



**HAL**  
open science

# Musique et correspondances transmodales : Une étude de l'imaginaire multisensoriel à l'œuvre dans l'écoute et la pratique musicales

Théo de Pinho

► **To cite this version:**

Théo de Pinho. Musique et correspondances transmodales : Une étude de l'imaginaire multisensoriel à l'œuvre dans l'écoute et la pratique musicales. Art et histoire de l'art. 2022. dumas-03720572

**HAL Id: dumas-03720572**

**<https://dumas.ccsd.cnrs.fr/dumas-03720572>**

Submitted on 12 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Département Arts  
Centre d'Étude des Arts Contemporains

**Musique et correspondances transmodales :**  
Une étude de l'imaginaire multisensoriel à l'œuvre  
dans l'écoute et la pratique musicales

Théo DE PINHO

Mémoire de Master 1

Master mention Arts

Parcours Musique, interprétation, invention

Directeur : Christian HAUER, Professeur des Universités

Année universitaire 2021-2022



# Remerciements

Je tiens à remercier Christian Hauer de m'avoir guidé dans la rédaction de ce mémoire. La bienveillance et la pertinence de ses conseils ont toute ma gratitude.

Je remercie également mes professeurs Francis Courtot et Iván Adriano Zetina pour leur cours de méthodologie.

Merci à ma famille pour son soutien et ses conseils.



# Table des matières

<b>Introduction.....</b>	<b>7</b>
<b>1. Cadre théorique.....</b>	<b>9</b>
1.1. État de l'art.....	9
1.2. Problématique.....	11
1.3. Méthodologie générale.....	12
<b>2. Introduction aux correspondances transmodales.....</b>	<b>15</b>
2.1. Histoire de l'étude des correspondances transmodales.....	15
2.2. Définitions.....	16
2.2.1. Modalité sensorielle, stimulus, attribut sensoriel.....	16
2.2.2. Correspondance transmodale, congruence.....	18
2.3. Quelques distinctions importantes.....	19
2.3.1. Correspondances transmodales et synesthésie.....	19
2.3.2. Correspondances transmodales et correspondances intramodales.....	20
2.3.3. Correspondances transmodales et intégration multisensorielle.....	21
2.4. Universalité des correspondances transmodales.....	21
2.5. Récapitulatif.....	24
<b>3. Les correspondances transmodales utilisables en musique.....</b>	<b>27</b>
3.1. Méthodologie.....	27
3.2. Ouïe – Vue.....	31
3.3. Ouïe – Toucher.....	33
3.4. Ouïe – Goût.....	34
3.5. Ouïe – Odorat.....	35
3.6. Récapitulatif.....	38
<b>4. Utiliser les correspondances transmodales en musique.....</b>	<b>41</b>
4.1. Méthodologie.....	41
4.2. Pourquoi ?.....	42
4.3. Évoquer.....	43

4.3.1. Où évoquer ?.....	43
4.3.2. Comment évoquer ?.....	44
<b>4.4. Lier.....</b>	<b>45</b>
4.4.1. Où lier ?.....	45
4.4.2. Comment lier ?.....	45
<b>4.5. Comprendre.....</b>	<b>48</b>
4.5.1. Où comprendre ?.....	48
4.5.2. Comment comprendre ?.....	48
<b>4.6. Quand ?.....</b>	<b>50</b>
4.6.1. Le problème de l'imprécision.....	50
4.6.2. Le problème de l'incohérence.....	52
4.6.3. Le problème de la concurrence des effets.....	57
<b>4.7. Récapitulatif.....</b>	<b>58</b>
<b>Conclusion.....</b>	<b>61</b>
<b>Table des figures.....</b>	<b>63</b>
<b>Table des tableaux.....</b>	<b>64</b>
<b>Références bibliographiques.....</b>	<b>65</b>

## **Introduction**

Lorsque nous écoutons de la musique, nous avons tendance à imaginer des objets/événements. Nous imaginons la forme de ces objets/événements, leur taille, leur hauteur, leur poids, et parfois même leur goût et leur odeur. Mais ces « images », ces stimuli non auditifs que nous imaginons ne se produisent pas de façon aléatoire : ils suivent certaines règles. Par exemple, pour un grand nombre d'humains, les sons aigus font souvent penser à des objets/événements hauts, pointus, fins et légers, et les sons graves à des objets/événements bas, arrondis, épais et lourds. Un jour, je me suis posé les questions suivantes : est-il possible de faire une liste de ces règles ? Sont-elles identiques pour tout le monde ? Est-il possible de les utiliser de façon systématique en musique ? Ce sont ces questions que j'ai décidé de traiter dans ce mémoire. Au cours de mes recherches, j'ai découvert que les règles en question portent un nom précis : « correspondances transmodales ». En un mot, donc, ce mémoire traitera des *liens entre la musique et les correspondances transmodales*.

Dans le terme « musique », nous incluons tous les domaines liés à la production de musique (composition, interprétation, improvisation). Nous parlerons aussi des liens entre les correspondances transmodales et l'analyse musicale. En principe, l'analyse musicale n'est pas incluse dans la définition du terme « musique ». Cela dit, étant donné que parler conjointement de « musique et d'analyse musicale » alourdit excessivement le texte, nous avons inclus l'analyse musicale dans notre définition du terme « musique ». Ainsi, ce terme regroupera à la fois la composition, l'interprétation, l'improvisation et l'analyse musicales.

Par « correspondance transmodale » (CT), nous désignons toute association intuitive entre un attribut sensoriel perçu par un sens (comme la hauteur sonore, la taille visuelle ou le poids tactile) et un attribut sensoriel perçu par un autre sens. L'adjectif « transmodales » vient du fait que ces correspondances *transcendent les modalités* sensorielles, c'est-à-dire les sens. Un exemple de CT est celle associant la hauteur sonore (perçue par l'ouïe) et la hauteur visuelle (perçue par la vue). Cette CT fait que la plupart des personnes associent intuitivement les sons aigus à des objets/événements visuellement hauts, et les sons graves à des objets/événement visuellement bas. Un autre exemple est la CT entre le tempo (perçu par l'ouïe) et la taille tactile (perçue par le toucher), suivant laquelle la plupart des personnes associent intuitivement les sons rapides à des objets/événements tactilement petits, et les sons lents à des objets/événements tactilement gros.

Ce mémoire n'abordera pas toutes les CT. En effet, étant donné que la musique est surtout perçue par le sens de l'ouïe, nous nous concentrerons sur les CT qui impliquent le sens de l'ouïe. Cela

inclut les CT audio-visuelles (reliant l'ouïe et la vue), audio-tactiles (reliant l'ouïe et le toucher), audio-gustatives (reliant l'ouïe et le goût) et audio-olfactives (reliant l'ouïe et l'odorat). Ainsi, ce mémoire porte sur les liens pouvant exister entre ces correspondances et la musique (en un sens qui inclut l'analyse musicale).

# 1. Cadre théorique

Dans cette partie, nous dresserons l'état de l'art, poserons la problématique du mémoire, et exposerons la méthodologie générale.

## 1.1. État de l'art

Voici comment nous avons effectué nos recherches et ce que nous avons trouvé. Dans un premier temps, nous avons cherché le terme utilisé dans la littérature scientifique pour désigner les associations qui nous intéressaient. Ayant trouvé que ce terme était « *crossmodal correspondences* » (« correspondances transmodales »), nous avons ensuite cherché des références utilisant cette expression. Nous avons ainsi trouvé un certain nombre d'articles. Parmi ces articles, nous avons d'abord lu ceux qui traitaient des CT d'un point de vue général en vue de bien comprendre ce phénomène. Les principaux que nous avons utilisés sont les suivants :

- SPENCE, « Crossmodal correspondences: A tutorial review », 2011. Cet article offre un état de l'art sur les CT. Voici en quoi il nous a semblé pertinent : il tente de définir les CT ; il retrace l'histoire de l'étude des CT ; il passe en revue les CT audio-visuelles établies *via* la tâche de classification accélérée<sup>1</sup>, une des expériences de psychologie utilisées pour mettre en évidence des CT ; il traite du rapport entre les CT et la synesthésie.
- MARKS, « Cross-modal interactions in speeded classification », 2004. Cet article traite des CT et de leur mise en évidence dans la tâche de classification accélérée. Il nous a informés sur le concept d'attribut sensoriel (qui désigne une caractéristique d'un stimulus) ainsi que sur l'effet de congruence (l'effet associé à la présence de correspondances transmodales).
- PARISE & SPENCE, « Audiovisual cross-modal correspondences in the general population », 2013. Cet article traite des CT audio-visuelles. Mais surtout, il offre une introduction aux CT et aborde la question de leur universalité.

Par la suite, nous avons davantage ciblé nos recherches. Nous avons cherché les références qui listent le plus grand nombre de CT impliquant l'ouïe que possible. Concernant les CT audio-visuelles, nous nous sommes principalement reposés sur l'article de Spence (voir ci-dessus), qui semblait déjà offrir une synthèse assez complète. Concernant les CT audio-tactiles, nous nous

---

<sup>1</sup> En anglais : *speeded classification task*.

sommes surtout appuyés sur l'article de Walker (voir ci-dessus). Pour ce qui est des CT audio-gustatives et audio-olfactives, les principales références que nous avons utilisées sont les suivantes :

- KNÖFERLE & SPENCE, « Crossmodal Correspondences between Sounds and Tastes », 2012. Cet article aborde plusieurs études empiriques visant à mettre en évidence des CT audio-gustatives.
- CRISINEL & SPENCE, « A Fruity Note: Crossmodal Associations between Odors and Musical Notes », 2012. Cet article est une étude empirique visant à mettre en évidence des CT audio-olfactives.

Enfin, nous avons cherché des références qui abordent l'utilisation des CT en musique. Les principales que nous avons utilisées sont les suivantes :

- GEORGIU, *Beyond narrative: A cross-modal approach to soundtrack composition*, 2014. Cette thèse tente d'établir une méthode systématique pour composer de la musique pour des films non narratifs. Cette méthode repose sur l'idée que les éléments visuels (présents dans la vidéo) peuvent être traduits en éléments sonores (pour composer la musique) par le biais de correspondances audio-visuelles. Cette thèse nous a informés sur les différentes façons de lier une musique à un médium non-auditif (vidéo, mais aussi image, danse, dispositif olfactif, etc.) *via* l'utilisation de CT.
- EITAN & TAMIR-OSTROVER, « The Sound of an Endless Column: How Music Imagines Unimaginable Space », 2019. Cet article traite des incohérences entre certaines CT et de la manière dont on peut en tirer profit en composition et en analyse musicale.

Alors que nous achevons la rédaction de ce mémoire, nous venons de prendre connaissance d'une communication scientifique de Chelsea Komschlies, dont l'enregistrement a été mis en ligne<sup>2</sup>. Nous n'avons pas pu intégrer cette communication à notre travail, mais ce dernier aurait bien pu en bénéficier, car les deux travaux ont le même sujet : les liens entre la musique et les CT. Il s'avère que Komschlies écrit actuellement une thèse de doctorat sur ce sujet à l'université McGill (École de musique Schulich) sous la direction du compositeur Jean Lesage. Son titre est : *How Music Encodes Shared Crossmodal Meaning: A Science-Based Guide for Composers, Analysts, and Listeners*. Komschlies compte la soutenir autour d'avril 2025. Cette thèse pourrait offrir une synthèse approfondie sur le sujet des liens entre la musique et les CT. Nous invitons vivement toute personne intéressée par le sujet à consulter la communication de Komschlies, ainsi que sa thèse lorsqu'elle sera soutenue et mise à disposition des chercheur·se·s.

---

2 KOMSCHLIES, Chelsea, « Research Alive: Crossmodal correspondences in composing and listening to music », université McGill (École de musique Schulich), 16 mars 2022, disponible *via* <https://www.youtube.com/watch?v=qpmJTaSiGKE> (consulté le 02/06/2022).

## 1.2. Problématique

Il semble que les musicien·ne·s — qu'ils soient compositeur·trice·s, interprètes, improvisateur·trice·s ou encore analystes — comprennent les CT et les utilisent souvent. Néanmoins, cette compréhension et cette utilisation sont souvent intuitives et inconscientes. Par conséquent, lorsque l'on est dans le milieu de la musique, il est *a priori* difficile d'avoir une image claire des liens entre la musique et les CT.

Consulter la littérature scientifique sur le sujet permet de surmonter partiellement cette difficulté. En effet, depuis des années, des chercheur·se·s étudient les liens entre la musique et les CT et ont déjà clarifié une partie des intuitions des musicien·ne·s sur ce sujet. À l'heure actuelle, la recherche a établi un certain nombre de CT utilisables en musique et a traité de leur utilisation musicale.

Toutefois, à notre connaissance, il n'existe pas de publications offrant une synthèse des connaissances sur les liens entre la musique et les CT. En d'autres termes, il n'existe pas de travaux établissant une liste de toutes les CT utilisables en musique et traitant de tous les aspects de leur utilisation en musique. Ce mémoire vise, non pas à constituer, mais au moins à contribuer à un tel travail<sup>3</sup>. La problématique sera donc la suivante : quelles sont les correspondances transmodales utilisables en musique, et quels sont les tenants et les aboutissants de leur utilisation musicale ?

Il est important d'insister sur la chose suivante. Lorsque nous parlons de « musique », nous parlons des domaines liés à la production (ou création) musicale (composition, interprétation, improvisation...), mais aussi du domaine de l'analyse musicale. Bien sûr, l'inclusion de l'analyse musicale dans la catégorie « musique » est contestable. Elle n'a ici qu'un but pratique : celui d'alléger la lecture en évitant l'emploi d'expressions du type « en musique et en analyse musicale ».

Revenons à la problématique. Celle-ci est composée de deux questions distinctes. La première question — quelles sont les correspondances transmodales utilisables en musique ? — est celle de la *nature* des CT utilisables en musique. C'est la question du « quoi ? ». Bien que la musique puisse impliquer plusieurs sens, et que les CT utilisables en musique pourraient inclure, par exemple, les CT tactilo-visuelles (reliant le toucher et la vue), ce mémoire se limitera aux CT impliquant le sens de l'ouïe. En effet, l'ouïe semble être le sens sur lequel la population se repose le plus pour faire l'expérience de la musique. La question sera donc :

---

3 Une contribution plus complète est actuellement en cours de rédaction. Il s'agit de la thèse de Chelsea Komschlies (cf. état de l'art).

1. **Quoi ?** Quels attributs (ou stimuli) auditifs ont tendance à être associés à quels attributs (ou stimuli) non auditifs ?

La deuxième question — quels sont les tenants et les aboutissants de l'utilisation musicale des CT utilisables en musique ? — recouvre plusieurs sous-questions :

2. **Pourquoi ?** Dans quel but les CT peuvent-elles être utilisées musicalement ?
3. **Où ?** Dans quels domaines les CT peuvent-elles être utilisées musicalement ?
4. **Comment ?** De quelle façon les CT peuvent-elles être utilisées musicalement ?
5. **Quand ?** Dans quelles circonstances les CT peuvent-elles être utilisées musicalement ?

Répondre à ces cinq questions, et donc à la problématique, pourrait avoir deux intérêts :

- Contribuer à une synthèse des connaissances sur les liens entre la musique et les CT, cette contribution ayant en soi une valeur épistémique (c'est-à-dire qu'elle contribue à la connaissance).
- Aider les musicien·ne·s désirant utiliser les CT en musique à mieux les comprendre et à maîtriser davantage leur utilisation. En effet, il est possible que leur compréhension et leur maîtrise actuelles des CT, qui semblent surtout intuitives, puissent être améliorées par un apprentissage explicite. Toutefois, ce point n'est pas certain : il se pourrait qu'un apprentissage explicite ne puisse constituer qu'une prise de conscience de principes déjà parfaitement intégrés.

### **1.3. Méthodologie générale**

Étudier les liens entre la musique et les CT suppose de comprendre les CT. C'est pourquoi, dans la première partie, nous procéderons à une introduction à ce phénomène. Cette introduction s'appuiera principalement sur des états de l'art et des dictionnaires de psychologie. Dans un premier temps, nous présenterons un résumé de l'histoire de l'étude des CT. Puis, nous essaierons de comprendre ce que ces associations sont et ce qu'elles ne sont pas. Enfin, nous traiterons la question de l'universalité des CT afin d'avoir une idée plus précise de l'ensemble d'individus (humains ou non) pouvant faire l'expérience de CT en écoutant de la musique.

Dans la deuxième partie, nous tenterons de répondre à la première partie de la problématique. Il s'agira de lister un certain nombre de CT impliquant le sens de l'ouïe en s'appuyant sur des études de psychologie expérimentale. Toutes les paires possibles de sens impliquant l'ouïe seront explorées : ouïe-vue, ouïe-toucher, ouïe-goût et ouïe-odorat.

La troisième et dernière partie traitera de la deuxième partie de la problématique. Il s'agira de traiter différentes questions relatives à l'utilisation des CT en musique : pourquoi ? comment ? où ? et quand ? Ces questions semblent avoir été peu traitées dans la littérature scientifique ; c'est pourquoi cette partie s'appuiera principalement sur des réflexions personnelles.

Pour des questions de praticité, chaque partie se terminera par un récapitulatif résumant les idées principales. Pour les parties 2 et 4, il s'agira d'une liste d'idées hiérarchisées. Pour la partie 3, il s'agira d'un ensemble de quatre tableaux regroupant les CT passées en revue, chaque tableau correspondant à une paire de sens.



## 2. Introduction aux correspondances transmodales

Cette partie consistera à introduire le phénomène des CT. Tout d'abord, nous résumerons l'histoire de l'étude de ce phénomène. Puis, nous donnerons quelques définitions, ce qui nous permettra de distinguer les CT d'autres phénomènes. Enfin, nous tenterons de répondre à la question de l'universalité des CT.

### 2.1. Histoire de l'étude des correspondances transmodales

L'un des premiers documents écrits à traiter des CT est le *Cratyle* de Platon, écrit entre le V<sup>ème</sup> et le IV<sup>ème</sup> siècle AEC. Cependant, l'étude scientifique de ces associations ne se développe pas avant la première moitié du XX<sup>ème</sup> siècle. En 1929, Edward Sapir<sup>4</sup> suggère qu'il existe une association entre les phonèmes /a/ et /i/ et la taille des objets/événements que l'on perçoit. Il observe que la plupart des personnes associent les mots (dénusés de sens) « mal » et « mil » à des objets/événements<sup>5</sup> gros et petits, respectivement.

La même année, Wolfgang Köhler<sup>6</sup> observe que, lorsqu'on montre à des personnes deux formes différentes — l'une arrondie, l'autre angulaire —, et qu'on leur demande d'associer les mots (à nouveau dénués de sens) « maluma » et « takete » à ces formes de la façon qui leur semble la plus naturelle, la plupart d'entre elles associent la forme arrondie au mot « maluma » et la forme angulaire au mot « takete », plutôt que l'inverse. En 2001, Ramachandran et Hubbard<sup>7</sup> répliquent l'expérience de Köhler en modifiant légèrement les formes (figure 1) et en remplaçant les mots « maluma » et

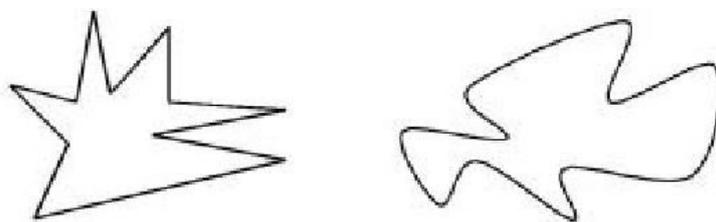


Figure 1. Formes utilisées par Ramachandran & Hubbard (2001) pour démontrer l'existence de l'effet boub/kiki. La majorité de la population associe la forme de gauche au mot « kiki » et la forme de droite au mot « boub ».

4 SAPIR, Edward, « A study in phonetic symbolism », *Journal of Experimental Psychology*, vol. 12, n°3, 1929, pp. 225-239.

5 À partir de maintenant, pour fluidifier la lecture, nous remplacerons l'expression « objet/événement » par le terme « objet ».

6 KÖHLER, Wolfgang, *Gestalt psychology* [1929], New York, Liveright, 1947, pp. 224-225.

7 RAMACHANDRAN, Vilayanur S. & HUBBARD, Edward M., « Synaesthesia--A Window Into Perception, Thought and Language », *Journal of consciousness studies*, vol. 8, n°12, 2001, p. 19.

« takete » par les mots « bouba » et « kiki ». Depuis, cette CT est désignée sous le nom d'effet bouba/kiki. L'effet bouba/kiki fonctionne pour plus de 90% de la population générale<sup>8</sup>.

Les travaux de ce type ont participé du développement de l'étude du *symbolisme phonétique*, qui peut être défini comme le phénomène d'association entre les sons du langage parlé et les significations des mots.

Parallèlement au développement de la recherche sur le symbolisme phonétique, les psychophysiciens ont commencé à étudier celles des CT qui n'impliquent pas les sons du langage parlé<sup>9</sup>. Dans ce mémoire, nous traiterons uniquement ce type de CT.

L'une des premières CT de ce type à avoir été établies empiriquement est celle entre l'intensité sonore et la luminosité visuelle, suivant laquelle la plupart des personnes associent les sons forts aux objets lumineux, et les sons faibles aux objets sombres<sup>10</sup>. Par la suite, la recherche sur les CT s'est beaucoup développée, particulièrement au cours de ces dernières années<sup>11</sup>. Un grand nombre de CT ont été étudiées, et ce entre de nombreuses paires de sens. D'après Spence<sup>12</sup>, il est probable que les CT existent entre toutes les paires possibles de sens.

## 2.2. Définitions

Dans cette sous-partie, nous allons tenter de définir le concept de correspondance transmodale, ainsi que celui de congruence, qui lui est associé. Mais avant cela, il peut être utile de commencer par définir les notions de modalité sensorielle, de stimulus et d'attribut sensoriel.

### 2.2.1. Modalité sensorielle, stimulus, attribut sensoriel

Une *modalité sensorielle*<sup>13</sup> est un sens, c'est-à-dire un moyen de percevoir le monde physique, qu'il s'agisse de notre environnement ou de notre propre corps<sup>14</sup>. La vue, l'ouïe, le toucher, le goût et l'odorat sont les sens les plus communément admis. Cependant, on y ajoute parfois le sens de la

---

8 Certaines communautés linguistiques sont néanmoins totalement insensibles à l'effet bouba/kiki. Pour une tentative d'explication, voir STYLES, Suzy J. & GAWNE, Lauren, « When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter », *I-Perception*, vol. 8, n° 4, 2017.

9 PARISE, C. V. & SPENCE, C., « Audiovisual cross-modal correspondences in the general population », in *Oxford handbook of synaesthesia*, Julia Simner & Edward Hubbard éd., Oxford, Oxford University Press, 2013, p. 798.

10 *Ibid.*

11 SPENCE, Charles, « Simple and Complex Crossmodal Correspondences Involving Audition », *Acoustical Science and Technology*, vol. 41, n°1, p. 6.

12 *Id.*, « Crossmodal correspondences: A tutorial review », *Attention, Perception, & Psychophysics*, vol. 73, n° 4, 2011, p. 975.

13 En anglais : *sensory/sense/stimulus modality*.

14 *S. v.* « sense », in VANDENBOS, Gary R. (éd.), *APA Dictionary of Psychology*, United States, American Psychological Association, 2<sup>nd</sup> édition, 2015, p. 962.

position de nos propres membres dans l'espace (la proprioception) et le sens de l'équilibre (l'équibrioception). Par ailleurs, le sens du toucher est parfois remplacé par trois sens distincts : le sens de la pression, le sens de la douleur (la nociception) et le sens de la température (la thermoception)<sup>15</sup>. Dans ce mémoire, nous adopterons néanmoins la conception classique des cinq sens, dans la mesure où elle est plus simple et suffit pour comprendre les liens entre la musique et les CT.

Un *stimulus* est un évènement, un agent ou une situation, interne ou externe, qui suscite une réponse de la part d'un organisme<sup>16</sup>. Il peut s'agir de la vue d'une maison, d'un son de tonnerre, du goût d'une pomme, de l'odeur d'un parfum, de la sensation d'une caresse ou même de la sensation des battements de notre propre cœur.

On distingue les stimuli simples des stimuli complexes. Un stimulus simple est un stimulus comportant un seul élément ou attribut, tandis qu'un stimulus complexe est un stimulus comportant plusieurs éléments ou attributs<sup>17</sup>. Dans le domaine auditif, un stimulus simple pourrait être un son pur, et un stimulus complexe un extrait d'une pièce de musique.

Un *attribut sensoriel*<sup>18</sup> est une caractéristique d'un stimulus<sup>19</sup>. La hauteur sonore, la luminosité visuelle et la fréquence vibrotactile sont des exemples d'attributs sensoriels. Chaque sens peut capter plusieurs attributs sensoriels. Par exemple, la vue peut capter la luminosité, la taille ou encore la hauteur d'un objet. Inversement, certains attributs sensoriels peuvent être captés par plusieurs sens. Par exemple, la taille peut être captée à la fois par la vue et le toucher. Cependant, il y a des attributs qui ne peuvent être perçus que par un seul sens. Lorsque l'on parle de « taille visuelle », par exemple, on parle de la taille en tant qu'elle est perçue par la vue : par définition, cet attribut ne peut donc être perçu que par la vue. Autre exemple plus simple : le goût amer. Il ne peut être perçu que par le goût<sup>20</sup>.

On distingue les attributs dimensionnels des attributs non dimensionnels<sup>21,22</sup>. Les attributs dimensionnels sont des attributs pouvant être mesurés quantitativement sur une échelle allant de « faible » à « élevé ». Ils incluent notamment les attributs cités plus haut : la hauteur sonore, la luminosité visuelle et la fréquence vibrotactile, qui varient entre son grave et son aigu, entre lumière faible et lumière intense, et entre fréquence faible et fréquence élevée, respectivement. Les CT trai-

---

15 MACPHERSON, Fiona, « Taxonomising the senses », *Philosophical Studies*, vol. 153, n°1, 2011, p. 125.

16 *S. v.* « stimulus », in VANDENBOS, Gary R. (éd.), *APA Dictionary of Psychology*, op. cit., p. 1032 sq.

17 SPENCE, Charles, « Simple and Complex Crossmodal Correspondences Involving Audition », art. cit., p. 6.

18 En anglais : *sensory/perceptual/stimulus attribute/feature/cue* ; le plus souvent : *stimulus attribute*.

19 *S. v.* « perceptual cues », in VANDENBOS, Gary R. (éd.), *APA Dictionary of Psychology*, op. cit., p. 776.

20 Le goût amer est cependant davantage un stimulus qu'un attribut sensoriel.

21 MARKS, Lawrence E., « Cross-modal interactions in speeded classification », in *The Handbook of Multisensory Processes*, Gemma A. Calvert, Charles Spence & Barry E. Stein éd., Cambridge, Mass: MIT Press, 2004, p. 88.

22 WALKER, Peter, « Cross-Sensory Correspondences: A Theoretical Framework and Their Relevance to Music », *Psychomusicology: Music, Mind, and Brain*, vol. 26, n° 2, 2016, p. 104.

tées dans ce mémoire impliquent le plus souvent des attributs dimensionnels. Cependant, il existe des attributs non dimensionnels. La couleur en est probablement un<sup>23</sup>.

## 2.2.2. Correspondance transmodale, congruence

Maintenant que nous avons défini les concepts de modalité sensorielle, de stimulus et d'attribut sensoriel, nous pouvons définir clairement ceux de correspondance transmodale et de congruence.

Selon Spence<sup>24</sup>, une *correspondance transmodale*<sup>25</sup> est une tendance pour un attribut sensoriel dans une modalité sensorielle donnée à être associé à un attribut sensoriel dans une autre modalité sensorielle. Nous proposons une définition similaire, quoique légèrement différente : une correspondance transmodale est une association intuitive entre un attribut sensoriel perçu par un sens et un attribut sensoriel perçu par un autre sens<sup>26</sup>. Pour illustrer cette définition, prenons l'exemple d'une CT mise en évidence chez un grand nombre de personnes : celle reliant la hauteur sonore à la hauteur visuelle<sup>27</sup>. L'existence de cette CT signifie qu'un grand nombre de personnes associent intuitivement la hauteur sonore (perçue par l'ouïe) à la hauteur visuelle (perçue par la vue).

Mais ce type d'information ne suffit pas à offrir une description complète d'une CT. Pour ce faire, il faut aussi répondre à la question suivante : quelles valeurs d'un attribut sont associées à quelles valeurs de l'autre attribut ? Pour ce qui concerne la CT reliant la hauteur sonore à la hauteur visuelle, la question peut être reformulée ainsi : comment les valeurs [sons aigus/sons graves] sont-elles associées aux valeurs [objets visuellement hauts/objet visuellement bas] ? En l'occurrence, il se trouve que les sons aigus sont associés aux objets visuellement hauts, et que les sons graves sont associés aux objets visuellement bas. Mais il est important de faire remarquer que cela aurait pu être l'inverse : les sons aigus auraient pu être associés aux objets visuellement bas, et les sons graves aux objets visuellement hauts. Ainsi, lorsqu'on décrit une CT, il est important de préciser la *façon* dont les valeurs des deux attributs sensoriels sont associées.

La *congruence* est l'état de deux valeurs d'attributs sensoriels associés dans le cadre d'une CT. L'inverse de la congruence est l'incongruence. Ainsi, les sons aigus et les objets visuellement hauts

---

23 SPENCE, Charles, « Crossmodal correspondences: A tutorial review », art. cit., p. 979, note 5.

24 *Id.*, « Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism », *Journal of Consumer Psychology*, vol. 22, n° 1, 2012, p. 37.

25 En anglais : *crossmodal/cross-modal/cross-sensory correspondence/association* ; le plus souvent : *crossmodal correspondence*.

26 Une limite de ces deux définitions est qu'elles omettent le fait qu'une CT peut aussi impliquer des stimuli. En effet, comme nous le verrons dans la partie 3, certaines CT reliant l'ouïe et le goût ou l'ouïe et l'odorat relient un attribut sensoriel et un stimulus, plutôt que deux attributs sensoriels. Toutefois, par souci de simplicité, nous allons ici supposer qu'une CT relie toujours deux attributs sensoriels.

27 *Id.*, « Crossmodal correspondences: A tutorial review », art. cit., p. 976.

(et leurs opposés) sont congruents, tandis que les sons aigus et les objets visuellement bas (et leurs opposés) sont incongruents. L'*effet de congruence* est le nom donné à l'effet présent chez un individu lorsque celui-ci perçoit deux valeurs congruentes d'attributs sensoriels<sup>28</sup>. Cet effet peut être mis en évidence *via* de multiples expérimentations en psychologie cognitive. L'une d'elles est la tâche de classification accélérée<sup>29</sup>. Dans cette tâche, l'effet de congruence se manifeste par le fait que les performances des participant·e·s sont plus rapides et plus souvent exactes lorsque les valeurs d'attributs sensoriels qui leur sont présentées sont congruentes que lorsqu'elles sont incongruentes<sup>30</sup>.

## 2.3. Quelques distinctions importantes

Nous venons d'expliquer ce que sont les correspondances transmodales. Nous allons maintenant expliquer ce qu'elles ne sont pas, en les distinguant de la synesthésie, des correspondances intramodales, et de l'intégration multisensorielle.

### 2.3.1. Correspondances transmodales et synesthésie

La *synesthésie* est une condition dans laquelle la stimulation d'un sens génère simultanément une sensation dans un autre sens<sup>31,32</sup>. Par exemple, lorsque les personnes douées d'une synesthésie son→couleur entendent un son (par exemple, un son de flûte), elles perçoivent simultanément ce son *et* une couleur (par exemple, la couleur bleu clair).

La littérature scientifique a parfois confondu les CT avec la synesthésie<sup>33</sup>. Néanmoins, ces deux phénomènes sont distincts. Voici quelques différences mises en avant par Eitan<sup>34</sup> :

- Avec la synesthésie, la stimulation d'un sens génère une sensation dans un autre sens, tandis qu'avec les CT, la stimulation d'un sens génère une *pensée* (consciente ou inconsciente) d'une sensation dans un autre sens. D'une certaine façon, on pourrait dire que la synesthésie est à la perception ce que les CT sont à l'imagination.

---

28 PARISE, C. V. & SPENCE, C., « Audiovisual crossmodal correspondences in the general population », art. cit., p. 792.

29 En anglais : *speeded classification task*.

30 MARKS, Lawrence E., « Cross-modal interactions in speeded classification », art. cit., p. 87.

31 *S. v.* « synesthesia », in VANDENBOS, Gary R. (éd.), *APA Dictionary of Psychology*, op. cit., p. 1060.

32 En fait, le terme « synesthésie » est aussi utilisé pour décrire une condition dans laquelle la stimulation d'un sens génère simultanément une sensation dans le *même* sens. Par exemple, on parle de « synesthésie graphème→couleur » lorsque la vue de graphèmes (par exemple, de lettres) génère simultanément la vue de couleurs (chaque lettre étant associée à une couleur) ; on voit bien que ces deux attributs sont perçus par un même sens : la vue.

33 DI STEFANO, Nicola, MURARI, Maddalena & SPENCE, Charles, « Crossmodal Correspondences in Art and Science: Odours, Poetry, and Music » in *Olfaction: An Interdisciplinary Perspective from Philosophy to Life Sciences*, Nicola Di Stefano & Maria Teresa Russo éd., Cham: Springer International Publishing, coll. « Human Perspectives in Health Sciences and Technology », vol. 4, 2022, p. 156.

34 EITAN, Zohar, « Musical connections: Cross-modal correspondences » in *The Routledge Companion to Music Cognition*, 1<sup>ère</sup> édition, Ashley Richard & Renee Timmers éd., New York, NY: Routledge, coll. « Routledge Companions », 2017, pp. 218-219.

- La synesthésie ne peut pas être altérée par l'apprentissage, tandis que les CT le peuvent.
- La synesthésie tend à être absolue (indépendante du contexte), tandis que les CT tendent à être relatives (dépendantes du contexte). Par exemple, une personne douée d'une synesthésie son→couleur pourrait en théorie associer la note *ré 5* à la couleur bleu azur. À l'inverse, une personne qui éprouve simplement des CT peut associer les sons aigus aux objets visuellement lumineux, mais elle ne peut pas associer une hauteur sonore précise à un degré précis de luminosité visuelle (du moins lorsque ces deux types de stimuli sont présentés isolément, ou dans des contextes différents).
- La synesthésie affecte un petit nombre d'êtres humains, tandis que les CT en affectent une majorité.
- Les synesthètes sont souvent affecté·e·s par des types très différents de synesthésie, tandis que les CT sont souvent de même nature d'un humain à l'autre.

Malgré les différences entre les CT et la synesthésie, certain·e·s chercheur·se·s ont proposé l'idée que ces deux phénomènes se situent sur un seul et même continuum. La synesthésie ne serait qu'une manifestation inhabituellement intense des CT (figure 2). Bien qu'élégante, cette hypothèse ne fait cependant pas consensus, les données disponibles ne permettant pas de trancher<sup>35</sup>.



Figure 2. Continuum sur lequel pourraient être regroupées les correspondances transmodales et la synesthésie, du phénomène le moins intense (à gauche) au plus intense (à droite).

### 2.3.2. Correspondances transmodales et correspondances intramodales

Les correspondances transmodales ne doivent pas non plus être confondues avec les *correspondances intramodales*. Les correspondances intramodales sont des associations intuitives entre deux attributs sensoriels perçus par un même sens<sup>36</sup>. Par exemple, l'association entre les sons aigus et les objets auditivement hauts (les objets dont l'oreille perçoit qu'ils sont hauts dans l'espace) et entre les sons graves et les objets auditivement bas relève d'une correspondance intramodale. En effet, les

35 SPENCE, Charles, « Crossmodal correspondences: A tutorial review », art. cit., p. 983.

36 EITAN, Zohar, « How pitch and loudness shape musical space and motion », in *The Psychology of Music in Multimedia*, Siu-Lan Tan, Annabel J. Cohen, Scott D. Lipscomb & Roger A. Kendall eds., 1<sup>re</sup> édition, Oxford: Oxford University Press, 2013, p. 166.

attributs à l'œuvre dans cette correspondance — la hauteur sonore et la hauteur spatiale perçue *via* l'ouïe — sont perçus par le même sens — l'ouïe.

Il ne s'agit donc pas d'une correspondance transmodale, car les correspondances transmodales sont des associations intuitives entre deux attributs sensoriels perçus par deux sens *différents*. En revanche, l'association entre les sons aigus et les objets *visuellement* hauts et entre les sons graves et les objets *visuellement* bas relève bien d'une correspondance transmodale, car les attributs à l'œuvre dans cette correspondance — la hauteur sonore et la hauteur spatiale perçue *via* la vue (ou hauteur visuelle) — sont perçus par deux sens différents — l'ouïe et la vue.

### 2.3.3. Correspondances transmodales et intégration multisensorielle

Enfin, les CT ne doivent pas être confondues avec l'*intégration multisensorielle*. L'intégration multisensorielle désigne la capacité du cerveau à synthétiser les informations sensorielles issues de plusieurs sens différents<sup>37</sup>. Elle est influencée par la proximité temporelle et spatiale des stimuli provenant des différents sens<sup>38</sup>. Les CT se distinguent de l'intégration multisensorielle en ce que ce sont des associations stockées dans la mémoire.

Néanmoins, il est possible que les CT influencent l'intégration multisensorielle (au même titre que la proximité temporelle et spatiale). Un certain nombre de chercheur·se·s ont proposé l'idée que deux stimuli faisant appel à deux sens différents sont d'autant plus regroupés ensemble (*via* l'intégration multisensorielle) qu'ils sont congruents, un effet parfois appelé « effet d'unité »<sup>39</sup>.

## 2.4. Universalité des correspondances transmodales

Les informations présentées jusqu'ici pourraient suffire pour commencer à proposer une réponse à la problématique. Mais il reste une question qui nous semble cruciale : à *quels individus* va pouvoir s'appliquer cette réponse ? Pour illustrer cette question avec un exemple, imaginons qu'un·e compositeur·trice veuille évoquer à son auditoire un objet visuellement petit. Iel choisit donc d'utiliser une des CT que nous allons voir dans la partie 3 et qui relie la taille visuelle à un attribut auditif, disons celle reliant la hauteur sonore à la taille visuelle. Par conséquent, iel compose une musique faite uniquement de sons aigus. La question que nous posons est : pour quels individus

---

37 STEIN, Barry E., LAURIENTI, Paul J., WALLACE, Mark T. & STANFORD, Terrence R., « Multisensory Integration », in *Encyclopedia of the Human Brain*, Vilayanur S. Ramachandran éd., Elsevier, 2002, p. 227.

38 SPENCE, Charles, « Crossmodal correspondences: A tutorial review », art. cit., p. 971.

39 *Ibid.*, pp. 971-972.

ce·tte compositeur·trice peut-iel espérer que sa composition faite uniquement de sons aigus évoquera un objet visuellement petit ?

Pour y répondre, il nous faut aborder la question de *l'universalité des CT* (différente de la question de leur innéité). Nous proposons de diviser celle-ci en trois sous-questions : les CT existent-elles chez tous les individus ? Sont-elles de même nature pour tous les individus qui en font l'expérience ? Ont-elles la même intensité pour tous les individus qui en font l'expérience ? Traitons ces questions une par une.

*Existence.* Tous les individus éprouvent-ils des CT ? Si l'on prend la définition la plus large du terme « individu », on peut supposer qu'un grand nombre d'individus n'en éprouvent aucune, à commencer par ceux qui n'ont ni conscience ni système sensoriel : une partie des êtres vivants, et probablement la totalité des robots ayant existé jusqu'à ce jour. En outre, les CT sont au moins partiellement absentes chez certains humains. Par exemple, certains enfants atteints du trouble du spectre autistique ne sont pas concernés par l'effet bouba/kiki<sup>40</sup>, qui affecte pourtant la majorité des humains. En revanche, au moins une partie des animaux — humains et non-humains<sup>41</sup> — éprouvent des CT, indépendamment de la nature et de l'intensité de ces dernières. De plus, il nous paraît très probable que la quasi-totalité des humains en éprouvent. En effet, les CT sont partiellement causées par l'observation et la mémorisation de corrélations (ou cooccurrences) dans le monde physique<sup>42</sup>. Par exemple, la CT reliant la taille (visuelle ou tactile) à la hauteur sonore est partiellement due à ce que les objets petits produisent généralement des sons aigus, et les objets gros des sons graves. Or, la quasi-totalité des êtres humains sont capables d'observer et de mémoriser de telles corrélations<sup>43</sup>.

*Nature.* Clarifions d'abord ce que nous voulons dire par « nature des correspondances transmodales ». Si la nature des CT variait, cela signifierait que certains individus font l'expérience d'une correspondance transmodale A mais pas d'une correspondance transmodale B (différente de A), et que d'autres font l'expérience de la correspondance B, mais pas de la correspondance A. Est-ce le cas ? Est-ce que la nature des CT varie parmi les individus qui en font l'expérience ? Nous n'avons pas trouvé d'éléments de preuve concernant la question de savoir si elle varie en fonction des espèces vivantes. En revanche, les études que nous allons passer en revue dans la partie 3 de ce mémoire suggèrent que, au sein d'une population humaine donnée, la majorité des individus éprouvera souvent les mêmes CT. De plus, un certain nombre d'études interculturelles ont montré que la

---

40 *Ibid.*, p. 974.

41 Voir par exemple KORZENIOWSKA, Anna T., ROOT-GUTTERIDGE, Holly, SIMNER, Julia & REBY, David, « Audio-Visual Crossmodal Correspondences in Domestic Dogs (*Canis Familiaris*) », *Biology Letters*, vol. 15, n°11, 2019.

42 SPENCE, Charles, « Crossmodal correspondences: A tutorial review », art. cit., p. 988.

43 Par exemple, la majorité des humains sont capables d'observer et de mémoriser la corrélation entre la taille (visuelle ou tactile) et la hauteur sonore, car, d'une part, ils sont doués des sens requis pour percevoir ces attributs sensoriels : la vue, le toucher et l'ouïe, et d'autre part, ils sont capables de remarquer et de mémoriser automatiquement les corrélations présentes dans leur environnement.

nature des CT relevant du symbolisme sonore varie peu entre les cultures humaines<sup>44</sup>, ce qui suggère que les CT *en général* sont de même nature pour une partie au moins non négligeable d'êtres humains. Toutefois, le pourcentage auquel renvoie l'expression « non négligeable » dépend peut-être de la CT considérée. Les CT qui reflètent directement des corrélations observées dans le monde physique — par exemple, la CT reliant la hauteur sonore à la taille — sont très probablement partagées par une grande majorité d'êtres humains. Pour d'autres CT — par exemple, celle reliant les timbres cuivrés au goût salé —, cela nous paraît moins certain.

*Intensité.* Clarifions d'abord ce que nous voulons dire par « intensité des correspondances transmodales ». Si l'intensité des CT variait, cela signifierait qu'une CT présente chez deux individus puisse être plus évidente ou saillante pour l'un que pour l'autre. Est-ce le cas ? Est-ce que l'intensité des CT varie parmi les individus qui en font l'expérience ? Nous n'avons pas trouvé d'éléments de preuve concernant la question de savoir si elle varie entre les espèces vivantes, et entre les humains d'une même culture. Mais il nous paraît probable qu'elle ne varie que peu entre les individus d'une même culture. De plus, certaines études interculturelles suggèrent que l'intensité des CT relevant du symbolisme sonore varie peu entre les cultures humaines<sup>45</sup>, ce qui suggère que les CT *en général* sont d'intensité semblable pour une partie au moins non négligeable d'êtres humains. Mais là encore, le pourcentage auquel renvoie l'expression « non négligeable » pourrait varier d'une CT à l'autre.

Ce modeste passage en revue révèle plusieurs choses. D'une part, ni l'existence, ni la nature, ni l'intensité des CT ne sont universelles au sens strict : que ce soit du point de vue de l'ensemble des individus, des êtres vivants, des animaux ou des humains, il y a toujours des exceptions. D'autre part, si l'on s'intéresse au *degré* d'universalité (c'est-à-dire au pourcentage d'individus de tel groupe possédant tel trait relatif aux CT), on peut affirmer les choses suivantes :

- Le pourcentage d'animaux non humains qui éprouvent des CT n'est pas nul. Cette information est pertinente par rapport à la problématique de ce mémoire, car beaucoup d'animaux non humains sont capables d'écouter et d'apprécier la musique. Par conséquent, l'utilisation de CT en musique pourrait servir à offrir une expérience imaginative particulière à plusieurs espèces animales, et pas uniquement à *Homo Sapiens*.
- Le pourcentage d'humains *qui éprouvent des CT* est probablement très élevé. Notre estimation personnelle est qu'il y a 95% de chances pour que ce pourcentage soit supérieur à 90%.

---

44 PARISE, C. V. & SPENCE, C., « Audiovisual cross-modal correspondences in the general population », art. cit., pp. 793-794.

45 *Ibid.*, p. 798.

- Le pourcentage d'humains qui éprouvent des CT *et qui éprouvent celles présentées dans la partie 3 de ce mémoire* (étayées par des études empiriques menées le plus souvent sur des populations occidentales) n'est probablement pas négligeable. Pour certaines de ces CT, il nous semble très probable (95% de chances) que ce pourcentage soit supérieur à 70%. L'utilisation de ces CT en musique est donc certainement profitable à une grande majorité d'humains.
- Le pourcentage d'humains qui éprouvent des CT et qui éprouvent celles présentées dans la partie 3 de ce mémoire *avec une intensité semblable* n'est probablement pas négligeable. Pour certaines de ces CT, il nous semble très probable (95% de chances) que ce pourcentage soit supérieur à 40%<sup>46</sup>.

Ainsi, pour reprendre l'exemple donné plus haut, si un·e compositeur·trice compose une musique aigüe en vue d'évoquer un objet visuellement petit, iel ne pourra pas être certain·e que sa composition évoquera un objet petit à la *totalité* des individus, des animaux, ou même des humains ; en revanche, ce dont iel pourra être à peu près sûr·e, c'est que cette évocation fonctionnera pour certains animaux, dont plus de 70% d'animaux humains (avec une intensité plus ou moins variable selon les personnes et les cultures). Néanmoins, si iel utilise une autre CT, une qui ne reflète pas directement des corrélations observées dans le monde physique, il semble moins certain qu'elle fonctionnera pour une grande majorité d'humains — même si cela est tout à fait envisageable.

## 2.5. Récapitulatif

Cette partie a consisté à proposer une introduction aux CT. Voici les différents aspects qui ont été abordés :

- **Histoire de l'étude des correspondances transmodales.** L'existence des correspondances transmodales est remarquée au moins depuis l'Antiquité. Mais l'étude scientifique de ce phénomène ne débute qu'au XX<sup>ème</sup> siècle, notamment avec Köhler. Depuis quelques années, elle connaît un rapide progrès.
- **Définitions.**
  - Modalité sensorielle : sens (ouïe, vue, toucher, goût ou odorat).
  - Stimulus : évènement, agent ou situation, interne ou externe, qui suscite une réponse de la part d'un organisme.

---

<sup>46</sup> Bien sûr, cette prédiction est moins facilement testable que les précédentes, car elle implique un terme vague (« intensité *semblable* »).

- ➔ Stimulus simple/stimulus complexe : stimulus comportant un/plusieurs élément(s) ou attribut(s) sensoriel(s).
- Attribut sensoriel : caractéristique d'un stimulus.
  - ➔ Attribut dimensionnel/attribut non dimensionnel : attribut sensoriel pouvant/ne pouvant pas être mesuré quantitativement sur une échelle allant de « faible » à « élevé ».
- Correspondance transmodale : association intuitive entre un attribut sensoriel perçu par un sens et un attribut sensoriel perçu par un autre sens.
- Congruence/incongruence : état de deux valeurs d'attributs sensoriels associés/dissociés dans le cadre d'une CT.
- **Quelques distinctions importantes.** Les correspondances transmodales se distinguent :
  - de la synesthésie : dans le cas de la synesthésie, la stimulation d'un sens génère une sensation dans un autre sens, tandis que dans le cas des CT, elle génère seulement une *pensée* d'une sensation dans un autre sens.
  - des correspondances intramodales : celles-ci sont des associations intuitives entre deux attributs sensoriels perçus par le même sens, tandis que les CT sont des associations intuitives entre deux attributs sensoriels perçus par deux sens *différents*.
  - de l'intégration multisensorielle : celle-ci est une capacité du cerveau à synthétiser les informations sensorielles issues de plusieurs sens différents, tandis que les CT sont des associations stockées dans la mémoire.
- **Universalité des CT.**
  - Le phénomène des CT est absent chez une partie des êtres vivants, des animaux et des humains.
  - Mais il est présent à la fois chez des humains et chez des animaux non humains.
  - Il est très probable que certaines CT présentées dans la partie 3 de ce mémoire soient éprouvées par plus de 70% des humains (quoiqu'avec des intensités potentiellement assez variables).



### **3. Les correspondances transmodales utilisables en musique**

S'il fallait dresser une liste de toutes les CT utilisables en musique, il faudrait, nous semble-t-il, procéder en deux étapes. En premier lieu, puisqu'il faudrait s'intéresser à toutes les CT impliquant au moins un des sens par lesquels la musique peut être perçue, il serait nécessaire de se pencher sur le nombre et la nature de ces sens. Car, outre le fait évident que la musique puisse être perçue par l'ouïe, on peut arguer qu'elle peut être perçue par d'autres sens, comme le toucher ou même la vue<sup>47</sup>. Une fois établis le nombre et la nature des sens permettant de percevoir la musique, il faudrait, en second lieu, lister toutes les CT faisant intervenir au moins un de ces multiples sens.

Tout ce travail serait fort utile et intéressant. Néanmoins, nous ne pouvons le réaliser dans le cadre de ce mémoire. Nous nous restreindrons donc aux CT utilisables en musique *et faisant intervenir le sens de l'ouïe*, pour la raison que l'ouïe est le sens sur lequel la population semble se reposer le plus pour faire l'expérience de la musique. Cependant, là encore, le nombre de ces CT est très élevé et il serait difficile d'être exhaustif. Par conséquent, nous nous contenterons de mettre en exergue *un certain nombre* de ces CT.

#### **3.1. Méthodologie**

Cette partie s'appuie sur une sélection d'articles en psychologie cognitive afférents aux CT. Nous avons sélectionné en priorité des articles qui résument plusieurs études empiriques. Lorsque nous n'en avons pas trouvé, comme dans le cas des CT reliant l'ouïe et l'odorat, nous nous sommes portés directement vers des études empiriques, en choisissant les plus citées.

Tous ces articles classent généralement les CT en fonction des sens qu'elles relient : il est question de CT audio-visuelles, tactilo-gustatives, visuo-olfactives, etc. Nous nous sommes conformés à cette logique et avons divisé cette 3<sup>ème</sup> partie en quatre sous-parties correspondant aux quatre paires de sens possibles entre l'ouïe et les autres sens. Ces paires de sens sont : ouïe-vue, ouïe-toucher, ouïe-goût et ouïe-odorat.

Les CT audio-visuelles, audio-tactiles et audio-gustatives sont présentées dans des tableaux à double entrée. À chaque article cité est consacré un tableau (sauf pour les articles ne testant qu'une seule CT).

---

<sup>47</sup> En effet, la perception de la musique implique le toucher *via* les vibrations tactiles induites par les ondes sonores. Le concept de « vusique » (en anglais : « *visic* », contraction de « *visual* » et « *music* »), inventé par le collectif d'artistes Sourd·e·s *Viscore*, suggère qu'elle implique aussi la vue.

Les tableaux des CT audio-visuelles et audio-tactiles sont organisés de la façon suivante. Chaque ligne correspond à un attribut sensoriel de l'ouïe (attribut auditif), et chaque colonne correspond à un attribut sensoriel de l'autre sens (attributs visuels ou tactiles). De là, pour chaque paire d'attributs sensoriels, il y a trois possibilités :

- Lorsque les deux attributs sensoriels forment ensemble une CT, l'intersection de la ligne et de la colonne correspondant à ces deux attributs contient une ou plusieurs expressions reliant par un tiret une valeur de l'attribut auditif et une valeur congruente de l'attribut non auditif. Par exemple, l'intersection de la ligne « hauteur sonore » et de la colonne « taille visuelle » dans le tableau 5 contient les expressions « aigu-petit » et « grave-gros » : cela signifie que ces deux attributs forment une CT et que, plus précisément, les sons aigus sont associés aux objets visuellement petits, et les sons graves aux objets visuellement gros.
- Lorsque deux attributs sensoriels ne forment pas de CT, et ce quelle que soit la combinaison de valeurs considérée, l'intersection de la ligne et de la colonne correspondant à ces deux attributs contient l'expression « ns » pour « non significatif ».
- Lorsque nous n'avons pas trouvé d'études testant une CT entre deux attributs sensoriels donnés, ou que l'étude dont nous résumons les résultats à tel ou tel endroit ne teste pas l'existence d'une CT entre ces deux attributs, l'intersection de la ligne et de la colonne correspondant à ces deux attributs est laissée vide.

Les tableaux des CT audio-gustatives sont organisés de la même façon, à deux changements près :

- Chaque colonne correspond, non pas à un *attribut* gustatif, mais à un *stimulus* gustatif (sucré, acide, salé et amer). De fait, les CT audio-gustatives que nous allons voir relient des attributs (auditifs) à des stimuli (gustatifs).
- Si un attribut auditif et un stimulus gustatif forment ensemble une CT, alors l'intersection de la ligne et de la colonne correspondant à ces deux éléments contient, non pas une expression reliant deux valeurs, mais simplement la valeur de l'attribut auditif associée au stimulus gustatif. Par exemple, les accords très consonants sont associés au goût sucré ; par conséquent, l'intersection de la ligne « consonance harmonique » et de la colonne « sucré » contient le mot « élevée », ce qui signifie « consonance harmonique élevée » (voir tableau 8). Ce changement se justifie par le fait que, contrairement aux attributs sensoriels, les stimuli ne sont pas des variables : ils ne peuvent pas prendre plusieurs

valeurs différentes. Et il aurait été inutile d'indiquer le stimulus gustatif impliqué dans chaque CT audio-gustative, puisque ce stimulus est déjà indiqué au sommet de la colonne.

Que ce soit pour les CT audio-visuelles, audio-tactiles ou audio-gustatives, une question reste en suspens : quelle confiance peut-on accorder dans l'existence et la nature de telle ou telle CT qui sera présentée ? Dans l'idéal, pour répondre à cette question, il aurait fallu pondérer chaque CT par une probabilité dépendante de la quantité et de la qualité des preuves disponibles à son sujet. Mais la tâche nous a paru difficile. C'est pourquoi nous nous sommes limités à trois niveaux, qui ne correspondent d'ailleurs pas exactement à des niveaux de confiance. Ces niveaux sont les suivants :

1. Lorsqu'une CT est évoquée dans un article et est explicitement jugée très étayée par les auteur·trice·s, la ou les expression(s) correspondant à cette CT est/sont soulignée(s).
2. Lorsqu'une CT est évoquée dans un article, mais n'est pas explicitement jugée très étayée par les auteur·trice·s, la ou les expression(s) correspondant à cette CT est/sont laissée(s) telle(s) quelle(s).
3. Lorsqu'une CT est raisonnablement déductible à partir d'une autre CT évoquée dans un article, mais qu'elle n'est pas évoquée dans cet article, la ou les expression(s) correspondant à cette CT est/sont mise(s) en italique.

Enfin, pour ce qui est des CT audio-olfactives, le tableau n'a pas été retenu comme moyen de représentation. En effet, les articles que nous allons citer présentent ces CT dans des graphiques, et ce moyen de représentation nous semble plus précis et judicieux que des tableaux. Nous nous sommes donc contentés de reproduire ces graphiques (avec leur légende) en traduisant ce qui devait être traduit.

À la fin de cette partie, dans le récapitulatif, les CT les plus importantes sont regroupées dans quatre tableaux, chaque tableau correspondant à une paire de sens. Cela concerne également les CT audio-olfactives, dont nous n'avons gardé que celles qui étaient les plus simples à représenter dans un tableau.

Les quatre tableaux suivants définissent les différents attributs sensoriels impliqués dans les CT que nous allons voir<sup>48</sup>. Les stimuli gustatifs ne sont pas définis puisqu'il s'agit simplement des quatre saveurs de base : sucré, acide, salé et amer.

---

<sup>48</sup> Lorsque nous définissons un attribut sensoriel, nous ne donnons pas nécessairement une définition consensuelle ni complète. Nous donnons simplement la définition qui a été utilisée ou semble avoir été utilisée dans les articles que nous citons.

Attribut auditif	Définition
Hauteur sonore	Fréquence fondamentale d'un son.
Direction d'un changement de hauteur sonore	Augmentation ou diminution de la fréquence fondamentale d'un son (le son devient de plus en plus aigu ou de plus en plus grave).
Consonance harmonique	Caractère d'un accord considéré comme agréable.
Consonance mélodique	Caractère d'une mélodie considérée comme agréable.
Ambitus	Intervalle entre le son le plus grave et le son le plus aigu d'une mélodie ou d'un accord.
Intervalles mélodiques	Distances entre les hauteurs successives d'une mélodie.
Timbre	Qualité sonore correspondant à un instrument de musique.
Centroïde spectral	Centre de masse d'un spectre sonore, ou moyenne des partiels d'un son pondérés par leur amplitude. Il s'agit d'une mesure de la « luminosité sonore » : toutes choses égales par ailleurs, plus le centroïde spectral d'un son est élevé, plus ce son semble lumineux.
Rugosité sonore	Impression d'aspérité résultant de la modulation de l'amplitude d'un son <sup>49</sup> .
Discontinuité sonore	Durée des silences séparant des sons.
Rythme	Organisation temporelle des positions et des durées des notes.
Attaque	Début d'un son, qui précède sa résonance.
Intensité sonore	Amplitude d'un son.
Tempo	Vitesse d'exécution d'une création musicale.

Tableau 1. Définitions de quelques attributs auditifs.

Attribut visuel	Définition
Hauteur visuelle <sup>50</sup>	Perception visuelle de la hauteur spatiale d'un objet.
Direction d'un mouvement visuel vertical	Perception visuelle de la direction d'un mouvement vertical.
Taille visuelle	Perception visuelle de la taille d'un objet.
Épaisseur visuelle	Perception visuelle de l'épaisseur d'un objet.
Angularité visuelle	Perception visuelle du caractère pointu ou arrondi d'un objet.
Fréquence visuo-spatiale	Mesure du nombre de répétitions par unité d'angle visuel. Cette grandeur caractérise les structures visuelles qui se reproduisent identiquement à des positions régulièrement espacées.
Luminosité visuelle	Perception visuelle de la quantité d'énergie qu'un objet rayonne par unité de temps sous forme d'ondes électromagnétiques.
Contraste visuel	Différence de luminosité visuelle entre les parties claires et les parties sombres d'un objet.
Couleur	Perception visuelle de la répartition spectrale de la lumière d'un objet.

Tableau 2. Définitions de quelques attributs visuels.

49 Pour des exemples sonores, consulter : <https://www.johndcook.com/blog/2016/08/17/acoustic-roughness-examples/> (consulté le 22/05/2022).

50 En anglais : *visual/visuospatial elevation*.

Attribut tactile	Définition
Hauteur tactile <sup>51</sup>	Perception tactile (plus précisément, digitale) de la hauteur spatiale d'un objet.
Taille tactile	Perception tactile de la taille d'un objet.
Poids tactile	Perception tactile du poids d'un objet.
Fréquence vibrotactile	Fréquence de vibrations perçues par le toucher.
Direction d'un mouvement tactile ressenti sur le doigt	Direction dans laquelle se déplace un stimulus tactile sur un doigt, le stimulus se déplaçant entre la base du doigt et le bout du doigt.

Tableau 3. Définitions de quelques attributs tactiles.

	Attribut olfactif	Définition
Descripteurs amodaux	Intensité	Concentration d'une odeur.
	Complexité	Caractère d'une odeur composée de plusieurs odeurs simples.
	Agréabilité <sup>52</sup>	Caractère plaisant d'une odeur.
Adjectifs olfactifs	Âcre	Qui a une odeur irritante.
	Floral	Qui a l'odeur d'une fleur.
	Fruité	Qui a l'odeur d'un fruit.
	Épicé	Qui a l'odeur d'une épice.

Tableau 4. Définitions de quelques attributs olfactifs.

### 3.2. Ouïe – Vue

Certaines CT audio-visuelles sont décelables dans les mots de la langue française. Le mot « hauteur », par exemple, peut désigner à la fois l'altitude d'un objet, altitude pouvant naturellement être perçue par la vue, et la fréquence fondamentale d'un son. De même, le mot « aigu » qualifie à la fois les formes visuellement pointues et les sons de fréquence fondamentale élevée. Nous allons voir que les CT suggérées par ces double-sens sont mises en évidence dans la littérature scientifique, en plus de bien d'autres CT audio-visuelles. En fait, il semble que les CT audio-visuelles soient de loin les plus étayées à ce jour<sup>53</sup>.

Spence<sup>54</sup> passe en revue un grand nombre de CT audio-visuelles. Nous avons regroupé ces CT dans le tableau suivant. D'après Spence, l'une des CT les plus étayées, non seulement parmi les CT

51 En anglais : *tactile elevation*.

52 Ce mot est rare, mais c'est la meilleure traduction que nous pouvons trouver du terme anglais « *pleasantness* ».

53 SPENCE, Charles, « Crossmodal correspondences: A tutorial review », *Attention, Perception, & Psychophysics*, vol. 73, n° 4, 2011, p. 975.

54 *Ibid.*, pp. 974, 975, 979.

audio-visuelles mais aussi parmi les CT en général, est celle entre la hauteur sonore et la hauteur visuelle<sup>55</sup>. Nous avons donc souligné les expressions correspondant à cette CT.

	Hauteur visuelle	Direction d'un mouvement visuel vertical	Taille visuelle	Angularité visuelle	Fréquence visuo-spatiale	Luminosité visuelle	Contraste visuel	Couleur
<b>Hauteur sonore</b>	<u>aigu-haut</u>	aigu-haut	aigu-petit	aigu-pointu	aigu-élevée	aigu-lumineux	ns	ns
	<u>grave-bas</u>	grave-bas	grave-gros	grave-arrondi	grave-faible	grave-sombre	ns	
<b>Intensité sonore</b>			fort-gros			fort-lumineux	fort-élevé	
			faible-petit			faible-sombre	faible-faible	

Tableau 5. Correspondances transmodales audio-visuelles évoquées dans Spence (2011).

Walker<sup>56</sup> énumère des CT audio-visuelles supplémentaires par rapport à Spence. Nous les avons regroupées dans le tableau suivant. Bien que nous avons inclus dans ce tableau la CT entre le tempo et l'angularité visuelle, Walker ne parle pas explicitement de cette CT. En effet, il écrit que « lorsqu'on demande aux personnes d'indiquer les attributs que leur évoquent des formes visuellement pointues, ou à l'inverse visuellement arrondies, les formes pointues sont jugées [...] plus rapides [...] que les formes arrondies »<sup>57</sup>. Autrement dit, les personnes associent les formes pointues au concept *général* de rapidité (et les formes arrondies au concept *général* de lenteur). Mais on ne peut totalement exclure que ce qui s'applique au concept général de rapidité ne s'applique guère au concept particulier de rapidité *sonore*, c'est-à-dire de tempo. Pour cette raison, la CT entre le tempo et l'angularité visuelle semble un peu moins certaine que les autres, c'est pourquoi nous l'avons mise en italique.

	Épaisseur visuelle	Angularité visuelle	Luminosité visuelle
<b>Hauteur sonore</b>	aigu-fin		
	grave-épais		
<b>Tempo</b>		<i>rapide-pointu</i>	rapide-lumineux
		<i>lent-arrondi</i>	lent-sombre

Tableau 6. Correspondances transmodales audio-visuelles évoquées dans Walker (2016).

55 *Ibid.*, p. 976.

56 WALKER, Peter, « Cross-Sensory Correspondences: A Theoretical Framework and Their Relevance to Music », *Psychomusicology: Music, Mind, and Brain*, vol. 26, n° 2, 2016, pp. 104-105.

57 *Ibid.*, p. 105. Traduction personnelle.

### 3.3. Ouïe – Toucher

Comme les CT audio-visuelles, les CT audio-tactiles sont parfois déductibles à partir de l’usage des mots. Le mot « doux » qualifie une surface tactilement lisse autant qu’un son agréable. Le mot « froid » désigne une sensation de température faible autant qu’une musique austère et désagréable. Le mot « léger » qualifie une sensation de poids faible autant que des notes aiguës, rapides et *staccatos*. Nous n’avons pas trouvé d’éléments de preuve pour toutes les CT que suggèrent ces mots, en partie parce que les CT impliquant le toucher ont été moins étudiées que celles impliquant la vue<sup>58</sup>, et qu’elles sont donc moins étayées.

Walker<sup>59</sup> passe en revue certaines CT audio-tactiles. Nous les avons regroupées dans le tableau suivant. Notons que la CT entre le tempo et le poids tactile revêt la même incertitude que celle entre le tempo et l’angularité visuelle ; nous renvoyons donc à la remarque qui a été faite au sujet de cette dernière dans la sous-partie 3.2.

	Hauteur tactile	Taille tactile	Poids tactile	Fréquence vibrotactile
Hauteur sonore	aigu-haut	aigu-petit	aigu-léger	aigu-élevée
	grave-bas	grave-gros	grave-lourd	grave-faible
Tempo		rapide-petit	<i>rapide-léger</i>	
		lent-gros	<i>lent-lourd</i>	

Tableau 7. Correspondances transmodales audio-tactiles évoquées dans Walker (2016).

Deroy et al.<sup>60</sup> mettent en évidence une autre CT audio-tactile : celle reliant la direction d’une variation de hauteur sonore et la direction d’un mouvement tactile ressenti sur le doigt. Cette CT signifie que les montées de hauteur sonore sont associées aux mouvements tactiles dirigés vers l’extérieur du corps (c’est-à-dire allant de la base du doigt vers le bout du doigt), et les descentes de hauteur sonore aux mouvements tactiles dirigés vers l’intérieur du corps (c’est-à-dire allant du bout du doigt vers la base du doigt). Elle a été mise en évidence dans deux conditions : lorsque le bras est à l’horizontale et lorsqu’il est à la verticale. Notons que tout cela concerne uniquement les personnes voyantes. En effet, aucun effet de congruence n’a été observé pour les personnes non voyantes, quel que soit l’âge auquel ces personnes sont devenues non-voyantes.

58 SPENCE, Charles & SATHIAN, Krish, « Audiovisual Crossmodal Correspondences: Behavioral Consequences and Neural Underpinnings », in *Multisensory Perception*, Krish Sathian & Vilayanur S. Ramachandran éd., Academic Press, 2020, p. 239.

59 WALKER, Peter, « Cross-Sensory Correspondences: A Theoretical Framework and Their Relevance to Music », art. cit., pp. 104-106.

60 DEROY, Ophelia, FASIELLO, Irène, HAYWARD, Vincent & AUVRAY, Malika, « Differentiated audio-tactile correspondences in sighted and blind individuals », *Journal of Experimental Psychology: Human Perception and Performance*, vol. 42, n°8, 2016, pp. 1204-1214.

### 3.4. Ouïe – Goût

Le vocabulaire utilisé dans le milieu de la musique savante occidentale peut nous aider à deviner certaines CT audio-gustatives. Le terme italien « *dolce* » est employé traditionnellement pour décrire une musique « douce », mais il signifie aussi « sucré ». Dans son *Grand traité d'orchestration et d'instrumentation modernes*, Hector Berlioz décrit le timbre du hautbois comme une « petite voix aigre-douce »<sup>61</sup>. On pourrait encore multiplier les exemples.

Comme dit dans la sous-partie 3.1, les CT audio-gustatives que nous allons présenter relient des attributs auditifs à des *stimuli* gustatifs, et non à des *attributs* gustatifs. En effet, les éléments gustatifs dont il est question sont simplement les quatre saveurs de base : le sucré, l'acide, le salé et l'amer<sup>62</sup>. Notons qu'il s'agit de stimuli gustatifs *simples* : de fait, à notre connaissance, peu d'études se sont penchées sur les CT audio-gustatives impliquant des stimuli gustatifs complexes (comme le goût de la pomme, de la tomate ou du tofu, par exemple).

Knöferle & Spence<sup>63</sup> offrent un état de l'art des CT audio-gustatives. Ils les ont regroupées dans un tableau que nous avons reproduit en y apportant de légères modifications. Précisons que la présence de plusieurs résultats dans certaines cases du tableau est due à ceci que des résultats contradictoires, voire parfois opposés (comme dans le cas de la CT reliant la discontinuité sonore et le goût

	Sucré	Acide	Salé	Amer
Hauteur sonore	<u>aigu</u> / médium	<u>aigu</u> / médium	médium / ns / grave	<u>grave</u> / ns
Consonance harmonique	<u>élevée</u>	<u>faible</u>	moyenne	moyenne
Consonance mélodique	élevée	faible	moyenne	moyenne
Ambitus	petit	grand		
Intervalles mélodiques	petits	grands		
Timbre	<u>piano</u>	cuivres	cuivres	cuivres / ns
Centroïde spectral	<u>bas</u> / ns	<u>haut</u>	ns	bas / ns
Rugosité sonore	faible	élevée	moyenne	élevée
Discontinuité sonore	<u>faible (legato)</u>	<u>élevée (staccato)</u> / moyenne	élevée (staccato) / ns	élevée (staccato) / faible (legato)
Rythme	valeurs égales	synopes		
Attaque	ns	ns	ns	ns
Intensité sonore	faible	moyenne	moyenne	moyenne
Tempo	lent / ns	<u>rapide</u>	ns	lent

Tableau 8. Correspondances transmodales audio-gustatives évoquées dans Knöferle & Spence (2012).

61 BERLIOZ, Hector, *Grand traité d'orchestration et d'instrumentation modernes* [1844], Allemagne, Bärenreiter, 2003, p. 156.

62 Il est parfois admis l'existence d'une cinquième saveur de base, l'*umami*, mais celle-ci semble avoir été peu étudiée dans les études sur les CT audio-gustatives.

63 KNÖFERLE, Klemens & SPENCE, Charles, « Crossmodal Correspondences between Sounds and Tastes », *Psychonomic Bulletin & Review*, vol. 19, n°6, 2012, p. 997.

amer), coexistent pour certaines CT. Pour autant, lorsqu'un des multiples résultats coexistant dans une même case est souligné, on peut considérer que celui-ci mérite davantage de crédit que les autres.

Postérieurement à Knöferle & Spence, Knöferle et al.<sup>64</sup> ont également mis en évidence des CT audio-gustatives. Une part de ces CT confirme les résultats rassemblés par Knöferle & Spence, l'autre part les complète. Nous avons regroupé le tout dans le tableau suivant.

	Sucré	Acide	Salé	Amer
Hauteur sonore	aigu	aigu	ns	grave
Centroïde spectral	bas	haut	bas	
Rugosité sonore	faible	élevée	moyenne	élevée
Discontinuité sonore	faible	élevée	élevée	élevée
Attaque	ns	ns	ns	ns
Tempo		rapide		lent

Tableau 9. Correspondances transmodales audio-gustatives évoquées dans Knöferle et al. (2015).

### 3.5. Ouïe – Odorat

Les parfumeur·se·s décrivent souvent les parfums avec des termes musicaux : iels disent chercher le « bon équilibre entre les notes aigües et graves », ils pensent en termes de « tonalité », etc<sup>65</sup>. On pourrait penser que ce jargon n'est que métaphorique et ne révèle aucune CT largement partagée, car, contrairement aux parfumeur·se·s, la majorité de la population (occidentale, du moins) met rarement en parallèle l'ouïe et l'odorat dans son usage du langage. Cependant, lorsque cette même majorité est amenée à effectuer une tâche à choix forcé, elle associe systématiquement certains attributs olfactifs à certains attributs auditifs<sup>66</sup>.

Deroy et al.<sup>67</sup> signalent qu'une CT a été mise en évidence entre l'intensité sonore et l'intensité olfactive : les sons forts sont associés aux odeurs fortes, et les sons faibles aux odeurs faibles.

Crisinel & Spence<sup>68</sup> mettent en évidence d'autres CT audio-gustatives. Celles-ci relient le timbre ou la hauteur sonore à trois espèces d'éléments olfactifs :

64 KNÖFERLE, Klemens M., WOODS, Andy, KÄPPLER, Florian & SPENCE, Charles, « That Sounds Sweet: Using Cross-Modal Correspondences to Communicate Gustatory Attributes », *Psychology & Marketing*, vol. 32, n°1, 2015, p. 110.

65 DEROY, Ophelia, CRISINEL, Anne-Sylvie & SPENCE, Charles, « Crossmodal Correspondences between Odors and Contingent Features: Odors, Musical Notes, and Geometrical Shapes », *Psychonomic Bulletin & Review*, vol. 20, n°5, 2013, p. 878.

66 *Ibid.*, p. 879.

67 *Ibid.*, p. 888.

68 CRISINEL, Anne-Sylvie & SPENCE, Charles « A Fruity Note: Crossmodal Associations between Odors and Musical Notes », *Chemical Senses*, vol. 37, n°2, 2012, pp. 151-158.

- des stimuli olfactifs, c'est-à-dire des odeurs particulières : amande, pomme, abricot, mûre, caramel, cèdre, chocolat noir, foin coupé, poivron vert, miel, citron, réglisse, champignon, musc, poivre, ananas, framboise, fumée, vanille, violette. Voir figures 3 et 4.
- des attributs amodaux, appelés « descripteurs amodaux » dans l'article : complexité, familiarité, intensité, agréabilité. Voir figure 5.
- des attributs plus spécifiquement olfactifs (bien qu'ils puissent également être perçus par le goût), appelés « adjectifs olfactifs » dans l'article : âcre, terreux, floral, fruité, noiseté<sup>69</sup>, épicé, boisé. Voir figure 6.

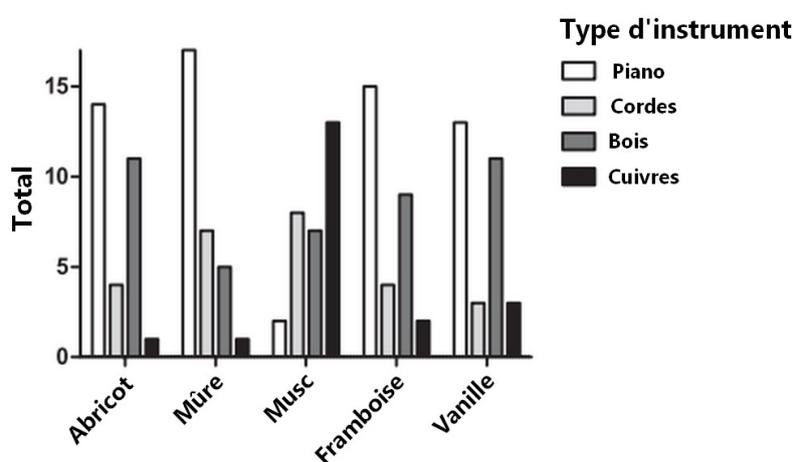


Figure 3. Choix d'instrument en fonction de l'odeur présentée. Seules les odeurs ayant conduit à des préférences statistiquement significatives sont représentées. (Figure et légende issues de Crisinel & Spence, 2012, figure 1.)

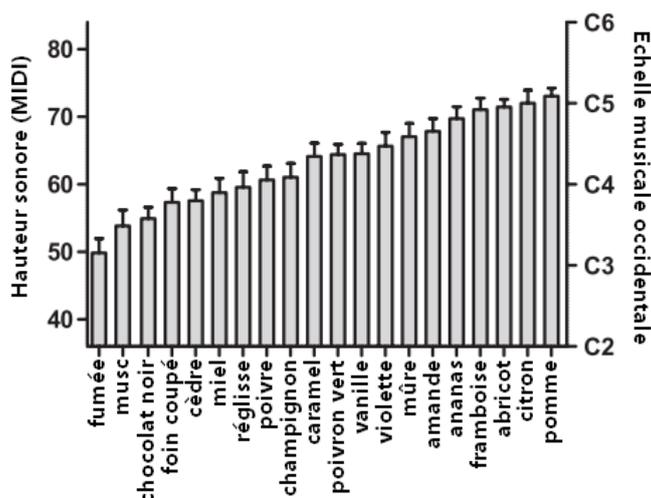


Figure 4. Hauteur sonore moyenne attribuée à chaque odeur testée. La hauteur est codée via des nombres MIDI (Musical Instrument Digital Interface). La notation de l'échelle musicale occidentale est montrée sur l'axe vertical de droite. Les notes aigües ont été privilégiées pour les odeurs de fruits. (Figure et légende issues de Crisinel & Spence, 2012, figure 5.)

69 En anglais : *nutty*. En français, la traduction littérale est : « qui a le goût ou l'odeur de noix ».

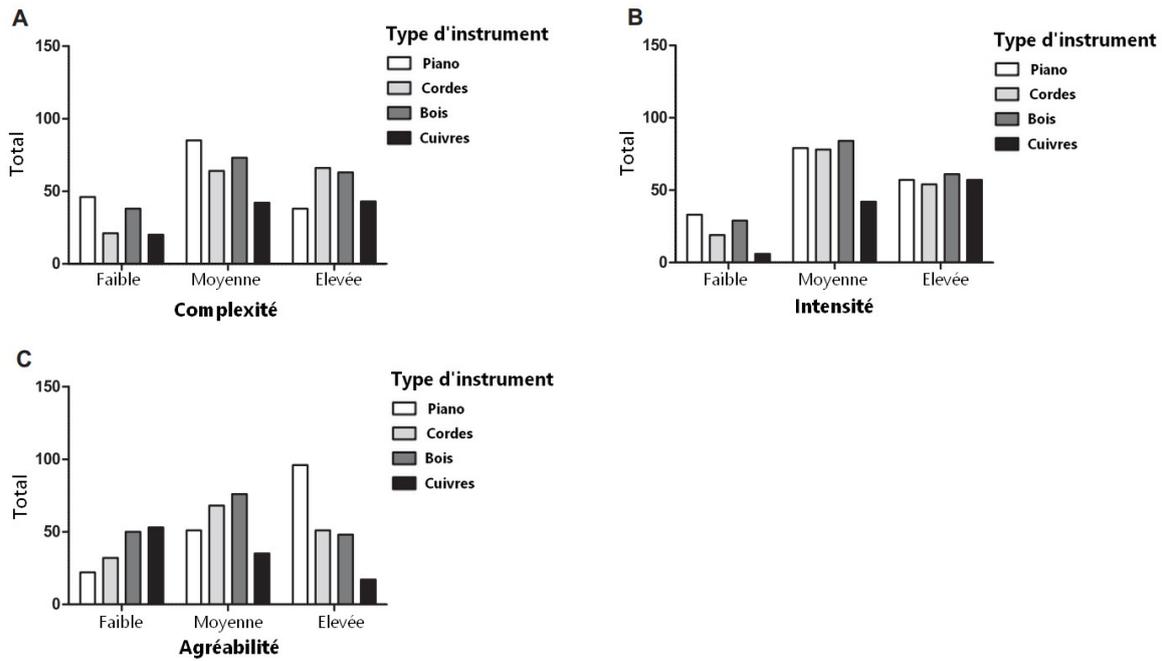


Figure 5. Choix d'instrument en fonction de jugements de descripteurs amodaux (seuls les adjectifs ayant eu un effet statistiquement significatif sur le choix d'instrument sont représentés) : complexité (A), intensité (B) et agréabilité (C), découpés en trois niveaux. La somme des totaux des trois niveaux est de 600 (30 participant·e·s × 20 stimuli). Le piano a été évité pour les odeurs jugées plus complexes. Plus l'intensité a été élevée, plus les participant·e·s ont choisi les cuivres. Les cuivres ont également été préférés pour les stimuli désagréables, tandis que le piano a été associé aux odeurs agréables. (Figure et légende issues de Crisinel & Spence, 2012, figure 2.)

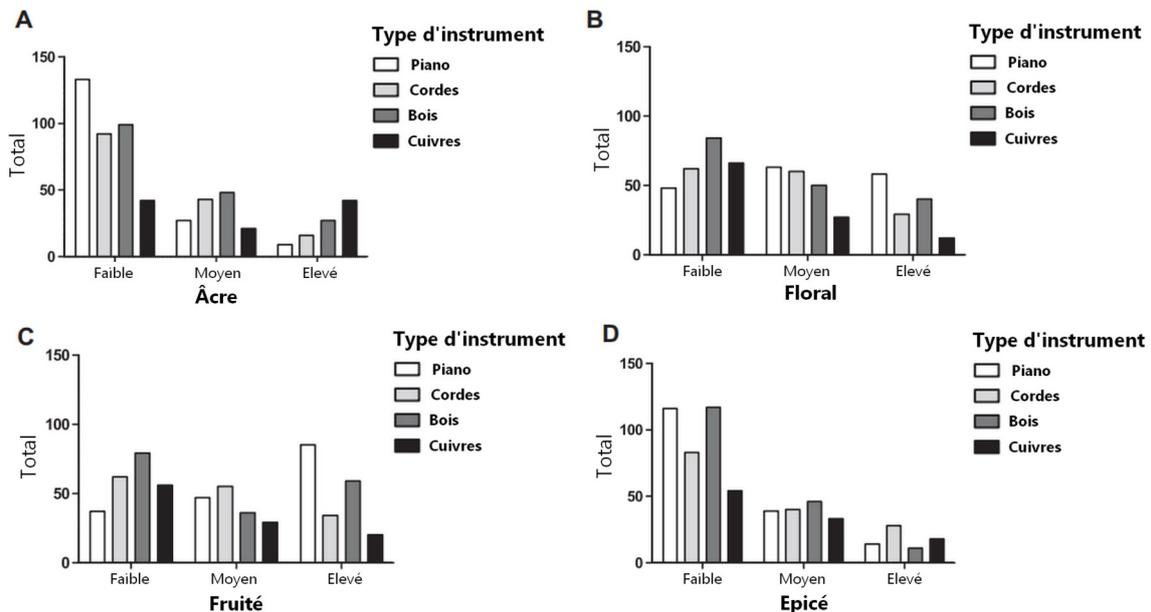


Figure 6. Choix d'instrument en fonction de jugements d'adjectifs olfactifs (seuls les adjectifs ayant eu un effet statistiquement significatif sur le choix d'instrument sont représentés) : âcre (A), floral (B), fruité (C) et épicé (D), découpés en trois niveaux. La somme des totaux des trois niveaux est de 600 (30 participant·e·s × 20 stimuli). (Figure et légende issues de Crisinel & Spence, 2012, figure 3.)

### 3.6. Récapitulatif

Les tableaux suivants récapitulent les principales CT passées en revue dans cette partie. Chaque tableau correspond à une paire de sens. Rappelons que ces tableaux ne sont pas exhaustifs. En effet, une partie des CT impliquant l'ouïe ont été établies mais n'ont pas été présentées dans ce mémoire. De plus, une partie d'entre elles restent à établir.

	Hauteur visuelle	Direction d'un mouvement visuel vertical	Taille visuelle	Épaisseur visuelle	Angularité visuelle	Fréquence visuo-spatiale	Luminosité visuelle	Contraste visuel
Hauteur sonore	<i>aigu-haut</i>	aigu-haut	aigu-petit	aigu-fin	aigu-pointu	aigu-élevée	aigu-lumineux	ns
	<i>grave-bas</i>	grave-bas	grave-gros	grave-épais	grave-arrondi	grave-faible	grave-sombre	ns
Intensité sonore			fort-gros				fort-lumineux	fort-élevé
			faible-petit				faible-sombre	faible-faible
Tempo					<i>rapide-pointu</i>		rapide-lumineux	
					<i>lent-arrondi</i>		lent-sombre	

Tableau 10. Correspondances transmodales audio-visuelles.

	Hauteur tactile	Taille tactile	Poids tactile	Fréquence vibrotactile	Direction d'un mouvement tactile ressenti sur le doigt
Hauteur sonore	aigu-haut	aigu-petit	aigu-léger	aigu-élevée	
	grave-bas	grave-gros	grave-lourd	grave-faible	
Direction d'un changement de hauteur sonore					ascendant-extérieur
					descendant-intérieur
Tempo		rapide-petit	<i>rapide-léger</i>		
		lent-gros	<i>lent-lourd</i>		

Tableau 11. Correspondances transmodales audio-tactiles.

	Sucré	Acide	Salé	Amer
Hauteur sonore	<u>aigu</u>	<u>aigu</u>	médium / ns / grave	<u>grave</u>
Consonance harmonique	<u>élevée</u>	<u>faible</u>	moyenne	moyenne
Consonance mélodique	élevée	faible	moyenne	moyenne
Ambitus	petit	grand		
Intervalles mélodiques	petits	grands		
Timbre	<u>piano</u>	cuivres	cuivres	cuivres / ns
Centroïde spectral	<u>bas</u>	<u>haut</u>	bas / ns	bas / ns
Rugosité sonore	faible	élevée	moyenne	élevée
Discontinuité sonore	<u>faible (legato)</u>	<u>élevée (staccato)</u>	élevée (staccato) / ns	élevée (staccato) / faible (legato)
Rythme	valeurs égales	syncopes		
Intensité sonore	faible	moyenne	moyenne	moyenne
Tempo	lent / ns	<u>rapide</u>	ns	lent

Tableau 12. Correspondances transmodales audio-gustatives.

	Descripteurs amodaux			Adjectifs olfactifs		
	Complexité	Intensité	Agréabilité	Âcre	Floral	Fruité
Hauteur sonore						aigu-très fruité
Timbre	cordes-élevée		piano-élevée	cuivres-très âcre	piano-très floral	piano-très fruité
	piano-moyenne		bois-moyenne			cordes-moyennement fruité
			cuivres-faible	piano-peu âcre	bois-peu floral	bois-peu fruité
Intensité sonore		fort-élevée				
		faible-faible				

Tableau 13. Correspondances transmodales audio-olfactives.



## **4. Utiliser les correspondances transmodales en musique**

Dans la partie précédente, nous avons abordé les *correspondances transmodales* utilisables en musique. Dans la présente partie, nous traiterons de l'*utilisation* de ces CT en musique. Nous n'aborderons pas leur utilisation dans des domaines *particuliers*, c'est-à-dire que nous n'aborderons pas séparément les domaines de la composition, de l'interprétation et de l'improvisation, par exemple. En revanche, nous adopterons une approche *générale* (ce qui ne nous empêchera pas de traiter les différents domaines en les regroupant dans des catégories plus larges). Cette approche générale sera déclinée en quatre questions :

1. **Pourquoi ?** Dans quel but les CT peuvent-elles être utilisées en musique ?
2. **Où ?** Dans quels domaines les CT peuvent-elles être utilisées en musique ?
3. **Comment ?** De quelle façon les CT peuvent-elles être utilisées en musique ?
4. **Quand ?** Dans quelles circonstances les CT peuvent-elles être utilisées en musique ?

Comme indiqué dans l'introduction et dans la sous-partie 1.2, nous traiterons non seulement l'utilisation des CT en musique, mais aussi en analyse musicale. Ainsi, cette partie ne se restreindra pas à la *production* de musique (composition, interprétation, improvisation...), mais traitera également de la *compréhension* de la musique (analyse musicale<sup>70</sup>). Et, par souci de simplicité, l'utilisation des CT en musique et/ou en analyse musicale sera simplement désignée par des expressions du type « utilisation des CT en musique » ou « utilisation musicale des CT ».

### **4.1. Méthodologie**

L'utilisation des CT en musique semble être un sujet peu traité par la littérature scientifique. C'est pourquoi cette partie s'appuiera principalement sur des réflexions personnelles. Toutefois, ces dernières ont en partie été alimentées par des articles scientifiques.

Voici comment sera organisée cette partie. Pour poser le cadre, nous commencerons par aborder, dans la sous-partie 4.2, la question du « pourquoi ». Nous nous demanderons quels buts peut avoir l'utilisation des CT en musique (autres qu'elle-même). Nous proposerons qu'elle puisse avoir trois objectifs principaux : évoquer, lier et comprendre.

Ensuite, les réponses aux questions du « où » et du « comment » étant différentes selon lequel de ces trois objectifs est visé, nous traiterons ces deux questions de manière conjointe pour chaque

---

<sup>70</sup> Bien sûr, la compréhension d'une musique ne se restreint pas à l'analyse musicale. Elle fait aussi appel à d'autres domaines, comme l'histoire de la musique. Mais nous nous limitons dans le cadre de ce mémoire à l'analyse musicale.

objectif. Ainsi, chacune des sous-parties 4.3, 4.4 et 4.5 aborde un objectif et contient deux parties : une partie « où ? » et une partie « comment ? ».

Dans les parties qui traitent la question du « où », nous tenterons, non pas d'établir une liste exhaustive des domaines dans lesquels les CT peuvent être utilisées musicalement, mais d'en relever au moins un certain nombre. Dans les parties qui traitent la question du « comment », nous proposerons des méthodes systématiques pouvant servir à utiliser les CT en musique.

Enfin, dans la sous-partie 4.6, nous aborderons la question du « quand ». Nous verrons dans quelles circonstances les CT peuvent être utilisées en musique, et surtout, dans quelles circonstances elles sont susceptibles de ne pas pouvoir l'être.

## 4.2. Pourquoi ?

Si utiliser les CT en musique peut être un objectif en soi, cela peut aussi être un moyen d'atteindre un autre objectif. Dans ce cas, plusieurs objectifs peuvent être visés. Nous suggérons qu'il y en a principalement trois : évoquer, lier et comprendre. Avant de les détailler, il est nécessaire de poser les définitions suivantes :

- **Création artistique** : toute chose pouvant relever de l'art, qu'il s'agisse d'une œuvre à proprement parler, d'une improvisation, d'une installation, d'une performance, etc.
- **Médium** : moyen par lequel est transmise une information. Nous distinguons les médiums auditifs (musique, parole, etc.), des médiums non auditifs (image/vidéo/scène visuelle, dispositif tactile, gustatif, olfactif, etc.). Une création artistique peut contenir un seul médium (par exemple, une symphonie ne contient que de la musique) ou plusieurs (par exemple, un opéra contient de la musique, des paroles et une scène visuelle).

Détaillons à présent ce que nous entendons par « évoquer », « lier » et « comprendre » :

- **Évoquer**. Il s'agit, dans une création artistique contenant de la musique, de suggérer chez l'auditoire des stimuli non auditifs *via* des stimuli auditifs contenus dans la musique. Nous nommons l'effet correspondant à cet objectif *effet d'évocation*, ou évocation.
- **Lier**. Il s'agit, dans une création artistique contenant de la musique et au moins un médium non auditif, de créer un rapport particulier entre des stimuli auditifs contenus dans la musique et des stimuli non auditifs contenus dans le/un des médium(s) non auditif(s). Nous nommons l'effet correspondant à cet objectif *effet de liaison*, ou liaison.

- **Comprendre.** Il s'agit de comprendre, par l'analyse musicale, les évocations et/ou les liaisons qui sont en jeu dans une musique.

Tentons d'illustrer ces trois objectifs avec des exemples. Pour ce qui est d'évoquer, on peut penser à un·e improvisateur·trice voulant suggérer *via* son improvisation un objet petit, fin et pointu. On peut également penser à un·e compositeur·trice de musique de film qui désirerait évoquer ce même objet dans une scène qui, par ailleurs, ne présente à l'image aucun objet possédant ses attributs sensoriels — voire présente à l'image un objet possédant des attributs sensoriels *opposés*, c'est-à-dire un objet gros, épais et arrondi.

Pour ce qui est de lier, il peut s'agir d'un·e compositeur·trice de musique de jeu vidéo qui doit composer une musique pour une séquence de jeu présentant à l'image un objet petit, fin et pointu, et qui se demande, d'une part, quel rapport iel veut établir entre la musique et l'image de cet objet, et d'autre part, quelles sortes de sons iel doit employer pour établir ce rapport. On pourrait penser que se demander quel rapport on veut établir entre deux médiums pour les lier est inutile, puisque lier deux médiums signifie nécessairement établir un rapport de congruence entre les deux, un rapport où les stimuli se correspondent. Mais, comme nous le verrons par la suite, le rapport de congruence pourrait ne pas être le seul rapport par lequel on puisse lier deux médiums, ce qui rendrait pertinente la question du choix du rapport à établir entre la musique et le médium non auditif.

Enfin, pour ce qui est de comprendre, on peut penser à un·e analyste cherchant à déduire les stimuli non auditifs évoqués par une création musicale à partir des sons qu'elle donne à entendre (ou cherchant à mettre à l'épreuve les intuitions qu'iel a à ce sujet, si iel a déjà une idée de ce qui est évoqué). On peut également penser à un·e analyste cherchant à comprendre le rapport existant entre une musique et un médium non auditif — les deux étant contenus dans une même création artistique — à partir des stimuli que chacun de ces deux médiums donne à percevoir.

## 4.3. Évoquer

### 4.3.1. Où évoquer ?

Dans quels domaines peut-on évoquer des stimuli non auditifs *via* de la musique ? Voici notre réponse :

- dans les domaines proprement musicaux, c'est-à-dire qui n'implique aucun autre médium que le médium musical (qu'il s'agisse de musique pure ou de musique à programme) : composition, improvisation, interprétation ;
- dans les domaines impliquant de la musique et au moins un médium non auditif : opéra, cinéma, jeu vidéo, installation impliquant de la musique et un dispositif tactile, gustatif et/ou olfactif, etc.

La raison pour laquelle il nous paraît que le premier type de domaine est capable de produire un effet d'évocation est qu'il réunit des conditions semblables à certaines des expériences ayant mis en évidence des CT<sup>71</sup>, c'est-à-dire qu'il présente des stimuli auditifs, mais aucun stimulus non auditif particulier. De fait, c'est dans ces conditions que l'effet d'évocation semble pouvoir être le plus efficace, car alors il n'est perturbé par aucun stimulus non auditif<sup>72</sup>.

Il s'ensuit qu'évoquer des stimuli non auditifs dans le deuxième type de domaines est probablement plus difficile, car ces domaines mettent en jeu des stimuli non auditifs susceptibles de gêner les évocations non auditives causées par la musique. Mais il n'en reste pas moins *possible* de provoquer un effet d'évocation dans ce type de domaines. De fait, à titre personnel, lorsque nous visionnons, par exemple, un film, la musique nous évoque parfois des stimuli sensoriels non contenus dans la vidéo. Ceci montre qu'il est possible qu'une musique produise un effet d'évocation tout en « cohabitant » avec un médium non auditif au sein de la même création artistique. De plus, il nous paraît plausible que cette possibilité ne se limite pas à nous et qu'elle subsiste chez un grand nombre d'humains, quoique des études empiriques soient nécessaires pour le confirmer.

### 4.3.2. Comment évoquer ?

La méthode que nous proposons pour évoquer musicalement un objet se compose des étapes suivantes :

1. Choisir l'objet que l'on veut évoquer musicalement.
2. Examiner les valeurs des attributs non auditifs de cet objet.
3. Consulter les tableaux des CT (cf. sous-partie 3.6) et déterminer les valeurs d'attributs auditifs correspondant aux valeurs des attributs non auditifs de l'objet.
4. Mettre en avant ces valeurs d'attributs auditifs dans la musique.

71 Par exemple, l'expérience n°2 dans cet article, déjà cité : KNÖFERLE, Klemens M., WOODS, Andy, KÄPPLER, Florian & SPENCE, Charles, « That Sounds Sweet: Using Cross-Modal Correspondences to Communicate Gustatory Attributes », *Psychology & Marketing*, vol. 32, n°1, 2015, pp. 107-120.

72 En tout cas, par aucun stimulus non auditif contenu dans la création artistique.

Voici un exemple d'application de cette méthode dans le domaine de l'interprétation :

1. Un·e interprète choisit d'évoquer une aiguille à travers son interprétation (possiblement parce que la pièce qu'il interprète s'intitule « Aiguille »).
2. Il considère qu'une aiguille est un objet petit, fin, pointu et léger.
3. Il consulte les tableaux des CT et constate que les stimuli non auditifs petits, fins, pointus et légers sont associés à des sons aigus, faibles et rapides.
4. Il met en avant les sons aigus, faibles et rapides contenus dans la pièce qu'il interprète, par exemple en les jouant plus forts, avec une articulation singulière ou un timbre particulier.

## 4.4. Lier

### 4.4.1. Où lier ?

Dans quels domaines peut-on lier une musique à un médium non auditif ? De toute évidence, on ne peut pas le faire dans les domaines qui ne font appel qu'à de la musique. L'objectif « lier » se limite donc aux domaines faisant appel à de la musique *et à au moins un médium non auditif*. Comme dit plus haut, ces domaines incluent l'opéra, le cinéma, le jeu vidéo, les installations impliquant de la musique et un dispositif tactile, gustatif et/ou olfactif, etc.

### 4.4.2. Comment lier ?

Comment créer un rapport particulier entre des stimuli auditifs contenus dans une musique et des stimuli non auditifs contenus dans un médium non auditif ? Avant de répondre à cette question, il convient de se pencher sur la notion de « rapport ». On peut en effet se demander : quels rapports sont-ils possibles ? Une réponse est proposée par Georgiou. Dans sa thèse<sup>73</sup>, elle propose une méthode systématique pour composer de la musique de films non narratifs<sup>74</sup>. Cette méthode repose sur l'établissement de rapports particuliers entre la musique et la vidéo par l'intermédiaire de correspondances audio-visuelles<sup>75</sup>. En s'appuyant sur la théorie générale des multimédias de Nicholas

---

73 GEORGIU, Chrystalla, *Beyond narrative: A cross-modal approach to soundtrack composition*, thèse, université du Hertfordshire, 2014.

74 C'est-à-dire de films n'ayant pas d'histoire, pas de récit.

75 Nous ne sommes pas sûr que les correspondances transmodales décrites dans la thèse de Georgiou soient étayées par des études empiriques, comme le sont celles dont nous parlons dans la partie 3 de ce mémoire. En effet, Georgiou n'évoque pas la manière dont ces correspondances ont été établies. Elle renvoie bien à l'ouvrage duquel elle les a tirées, *Analyzing Musical Multimedia* de Nicholas Cook, mais nous n'avons pas eu la possibilité de le consulter. Toutefois, l'aspect de la thèse de Georgiou qui nous intéresse ici n'est pas les correspondances transmodales en tant que telles, mais la méthode proposée pour les utiliser dans l'objectif de lier une musique à une vidéo.

Cook et le modèle « encodage-décodage » de la communication de Stuart Hall, Georgiou propose trois rapports possibles entre la musique et la vidéo :

<i>Théorie générale des multimédias (Cook)</i>	<i>Modèle « encodage-décodage » de la communication (Hall)</i>	<b>Les trois rapports possibles entre une musique et une vidéo selon Georgiou</b>	<b>Explication</b>
<i>Conformité</i>	<i>Lecture dominante</i>	Reconnaissance des événements visuels par similarité	La musique et la vidéo sont congruentes.
<i>Complémentation</i>	<i>Lecture négociée</i>	Absence de reconnaissance des événements visuels (compromis)	La musique et la vidéo sont en partie congruentes et en partie incongruentes.
<i>Conflit</i>	<i>Lecture oppositionnelle</i>	Reconnaissance des événements visuels par opposition	La musique et la vidéo sont incongruentes.

Tableau 14. Les trois rapports possibles entre une musique et une vidéo dans un film non narratif selon Georgiou (et leur similitude avec les théories de Cook et de Hall).

Le premier rapport possible, la reconnaissance des événements visuels par *similarité*, est le plus simple. Il consiste en une relation de congruence entre les stimuli auditifs ou sons (qui forment la musique) et les stimuli visuels (qui forment la vidéo). Il peut se manifester, par exemple, par l'usage de sons faibles et graves pour accompagner des événements visuels sombres. La reconnaissance des événements visuels par *opposition* est le rapport inverse. Ici, les stimuli auditifs et visuels sont incongruents. Cela peut conduire à l'usage de sons forts et aigus pour accompagner une vidéo sombre. Enfin, l'absence de reconnaissance des événements visuels est un rapport intermédiaire, qui relève d'un *compromis* entre les rapports de similarité et d'opposition. Dans ce cas, les stimuli auditifs et visuels sont en partie congruents et en partie incongruents. Une vidéo sombre pourra par exemple être accompagnée par des sons faibles et aigus. Georgiou note que ce rapport offre beaucoup plus de possibilités que les deux autres. En effet, il constitue un continuum séparant le pôle « similarité » et le pôle « opposition » : il y a une multitude de possibilités entre le rapport d'absolue congruence et le rapport d'absolue incongruence<sup>76</sup>.

Nous pensons que le modèle de Georgiou est généralisable à des paires de médiums autres que la paire musique-vidéo, et à des domaines autres que le cinéma non narratif. Nous avons représenté cette généralisation dans le tableau 15. Pour simplifier la lecture, nous y nommons chacun des trois rapports par un unique mot : similarité, compromis et opposition.

76 GEORGIU, Chrystalla, *Beyond narrative*, op. cit., p. 26.

Les trois rapports possibles entre une musique et un médium non auditif	Explication
Similarité	La musique et le médium non auditif sont congruents.
Compromis	La musique et le médium non auditif sont en partie congruents et en partie incongruents.
Opposition	La musique et le médium non auditif sont incongruents.

Tableau 15. Les trois rapports possibles entre une musique et un médium non auditif.

Une incertitude demeure cependant concernant ce modèle généralisé. Car s'il prétend décrire les différents rapports *objectifs* par lesquels on peut lier une musique et un médium non auditif, il ne dit pas si chacun de ces rapports produira concrètement un effet *subjectif* sur l'auditoire, ce qui est pourtant ce qui nous intéresse ici<sup>77</sup>. Se pose donc la question suivante : chacun de ces rapports produira-t-il à lui seul un effet de liaison sur l'auditoire ? Nous ne pouvons apporter de réponse complète et certaine à cette question. Ce que nous pouvons dire, c'est qu'il est possible que le rapport de similarité produise un *effet d'unité* (effet suivant lequel deux stimuli faisant appel à deux sens différents sont d'autant plus unifiés qu'ils sont congruents — ou dans un rapport de similarité, pour reprendre les termes de Georgiou)<sup>78</sup>. Néanmoins, des études empiriques sont nécessaires pour confirmer l'existence de cet effet de liaison et tester l'existence d'effets de liaison pour les deux autres rapports.

Mais supposons que les trois rapports produisent chacun un effet particulier. La méthode que nous proposerions pour lier une musique et un médium non auditif dans une création artistique se composerait des étapes suivantes :

1. Si la création artistique contient plus d'un médium non auditif, choisir celui avec lequel on veut lier la musique.
2. Examiner les valeurs d'attributs non auditifs utilisées dans ce médium (ces valeurs pouvant bien évidemment varier au cours du temps).
3. Choisir le rapport que l'on veut établir entre le médium non auditif et la musique (similarité, compromis ou opposition)<sup>79</sup>.
4. Consulter les tableaux des CT (cf. sous-partie 3.6) et déterminer les valeurs d'attributs auditifs permettant d'établir ce rapport.

<sup>77</sup> Bien évidemment, rien n'empêche de vouloir utiliser ces trois rapports sans se soucier du fait qu'ils auront un effet ou non sur l'auditoire.

<sup>78</sup> Cf. la sous-partie 2.3.3 qui traite de l'intégration multisensorielle.

<sup>79</sup> Si un ou plusieurs de ces rapports ne produisent aucun effet de liaison sur l'auditoire (ou un effet de liaison trop négligeable), alors l'éventail des choix de rapports que peut faire une personne cherchant à produire un effet de liaison sur l'auditoire est plus réduit (voire inexistant si un seul des trois rapports produit un effet de liaison, ou si aucun n'en produit).

5. Mettre en avant ces valeurs dans la musique.

Voici un exemple d'application de cette méthode pour une installation artistique impliquant de la musique, un médium tactile et un médium visuel :

1. Le·la compositeur·trice qui compose la musique de cette installation choisit de lier la musique au médium tactile.
2. I·el observe que le médium tactile utilise essentiellement des stimuli tactiles hauts, petits et légers.
3. I·el choisit d'établir un rapport de similarité entre le médium tactile et la musique.
4. I·el consulte les tableaux des CT et observe que les stimuli tactiles hauts, petits et légers sont associés à des sons aigus et rapides.
5. I·el met en avant des sons aigus et rapides dans sa composition, par exemple en y mettant davantage de sons de cette sorte que de sons graves et lents.

## 4.5. Comprendre

### 4.5.1. Où comprendre ?

Dans quels domaines peut-on tenter de comprendre, par l'analyse musicale, les évocations et/ou les liaisons qui sont en jeu dans une musique ? Dans les domaines impliquant de la musique et dans lesquels des CT sont utilisées en lien avec la musique — soit pour *évoquer* musicalement, soit pour former une *liaison* entre la musique et un médium non auditif. Cela inclut tous les domaines évoqués jusqu'à présent : à la fois les domaines qui ne font appel qu'à de la musique (composition, improvisation, interprétation) et les domaines qui font appel à de la musique et à au moins un médium non auditif (opéra, cinéma, jeu vidéo, installation impliquant de la musique et un dispositif tactile, gustatif et/ou olfactif, etc.).

### 4.5.2. Comment comprendre ?

Pour analyser des *évocations*, nous proposons la méthode suivante :

1. Examiner les valeurs d'attributs auditifs utilisés dans la musique (ces valeurs pouvant bien évidemment varier au cours du temps).

2. Consulter les tableaux des CT (cf. sous-partie 3.6) et déterminer les valeurs d'attributs non auditifs correspondant à ces valeurs d'attributs auditifs.
3. Conclure que la musique aura tendance à évoquer ces valeurs d'attributs non auditifs.

Voici un exemple d'application :

1. Un·e analyste observe qu'une musique utilise essentiellement des sons de piano aigus, legato et très consonants.
2. Iel consulte le tableau des CT audio-gustatives et observe que les sons de piano aigus, legato et très consonants sont associés au goût sucré.
3. Iel conclut que, pour ce qui est des évocations gustatives, la musique aura tendance à évoquer le goût sucré (du moins, davantage que les autres saveurs de base).

Pour analyser des *liaisons*, nous proposons la méthode suivante :

1. Si la création artistique contient plus d'un médium non auditif, choisir celui dont on veut analyser la liaison avec la musique.
2. Examiner les valeurs d'attributs sensoriels mises en avant dans la musique et dans le médium non auditif choisi.
3. Comparer ces valeurs à la lumière des tableaux des CT (cf. sous-partie 3.6).
4. Tirer une des conclusions suivantes :
  - Si les valeurs sont totalement congruentes, conclure que les deux médiums ont un rapport de similarité (et pourraient produire un effet d'unité).
  - Si les valeurs sont partiellement congruentes et partiellement incongruentes, conclure que les deux médiums ont un rapport de compromis.
  - Si les valeurs sont totalement incongruentes, conclure que les deux médiums ont un rapport d'opposition.

Voici un exemple d'application de cette méthode pour une installation artistique impliquant de la musique, un médium visuel et un médium olfactif :

1. Un·e analyste choisit d'analyser la liaison entre la musique et le médium olfactif.
2. Iel examine les valeurs d'attributs sensoriels mises en avant dans ces deux médiums. Iel constate que la musique met en avant des sons de piano, et que le médium olfactif met en avant des odeurs moyennement complexes, très agréables, très peu âcres, moyennement florales et très peu fruitées.

3. Iel compare ces valeurs à la lumière des tableaux des CT. Iel constate que le timbre de piano est congruent avec la complexité, l'agréabilité et le caractère âcre des odeurs ; incongruent avec leur caractère floral et fruité.
4. Iel constate que les valeurs d'attributs auditifs et olfactifs sont partiellement congruentes et partiellement incongruentes. Iel en conclut que la musique et les odeurs ont un rapport de compromis.

## 4.6. Quand ?

Dans cette sous-partie, nous allons traiter la question du « quand ? », à savoir : quand est-ce que les CT peuvent ou ne peuvent pas être utilisées en musique ? En pratique, nous nous pencherons surtout sur la question de savoir quand elles *ne* peuvent *pas* être utilisées en musique. Autrement dit, nous nous interrogerons sur les problèmes éventuels auxquels pourrait se heurter une personne désirant utiliser les CT en musique. Nous n'affirmons pas que tous ces problèmes existent, mais simplement qu'ils pourraient exister et qu'ils méritent considération.

Les deux premiers problèmes que nous allons aborder sont liés à l'utilisation des CT pour *évoquer*, et pour *comprendre ce qui est évoqué*. Nous les appelons le problème de l'imprécision et le problème de l'incohérence. Le troisième et dernier problème est lié à l'utilisation des CT pour *évoquer*, pour *lier*, et pour *comprendre ce qui est évoqué ou lié*. Nous l'appelons le problème de la concurrence des effets.

### 4.6.1. Le problème de l'imprécision

Le *problème de l'imprécision* est qu'il semble difficile d'évoquer musicalement des valeurs précises d'attributs non auditifs, ou de comprendre quelles valeurs précises d'attributs non auditifs sont évoquées par une pièce de musique donnée. Par exemple, il semble facile d'évoquer un objet petit grâce aux CT : il suffit de mettre en avant des sons aigus, faibles et rapides. Mais il semble difficile d'évoquer un objet dont la taille est de 5,72 cm. De même, il est aisé de comprendre qu'une musique qui met en avant des sons aigus, faibles et rapides évoque un objet petit, mais il est plus difficile de deviner la taille exacte de cet objet (s'il a une taille exacte).

Crisinel et al.<sup>80</sup> suggèrent que le problème de l'imprécision existe bel et bien, du moins en ce qui concerne les CT audio-olfactives. Ils ont réalisé deux expériences. Dans la première expérience, les participant·e·s devaient humer des odeurs particulières (dont les odeurs de cookie au gingembre, d'orange confite et de crème brûlée) puis les faire correspondre à des hauteurs sonores et types d'instruments particuliers. Cette première expérience a mis en évidence un certain nombre de CT. Pour la seconde expérience, ces CT ont été utilisées pour composer trois morceaux de musique, chacun ayant un rapport de congruence avec une des trois odeurs sus-citées (cookie au gingembre, orange confite et crème brûlée). Dans cette deuxième expérience, on donnait à humer une de ces trois odeurs aux participant·e·s<sup>81</sup>, puis on leur faisait écouter les trois morceaux de musique. Ils devaient choisir le morceau qui correspondait le plus à l'odeur qu'ils venaient de humer. Cette procédure était répétée pour les deux autres odeurs.

Quels sont les résultats de cette deuxième expérience ? Parmi les trois odeurs présentées, seules deux ont donné lieu à des préférences significatives dans le choix de morceau : les odeurs d'orange confite et de cookie au gingembre. De plus, là où l'odeur d'orange confite a bien été associée à la pièce censée lui correspondre, ce n'était pas le cas de l'odeur de cookie au gingembre, qui a été majoritairement associée à la pièce censée correspondre à l'odeur de crème brûlée.

Comme le suggèrent les auteur·trice·s de l'étude :

ces résultats mitigés soulignent la difficulté de composer une pièce de musique que la majorité de la population associera à une ou des odeurs spécifiques. Ici, l'intuition des compositeur·trice·s ne semble pas suffisante. [...] Il se peut que les associations entre, d'une part les goûts ou les odeurs, et d'autre part la musique, ne soient pas suffisamment spécifiques pour représenter un seul stimulus, et qu'elles ne puissent donc être utilisées que pour évoquer des catégories générales comme les saveurs de base<sup>82</sup>.

(La dernière phrase de cette citation met les goûts et les odeurs dans la même catégorie. De fait, ces stimuli sont considérés par certain·e·s chercheur·se·s comme ayant probablement les mêmes propriétés psychologiques. Cela s'explique par le fait que les goûts et les odeurs sont souvent fortement associés, notamment en raison de leur fréquente cooccurrence dans la vie quotidienne, lors de repas<sup>83</sup>.)

---

80 CRISINEL, Anne-Sylvie, JACQUIER, Caroline, DEROY, Ophelia & SPENCE, Charles, « Composing with Cross-Modal Correspondences: Music and Odors in Concert », *Chemosensory Perception*, vol. 6, n°1, 2013, pp. 45-52.

81 Les participant·e·s de la seconde expérience étaient différent·e·s de ceux de la première.

82 *Ibid.*, pp. 50-51. Traduction personnelle.

83 Voir à ce sujet : STEVENSON, Richard J., BOAKES, Robert A., « Sweet and Sour Smells: Learned Synesthesia Between the Senses of Taste and Smell » in *The Handbook of Multisensory Processes*, Gemma A. Calvert, Charles Spence & Barry E. Stein édés., Cambridge, Mass: MIT Press, 2004, pp. 69-83.

À la fin de cette citation, Crisinel et al. suggèrent que, si des stimuli non auditifs très spécifiques ne peuvent pas (ou peuvent difficilement) être évoqués musicalement, des stimuli non auditifs plus généraux ou « basiques » le peuvent. Et c'est effectivement ce qui a été suggéré par certaines études empiriques, comme celle de Mesz et al.<sup>84</sup>. Dans cette étude, il a été demandé à des pianistes professionnel·le·s d'improviser en s'inspirant d'un des noms de saveurs de base (« sucré », « acide », « salé » ou « amer »). Il a été observé que, lorsque des non-musicien·ne·s écoutent ces improvisations, iels arrivent à retrouver le nom de saveur duquel elles sont inspirées (c'est-à-dire qu'iels arrivent à deviner que telle improvisation est inspirée du nom de saveur « sucré », que telle autre est inspirée du nom de saveur « amer », etc.).

Les articles de Crisinel et al. et de Mesz et al. suggèrent que le problème de l'imprécision existe bel et bien en ce qui concerne les CT audio-gustatives et audio-olfactives : si des stimuli basiques ou généraux peuvent facilement être évoqués musicalement, des stimuli précis le peuvent plus difficilement. Néanmoins, des études empiriques supplémentaires sont nécessaires pour le confirmer, et pour déterminer si le problème existe également pour les CT audio-visuelles et audio-tactiles.

#### 4.6.2. Le problème de l'incohérence

Comme le problème de l'imprécision, le *problème de l'incohérence* pourrait valoir autant lorsque l'on cherche à évoquer, que lorsqu'on cherche à comprendre par l'analyse musicale ce qui est évoqué. Pour ce qui concerne l'évocation, le problème de l'incohérence peut être formulé ainsi : il est difficile de ne pas créer d'incohérence lorsqu'on emploie ou évoque des valeurs différentes, voire opposées, d'un même attribut sensoriel (qu'il soit auditif ou non-auditif). Pour ce qui est de l'analyse musicale d'évocations, le problème peut être formulé de la façon suivante : il est difficile de comprendre ce qui est évoqué par une musique lorsque des valeurs différentes, voire opposées, d'un même attribut non auditif semblent être évoquées. Bien que certaines affirmations présentes dans cette sous-partie sont valables pour l'analyse d'évocations, nous allons ici nous concentrer sur le problème de l'incohérence du point de vue de l'évocation.

---

84 MESZ, Bruno, TREVISAN, Marcos A. & SIGMAN, Mariano, « The Taste of Music », *Perception*, vol. 40, n°2, 2011, pp. 209-219.

Une première façon d'interpréter ce problème est de dire qu'il semble difficile de se reposer sur un unique attribut auditif pour évoquer un ou des objet(s) lorsqu'on ne peut employer qu'une seule valeur de cet attribut à la fois, et que les valeurs des attributs non auditifs de cet ou ces objet(s) sont associées à des valeurs différentes, voire opposées, dudit attribut auditif. Par exemple, si l'on ne peut utiliser qu'une seule hauteur sonore à la fois (par exemple, dans une pièce monodique), il paraît difficile de se reposer uniquement sur l'attribut de la hauteur sonore pour évoquer un objet simultanément

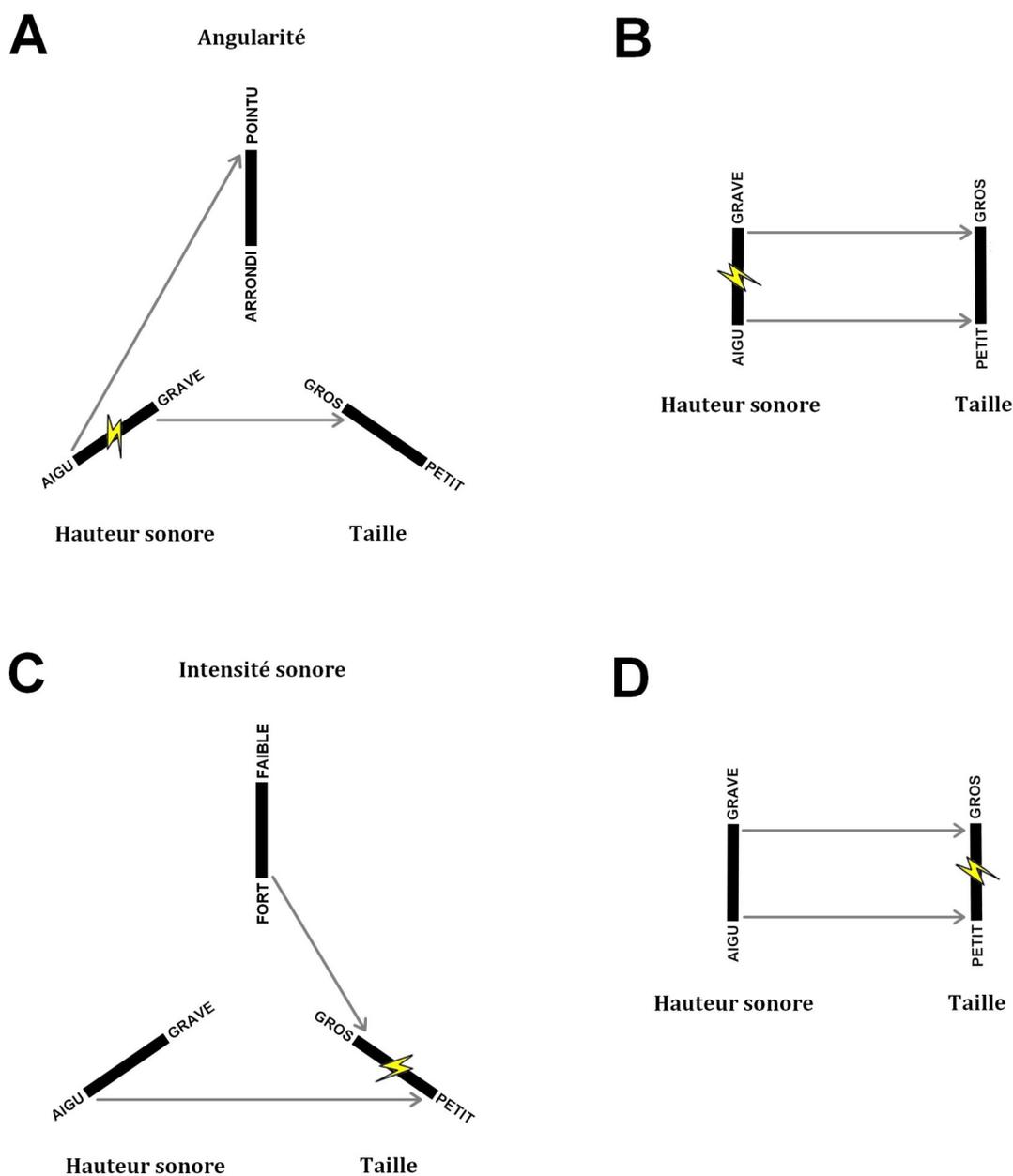


Figure 7. Plusieurs incohérences possibles. A : si l'on veut évoquer un objet gros et pointu uniquement à travers la hauteur sonore d'un son musical, ce son devra-t-il être aigu ou grave ? B : si l'on veut évoquer deux objets, l'un étant gros, l'autre petit, uniquement à travers la hauteur sonore d'un son musical, ce son devra-t-il être aigu ou grave ? C : si l'on veut utiliser un son aigu et fort, l'objet évoqué sera-t-il gros ou petit ? D : si l'on veut utiliser deux sons, l'un aigu, l'autre grave, l'objet évoqué sera-t-il gros ou petit ?

ment gros et pointu, car la grosseur est associée à des sons graves tandis que le caractère pointu est associé à des sons aigus (figure 7A). Si l'on ne peut utiliser qu'une seule hauteur sonore à la fois, il ne paraît pas non plus évident d'utiliser uniquement la hauteur sonore pour évoquer simultanément deux objets, l'un étant gros, l'autre petit — car la grosseur est associée à des sons graves tandis que la petitesse est associée à des sons aigus (figure 7B).

Une solution potentielle pour remédier à cette première version du problème (et donc à ses deux sous-versions) serait de faire en sorte que chaque valeur d'attribut non auditif que l'on veut évoquer le soit, non pas *via* un seul attribut auditif, mais *via* plusieurs attributs auditifs différents. Dans le cas où l'on veut évoquer un objet gros et pointu, cela pourrait se traduire par l'utilisation de sons simultanément forts (évoquant ainsi un objet gros) et aigus (évoquant ainsi un objet pointu). Toutefois, cette solution pose un autre problème. Les sons aigus sont certes associés à des objets pointus, mais ils sont aussi associés à des objets petits. Or, comme nous venons de le voir, les sons forts, eux, sont associés à des objets gros. Ceci produit donc une nouvelle incohérence.

Cela nous amène à la deuxième version du problème de l'incohérence. Celle-ci consiste à affirmer qu'il semble difficile de prévoir quelle valeur d'un attribut non auditif sera évoquée lorsque des valeurs d'attributs auditifs utilisées simultanément sont associées à des valeurs différentes, voire opposées, de cet attribut non auditif. Pour reprendre l'exemple précédent, lorsque des sons simultanément forts et aigus sont utilisés, il paraît difficile de prévoir si l'objet évoqué sera gros ou petit, car les sons forts sont associés à des objets gros, tandis que les sons aigus sont associés à des objets petits (figure 7C). Plus simplement, lorsque des sons graves et des sons aigus sont employés en même temps, il semble difficile de prévoir la taille de l'objet évoqué, car les sons graves sont associés à des objets gros, tandis que les sons aigus sont associés à des objets petits (figure 7D).

Il est probable que l'incohérence présente dans ces deux derniers exemples puisse être résolue facilement. En effet, on peut imaginer qu'au lieu d'évoquer un seul objet, la musique en évoque deux, auquel cas chacune des différentes valeurs de l'attribut non auditif peut être assignée à un objet différent, ce qui fait disparaître l'incohérence. Un moyen efficace de renforcer l'impression que la musique évoque deux objets différents au lieu d'un seul, et donc d'estomper l'éventuelle incohérence de l'évocation, est d'accentuer les différences entre les sons. Dans le cas où l'on emploie simultanément des sons forts et aigus, cela impliquerait de faire en sorte qu'il n'y ait aucun son qui soit à la fois fort et aigu, et que les sons forts soient très différents des sons aigus. De fortes différences de durée et de timbre peuvent être utilisées : par exemple, les sons forts peuvent être des notes longues au violoncelle, et les sons aigus des notes brèves à la flûte. Dans le cas où l'on utilise en même temps

des sons graves et des sons aigus, le principe s'applique de façon similaire : on accentue les différences entre les sons graves et les sons aigus.

À travers ces deux versions et ces quatre sous-versions du problème de l'incohérence, nous voyons qu'à première vue, l'utilisation ou l'évocation de valeurs différentes, voire opposées, d'un attribut sensoriel est dans certains cas ardue et difficilement contournable. Les CT sont si nombreuses et interconnectées que les utiliser en musique semble être un vrai défi. Eitan & Tamir-Ostrover<sup>85</sup> reconnaissent l'existence et la difficulté apparente du problème de l'incohérence<sup>86</sup>, du moins en ce qui concerne les CT audio-visuelles relatives à l'espace et au mouvement<sup>87</sup> :

Les analogies auditives de l'espace et du mouvement ne génèrent pas toujours des correspondances audio-visuelles univoques. Au contraire, elles créent souvent des incohérences, des correspondances opposées impliquant des dimensions statiques et dynamiques, des perspectives différentes, ou des aspects différents d'un attribut spatial ou cinétique (par exemple, lorsqu'il s'agit de représenter un objet loin de nous — et donc subjectivement *petit* —, ou bien la *grande* distance qui nous sépare de cet objet). Comment un·e compositeur·trice cherchant à évoquer musicalement des aspects de l'espace et du mouvement physiques pourrait-iel résoudre un casse-tête aussi ardu en apparence ?

Dans la suite de leur article, Eitan & Tamir-Ostrover tentent de montrer que le problème de l'incohérence n'est pas véritablement un problème. En effet, iels affirment que, plutôt que de compromettre les effets d'évocation, les incohérences offrent au contraire de nouvelles opportunités, puisqu'elles peuvent être utilisées pour évoquer des espaces et mouvements paradoxaux, impossibles à percevoir dans le monde physique (un des exemples donnés est celui d'un objet à la fois extrêmement proche et extrêmement éloigné). Toutefois, il nous semble que les arguments invoqués par Eitan & Tamir-Ostrover, d'ordre théorique, ne suffisent pas à démontrer que les incohérences peuvent évoquer de tels paradoxes. Nous croyons que la bonne manière de le démontrer serait d'observer *concrètement*, *via* des études empiriques, que les personnes imaginent des espaces et mouvements paradoxaux.

Bien qu'Eitan & Tamir-Ostrover ne proposent pas d'autres « solutions » au problème de l'incohérence, on peut en déduire une de leur exposé, quoiqu'il ne s'agisse que d'une solution potentielle. C'est cette solution que nous allons maintenant tenter d'expliquer.

---

85 EITAN, Zohar & TAMIR-OSTROVER, Hila, « The Sound of an Endless Column: How Music Imagines Unimaginable Space », in *The Oxford Handbook of Sound and Imagination*, vol. 1, Mark Grimshaw-Aagaard, Mads Walther-Hansen & Martin Knakkegaard eds., New York, NY: Oxford University Press, coll. « Oxford Handbooks », 2019, p. 14. Traduction personnelle. Les italiques sont de notre fait.

86 Bien qu'iels ne nomment pas ce problème, ni par l'expression « problème de l'incohérence », ni par une autre expression.

87 Un certain nombre de ces CT manquent d'ailleurs à la partie 3 de ce mémoire, et pourraient la compléter.

Une partie importante du propos d'Eitan & Tamir-Ostrover consiste à analyser une étude pour piano de György Ligeti intitulée *Colonne sans fin*<sup>88</sup>, sensée représenter musicalement une colonne infiniment grande. Cette étude consiste en plusieurs montées (de hauteur sonore) qui se chevauchent et qui, pour une majorité d'entre elles, partent du registre grave. Au cours de leur analyse, Eitan & Tamir-Ostrover tentent de montrer, entre autres, que la pièce évoque deux choses : d'une part, la taille visuellement infinie de la colonne (grande taille visuelle), et d'autre part, l'ascension infinie du regard d'un·e observateur·trice qui, fixant d'abord la base de la colonne, lèverait progressivement les yeux pour tenter d'en percevoir le sommet (augmentation de la hauteur visuelle). Selon eux, la grande taille visuelle serait évoquée (notamment) *via* le registre grave, maintenu pendant la majorité de la pièce par des retours à ce registre au début de chaque montée ; quant à l'augmentation de la hauteur visuelle, elle serait suscitée (notamment) *via* les montées elles-mêmes.

Bien qu'Eitan & Tamir-Ostrover ne le précisent pas, cette façon d'analyser les choses procède d'un choix. Un choix qui a consisté à considérer que le retour constant au registre grave évoque une grande *taille* visuelle, et que l'augmentation de la hauteur sonore évoque une augmentation de la *hauteur* visuelle. Or, il aurait été possible de faire le choix inverse, c'est-à-dire de considérer que le retour constant au registre grave évoque une faible *hauteur* visuelle, et que l'augmentation de la hauteur sonore évoque une diminution de la *taille* visuelle (ce qui aurait évidemment amené à remettre en question l'objet et les événements évoqués par la pièce). Ainsi, Eitan & Tamir-Ostrover auraient pu présenter les deux options sans affirmer que l'une est plus probable que l'autre, sans dire que l'une a plus de chances de refléter l'« expérience imaginative » des auditeur·trice·s que l'autre<sup>89</sup>. Néanmoins, ils ont choisi de présenter la première analyse plutôt que la seconde, précisément parce qu'*iels l'ont trouvée plus probable*.

C'est ici que transparait une solution potentielle au problème de l'incohérence. Qu'est-ce qui a conduit Eitan & Tamir-Ostrover à considérer que les CT entre le registre grave et une grande taille et entre une augmentation de la hauteur sonore et une augmentation de la hauteur visuelle étaient plus probables que celles entre le registre grave et une faible hauteur visuelle et entre une augmentation de la hauteur sonore et une diminution de la taille ? Une réponse plausible est que les premières CT sont plus intuitives, plus saillantes, plus *intenses* que les secondes. De fait, si Ligeti est parvenu à utiliser des valeurs différentes de la hauteur sonore sans créer d'incohérence dans les évocations, on

---

88 Il s'agit de l'étude n°14 du livre II de ses études pour piano.

89 Ce que nous voulons souligner ici est simplement ceci : si un algorithme d'analyse musicale tentait de déduire ce qu'évoque la *Colonne sans fin* avec pour seules données en entrée l'enregistrement de la pièce (ou la partition) et les informations présentées dans notre mémoire, elle ne pourrait pas trancher entre l'un ou l'autre choix d'analyse. Cela ne remet pas en cause la possibilité que, si un groupe d'humains écoute la pièce ou lit la partition (même sans connaître le titre de la pièce ni l'intention de Ligeti), la majorité d'entre eux trouve le choix d'analyse d'Eitan & Tamir-Ostrover plus évident, plus intuitif que le choix alternatif. Mais si l'on fait fi de nos intuitions, et que l'on se base uniquement sur notre raison et sur les informations présentées dans ce mémoire, rien ne nous permet de trancher.

peut aisément supposer que c'est en jouant sur cette différence d'intensité. On peut donc supposer que, *de manière générale*, il soit possible de jouer sur les différences d'intensité entre les CT impliquant l'ouïe pour utiliser (ou évoquer) des valeurs différentes, voire opposées, d'un même attribut sensoriel, et ce sans créer d'incohérence flagrante. Si cela est vrai, cela pourrait résoudre toutes les versions du problème de l'incohérence.

Mais la validité de cette solution (si toutefois le terme « solution » est approprié, puisque l'existence du problème de l'incohérence n'est pas établie avec certitude) nécessite d'être étudiée empiriquement. Plus généralement, le problème de l'incohérence pourra être mieux compris et théorisé avec une approche empirique, mesurant les effets réels et concrets des incohérences sur l'imaginaire des auditeur·trice·s.

### 4.6.3. Le problème de la concurrence des effets

Supposons une création artistique contenant de la musique et un médium non auditif. Comme nous l'avons vu dans les sous-parties 4.3.1 et 4.4.1, cette création artistique peut produire sur l'auditoire un effet de liaison autant qu'un effet d'évocation. On pourrait donc penser qu'un·e musicien·ne qui souhaiterait utiliser des CT dans cette création artistique aurait l'entière liberté de produire l'un ou l'autre effet. Mais cela serait imprudent, car il n'est pas impossible que, au moins sous certaines conditions, l'un des deux effets domine l'autre, c'est-à-dire qu'il soit plus intense ou saillant que l'autre. Par exemple, il se pourrait que, lorsque la musique et le médium non auditif ont un rapport d'opposition, l'effet d'évocation supplante l'effet de liaison, ou que, lorsqu'ils ont un rapport de similitude, l'effet de liaison supplante l'effet d'évocation.

Si ce risque est vérifié, cela créerait un problème que nous nommerons *problème de la concurrence des effets*. Ce problème résiderait dans le fait qu'un·e musicien·ne souhaitant produire chez l'auditoire l'un ou l'autre des effets d'évocation et de liaison se heurte à ce que cet effet soit atténué voire supplanté par l'autre effet, plus intense. Le problème de la concurrence des effets pourrait donc se manifester de deux manières :

- Alors que l'on cherche à *lier* la musique et le médium non auditif, chacun des deux médiums, ou un seul, produit un effet d'évocation, et ce dernier atténue voire supplante l'effet de liaison que l'on cherchait à produire.

- Alors que l'on cherche à *évoquer* un objet *via* la musique, celle-ci et le médium non auditif produisent un effet de liaison, et ce dernier atténue voire supplante l'effet d'évocation que l'on cherchait à produire.

Le problème de la concurrence des effets pourrait aussi valoir lorsqu'on cherche à *comprendre*. En effet, lorsqu'on analyse une création artistique contenant une musique et au moins un médium non auditif, il pourrait être difficile de déterminer lequel des deux effets a le plus d'impact sur l'auditoire.

L'existence du problème de la concurrence des effets nécessite d'être mise à l'épreuve par des études empiriques. De plus, si son existence est vérifiée, la recherche empirique pourrait également permettre de mettre au point des solutions pour le résoudre.

## 4.7. Récapitulatif

Dans cette partie, nous avons abordé la seconde partie de la problématique, qui traite de l'utilisation des CT en musique (le mot « musique » étant compris comme incluant l'analyse musicale), et ce à travers quatre questions : Pourquoi ? Où ? Comment ? Quand ? Voici les réponses qui ont été proposées :

- **Pourquoi ?** L'utilisation des CT peut avoir principalement trois objectifs :
  - évoquer des stimuli non auditifs *via* de la musique (évocation)
  - lier une musique à un médium non auditif (liaison)
  - comprendre, par l'analyse musicale, les évocations et liaisons qui sont en jeu dans une musique (compréhension).
- **Évoquer**
  - **Où évoquer ?** Dans les créations artistiques ne contenant que de la musique, et dans les créations artistiques contenant de la musique et au moins un médium non auditif.
  - **Comment évoquer ?** *Via* la méthode que nous avons proposée (cf. sous-partie 4.3.2).
- **Lier**
  - **Où lier ?** Dans les créations artistiques contenant de la musique et au moins un médium non auditif.
  - **Comment lier ?** *Via* la méthode que nous avons proposée (cf. sous-partie 4.4.2). Il y a trois manières de lier une musique à un médium non auditif : établir un rapport de similarité (congruence totale), de compromis (congruence partielle/incongruence

partielle) ou d'opposition (incongruence totale). Mais il n'est pas certain que chacun de ces rapports produise un effet de liaison spécifique sur l'auditoire.

- **Comprendre**

- **Où comprendre ?** Les créations artistiques susceptibles d'être analysées musicalement à la lumière des CT sont celles ne contenant que de la musique et celles contenant de la musique et au moins un médium non auditif.
- **Comment comprendre ?** *Via* la méthode que nous avons proposée (cf. sous-partie 4.5.2).

- **Quand ?** Plusieurs problèmes pourraient se poser lors de l'utilisation des CT en musique :

- **Le problème de l'imprécision.** Comment évoquer musicalement une valeur précise d'un attribut non auditif ? Et comment déterminer quelle valeur précise d'un attribut non auditif est évoquée par une pièce de musique donnée ?
- **Le problème de l'incohérence.** Comment ne pas créer d'incohérence lorsqu'on emploie ou évoque musicalement des valeurs différentes, voire opposées, d'un même attribut sensoriel (qu'il soit auditif ou non-auditif) ? Et comment comprendre ce qui est évoqué par une pièce de musique lorsque des valeurs différentes, voire opposées, d'un même attribut non auditif semblent être évoquées ?
- **Le problème de la concurrence des effets.** (Ce problème s'applique spécifiquement aux créations artistiques contenant de la musique et au moins un médium non auditif.) Lorsqu'on souhaite produire chez l'auditoire l'un ou l'autre des effets d'évocation et de liaison, comment s'assurer que cet effet ne soit pas atténué voire supplanté par l'autre effet, plus intense ? Et comment déterminer lequel des deux effets a le plus d'impact sur l'auditoire ?



## **Conclusion**

Le sujet de ce mémoire est les liens entre la musique et les correspondances transmodales (CT). Dans une première partie, nous avons présenté le cadre théorique : l'état de l'art, la problématique et la méthodologie générale. La problématique était : quelles sont les CT utilisables en musique, et quels sont les tenants et les aboutissants de leur utilisation musicale ? Les parties 2, 3 et 4 visaient à répondre à cette problématique.

La partie 2 était une introduction aux CT. En premier lieu, nous avons abordé l'histoire de l'étude des CT. Puis, nous les avons définies comme des associations intuitives entre un attribut sensoriel perçu par un sens (comme la hauteur sonore, la taille visuelle ou le poids tactile) et un attribut sensoriel perçu par un autre sens. Nous les avons ainsi distinguées de la synesthésie, des correspondances intramodales et de l'intégration multisensorielle. Par ailleurs, nous avons également traité la question de l'universalité des CT (différente, encore une fois, de la question de leur innéité) et avons suggéré qu'il était très probable que certaines des CT présentées dans la partie 3 soient éprouvées par une majorité d'humains.

Dans la partie 3, nous avons tenté de répondre (partiellement) à la première partie de la problématique — quelles sont les CT utilisables en musique ? — en listant un certain nombre de CT utilisables en musique et impliquant le sens de l'ouïe. Nous les avons classées en quatre catégories : correspondances audio-visuelles (tableau 10), audio-tactiles (tableau 11), audio-gustatives (tableau 12) et audio-olfactives (tableau 13).

Enfin, dans la partie 4, nous nous sommes penchés sur la seconde partie de la problématique, qui traite de l'utilisation de ces CT en musique (le mot « musique » étant compris comme incluant l'analyse musicale). Nous avons fait l'hypothèse que cette utilisation pouvait avoir principalement trois objectifs : évoquer des stimuli non auditifs *via* de la musique (évoquant), lier une musique à un médium non auditif (liaison), et comprendre, par l'analyse musicale, les évocations et liaisons qui sont en jeu dans une musique (compréhension). Nous avons exploré la question de savoir dans quels domaines ces objectifs peuvent être réalisés, et de quelle manière. Pour finir, nous avons tenté de comprendre les problèmes potentiels qui pourraient entraver l'utilisation musicale des CT, en relevant trois : le problème de l'imprécision, le problème de l'incohérence et le problème de la concurrence des effets.

Quelles pistes de recherche sont-elles prometteuses pour la recherche sur les liens entre la musique et les CT ? Tout d'abord, certains aspects des CT (cf. partie 2) n'ont pas été abordés dans ce

mémoire et pourraient pourtant contribuer à la recherche sur les liens entre la musique et les CT. C'est notamment le cas des *propriétés* des CT. Certaines des propriétés (avérées ou potentielles) des CT sont : l'amodalité (la non-dépendance à une modalité sensorielle spécifique), la bidirectionnalité (l'idée que, si un attribut A fait penser à un attribut B, alors l'attribut B fera penser à l'attribut A), la transitivité (l'idée que si un attribut A est associé à un attribut B, et que l'attribut B est associé à un attribut C, alors l'attribut A est associé à l'attribut C) et la relativité<sup>90</sup> (l'idée que, lorsqu'une CT associe des valeurs d'un attribut aux valeurs d'un autre attribut, elle « sélectionne » les valeurs relativement à un ensemble de valeurs pouvant varier selon le contexte, et « ignore » les valeurs absolues). Ces propriétés sont traitées par Walker<sup>91</sup>. Un autre point concernant les CT qui pourrait être étudié davantage est le degré d'universalité des CT.

Concernant la question de savoir quelles sont les différentes CT utilisables en musique (cf. partie 3), on pourrait chercher des études empiriques supplémentaires, à la fois pour éprouver la réalité de l'existence des CT que nous avons mises en lumière (certaines d'entre elles sont beaucoup moins étayées que d'autres), et pour mettre en lumière davantage de CT impliquant l'ouïe. On pourrait également se demander quels sens autres que l'ouïe sont impliqués dans la perception de la musique (un bon candidat est le sens du toucher, suivi du sens de la vue), et quels sont les CT impliquant ces sens.

Enfin, concernant l'utilisation des CT en musique (cf. partie 4), plusieurs points sont également à approfondir. D'abord, cette utilisation pourrait avoir d'autres buts que l'évocation, la liaison et la compréhension. Ensuite, on pourrait rechercher des études empiriques pour vérifier s'il est possible d'évoquer des stimuli non auditifs *via* une musique dans le cas où cette musique coexiste avec un médium non auditif au sein de la même création artistique. On pourrait également chercher des études empiriques pour savoir si les rapports de similarité, de compromis et d'opposition qui peuvent exister entre une musique et un médium non auditif produisent chacun un effet de liaison spécifique. Enfin, la recherche empirique pourrait nous informer sur l'existence des problèmes de l'imprécision, de l'incohérence, et de la concurrence des effets, et sur l'existence de potentielles solutions pour les résoudre.

---

90 Autrement appelée « sensibilité au contexte » ; voir WALKER, Peter, « Cross-Sensory Correspondences: A Theoretical Framework and Their Relevance to Music », *Psychomusicology: Music, Mind, and Brain*, vol. 26, n° 2, 2016, pp. 103-116.

91 *Ibid.*

# Table des figures

Figure 1. Formes utilisées par Ramachandran & Hubbard (2001) pour démontrer l'existence de l'effet bouba/kiki.....	15
Figure 2. Continuum sur lequel pourraient être regroupées les correspondances transmodales et la synesthésie.....	20
Figure 3. Choix d'instrument en fonction de l'odeur présentée. (Figure issue de Crisinel & Spence, 2012, figure 1.).....	36
Figure 4. Hauteur sonore moyenne attribuée à chaque odeur testée. (Figure issue de Crisinel & Spence, 2012, figure 5.).....	36
Figure 5. Choix d'instrument en fonction de jugements de descripteurs amodaux. (Figure issue de Crisinel & Spence, 2012, figure 2.).....	37
Figure 6. Choix d'instrument en fonction de jugements d'adjectifs olfactifs. (Figure issue de Crisinel & Spence, 2012, figure 3.).....	37
Figure 7. Plusieurs incohérences possibles.....	53

## Table des tableaux

Tableau 1. Définitions de quelques attributs auditifs.....	30
Tableau 2. Définitions de quelques attributs visuels.....	30
Tableau 3. Définitions de quelques attributs tactiles.....	31
Tableau 4. Définitions de quelques attributs olfactifs.....	31
Tableau 5. Correspondances transmodales audio-visuelles évoquées dans Spence (2011).....	32
Tableau 6. Correspondances transmodales audio-visuelles évoquées dans Walker (2016).....	32
Tableau 7. Correspondances transmodales audio-tactiles évoquées dans Walker (2016).....	33
Tableau 8. Correspondances transmodales audio-gustatives évoquées dans Knöferle & Spence (2012).....	34
Tableau 9. Correspondances transmodales audio-gustatives évoquées dans Knöferle et al. (2015).....	35
Tableau 10. Correspondances transmodales audio-visuelles.....	38
Tableau 11. Correspondances transmodales audio-tactiles.....	38
Tableau 12. Correspondances transmodales audio-gustatives.....	39
Tableau 13. Correspondances transmodales audio-olfactives.....	39
Tableau 14. Les trois rapports possibles entre une musique et une vidéo dans un film non narratif selon Georgiou (et leur similitude avec les théories de Cook et de Hall).....	46
Tableau 15. Les trois rapports possibles entre une musique et un médium non auditif.....	47

## Références bibliographiques

- BERLIOZ, Hector, *Grand traité d'orchestration et d'instrumentation modernes* [1844], Allemagne, Bärenreiter, 2003.
- CRISINEL, Anne-Sylvie & SPENCE, Charles, « A Fruity Note: Crossmodal Associations between Odors and Musical Notes », *Chemical Senses*, vol. 37, n°2, 2012, pp. 151-158. DOI : [10.1093/chemse/bjr085](https://doi.org/10.1093/chemse/bjr085)
- CRISINEL, Anne-Sylvie, JACQUIER, Caroline, DEROY, Ophelia & SPENCE, Charles, « Composing with Cross-Modal Correspondences: Music and Odors in Concert », *Chemosensory Perception*, vol. 6, n°1, 2013, pp. 45-52. DOI : [10.1007/s12078-012-9138-4](https://doi.org/10.1007/s12078-012-9138-4)
- DEROY, Ophelia, CRISINEL, Anne-Sylvie & SPENCE, Charles, « Crossmodal Correspondences between Odors and Contingent Features: Odors, Musical Notes, and Geometrical Shapes », *Psychonomic Bulletin & Review*, vol. 20, n°5, 2013, pp. 878-896. DOI : [10.3758/s13423-013-0397-0](https://doi.org/10.3758/s13423-013-0397-0)
- DEROY, Ophelia, FASIELLO, Irène, HAYWARD, Vincent & AUVRAY, Malika, « Differentiated audio-tactile correspondences in sighted and blind individuals », *Journal of Experimental Psychology: Human Perception and Performance*, vol. 42, n°8, 2016, pp. 1204-1214. DOI : [10.1037/xhp0000152](https://doi.org/10.1037/xhp0000152)
- DI STEFANO, Nicola, MURARI, Maddalena & SPENCE, Charles, « Crossmodal Correspondences in Art and Science: Odours, Poetry, and Music » in *Olfaction: An Interdisciplinary Perspective from Philosophy to Life Sciences*, Nicola Di Stefano & Maria Teresa Russo édés., Cham: Springer International Publishing, coll. « Human Perspectives in Health Sciences and Technology », vol. 4, 2022, pp. 155-189. DOI : [10.1007/978-3-030-75205-7\\_10](https://doi.org/10.1007/978-3-030-75205-7_10)
- EITAN, Zohar, « How pitch and loudness shape musical space and motion », in *The Psychology of Music in Multimedia*, Siu-Lan Tan, Annabel J. Cohen, Scott D. Lipscomb & Roger A. Kendall édés., 1<sup>ère</sup> édition, Oxford: Oxford University Press, 2013, pp. 165-191. DOI : [10.1093/acprof:oso/9780199608157.001.0001](https://doi.org/10.1093/acprof:oso/9780199608157.001.0001)
- , « Musical connections: Cross-modal correspondences » in *The Routledge Companion to Music Cognition*, 1<sup>ère</sup> édition, Ashley Richard & Renee Timmers édés., New York, NY: Routledge, coll. « Routledge Companions », 2017, pp. 213-224. DOI : [10.4324/9781315194738](https://doi.org/10.4324/9781315194738)
- EITAN, Zohar & TAMIR-OSTROVER, Hila, « The Sound of an Endless Column: How Music Imagines Unimaginable Space », in *The Oxford Handbook of Sound and Imagination*, vol. 1, Mark Grimshaw-Aagaard, Mads Walther-Hansen & Martin Knakkegaard édés., New York, NY: Oxford University Press, coll. « Oxford Handbooks », 2019. DOI : [10.1093/oxfordhb/9780190460167.013.36](https://doi.org/10.1093/oxfordhb/9780190460167.013.36)
- GEORGIU, Chrystalla, *Beyond narrative: A cross-modal approach to soundtrack composition*, thèse, université du Hertfordshire, 2014, texte disponible via <http://hdl.handle.net/2299/18621> (consulté le 31/05/2022).
- KNÖFERLE, Klemens M. & SPENCE, Charles, « Crossmodal Correspondences between Sounds and Tastes », *Psychonomic Bulletin & Review*, vol. 19, n°6, 2012, pp. 992-1006. DOI : [10.3758/s13423-012-0321-z](https://doi.org/10.3758/s13423-012-0321-z)
- KNÖFERLE, Klemens M., WOODS, Andy, KÄPPLER, Florian & SPENCE, Charles, « That Sounds Sweet: Using Cross-Modal Correspondences to Communicate Gustatory Attributes », *Psychology & Marketing*, vol. 32, n°1, 2015, pp. 107-120. DOI : [10.1002/mar.20766](https://doi.org/10.1002/mar.20766)
- KÖHLER, Wolfgang, *Gestalt psychology* [1929], New York, Liveright, 1947.
- KOMSCHLIES, Chelsea, « Research Alive: Crossmodal correspondences in composing and listening to music », université McGill (École de musique Schulich), 16 mars 2022, disponible via <https://www.youtube.com/watch?v=qpmJTaSiGKE> (consulté le 02/06/2022).
- KORZENIOWSKA, Anna T., ROOT-GUTTERIDGE, Holly, SIMNER, Julia & REBY, David, « Audio-Visual Crossmodal Correspondences in Domestic Dogs (*Canis Familiaris*) », *Biology Letters*, vol. 15, n°11, 2019. DOI : [10.1098/rsbl.2019.0564](https://doi.org/10.1098/rsbl.2019.0564)

- MACPHERSON, Fiona, « Taxonomising the senses », *Philosophical Studies*, vol. 153, n°1, 2011, pp. 123-142. DOI : [10.1007/s11098-010-9643-8](https://doi.org/10.1007/s11098-010-9643-8)
- MARKS, Lawrence E., « Cross-modal interactions in speeded classification », in *The Handbook of Multisensory Processes*, Gemma A. Calvert, Charles Spence & Barry E. Stein éd., Cambridge, Mass: MIT Press, 2004, pp. 85-105.
- MESZ, Bruno, TREVISAN, Marcos A. & SIGMAN, Mariano, « The Taste of Music », *Perception*, vol. 40, n°2, 2011, pp. 209-219. DOI : [10.1068/p6801](https://doi.org/10.1068/p6801)
- PARISE, Cesare V. & SPENCE, Charles, « Audiovisual Cross-Modal Correspondences in the General Population », in *Oxford handbook of synaesthesia*, Julia Simner & Edward Hubbard éd., Oxford, Oxford University Press, 2013, pp. 790-815. DOI : [10.1093/oxfordhb/9780199603329.013.0039](https://doi.org/10.1093/oxfordhb/9780199603329.013.0039)
- RAMACHANDRAN, Vilayanur S. & HUBBARD, Edward M., « Synaesthesia--A Window Into Perception, Thought and Language », *Journal of consciousness studies*, vol. 8, n°12, 2001, pp. 3-34.
- SAPIR, Edward, « A study in phonetic symbolism », *Journal of Experimental Psychology*, vol. 12, n°3, 1929, pp. 225-239. DOI : [10.1037/h0070931](https://doi.org/10.1037/h0070931)
- SPENCE, Charles, « Crossmodal correspondences: A tutorial review », *Attention, Perception, & Psychophysics*, vol. 73, n° 4, 2011, pp. 971-995. DOI : [10.3758/s13414-010-0073-7](https://doi.org/10.3758/s13414-010-0073-7)
- , « Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism », *Journal of Consumer Psychology*, vol. 22, n° 1, 2012, pp. 37-54. DOI : [10.1016/j.jcps.2011.09.004](https://doi.org/10.1016/j.jcps.2011.09.004)
- , « Simple and Complex Crossmodal Correspondences Involving Audition », *Acoustical Science and Technology*, vol. 41, n°1, 2020, pp. 6-12. DOI : [10.1250/ast.41.6](https://doi.org/10.1250/ast.41.6)
- SPENCE, Charles & SATHIAN, Krish, « Audiovisual Crossmodal Correspondences: Behavioral Consequences and Neural Underpinnings », in *Multisensory Perception*, Krish Sathian & Vilayanur S. Ramachandran éd., Academic Press, 2020, pp. 239-258. DOI : [10.1016/B978-0-12-812492-5.00011-5](https://doi.org/10.1016/B978-0-12-812492-5.00011-5)
- STEIN, Barry E., LAURIENTI, Paul J., WALLACE, Mark T. & STANFORD, Terrence R., « Multisensory Integration », in *Encyclopedia of the Human Brain*, Vilayanur S. Ramachandran éd., Elsevier, 2002, pp. 227-241. DOI : [10.1016/B0-12-227210-2/00225-9](https://doi.org/10.1016/B0-12-227210-2/00225-9)
- STEVENSON, Richard J., BOAKES, Robert A., « Sweet and Sour Smells: Learned Synesthesia Between the Senses of Taste and Smell » in *The Handbook of Multisensory Processes*, Gemma A. Calvert, Charles Spence & Barry E. Stein éd., Cambridge, Mass: MIT Press, 2004, pp. 69-83.
- STYLES, Suzy J. & GAWNE, Lauren, « When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter », *I-Perception*, vol. 8, n° 4, 2017. DOI : [10.1177/2041669517724807](https://doi.org/10.1177/2041669517724807)
- WALKER, Peter, « Cross-Sensory Correspondences: A Theoretical Framework and Their Relevance to Music », *Psychomusicology: Music, Mind, and Brain*, vol. 26, n° 2, 2016, pp. 103-116. DOI : [10.1037/pmu0000130](https://doi.org/10.1037/pmu0000130)

