors Investigated Using Functional Magnetic Resonance Imaging”, *Journal of Neuroscience*, n° 31, p. 3843-3852.
- ROHRMEIER Martin, 2011, “Towards a Generative Syntax of Tonal Harmony”, *Journal of Mathematics and Music*, n° 5, p. 35-53.
- SCHOENBERG Arnold and STEIN Leonard, 1969, *Structural Functions of Harmony*, New York, Norton.
- SLEVC Robert, ROSENBERG Jason and PATEL Aniruddh, 2009, “Making Psycholinguistics Musical: Self-paced Reading Time Evidence for Shared Processing of Linguistic and Musical Syntax”, *Psychonomic Bulletin & Review*, n° 16, p. 374-381.
- SNIJDERS Tineke, VOSSE Theo, KEMPEN Gerard, VAN BERKUM Jos, PETERSSON Karl Magnus and HAGOORT Peter, 2009, “Retrieval and Unification of Syntactic Structure in Sentence Comprehension: An fMRI Study Using Word-category Ambiguity”, *Cerebral Cortex*, n° 19, p. 1493-1503.
- TILLMANN Barbara, BHARUCHA Jamshed and BIGAND Emmanuel, 2000, “Implicit Learning of Tonality: A Self-organizing Approach”, *Psychological Review*, n° 107, p. 885-913.