

Philippe LALITTE

(Chargé de recherches au LEAD-CNRS, UMR 5022, chercheur associé au MINT)

LA FORME MUSICALE AU REGARD DES SCIENCES COGNITIVES

Les psychologues définissent souvent la notion de forme musicale en l'opposant à celle de syntaxe¹. Selon Francès, qui assimile la notion de rhétorique à celle de forme, la rhétorique musicale désigne les rapports « qui s'établissent entre les parties d'une œuvre musicale », alors que la syntaxe musicale désigne ceux « qui régissent la composition de ces parties ».² La fonction de la syntaxe serait de régir les relations entre les sons à un niveau local alors que celle de la forme se situerait pour organiser les structures à un niveau plus global (phrases, sections, parties). Cependant, d'autres psychologues, notamment anglo-saxons, ont une conception de la forme qui s'appuie moins sur les éléments thématiques que sur les structures harmoniques de la musique tonale. D'une part, la théorie des attentes perceptives de L. B. Meyer³, et d'autre part, le concept de prolongation développé par Lerdhal et Jackendoff dans leur *Théorie Générative de la Musique Tonale*⁴ (TGMT) ont majoritairement servi de cadre théorique aux recherches en cognition auditive. Ces auteurs ont en commun l'idée qu'une pièce de musique tonale possède une structure fortement hiérarchisée où chaque événement s'intègre dans un réseau de relations. Grâce à cette structure sous-jacente, les différents événements de la pièce étant subordonnés à la tonalité principale, la musique acquiert unité et cohérence.

La conception hiérarchique est attrayante pour la psychologie cognitive parce que les structures hiérarchiques sont des formes « économiques » de représentation. Elle possède de nombreux avantages en termes de traitement et de stockage en mémoire. Ces avantages semblent être particulièrement importants au moment où l'auditeur découvre les événements musicaux successifs. L'expérience perceptive de l'auditeur étant irréversible, celui-ci aura donc à traiter, au fur et à mesure, l'information en perpétuel changement. De ce fait, la perception des structures musicales à grande échelle dépend en premier lieu de la façon dont les informations sont traitées à un niveau local et en second lieu des relations entre les unités locales.

¹ Pour une définition critique de ces deux notions voir : Nicholas Cook, « Forme et syntaxe », in *Musiques, une encyclopédie pour le XXI^e siècle*, sous la dir. de Jean-Jacques Nattiez, Tome 2, *Les savoirs musicaux*, Paris, Actes Sud/Cité de la Musique, 2004, p. 162-188.

² Robert Francès, *La perception de la musique*, Paris, Vrin, 1958, 2^e édition 1984, p. 152-153.

³ Leonard B. Meyer, *Emotion and Meaning in Music*, Chicago and London, The University of Chicago Press, 1956.

⁴ Fred Lerdhal et Ray Jackendoff, *A Generative Theory of Tonal Music*, Cambridge Massachusetts, The MIT Press, 1983.

1 - TRAITEMENT DE L'INFORMATION MUSICALE DANS LE TEMPS

1.1 - TRAITEMENT DE L'INFORMATION AU NIVEAU LOCAL

Les processus de traitement de l'information musicale à une échelle temporelle locale sont aujourd'hui bien connus. La première étape est celle de la reconnaissance de sources sonores et de micro-événements sonores. Cette étape, d'une durée extrêmement courte, est constituée de processus automatiques et implicites qui permettent l'intégration du flux auditif en catégories discrètes. Nous nous contenterons simplement de résumer les cinq processus de traitement de la reconnaissance auditive et renvoyons le lecteur à la description précise de S. McAdams⁵ :

1. Transduction sensorielle (représentation du signal acoustique dans le système nerveux auditif périphérique),
2. Groupement auditif (séparation des différentes sources sonores),
3. Analyse des propriétés spectrales et/ou traits auditifs,
4. Ajustement au lexique auditif (comparaison à des catégories de sources en MLT),
5. Reconnaissance et identification (activation du lexique verbal et de structures sémantiques associées).

La deuxième étape de traitement correspond à une échelle de temps dénommée « présent perceptuel » par P. Fraisse. Celui-ci le définit comme un acte mental unifié indispensable à la perception du successif : « Nous ne percevons le successif que parce que, dans certaines limites, un acte mental unifié est possible. Cette unité perceptive du successif — le tic-tac de notre pendule — a comme conséquence l'existence d'un présent perçu qui ne se ramène pas à l'évanescence de ce qui n'était pas encore dans ce qui n'est plus »⁶. Dans ce présent, il y a une forme de simultanéité décomposable en un certain nombre d'éléments et d'intervalles entre ces éléments qui tient à l'unité même de l'acte de perception. Selon Fraisse, cette fenêtre a une durée moyenne de 2 à 3 sec., sa limite supérieure étant de 7 à 8 sec. Bien qu'il admette la même durée moyenne de 2 à 3 sec., Michon élargit considérablement la limite supérieure puisque, selon lui, le présent occupe « un laps de temps très variable avec une durée maximum impossible à déterminer avec précision, mais qui peut atteindre environ 30 sec »⁷. Le présent perceptuel peut être représenté sous la forme d'une fenêtre coulissante à laps de temps variable à travers laquelle le sujet perçoit le flux musical et opère des groupements d'événements (Fig. 1). Les processus de segmentation de la surface musicale concourent à définir la place et la taille de cette fenêtre coulissante. Comme le précise E. Bigand : « Plusieurs recherches ont montré qu'ils opèrent de façon conforme aux principes de groupement de la théorie de la forme (*Gestalttheorie*). Toute frontière de groupe constitue alors un point de

⁵ Stephen McAdams, « La reconnaissance de sources et d'événements sonores », in S. McAdams et E. Bigand (éd.), *Penser les sons : La psychologie auditive de l'audition*, Paris, PUF, 1994, p. 157-213.

⁶ Paul Fraisse, *Psychologie du temps*, Paris, Presses Universitaires de France, 1957, p. 84.

⁷ John A. Michon, « Le traitement de l'information temporelle », in *Du temps biologique au temps psychologique*, Symposium de l'Association de psychologie scientifique de langue française, 1977, p. 259.

segmentation de la surface musicale »⁸. La difficulté d'établir une limite de durée précise de cette fenêtre provient certainement du fait qu'une telle limite est déterminée autant par les propriétés structurelles de la musique que par d'autres facteurs comme le focus attentionnel, la familiarité ou même l'expertise musicale.

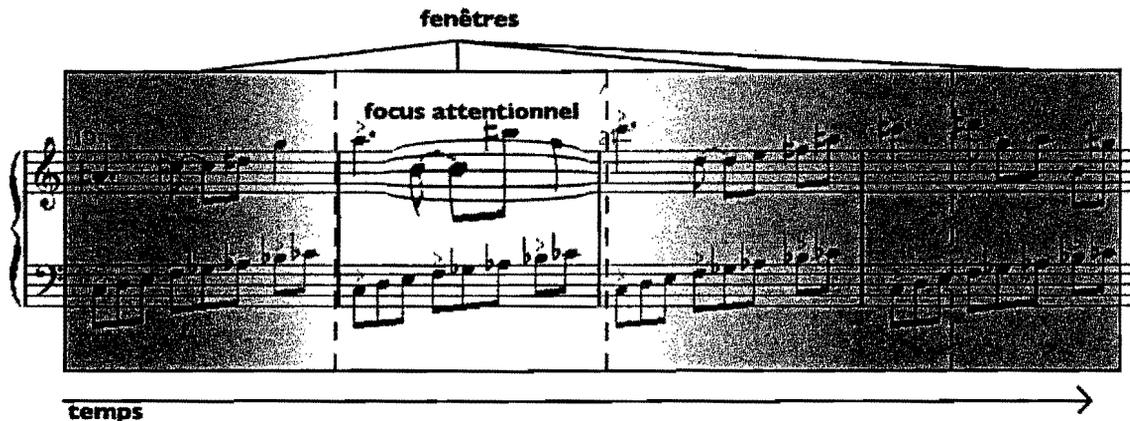


Figure 1. Perception au niveau local à travers un ensemble de centrations perceptives

1.2 – TRAITEMENT DE L'INFORMATION AU NIVEAU GLOBAL

Si les processus de traitement de l'information dans le temps, tels qu'ils sont décrits plus haut, sont en général bien acceptés par la communauté scientifique, il n'en va pas de même en ce qui concerne la perception à une échelle de temps plus large. Les avis divergent et les résultats des études expérimentales sont plutôt contradictoires. Si l'on accepte l'idée de la fenêtre coulissante comme cadre théorique de la perception à un niveau local, l'étape suivante devrait correspondre à un ensemble de présents perceptuels regroupés par l'auditeur. A ce stade, l'information est toujours disponible, mais plus par la perception. La mémoire est alors mise à contribution pour établir des relations entre les groupements.

Si la notion de forme a une réalité perceptive, le traitement de l'information dans le temps ne devrait pas constituer une pure succession de segments relativement courts. L'information relative à plusieurs segments devrait être reliée à un niveau supérieur d'organisation. Les psychologues ont donc émis l'hypothèse qu'à ce stade, l'intégration de différents segments locaux dans une structure hiérarchique globale devraient être rendue possible par un ensemble de relations de niveau plus élevé établi entre les différentes structures locales. Une bonne perception des structures locales devrait également contribuer à une meilleure perception des relations à haut niveau.

Cependant, cette vision des choses est contestée, notamment par le philosophe américain J. Levinson auquel se rallient plusieurs psychologues. Celui-ci défend l'idée

⁸ Emmanuel Bigand, « Contributions de la musique aux recherches sur la cognition auditive humaine », in S. McAdams et E. Bigand (éd.), *op. cit.*, p. 271. Les principales règles de groupement sont : les règles de proximité en fonction des intervalles de hauteur et de temps, du phrasé et des silences, les règles de similarité en fonction du registre, de la dynamique, des durées et des timbres, et par les règles de continuité (contour mélodique).

que la perception des grandes formes n'a pratiquement aucune importance pour l'appréciation et la compréhension de la musique, puisque l'écoute musicale s'inscrit uniquement dans la concaténation de petits moments individuels. L'essentiel du processus d'écoute se passe à un niveau local, d'une phrase à une autre, sans qu'il y ait la possibilité d'intégrer ces moments différents dans un tout. Selon Levinson, « la musique, quelle que soit sa durée, consiste en une série d'événements successifs qui ne peuvent être appréhendés simultanément dans un seul acte perceptif. Les parties d'une façade architecturale peuvent être appréhendées plus ou moins dans un seul regard ; les parties d'une symphonie non »⁹. Levinson décrit l'écoute comme un processus à trois faces nommé *quasi-hearing* : l'audition réelle (*actual hearing*) d'un instant de musique, la remémoration vive (*vivid remembering*) d'un moment musical qui vient juste de passer, l'anticipation vive (*vivid anticipation*) d'un moment juste à venir. Remémoration vive et anticipation vive sont des images mentales sonores qui existent simultanément dans la conscience auditive à la manière dont les objets aux extrêmes du champ de vision sont incorporés à ceux qui sont au centre. La largeur de la fenêtre de ce processus à trois faces est proportionnelle à la capacité de la remémoration vive et de l'anticipation vive. Puisqu'elle déborde l'acte perceptif, la forme est considérée par les concaténationnistes comme quelque chose d'extérieur à la musique et par conséquent de superflu. Cette position conduit Levinson à mettre en avant le contenu sensoriel et émotionnel de la musique au détriment de l'aspect formel et sémantique. L'écoute focalisée sur l'instant présent suffit pour apporter le plaisir musical et les effets émotionnels d'une pièce sont suffisamment portés par les moments individuels qui se succèdent.

Un ensemble de questions émerge de ces premières constatations : comment les auditeurs segmentent-ils une pièce de musique, et quels facteurs infèrent cette segmentation ? L'écoute de la musique est-elle limitée à des connexions locales de proche en proche ? Dans le cas contraire, dans quelle mesure et comment se font les connexions à une échelle temporelle plus vaste ? Quelle représentation mentale de la structure formelle les auditeurs possèdent-ils pour localiser un extrait ou pour reconnaître un élément thématique ? Enfin, existe-t-il une différence dans la perception des structures temporelles entre auditeurs expérimentés (musiciens) et novices (non musiciens) ? Ces questions ont été abordées par les psychologues, parfois en collaboration avec des musicologues, à l'aide de différentes méthodes. On peut regrouper ces différentes méthodes dans deux grandes approches : d'une part une approche qui étudie la perception des structures à grande échelle à partir de capacités explicites (segmentation, reconnaissance, localisation) et d'autre part une approche qui s'appuie sur le ressenti des auditeurs à travers des jugements subjectifs, des réponses introspectives. Dans les pages qui suivent, nous allons donc présenter quelques études expérimentales de cognition musicale qui nous semblent représentatives de ces deux approches.

⁹ Jerrold Levinson, *Music in the Moment*, Ithaca and London, Cornell University Press, 1997, p. 2 (notre traduction).

2 - PROCESSUS DE GROUPEMENTS ET DE MEMORISATION

2.1 - SEGMENTATION DE LA SURFACE MUSICALE

M. Imberty postule que les opérations mentales de segmentation et de groupement effectuées par l'auditeur pour maîtriser le flux des événements vont constituer des indices pour observer le rapport entre la structure de l'œuvre et sa perception. Selon Imberty : « On peut donc penser que la netteté plus ou moins grande de ces parties, l'aisance avec laquelle les contrastes, les ruptures, les entrées des thèmes sont perçues, la prégnance de tel ou tel élément par rapport aux autres, constituent une hiérarchie plus ou moins forte d'unités s'emboîtant les unes dans les autres ou se juxtaposant, qui témoigne de l'organisation intrinsèque de la pièce musicale, ou tout au moins de la part de cette organisation qui donne lieu à des comportements réactifs observables chez le sujet »¹⁰. Une forte hiérarchie implique la perception de quelques changements très prégnants, qui constituent un premier niveau, et de niveaux intermédiaires dont les changements sont moins prégnants, mais toujours en petit nombre à chaque niveau. Une faible hiérarchie suppose un premier niveau avec de nombreux changements perçus d'emblée et un ou plusieurs niveaux internes avec des changements faiblement perçus.

La deuxième hypothèse d'Imberty est que la connaissance, implicite ou non, de certains types de structuration du temps fortement représentés dans la culture musicale occidentale, comme les formes tripartites, constituent des schémas structurels qui vont plus ou moins conditionner les stratégies de segmentation auditive. Ces schémas, qu'il nomme macrostructures, seraient employés par l'auditeur pour stocker les informations et, lors de leur réactivation, serviraient alors à reconstituer la forme globale de la pièce. Les indices qui déterminent le schéma de la macrostructure sont de deux ordres : d'une part des indices structuraux objectifs qui correspondent à des éléments discrets (repérage d'un thème ou d'une cellule rythmique, changement de registre ou d'harmonie) et qui concernent la cohérence formelle, d'autre part des indices subjectifs qui marquent l'évolution dynamique d'un élément à un autre ou d'un moment à un autre. Ces deux types d'indices sont nommés respectivement schèmes de relation d'ordre et schèmes d'ordre. Imberty a ainsi pu montrer les différences perceptives qui ressortent à l'écoute de l'*Intermezzo en mi bémol mineur* op. 118 n° 6 de Brahms et de *La Puerta del Vino* (Second livre des *Préludes*) de Debussy. La macrostructure des participants (musiciens et non musiciens) pour la pièce de Brahms était beaucoup plus hiérarchisée et dépendante des schèmes d'ordre¹¹ que la macrostructure découlant de la pièce de Debussy dont le nombre de segmentations était beaucoup plus important et la répartition plus diffuse.

L'approche théorique d'I. Deliège est fondée sur les théories gestaltistes de formation de groupes et sur la Théorie générative de la musique tonale de Lerdahl et Jackendoff. La perception de la forme est pensée par I. Deliège comme l'élaboration progressive d'un schéma simplifié par rapport au contenu objectif de l'œuvre. Il s'agit alors d'évaluer

¹⁰ Michel Imberty, *Les écritures du temps, Sémantique psychologique de la musique*, tome 2, Paris, Dunod, 1981, p. 87.

¹¹ Les segmentations ont fait apparaître une structure tripartite (AABA') typique de la forme d'un intermezzo.

la capacité des auditeurs à « opérer des groupements de groupes en fonction de la pertinence qu'il dégage des structures, de son aptitude à éprouver les liens qui existent entre ces structures et à les faire apparaître dans le schéma qui résulte de son opération de groupement »¹². Les opérations de groupement sont facilitées par un processus mental d'extraction d'indices pertinents tirés de la surface musicale. Ces indices sont comme des étiquettes collées à un « objet musical », mais seuls les indices les plus forts sont conservés en mémoire. Lorsqu'ils reviennent dans le déroulement de la pièce, sous une forme littérale ou variée, leur rappel permet à l'auditeur d'établir des comparaisons, des catégorisations, des groupements à plus haut niveau, des hiérarchies, et donc de se construire une représentation mentale de la pièce. Deliège a testé ses hypothèses sur la *Sequenza VI* pour alto solo (1970) de Berio et *Eclat* (1965) pour ensemble instrumental de Boulez. L'hypothèse était que dans la pièce de Berio, le fort taux d'invariance devrait favoriser le groupement de groupes en fonction de la similitude des indices alors que dans la pièce de Boulez l'enchaînement de trames sonores fluctuantes devrait inférer une extraction d'indices à un niveau plus local sans qu'il y ait de groupements à niveau supérieur.

Pour la *Sequenza VI*, musiciens et non musiciens ont effectué les groupements de groupes d'une manière très semblable. Six sections principales ont été délimitées, dans chacune d'elles on peut repérer des indices invariants. Pour *Eclat*, quelques différences entre musiciens et non musiciens sont apparues dans la manière d'effectuer les groupements de groupes. La pièce a été segmentée en quatre sections principales qui correspondent à des contrastes marquants. Concernant la différence entre les deux pièces, l'hypothèse selon laquelle l'extraction d'indice serait facilitée dans la *Sequenza VI* n'a pas été confirmée. Dans *Eclat*, « il s'est opéré, selon Deliège, une sorte d'assimilation sur base des effets de hauteur et de timbre, perçus ici plutôt comme indices »¹³. Il semble donc que les caractéristiques musicales qui infèrent les segmentations ne sont pas fixes. Dans le cas de la *Sequenza*, il semble que les segmentations s'établissent à partir d'un ensemble de figures, alors que pour *Eclat* ce sont plutôt des changements de densité, de texture et de timbre.

Un projet scientifique récent conduit sous la direction de S. McAdams à l'IRCAM en collaboration avec le compositeur américain et l'équipe de E. Bigand au LEAD a permis de réaliser un ensemble important d'études sur une pièce spécialement composée par le compositeur américain R. Reynolds¹⁴. Une des études (Lalitte, Bigand, et coll., 2004)¹⁵ a examiné la structure perceptive du matériau thématique de la pièce. Les participants devaient, dans une première phase, indiquer les changements d'idées musicales en continu, et, dans une deuxième phase, juger si des paires d'extraits appartiennent au même thème et, dans le cas d'une réponse positive, lequel de ces deux extraits apparaît

¹² Irène Deliège, « Approche perceptive des formes musicales contemporaines », in *La musique et les sciences cognitives*, S. McAdams et I. Deliège (éd.), Liège/Bruxelles, Pierre Margada Editeur, 1989, p. 306.

¹³ Irène Deliège, *Ibid.*, p. 322.

¹⁴ *The Angel of Death* (1998-2001) pour piano, ensemble instrumental et sons générés par ordinateur.

¹⁵ Philippe Lalitte, E. Bigand, B. Poulin, S. McAdams, C. Delbé et D. D'Adamo, « The perceptual structure of materials in *The Angel of Death* », *Music Perception, Special issue Creation and Perception of a Contemporary Musical Work : The Angel of Death Project*, vol. 22, 2004, p. 265-296.

en premier dans le thème. Les structures perçues par les auditeurs (musiciens et non musiciens) correspondent globalement à celles formulées par le compositeur, mais les sujets semblent particulièrement influencés par la qualité temporelle des thèmes. Ainsi, lorsque les thèmes sont constitués de sections contrastées, les structures perceptives sont particulièrement corrélées aux structures décrites par le compositeur, mais lorsque les thèmes suivent une trajectoire continue, la structure perceptive correspond à l'évolution musicale, mais reflète peu la structure compositionnelle. Un des résultats intrigants de l'étude concerne l'effet de l'expertise musicale sur la durée des idées musicales perçues. La durée moyenne des idées musicales perçues par les non musiciens (7 à 8 sec.) est plus courte que celle des musiciens (15 sec.). Globalement, la durée des idées musicales perçues s'étend jusqu'à un maximum de 24 sec., ce qui correspond assez bien à la notion de *quasi-hearing* de Levinson¹⁶. En ce qui concerne les jugements d'appartenance et d'ordre temporel (2^e tâche), les sujets ont rencontré des difficultés. Ceci s'explique en raison de la nature plus abstraite de la tâche (jugements hors contexte) et par la difficulté de mémoriser des matériaux peu familiers. Cette difficulté de détection et de localisation semble confirmée par d'autres études.

2.2 - JUGEMENTS DE DETECTION ET DE LOCALISATION D'ELEMENTS THEMATIQUES

Selon R. Francès, « La perception de la forme d'une œuvre suppose avant tout la ségrégation du ou des thèmes et leur identification chaque fois qu'ils réapparaissent plus ou moins modifiés, à la suite d'une activité *assimilative* »¹⁷. L'activité assimilative correspond à l'appréhension des principales caractéristiques morphologiques du thème. Cette assimilation, dans la mesure où elle est réussie, permet la distinction des éléments de nouveauté introduits (variation, amplification, changements d'harmonie ou de rythme) et la différenciation de tout ce qui n'appartient pas au thème (nouveau thème, pont, transition, divertissement). Afin de mettre en évidence les processus d'assimilation, de distinction et de différenciation, Francès a réalisé une expérience avec le scherzo du *Trio* op. 97 de Beethoven qui comporte deux éléments thématiques bien contrastés¹⁸. Les participants ont été informés de la présence de deux idées thématiques principales présentes dans la pièce. Après une première audition de familiarisation, les participants devaient signaler l'apparition des thèmes et de leurs dérivés. Les résultats ont montré une influence de l'éducation musicale¹⁹ lorsque les processus de reconnaissance et d'assimilation demandaient une analyse assez fine du contenu musical (par exemple reconnaître les éléments apparentés à l'idée principale). Une seconde expérience, où les participants n'ont pas eu connaissance du nombre de thèmes, ni de leur mode d'apparition, a révélé une grande difficulté de distinction et de localisation des thèmes. Francès conclut que, dans une situation non guidée, l'assimilation des éléments thématiques devient encore plus difficile.

¹⁶ Selon Levinson, la durée de cette fenêtre est équivalente à celle d'une longue mélodie.

¹⁷ Robert Francès, Michel Imberty et Arlette Zenatti, « Le domaine musical », in *Psychologie de l'art et de l'esthétique*, sous la direction de R. Francès, Paris, Presses Universitaires de France, 1979, p. 153.

¹⁸ Robert Francès, *La perception de la musique*, Paris, Vrin, 1958, 2^e édition 1984.

¹⁹ L'expérience s'est déroulée avec trois groupes de sujets : des élèves de lycée non musiciens, des élèves de lycée ayant une pratique musicale et des élèves de conservatoire.

Clarke et Krumhansl (1990)²⁰ ont abordé la question de la perception des grandes formes musicales en tenant compte à la fois de la perception des éléments thématiques et de la trajectoire interne à chaque pièce. L'étude a été menée sur deux pièces historiquement et stylistiquement contrastées : le *Klavierstück IX* pour piano solo (1961) de K. Stockhausen et la *Fantaisie en ut mineur* pour piano K. 475 de Mozart. Les expériences comportaient des tâches de segmentation et de localisation d'extraits. Les extraits pour la tâche de localisation ont été choisis en fonction des frontières établies lors de l'expérience de segmentation. Les participants (non musiciens) ont entendu deux fois la pièce entière puis les dix-huit extraits dans un ordre aléatoire. Ils ont dû localiser les extraits en donnant leur réponse en traçant, sur une ligne horizontale, deux traits verticaux pour indiquer le début et la fin de l'extrait. D'après les résultats, les valeurs moyennes de jugements de localisation sont corrélées de façon significative avec les positions réelles des extraits, bien que ceux situés près du centre aient dévié de leur position réelle (ceci particulièrement pour le *Klavierstück IX*). D'après les auteurs, cette différence entre les deux pièces s'explique par la structure métrique (absente de la pièce de Stockhausen) qui procure un cadre temporel permettant aux auditeurs de maintenir un sens relativement robuste de la localisation et par les proportions équilibrées de la pièce de Mozart qui facilitent leur identification et leur mémorisation.

Tillmann et Bigand (1998)²¹ ont adapté un paradigme utilisé en perception visuelle et verbale pour examiner l'influence du contexte global sur la détection et la reconnaissance de cibles. Pour modifier la cohérence globale du contexte, des menuets ont été découpés en unités de 4, 2 ou 1 mesures et réorganisés de façon à briser la cohérence globale des pièces. Pour chaque menuet, un extrait cible (d'environ 2 sec.) a été sélectionné. Les participants (musiciens et non musiciens) ont eu deux types de tâches à effectuer : une tâche de détection de cible et une tâche de reconnaissance de cible. Dans le premier cas, les participants ont écouté la cible deux fois, puis devaient détecter la cible en écoutant un menuet entier (version originale et versions modifiées). Dans le deuxième cas, le menuet, écouté d'abord en entier, était suivi de l'extrait cible. Les participants devaient signaler si l'extrait cible appartenait ou non au menuet. L'hypothèse était que les cibles seraient mieux détectées ou reconnues avec les versions cohérentes des menuets. Cependant, la cohérence globale du contexte n'a que peu affecté l'identification des extraits. Seule la version la plus incohérente des menuets (découpage au niveau de la mesure) a eu un effet significatif. Dans la tâche de reconnaissance, les temps de réponse ont été plus importants pour les leurres que pour les bons extraits cible et la cohérence du contexte n'a affecté que le rejet des leurres. Globalement, les musiciens ont eu plus de réponses correctes et ont eu des temps de réponses plus courts que les non musiciens, mais de façon marginalement significative.

Les tâches employées dans ces études sollicitaient les auditeurs avec des réponses explicites. D'un certain point de vue, on peut se demander si les tâches explicites de segmentation, de reconnaissance ou de détection de cibles rendent véritablement compte des processus d'écoute car ces tâches induisent une distorsion par rapport à l'écoute

²⁰ Eric F. Clarke et Carol L. Krumhansl, « Perceiving Musical Time », *Music Perception*, Vol. 7, 1990, p. 213-253.

²¹ Barbara Tillmann et Emmanuel Bigand, « Influence of Global Structure on Musical Target Detection and Recognition », *International Journal of Psychology*, Vol. 33-2, 1998, p. 107-122.

« naturelle ». À l'audition d'une pièce de musique, l'auditeur *lambda* cherche rarement de façon aussi explicite à décortiquer les structures musicales, il se laisse plutôt porter par le flux musical et les émotions qui en découlent. C'est pourquoi les psychologues ont cherché d'autres méthodes permettant de tester la sensibilité aux structures à grande échelle d'une manière indirecte.

3 - IMPLICATION DES STRUCTURES A GRANDE ECHELLE SUR LE RESENTI DES AUDITEURS

3.1 - PARADIGME DE LA VIOLATION DE CLOTURE TONALE

Une des méthodes pour tester la réalité perceptive des structures à grande échelle consiste à manipuler la structure harmonique ou rhétorique des pièces musicales et ensuite à évaluer de manière plus ou moins directe combien les auditeurs perçoivent ces manipulations. De nombreuses recherches ont prouvé que les auditeurs perçoivent les fonctions cadentielles présentées isolément ou dans des contextes très courts. Mais dans quelle mesure les hiérarchies tonales sont-elles prises en compte à plus grande échelle ? N. Cook (1987)²² a manipulé des extraits musicaux de différentes durées (entre 30 sec. et 6 min.) de sorte que la tonalité à la fin de l'extrait soit différente de la tonalité principale (afin de violer la clôture tonale). Les participants (étudiants débutants en musique) ont jugé les extraits sur des échelles subjectives (cohérence, accomplissement, plaisir et expressivité). Les résultats ont révélé que l'effet de la clôture tonale n'a été ressenti que sur les extraits courts (30 sec.) et seulement pour les échelles de cohérence et d'accomplissement. Des données convergentes ont été rapportées par West-Marvin et Brinkman (1999)²³ avec des jugements explicites sur la tonalité de la cadence finale. Les participants (musiciens) ont dû indiquer si la fin des extraits (de moins de 2 minutes) étaient dans la même tonalité que la tonalité initiale de l'extrait. Globalement, les participants ont réussi à détecter le changement de tonalité avec 64 % de réponses correctes. Il est intéressant de noter que les auteurs ont rapporté des différences selon le degré d'expertise musicale (71 % pour les étudiants en musicologie, 61 % pour les interprètes, 48 % pour les compositeurs). Dans une seconde expérience, les extraits musicaux (\pm 1min.) ont été segmentés en quatre fragments (ABCD) qui ont été réorganisés afin de violer la clôture tonale (CDAB, DCAB, DCBA). Les participants ont répondu pratiquement au hasard (55 %) et ont semblé complètement ignorants de la violation de la structure harmonique par la réorganisation temporelle.

Tillmann, Bigand et Madurell (1998)²⁴ ont choisi une autre méthode pour aborder cette question. Le matériel expérimental était constitué de menuets de seize mesures, de J.-S. Bach, J.-Ch. Bach, Haydn, Mozart et Pleyel, divisés en deux moitiés et transposés dans différentes tonalités plus ou moins proches. La première moitié se terminait soit par

²² Nicholas Cook, « The perception of large-scale tonal closure », *Music Perception*, Vol. 5, 1987, 197-205.

²³ Elizabeth West Marvin et Alexander Brinkman, « The effect of Modulation and Formal Manipulation on Perception of Tonic Closure by Expert Listeners », *Music Perception*, vol. 16-4, 1999.

²⁴ Barbara Tillmann, Emmanuel Bigand et François Madurell, « Local versus global processing of harmonic cadences in the solution of musical puzzles », *Psychological Research*, n° 61, 1998, p. 157-174.

une demi-cadence dans le ton principal, soit par une cadence parfaite dans le ton de la dominante, soit par une cadence parfaite dans le ton principal. La deuxième moitié se terminait toujours par une cadence parfaite dans le ton principal. Les participants (musiciens, mélomanes et non-musiciens) ont eu à joindre les moitiés de menuets de façon cohérente en fonction des tonalités et des cadences. L'hypothèse était que si la cadence est traitée localement, elle devrait être interprétée par les participants comme une cadence parfaite, si elle est traitée globalement, elle devrait être interprétée comme une demi-cadence. Dans le premier cas, les participants ne devraient pas être capable de replacer les moitiés dans le bon ordre. Les résultats ont indiqué que les auditeurs (musiciens et non musiciens) n'ont pas réussi à reconstruire correctement les menuets lorsque la première moitié se terminait par une cadence parfaite dans le ton de la dominante. Ils ont aussi mis plus de temps pour réaliser la tâche. Les auteurs ont conclu que le traitement local des structures harmoniques prévaut sur le traitement global et que les auditeurs ne semblent pas capables d'intégrer ces structures locales dans une structure globale. Cependant, il faut noter un effet de l'expertise musicale car les performances des musiciens ont été supérieures à celles des non musiciens, cela autant pour les temps de réponse que pour la tâche de reconstruction des menuets. De plus, les auteurs ont observé que les musiciens ont écouté et comparé plus souvent les extraits. Les musiciens ont donc expérimenté les mêmes difficultés perceptives, mais à un moindre degré.

3.2 - PARADIGME DES JUGEMENTS SUBJECTIFS SUR DIFFERENTES ECHELLES

H. Gotlieb et V. Konecni (1985)²⁵ ont étudié l'impact de manipulations de la structure d'une pièce sur le plaisir, l'intérêt et l'émotion. Le matériel expérimental employé était constitué de trois arrangements différents des *Variations Goldberg* de Bach : l'enregistrement original de G. Gould (1955) ; une version dans laquelle les variations ont été aléatoirement mélangées (les arias étant maintenues dans leurs positions originales) ; une version comprenant le même ordre aléatoire utilisé dans la version 2, mais avec les arias placées après les variations 14 et 15. Les participants (non musiciens) ont dû évaluer la pièce à l'aide de quinze échelles bipolaires²⁶. Sur aucune des échelles, excepté l'échelle froid/chaud, la manipulation de structure n'a influencé l'appréciation des participants.

Une deuxième étude a prolongé cette investigation en se focalisant uniquement sur l'organisation en groupes de trois variations. Les auditeurs n'avaient donc plus le contexte global de la pièce. Deux triplettes ont été utilisées (variations 1-3 et 16-18). Les sujets ont entendu trois pièces : les deux triplettes originales et deux triplettes dont l'ordre interne a été bouleversé (1, 3, 2 et 18, 17, 16). Les participants (non musiciens) ont dû estimer les pièces à l'aide de huit échelles bipolaires. Les résultats n'ont montré aucun effet des modifications de structure. Il semblerait, d'après les auteurs, que le plaisir produit par une pièce musicale résulte de quelque chose d'autre que la structure,

²⁵ Heidi Gotlieb et Vladimir J. Konecni, « The Effect of Instrumentation, Playing Style, and Structure in the *Goldberg Variations* by Johann Sebastian Bach », *Music Perception*, Vol. 3, 1985, p. 87-102.

²⁶ À titre d'exemple : précis/imprécis, laid/beau, plaisant/déplaisant, simple/complexe, froid/chaud, excitant/calmand, étonnant/pas étonnant, etc.

contrairement à ce qu'enseignent les théories de la musique. Les auteurs ont alors préconisé « qu' un plus grand degré de circonspection, de modération, et d'humilité chez les musicologues et les critiques déclarant souvent des généralités — purement spéculatives en réalité — serait une conséquence bienvenue de ce type de recherche en psychologie de la musique »²⁷ !

Karno et Konecni (1992)²⁸ ont appliqué la même méthode en manipulant les relations entre les différentes sections du premier mouvement de la *Symphonie n° 40 en sol mineur* K. 550 de Mozart. Les jugements subjectifs sur le même type d'échelles n'ont dégagé aucune différence significative entre l'original et les versions modifiées. Selon les auteurs, ces résultats remettent en cause définitivement l'impact perceptif des structures musicales pour l'auditeur. Cependant un examen plus poussé des résultats montre que l'évaluation des participants tend systématiquement à favoriser les versions originales et que les auteurs ont minimisé ces résultats. Bien que ces différences n'aient pas atteint un niveau statistiquement important, elles suggèrent néanmoins que la sensibilité des auditeurs à la forme globale était très faible, mais pas totalement nulle.

Bigand et Tillmann (1996)²⁹ ont infligé des modifications de structure encore plus fortes à des pièces pour piano de Bach, Mozart et Schönberg³⁰. Chaque pièce a été segmentée et réorganisée en plusieurs fragments selon trois critères : l'ordre d'apparition des sections a été inversé, les transitions entre fragments ont été réalisées sans coupure acoustique audible, les fragments duraient environ six secondes. Après une première écoute de familiarisation, les participants (non musiciens) ont jugé les versions originales ou rétrogrades sur des échelles bipolaires (plus particulièrement l'expressivité et de la cohérence). La pièce de Mozart a évoqué des sentiments plus positifs que celles de Bach et Schönberg, tandis que la pièce de Schönberg a évoqué des sentiments de dynamisme plus forts que les pièces de Mozart et de Bach. La pièce de Schönberg, cependant, a été jugée davantage bizarre et mystérieuse que celles de Bach et de Mozart. La destruction de la cohérence au niveau global n'a pas affecté les jugements d'expressivité et de cohérence. À la fin de l'expérience, les participants devaient identifier quelle version (originale ou inversée) ils avaient écouté. Pour la version originale, 77 % des sujets ont répondu correctement. Seulement 43 % des sujets ayant écouté la version inversée l'ont identifiée en tant que telle. Seuls les auditeurs ayant écouté la version originale ont répondu significativement au-dessus du hasard. Les auteurs ont expliqué ces résultats en faisant appel à la différence entre musique et langage. Si l'on opère des modifications semblables dans l'ordre des phrases d'un texte, la compréhension du texte diminue fortement. La raison en est que le texte narratif se rapportant à une réalité externe, la désorganisation des phrases va provoquer des incongruités entre l'organisation des structures d'événement dans le monde réel et l'arrangement de ces événements dans le texte. En revanche, la musique ne se rapportant à aucune réalité externe, réordonner les

²⁷ Heidi Gotlieb et Vladimir Konecni, *op. cit.*, p. 98 (notre traduction).

²⁸ Mitchell Karno et Vladimir Konecni, « The effect of structural interventions in the First Movement of Mozart's Symphony in G Minor K. 550, on aesthetic preference », *Music Perception* Vol. 10, 1992, p. 63-72.

²⁹ Emmanuel Bigand et Barbara Tillmann, « Does Formal Musical Structure Affect Perception of Musical Expressiveness ? », *Psychology of Music*, n° 24, 1996, p. 3-17.

³⁰ La gigue de la première *Suite française* en ré mineur BWV 812 de J.-S. Bach, l'allegretto de la *Sonate en sib* majeur KV 570 de W. A. Mozart et la gigue de la *Suite pour piano* op. 25 d'A. Schönberg.

fragments d'une pièce musicale ne peut pas affecter drastiquement l'expression ressentie par les auditeurs.

3.3 - PARADIGME DE REPOSE CONTINUE A LA TENSION ET A LA SIMILARITE PERCEPTIVE

C. Krumhansl (1996)³¹ a testé la perception de la rhétorique musicale à différentes échelles, non seulement avec une tâche de segmentation, mais aussi avec un nouveau type de tâche : le pistage de la tension musicale en temps réel. Les participants avaient à manipuler un curseur pour indiquer le degré de tension produit par la *Sonate* pour piano en *mi* bémol majeur K. 282 de Mozart. L'une des conclusions intéressantes de l'étude est que, dans la structure perçue, la fin des grandes sections est corrélée avec les pics de tension (et les ralentissements de tempo). Les pics de tension tendent à être asymétriques : la tension croît progressivement, puis diminue rapidement. Les variations locales de tension co-varient avec les contours mélodiques et les dynamiques, tandis que les variations de tension à grande échelle sont associées aux instabilités harmoniques et aux changements de tonalité.

Cette tâche a été employée différemment dans une étude de S. McAdams et ses collaborateurs (2004)³². L'étude a exploré les aspects structuraux et affectifs de l'expérience temporelle à l'écoute de *The Angel of Death* lors des créations française (à Paris) et américaine (à San Diego, Californie). Les réponses des sujets ont été enregistrées en temps réel à l'aide de boîtiers munis d'un curseur reliés à un ordinateur. Un groupe de sujets devaient suivre la progression musicale en indiquant son sentiment de familiarité (reconnaissance d'éléments déjà entendus dans la pièce). Un autre groupe devait indiquer son ressenti par rapport à la force expressive dégagée par la musique. Le but de l'étude était donc de pister l'évolution temporelle de l'écoute musicale à travers des processus liés à la mémoire et aux émotions. La stratégie formelle de la pièce avait été conçue en fonction de contraintes déterminées par les psychologues pour les expérimentations projetées³³. Selon les résultats, l'évolution des réponses sur la dimension de familiarité a été assez précisément inférée par la structure formelle de la pièce et le matériau thématique. Bien que le profil des réponses émotionnelles

³¹ Carol Krumhansl « A Perceptual Analysis of Mozart's Piano Sonata K. 282 : Segmentation, Tension, and Musical Ideas », *Music Perception*, vol. 13, n° 3, 1996, p. 401-432.

³² Steve McAdams, B. Vines, S. Vieillard, B. Smiths et R. Reynolds, « Influences of Large-Scale Form on Continuous Ratings in Response to a Contemporary Piece in a Live Concert Setting », *Music Perception*, Special issue Creation and Perception of a Contemporary Musical Work: *The Angel of Death* Project, vol. 22, 2004, p. 297-350.

³³ La pièce est constituée de deux moitiés, nommées *Section* (S) et *Domaine* (D), qui ont en commun les mêmes matériaux thématiques, présentés dans une même chronologie mais traités différemment. Un premier niveau de transformation concerne l'instrumentation : les thèmes joués par le pianiste dans *Section*, le sont par l'ensemble instrumental dans *Domaine* et vice-versa. Un deuxième niveau concerne le type d'écriture : dans *Section* les frontières des éléments thématiques sont clairement délimitées, alors que dans l'identité des éléments thématiques de *Domaine* est remise en cause à plusieurs niveaux. Enfin, la pièce a été entendue dans deux versions suivant l'ordre d'apparition des moitiés (S-D ou D-S). Quelle que soit la version, une partie électronique commence peu avant la fin de la première partie et intervient pendant toute la deuxième partie.

corresponde globalement aux sections de la pièce, cette dimension n'a pas montré une relation aussi précise à la forme. Elle a été beaucoup plus dépendante de variables externes à la pièce telles que l'interprétation du soliste et de l'ensemble, le public, etc. En ce qui concerne les performances de mémoire, les moyennes des profils de familiarité ont été globalement plus élevées dans la deuxième moitié de la pièce (les auditeurs ont reconnu les thèmes entendus dans la première moitié de la pièce). Le choix de la version (S-D ou D-S) a eu des effets plutôt surprenants. L'hypothèse était que la version S-D faciliterait la perception et serait préférée par les auditeurs. Cependant, les résultats se sont avérés inverses. La version D-S a eu une structure perceptive beaucoup plus claire et plus stable que la version S-D. Ainsi le passage d'une structure diffuse à une structure plus nette semble satisfaire davantage l'expérience perceptive dans son aspect dynamique. Un autre résultat intéressant de cette étude a consisté à examiner dans quelle mesure la relation entre les dimensions de familiarité et d'émotion a contribué à former l'expérience de l'auditeur. Si l'on considère la théorie de Meyer qui postule que la violation des attentes musicales joue un rôle important pour les réponses émotionnelles, la force émotionnelle devrait croître alors que la familiarité décroît (apparition d'un nouvel élément ou d'un contraste). Cependant, la corrélation négative attendue entre les deux dimensions a été très faible. La relation entre attente et réponse émotionnelle ne semble pas avoir joué un rôle crucial dans la réception de cette pièce.

CONCLUSION

Plusieurs conclusions peuvent être tirées de cet aperçu de la littérature sur la perception des formes musicales. Dans toutes les études portant sur les processus de groupement, les structures perceptives obtenues par segmentation du continuum musical correspondent bien aux structures rhétoriques conçues par les compositeurs. Peu de différences entre musiciens et non-musiciens ont été observées dans la manière de segmenter et de grouper les structures musicales et ceci quels que soient le style et l'époque des pièces auditionnées (tonales et atonales). Les facteurs qui infèrent les groupements sont, en revanche, largement dépendant des pièces, des styles et des époques. Certains de ces facteurs correspondent aux règles préférentielles de la TGMT, mais d'autres facteurs apportent leur contribution suivant le contexte (figures thématiques, traits de surface, changements de densité ou de texture, etc.). D'une manière générale, le retour ou la répétition d'un élément précédemment entendu semble être un facteur important pour le groupement. Certains auteurs mentionnent parfois les effets de primauté ou de récence comme facteurs importants pour la mémorisation des matériaux. Les éléments entendus en début et en fin d'écoute sembleraient être catégorisés et mémorisés plus facilement. Il semble donc que les tâches de segmentation révèlent que d'une part les processus de traitement sont largement partagés par les experts et les novices et d'autre part que les structures perçues correspondent plutôt bien aux formes écrites.

Si les résultats semblent converger en ce qui concerne les processus de groupement, la reconnaissance et la localisation dans le temps d'éléments musicaux semblent poser plus de problèmes aux auditeurs. Ces tâches, qui nécessitent des processus de jugements de similarité, de catégorisation et de rappel en mémoire, semblent dépendre plus de la familiarité avec les styles et les pièces. Lorsqu'il y a absence de familiarité avec le

matériau musical, les jugements de similarité sont effectués en fonction de caractéristiques musicales de surface plutôt qu'en fonction de caractéristiques structurelles ou formelles (Lamont et Dibben, 2001)³⁴. Il a été également montré (Pollard-Gott, 1983)³⁵ que des écoutes répétées améliorent la perception de la similarité sur la base de structures profondes et non de caractéristiques de surface. L'exposition répétée à la musique favorise les musiciens qui semblent apprendre et intégrer plus rapidement les structures musicales. Une meilleure familiarité avec le matériau devrait donc favoriser la représentation mentale de l'articulation temporelle et réduire sensiblement les capacités de stockage nécessaires à la mémorisation (que ce soit pour la musique tonale ou atonale).

Un certain nombre d'études pointent la difficulté d'intégrer les structures à grande échelle. Les auditeurs, qui montrent de réelles capacités à percevoir les fonctions harmoniques de la musique tonale à une échelle locale, ne semblent pas les intégrer lorsque le contexte temporel dépasse 30 secondes. Les ordonnancements aberrants de pièces réorganisées dans leur structure formelle ne semblent avoir aucun effet sur le ressenti des auditeurs en termes de cohérence et d'appréciation esthétique. Ces résultats provocateurs remettent en question autant les conceptions formelles qui s'appuient sur le concept de prolongation (Schenker, Lerdahl et Jackendoff) que celles qui ont une vision plus architecturale. Le rôle perceptif des hiérarchies tonales, si déterminant à un niveau local, serait-il surévalué pour les structures à grande échelle ? Bigand et Tillmann expliquent ce paradoxe en considérant qu'à un niveau local, les unités musicales à petite échelle sont si riches pour l'expérience esthétique que le traitement de plus grandes unités musicales n'a aucun besoin crucial à combler.

Cependant, une autre interprétation peut être avancée. À l'échelle locale de la syntaxe, les processus de traitement sont automatiques et implicites, alors qu'au niveau global de la forme la compréhension des structures nécessite une activité consciente et active de la part de l'auditeur. À cela, il faut ajouter que les structures syntaxiques appartiennent, dans une culture donnée, à un langage prédéfini avec son lexique et ses règles appris implicitement par les auditeurs par simple exposition à la musique. L'intégration formelle, au contraire, consiste à saisir une organisation nouvelle à chaque œuvre et à chaque moment d'une œuvre. Cette organisation est dans une certaine mesure imprévisible, d'autant plus que les schémas censés décrire les formes musicales sont souvent triviaux si on les compare aux produits de l'imagination des compositeurs. Comme l'a montré C. Rosen à propos de la forme sonate³⁶, la manifestation des œuvres est bien plus complexe et bien plus diversifiée que ce que laissent croire les schémas. Il apparaît donc que l'expérience perceptuelle des formes n'est pas si normative qu'il n'y paraît et qu'elle laisse une large place à la subjectivité de l'auditeur. Cette expérience est en tout cas beaucoup moins contrainte qu'à un niveau local. Il en résulte que l'expérience esthétique et émotionnelle des auditeurs peut ne pas être bouleversée par des réorganisations drastiques des structures formelles. Cette interprétation semble être

³⁴ Alexandra Lamont, et Nicola Dibben, « Motivic structure and the perception of similarity », *Music Perception*, Vol. 18-3, 2001, p. 245-274.

³⁵ Lucy Pollard-Gott, « The emergence of thematic concepts in repeated listening to music », *Cognitive Psychology*, n°15, 1983, p. 66-94.

³⁶ Charles Rosen, *Sonata Forms*, New York, London, W. W. Norton, 1988.

appuyée par un ensemble récent d'expériences menée au LEAD³⁷ dans lesquelles des pièces choisies dans des styles très divers (rock, jazz, musique contemporaine, etc.) ont été complètement réorganisées dans leur structure formelle. Les auditeurs ont relativement bien perçu l'incohérence des pièces réorganisées tout en ne montrant aucune différence dans leur appréciation esthétique des pièces cohérentes et incohérentes. Cette différence entre perception et appréciation esthétique n'est, après tout, pas si étonnante. Il existe de nombreux exemples de « plasticité » formelle comme certaines musiques de tradition orale (par exemple en Roumanie) ou les pots-pourris très en vogue aux XVIII^e et XIX^e siècles. À cela il faut ajouter qu'une grande partie de l'activité artistique du XX^e siècle a souvent franchi les limites de la causalité que ce soit dans les arts plastiques (cubisme, dadaïsme, surréalisme, art conceptuel, *happenings*, etc.) ou la littérature (Soupault, Joyce, Beckett, Burroughs, Borges, etc.). La musique savante du XX^e siècle n'est d'ailleurs pas en reste avec les formes mobiles ou indéterminées, certaines pièces de musique minimaliste, ainsi que les formes à déroulement non linéaire inspirées par les théories scientifiques (théorie des catastrophes, théorie du chaos, etc.). Ces nouveaux types de cohérence non causale ont sans doute des répercussions sur notre façon de percevoir et d'apprécier la musique, rendant moins « étranges » les structures temporelles non causales.

Les études sur la perception des structures à grande échelle remettent moins en cause la validité perceptive des formes musicales que les conceptions académiques des formes, qu'elles se fondent sur une structuration hiérarchique ou sur une architectonique normative. Il est possible que l'intégration à plus haut niveau des événements perçus dans les courtes fenêtres perceptives ne soit pas contrainte par une quelconque hiérarchie prédéterminée, mais dépende plus de l'interaction entre le sujet et le contexte. Sans minimiser l'impact perceptif de schémas abstraits encodés (répétition, contraste, alternance, tension/détente, etc.), la perception des formes musicales semble plus dépendre de stratégies adaptatives en fonction de la familiarité avec la musique, de l'acquisition de connaissances abstraites, de la complexité des structures, du contexte d'écoute, etc. Dès lors la forme musicale perçue est moins une information donnée à reconnaître ou à déchiffrer qu'un processus d'apprentissage et de recreation à partir d'une succession de « mises en formes », plus ou moins relié aux intentions formelles du compositeur.

³⁷ Etude en cours, non encore publiée.