



Figures multimodales

DENIS SERON

FNRS – Université de Liège

✉ D.Seron@uliege.be

Résumé La question posée dans cette étude est de savoir s'il peut exister des figures multimodales (par exemple audiovisuelles) et, si oui, à quelles conditions. L'auteur met en lumière des conditions nécessaires de la multimodalité figurale en partant de la définition de la figure de Grelling et Oppenheim. Les conditions incluses dans cette définition l'amènent à déterminer à quelles contraintes devrait se plier une figure multimodale.

Mots-clefs Gestalt, modalités sensorielles, Grelling, Oppenheim, musique.

Nous pouvons supposer, avec Koffka (1936 : 200–201), que la distinction entre fond et figure (*Gestalt*) s'applique à toutes les modalités sensorielles. De même qu'un nuage blanc se détache du ciel azuré, de même une figure mélodique jouée au soprano se détache d'un fond joué à la basse. Cela est vrai, semble-t-il, même dans le cas des « sens inférieurs » comme l'odorat et le toucher. Une odeur particulière à une pièce — celle de poussière humide dans la cave, par exemple — imprègne et détermine, à la manière d'un fond, toutes les expériences olfactives que je peux y avoir. Là où une telle « atmosphère » fait défaut, poursuit Koffka, l'état de repos (*stillness*) peut faire fonction de fond pour des figures. De même, quand je touche les poils d'un balai, l'espace entre eux m'apparaît comme un « espace tactile non recouvert de matière », qui fait fonction de fond pour une figure tactile (Katz 1925 : 42–43). Dans tous les cas, il semble que le schéma figure-fond soit plus général que les distinctions entre modalités sensorielles : « Notre fond le plus général, dit encore Koffka, est suprasensoriel au sens où il doit son existence aux contributions de possiblement tous les sens » (Koffka 1936 : 201).

Mais s'il existe des figures dans toutes les modalités sensorielles, en existe-t-il qu'on pourrait rattacher à plusieurs modalités distinctes ? Une unique figure pourrait-elle être perçue visuellement et auditivement ? Une mélodie, par exemple, peut-elle être transposée en une complexion de couleurs de manière à former avec elle une unique figure à la fois auditive et visuelle ?

Dans un texte de 1925 publié sous le titre « L'unité des sens », le musicologue viennois Erich Moritz von Hornbostel, représentant périphérique de l'école gestaltiste de Berlin, racontait que le peintre Lyonel Feininger s'était mis subitement, à l'âge de cinquante ans, à composer des fugues pour orgue. Voici comment il décrivait le fait :

Jusqu'ici, des fugues, il n'en avait que peintes. Maintenant, ses images sont visibles également pour l'aveugle. Dans l'art non plus, la sphère délimitée par un sens particulier (*die Sinnessphäre*) n'est pas décisive ; la transposition de l'une dans l'autre est possible. (Hornbostel 1925 : 295)

Une fugue, pour cet auteur, est donc quelque chose qui, non strictement auditif, peut aussi faire l'objet d'expériences visuelles. Autrement dit, percevoir une fugue n'est pas nécessairement entendre. Il existe, clame Hornbostel, des « perceptions sensibles suprasensorielles » (*übersinnliche Sinneswahrnehmungen*), c'est-à-dire non limitées à une unique modalité. Cette idée a été discutée notamment au sujet du mouvement. Il semble que le mouvement en arc d'un objet touchant successivement la peau en plusieurs points rapprochés soit perçu tactilement comme un mouvement, en tous points analogue au mouvement stroboscopique de la perception visuelle (Köhler 1947 : 77–78 et 100–101). Comme conclut Hornbostel (1925 : 293), « on peut voir, entendre ou toucher le mouvement ».

Sous-jacente à la question des figures multimodales, il y a celle plus générale de l'unité des sens, sur laquelle je reviendrai en conclusion. Cette question se ramène à ceci : que faisons-nous au juste, quand nous analysons l'expérience en phénomènes visuels, auditifs, tactiles, etc. ? Restituons-nous par là, comme nous le faisons en différenciant anatomiquement les organes des sens, des différences réelles ? Ou bien ces différences sont-elles conceptuelles, dépendantes des descripteurs et donc du langage ? Quand nous employons le mot « clair » pour décrire ici un son, là une couleur, le mot est-il équivoque ou univoque ? Ne faisons-nous qu'appliquer métaphoriquement un terme de l'idiome visuel à la perception auditive — ou bien l'expérience décrite est-elle primitivement transmodale ? La différence entre « rouge » et « sucré » est-elle vraiment d'une autre nature que celle entre « rouge » et « bleu » ? Pourquoi devrait-elle l'être ?

L'objectif de la présente étude est plus spécial. La question initiale en est celle-ci : peut-il exister des figures multimodales, par exemple des fugues audiovisuelles ? Et, si oui, à quelles conditions ? Je ne vise pas à apporter à cette question une réponse complète ni même à décider si une figure multimodale est possible ou non, mais seulement à mettre en lumière des conditions nécessaires de la multimodalité figurale. À cette fin, je partirai d'une certaine définition de ce qu'est une figure. Les conditions incluses dans cette définition nous amèneront ensuite à déterminer à quelles contraintes devrait se plier une figure multimodale, quelle qu'elle soit.

Qu'est-ce qu'une figure ?

Dans leur article de 1938 « Le concept de figure à la lumière de la nouvelle logique », dont il sera longuement question dans la suite, Kurt Grelling et Paul Oppenheim distinguaient deux types de figures décrits par les gestaltistes.

Premièrement, on entend par « figure » quelque chose qui répond aux deux critères fixés par Ehrenfels dans son fameux article de 1890 « Sur les qualités figurales » (Ehrenfels 1890 ; Grelling & Oppenheim 1938 : 218–219 ; Schlick 2020 : 311). D'une part, une telle figure doit être l'invariant d'une transposition. Par exemple, une mélodie est invariante si on la transpose dans un autre mode harmonique, un rectangle est invariant pour certaines transformations géométriques, etc. D'autre part, une figure au premier sens doit être « plus que la somme des parties ». Par exemple, une mélodie peut être jouée à l'envers, en partant de la dernière note. Dans ce cas, la mélodie originale et la mélodie inversée contiennent exactement les mêmes notes — et pourtant elles sont deux figures mélodiques différentes. En bref, une figure au premier sens satisfait ces deux conditions : il est faux que dans tous les cas le tout change si on change les parties ; il est faux que dans tous les cas les parties changent si on change le tout.

Deuxièmement, on appelle aussi « figures » des « systèmes d'effectuation » (*Wirkungssysteme*), c'est-à-dire des totalités structurales qui peuvent être ou non en équilibre. À cette acception se rattache la définition de Kurt Lewin selon laquelle une figure est « un système dont les parties sont connectées dynamiquement de telle manière qu'un changement d'une partie a pour effet un changement de toutes les autres parties » (Lewin 1936 : 218). Fait remarquable, les « systèmes d'effectuation » ne satisfont pas nécessairement le premier critère d'Ehrenfels ci-dessus (Grelling & Oppenheim 1938 : 224), d'où il est possible de caractériser les figures ehrenfelsiennes comme des invariants de transposition. Selon Grelling et Oppenheim, l'usage du terme

« *Gestalt* » pour de telles totalités est l'effet d'une confusion dommageable¹. C'est pourquoi, après ces deux auteurs, j'emploierai désormais le mot « figure » exclusivement au sens des deux critères d'Ehrenfels, donc en référence à l'invariant d'une transposition. Ce choix se justifie également par le fait que, pour des raisons que je préciserai en conclusion, la possibilité de « systèmes d'effectuation » multimodaux est selon moi triviale et n'a pas besoin d'être établie ni discutée.

Notre question peut dès lors recevoir la formulation plus restrictive suivante : peut-il exister des figures ehrenfelsiennes multimodales ? Et si oui, à quelles conditions ? À la première question, Ehrenfels lui-même a apporté une réponse affirmative :

Nous voyons émerger ici un nouveau problème. L'étroite fusion (*Verschmelzung*) des sensations tactiles et thermiques, et parfois aussi des sensations gustatives et olfactives en une impression d'ensemble unitaire soulève la question de savoir si des qualités figurales ne seraient pas présentes ici aussi, lesquelles s'édifieraient sur un fondement appartenant à *plusieurs domaines des sens*. *A priori*, comme on peut le voir facilement, un complexe de sensations tactiles et thermiques peut tout aussi bien servir de fondement à une qualité figurale qu'un complexe de sensations sonores. De même, on ne peut *a priori* rien objecter contre l'existence de figures chromatico-sonores (*Ton-Farbengestalten*) qui relieraient les *data* des deux sens à la manière d'un pont. (Ehrenfels 1890 : 267–268)

Qu'est-ce qu'une figure au sens d'Ehrenfels ? De nombreuses définitions ont été proposées par le passé. Certaines sont méréologiques et non relationnelles (Ehrenfels 1890 ; Husserl 1970 ; Husserl 1984 : 237), d'autres méréologiques et relationnelles (Meinong 1894 ; Marty 1908), d'autres enfin purement relationnelles (structurales) (Köhler 1925 ; Helson 1933 ; Grelling & Oppenheim 1938 ; Ajdukiewicz d'après Grelling & Oppenheim 1938 : 216)². Bien qu'elle ne soit pas exempte de défauts comme je le suggérerai plus loin, la plus aboutie me semble celle énoncée par Grelling et Oppenheim dans l'article cité. Cette définition est la suivante (Grelling & Oppenheim 1938 : 216) : « Les figures sont les cercles de similitude (*Gleichheitskreise*) de correspondances. » C'est elle que je me propose de clarifier et de prendre pour point de départ en vue de

¹ Sur cette confusion, voir aussi (Spearman 1926 ; Scheerer 1931 : 4 suiv. ; Spearman 1937a : 427–429 ; Spearman 1937b ; Grelling & Oppenheim 1988b). La seconde acception est développée en particulier dans (Grelling & Oppenheim 1988b) et à sa suite dans (Rescher & Oppenheim 1955).

² Voir sur ce point (Chrudzimski 2013 : 249).

répondre à notre question initiale : à quelles conditions peut-il exister des figures (ehrenfelsiennes) multimodales ?

Dans le prolongement direct de la *Gestalttheorie* de l'école de Berlin, l'ambition de Grelling et Oppenheim était d'élaborer une définition non méréologique de la figure, à savoir une définition en termes de « complexes » (au sens carnapien). Ces auteurs entreprennent de caractériser une figure en se passant totalement de notions issues du vocabulaire méréologique : « tout », « partie », « appartenir », « contenir », etc. Voilà l'enjeu¹.

Nous commencerons par caractériser un objet paradigmatique, et cette caractérisation pourra ensuite nous servir pour définir ce qu'est une figure ehrenfelsienne. Nous prendrons pour exemple une figure mélodique, mettons le thème du chat dans *Pierre et le loup* de Prokofiev, qui est joué sur une clarinette soprano en la :



Comment caractériser le thème du chat en termes non méréologiques ? Une telle figure, posent Grelling et Oppenheim, est caractérisée par les traits suivants : d'abord un « domaine de places » (*Stellengebiet*), ensuite une classe de « classificateurs d'états » ou « E-classificateurs » (*Zustands-Klassifikatoren*, *Z-Klassifikatoren*), enfin une classe de « classificateurs de parcours de valeurs » ou « P-classificateurs » (*Wertverlaufs-Klassifikatoren*, *W-Klassifikatoren*).

Une mélodie est assurément quelque chose dont on a des expériences sensorielles. Cependant, ces expériences ne sont pas n'importe quelles expériences. Par exemple, entendant le thème du chat, l'auditeur voit un musicien souffler dans une clarinette, touche l'accoudoir de son fauteuil, imagine un chat, etc. : aucune de ces données n'est *constitutive* de la mélodie elle-même.

¹ On trouve quelque chose d'assez semblable dans la théorie des couleurs de Meinong, qui entend substituer à la composition des couleurs (le vert est un mélange de bleu et de jaune) leurs coordonnées sur un octaèdre chromatique. Voir (Meinong 1903).

Pour décrire ce qui constitue la mélodie, nous avons besoin de types déterminés de données sensorielles, que Grelling et Oppenheim, après Carl Hempel, appellent des « classificateurs »¹.

Par exemple, trois enfants peuvent être classés selon leur âge ou selon qu'ils portent ou non des lunettes. Le classificateur « porter des lunettes » est une fonction qui associe à chaque enfant la valeur « porte des lunettes » ou « ne porte pas des lunettes ». Le classificateur « âge » est une fonction qui associe à chaque enfant une valeur numérique, par exemple « six ans », « sept ans », etc. Maintenant, au lieu d'enfants, rien n'empêche de considérer des places à l'intérieur d'une pure structure ou « domaine de places », par exemple des coordonnées dans le plan euclidien. Dans ce cas, poursuivent Grelling et Oppenheim, nous avons affaire à un « classificateur d'état » ou E-classificateur. Ainsi, tel point du plan euclidien peut se voir associer la valeur « rouge », tel autre la valeur « bleu », etc. Le classificateur « couleur » est alors un E-classificateur, c'est-à-dire une fonction qui, à chaque couple de coordonnées sur le plan euclidien, assigne une valeur « rouge », « bleu », etc.

En l'occurrence, dans le cas de la mélodie, nous avons d'abord besoin de la classe d'E-classificateurs suivante, supposée suffisante pour caractériser le thème du chat : hauteur sonore, timbre, intensité. Ensuite, il nous faut un domaine de places qui fournisse les arguments des E-classificateurs. Nous supposons dans la suite que celui-ci, dans le cas du thème du chat, est le continuum temporel, c'est-à-dire une structure unidimensionnelle. Par exemple, l'E-classificateur « hauteur sonore » fait correspondre à l'instant t_0 la valeur « ré », à l'instant t_1 la valeur « sol », à l'instant t_2 la valeur « si », etc. ; l'E-classificateur « timbre » fait correspondre à t_0 la valeur « clarinette en la », à t_1 la valeur « clarinette en la », etc.

Ainsi décrit, le thème du chat est ainsi caractérisé par ce que Grelling et Oppenheim dénomment (au sens de Carnap) un *complexe*, c'est-à-dire une « relation entre une classe d'E-classificateurs et un domaine de places, telle qu'à une place du domaine tout E-classificateur associe à chaque fois une valeur » (Grelling & Oppenheim 1938 : 213)². Nous dirons que le thème du chat

¹ Grelling et Oppenheim préfèrent « classificateur » au terme carnapien de « foncteur » parce que celui-ci, à la différence de celui-là, n'admet que des valeurs numériques (Grelling & Oppenheim 1988a).

² Cf. (Carnap 1998 : 5) : « Nous distinguerons entre “tout” et “complexe logique” : le tout est composé de ses éléments, ils en sont des parties ; un complexe logique indépendant n'a pas ce type de rapport avec ses éléments, mais il se caractérise par le fait que tous les énoncés sur lui peuvent être tournés comme des énoncés sur ses éléments. » Il est à remarquer que Carnap, en 1928, distinguait entre les complexes logiques et les « tous véritables » que sont les figures (Carnap 1998 : 48).

est le parcours de valeurs du complexe $C((t_0, t_1, t_2...), \{HAUTEUR, TIMBRE, INTENSITÉ\})$, lequel peut être représenté dans une partition. L'axe horizontal de la portée figure le domaine de places ; l'axe vertical ainsi que les altérations représentent l'E-classificateur « hauteur », etc.

Cependant, notre description en termes de complexes est encore incomplète. Car nous ne voulons pas décrire simplement une mélodie donnée. Nous voulons décrire le thème du chat en tant que *figure* mélodique, c'est-à-dire en tant que quelque chose qui peut être joué à différents moments, dans une autre tonalité harmonique, avec un autre instrument, etc. Pour décrire une figure mélodique, il est nécessaire que, dans notre formalisme, elle satisfasse les deux conditions d'Ehrenfels. Il faut qu'elle soit transposable dans une autre tonalité (c'est-à-dire dans certains cas invariante si on change les valeurs de l'E-classificateur « hauteur sonore »), exécutable sur un autre instrument (c'est-à-dire dans certains cas invariante si on change les valeurs de l'E-classificateur « timbre »), etc. Par exemple, la transposition du thème du chat à la tierce mineure supérieure donne le même thème en Mib majeur :



Le point important, ici, est le suivant : toutes les valeurs de l'E-classificateur « hauteur » pour un continuum temporel donné ont été modifiées — et pourtant nous sommes en présence de *la même* figure mélodique. En vue de décrire de telles figures transposables, Grelling et Oppenheim introduisent un nouveau type de classificateur, qu'ils appellent « classificateur de parcours de valeurs » ou « P-classificateur ».

On procède en trois étapes. Premièrement, on considère les parcours de valeurs de chaque E-classificateur séparément. Par exemple, le parcours de valeurs de l'E-classificateur « hauteur sonore » est (ré, sol, si...) ; celui de l'E-classificateur « timbre » est (clarinette, clarinette, clarinette...). Deuxièmement, on compare les parcours de valeurs de l'E-classificateur « hauteur » pour le thème du chat et sa transposition, d'un côté donc le parcours de valeurs (ré, sol, si...), de l'autre le parcours de valeurs (fa, sib, ré...). Troisièmement, l'idée de Grelling et Oppenheim est que, si la figure mélodique est préservée, c'est parce qu'il existe une relation d'homologie entre les deux parcours de valeurs. En l'occurrence, la suite d'intervalles est identique, à savoir (+ 2½, + 2...). C'est précisément cette homologie que capture le P-classificateur. Le

P-classificateur est une fonction dont les arguments sont les parcours de valeurs d'un E-classificateur. Il est invariant pour la transposition, au sens où le thème du chat et ses transpositions ont le même P-classificateur pour la hauteur.

Grelling et Oppenheim comparent le P-classificateur à une courbe de température. La courbe de température d'un malade — correspondant au parcours de valeurs — présente quelquefois un tracé, une signature, un certain « caractère » — correspondant au P-classificateur — qui révèle la présence de malaria. De même, une certaine suite d'intervalles révèle la présence du thème du chat.

Fait important, la distinction entre E-classificateurs et P-classificateurs neutralise par avance une objection de William James à l'encontre de la conception relationnelle de la sensation¹. Si seules étaient connues les relations, arguait cet auteur, alors il serait impossible de faire la différence entre une mélodie et sa transposition à l'octave supérieure :

Si tout ce que nous sentons était la *différence* sur l'échelle musicale entre do1 et ré1 ou entre do2 et ré2, alors, comme la différence est la même dans les deux couples de notes, les couples eux-mêmes seraient les mêmes, et le langage pourrait se passer de substantifs. (James 1950 : 12)

À cette objection nous pouvons maintenant répondre ceci : alors que le P-classificateur du thème du chat a les mêmes valeurs (c'est-à-dire la même suite d'intervalles) pour toutes ses transpositions, en revanche, les valeurs de l'E-classificateur HAUTEUR sont à chaque fois différentes. Tout dépend, en définitive, de ce que nous souhaitons décrire : soit le thème du chat en do1 majeur, soit le thème du chat en tant qu'invariant pour la transposition. Dans le second cas, il est évident que la différence entre les deux suites de notes (ré1, sol1, si1...) et (ré2, sol2, si2...) est précisément ce dont il convient de faire abstraction. Ce qui demeure vrai dans la remarque de James, cependant, c'est que nous pouvons nous passer de constantes individuelles. Le complexe mélodique a autant d'instanciations qu'on veut.

Nous sommes maintenant parvenus à un résultat important. Nous avons réussi à caractériser une mélodie sans jamais suggérer que la mélodie serait un tout composé de notes, donc *en termes non méréologiques*. Le thème du chat a été décrit non pas comme un tout composé de notes qui seraient ses parties,

¹ Cf. (Trémault 2020 : 161–162), qui adresse avec profondeur l'objection de James au structuralisme gestaltiste.

mais exclusivement comme une relation entre la structure temporelle et un ensemble déterminé d'E-classificateurs, en l'occurrence {HAUTEUR, TIMBRE, INTENSITÉ}. Le résultat obtenu est important parce que, comme le prescrit le premier critère d'Ehrenfels, une mélodie doit précisément demeurer identique si on la transpose, c'est-à-dire si on change toutes ses parties.

Pourtant, ce résultat est encore insuffisant. Si le second critère d'Ehrenfels — la figure est plus que la somme des parties — est maintenant clarifié dans une large mesure, en revanche le second ne l'est pas suffisamment et cette lacune ne peut que faire obstacle à notre tentative de définir ce qu'est une figure. Considérons la combinaison *C* du thème du chat et d'un carré rouge. D'après la conception de Grelling & Oppenheim, *C* a ses E-classificateurs et ses P-classificateurs, et rien n'empêche de le transposer dans une autre combinaison, mettons celle formée par le thème du chat en Mib majeur et le carré rouge tourné de quarante-cinq degrés. Mais dirons-nous que *C* est une figure transposée ? Visiblement, comme *C* n'est pas le genre de chose que nous qualifions de « figure », nous avons besoin de caractérisations plus fines et plus restrictives.

Revenons au premier critère d'Ehrenfels : que veut dire « transposition » ? Le mot doit être compris, naturellement, en un sens élargi, incluant non seulement les transpositions musicales, mais aussi certaines transformations géométriques, etc., et cela dans toutes les modalités sensorielles (s'il existe des figures dans toutes les modalités sensorielles). Par ailleurs, si nous voulons définir ce qu'est une transposition, nous devons le faire sans utiliser le concept de figure, puisque la figure est précisément ce que nous cherchons à définir. Il est en conséquence impossible de définir la transposition comme étant un changement qui préserve une figure.

Grelling et Oppenheim énoncent trois conditions auxquelles quelque chose *A* est la transposition de quelque chose d'autre *B* — auxquelles, en d'autres termes, il existe entre les deux une relation de « correspondance » (*Korrespondenz*) (Grelling & Oppenheim 1938 : 214, se référant à Carnap 1929 : 85–86) :

[1] *Les domaines de places de A et B doivent être isomorphes.* Par exemple, le thème du chat peut être joué à des moments différents, mais les relations de succession doivent rester identiques. Autrement dit, il existe entre les deux exécutions une application bijective qui préserve les « relations de position » (*Lagerrelationen*).

[2] *Les E-classificateurs de A et de B doivent être identiques.* Par exemple, le thème du chat et ses transpositions ont les mêmes E-classificateurs « hauteur », « timbre » et « intensité ».

[3] *Certains P-classificateurs de A et B sont identiques*. Comme l'indiquent Grelling et Oppenheim, « les parcours de valeurs d'E-classificateurs correspondants, ou homologues pour ainsi dire, sont semblables (*gleich*) ». Par exemple, le thème du chat et ses transpositions sont semblables au sens où l'E-classificateur « hauteur sonore » présente des deux côtés une même suite d'intervalles, c'est-à-dire des parcours de valeurs qui tombent sous un même P-classificateur. Il n'est pas nécessaire que tous les parcours de valeurs soient semblables deux à deux — comme c'est le cas lorsque le thème du chat est joué à deux moments différents sur le même instrument, dans la même tonalité et avec la même force. Deux mélodies peuvent se correspondre seulement sous le rapport de la hauteur sonore (Grelling & Oppenheim 1938 : 215).

Par ce biais, nous avons caractérisé ce que sont une transposition et une relation de correspondance. La correspondance est une relation de similitude sous le rapport des domaines de places, des E-classificateurs et des P-classificateurs.

Si maintenant on considère à nouveau le thème du chat, on pourra décrire une multitude de relations de correspondance avec des transpositions du même thème. Par exemple, si on considère le thème original et deux de ses transpositions, on aura en tout trois relations de correspondance. Par ce biais, on obtient ce qu'on désigne, dans la tradition empiriste, sous le terme de « cercle de similitude ». Or, pour bon nombre d'empiristes et en particulier pour Carnap — c'est là l'essentiel de la méthode d'analyse préconisée au § 70 de son *Aufbau* (Carnap 1998 : 95–97 ; Leclercq 2020 : 126–128) —, la présence de cercles de similitude est suffisante pour construire des classes et, par conséquent, pour définir un concept et constituer un (quasi-)objet. Partant, nous pouvons considérer que nous avons défini la figure « thème du chat » : « Les figures sont les cercles de similitude de correspondances » (Grelling & Oppenheim 1938 : 216).

Ce que nous avons obtenu par là, c'est une description conceptuelle de la figure mélodique « thème du chat » — en tant qu'« individu figural » (*Gestalt-Individuum*) comme disent Grelling et Oppenheim (1938 : 216) —, dont les arguments sont des places du continuum temporel. Ce résultat permet de définir aisément ce qu'est en général une figure. Si nous choisissons comme classificateur le concept « figure mélodique », alors les arguments du classificateur seront des suites (c'est-à-dire des complexes) de sons individuels et ses valeurs des figures mélodiques particulières, par exemple le thème du chat. Une figure en général pourra ainsi être définie comme un concept dont les valeurs sont des figures particulières telles qu'on vient de les définir.

À quelles conditions peut-il exister des figures multimodales ?

Il n'est pas sûr que la définition de Grelling et Oppenheim soit tout à fait satisfaisante. Une question importante est de savoir si elle n'est pas trop générale. Premièrement, elle pose problème dans les cas de silences prolongés et de bruits blancs. Elle nous oblige à dire qu'un silence prolongé est une figure, puisqu'on peut le transposer dans un autre silence prolongé dont toutes les valeurs d'E-classificateurs sont identiques, à savoir « nul ». Ce fait est embarrassant, car on souhaiterait une définition plus restrictive et ciblée de ce qu'est une figure — en particulier une définition qui inclue comme condition la différence figure-fond. Une deuxième difficulté vient du fait que la relation de correspondance, telle que l'ont définie Grelling et Oppenheim, est manifestement réflexive. Si toute donnée est transposable en elle-même, alors toute donnée est l'invariant d'une transposition — ce qui ôte à la notion de figure ehrenfelsienne une bonne part de son contenu. Notre objectif étant d'énoncer des conditions nécessaires, et non des conditions nécessaires et suffisantes, ces difficultés peuvent néanmoins être mises entre parenthèses en vue des problèmes qui nous occupent.

Nous pouvons maintenant revenir à notre question initiale : à quelles conditions peut-il exister des figures multimodales ? À cette fin, je distinguerai deux cas. D'une part, je regroupe sous le terme « intermodalité » les cas où une unique figure présente plusieurs modalités sensorielles distinctes, par exemple auditive et visuelle. Une telle figure peut être transposée dans une autre figure présentant les mêmes modalités sensorielles, selon le schéma : $Fmn * Gmn$ (où les majuscules, les minuscules et l'astérisque symbolisent respectivement la figure, les modalités et la relation de correspondance). Ce cas est celui évoqué par Ehrenfels en termes de « fusion » (voir *supra*). D'autre part, j'entends par « synesthésie » le cas où une figure d'une certaine modalité sensorielle correspond à une autre figure d'une autre modalité, selon le schéma : $Fm * Gn$. Le cube et la sphère visiotactiles de Molyneux pourraient être rangés dans cette catégorie, si l'aveugle de naissance qui a recouvré la vue est capable de les reconnaître visuellement (Locke 1997 : 144)¹.

Voyons d'abord ce qui se passe au niveau des classificateurs. Dans les cas d'intermodalité, rien n'empêcherait *a priori* d'avoir, au sein d'un même

¹ Je distingue ces deux cas à des fins didactiques. Si une figure est définie par une relation de correspondance, alors F et G sont dans tous les cas des instanciations individuelles d'une même figure Γmn .

complexe, des E-classificateurs de plusieurs modalités distinctes, mais unimodaux (relevant chacun d'une unique modalité)¹, par exemple « hauteur sonore » et « couleur », « acidité » et « chaleur ». En revanche, dans les cas de synesthésie, on devrait impérativement recourir à des E-classificateurs transmodaux. Cette exigence tient simplement à la contrainte d'identité des E-classificateurs qui pèse sur les relations de correspondance (voir *supra*) : le thème du chat et ses transpositions doivent présenter les mêmes E-classificateurs. Si la figure transposée est d'une modalité différente, alors chaque E-classificateur doit se rattacher à deux modalités distinctes au moins.

Il est facile de concevoir de tels E-classificateurs transmodaux, par exemple la rugosité pour l'ouïe et le toucher (Köhler 1947 : 174), la chaleur et la froideur pour le toucher et la vision (couleurs) (Buytendijk 1970) ou encore la clarté qui semble s'étendre à tous les sens (Brentano 1979 : 161–162 ; Hornbostel 1925 : 290–291 ; Straus 1978). Ce que Daniel Stern (1998 : 53 suiv.) appelle le « *vitality affect* » entre dans cette catégorie. Le thème du chat doit être joué avec des *staccati*. Or un *staccato* est aussi un certain mouvement qu'on voit : le mouvement des doigts du clarinettiste est *saccadé* comme le sont les notes entendues. Une conséquence de cela est qu'on peut envisager, à ce niveau tout au moins, que la partition du thème du chat soit la transposition synesthésique de sa version sonore : tout dépend du choix des E-classificateurs.

En réalité, le vrai problème ne se situe pas au niveau des E-classificateurs, mais à celui du domaine de places. Toute transposition, on l'a vu, exige l'isomorphie du « domaine des places ». Dans les cas d'intermodalité, cela signifie que tous les E-classificateurs d'une même figure doivent prendre leurs arguments dans un même domaine de places (sinon ils ne constituent pas une unique figure). En conséquence, tous les E-classificateurs d'une figure multimodale ont le même domaine de places, par exemple le plan euclidien ou le continuum temporel. Dans les cas de synesthésie, nous dirons que le domaine de places doit être invariant pour la transposition, c'est-à-dire que le domaine de places d'une figure donnée doit être isomorphe au domaine de places de sa transposition : le thème du chat et sa transposition ont tous deux pour domaine de places le continuum temporel.

Seulement, il est permis de douter que la contrainte d'isomorphie du « domaine des places » puisse toujours être respectée. Cette question nous renvoie à la difficile question de la *n*-dimensionnalité des modalités sensorielles — que nous assimilerons provisoirement et approximativement à l'arité des E-

¹ Et non amodaux, le problème posé ici étant la multimodalité et non l'amodalité.

classificateurs¹. À première vue, certains sens, comme le goût et l'odorat, semblent unidimensionnels, d'autres, comme la vue, bidimensionnels ou tridimensionnels. Or, si le nombre de dimensions est différent, alors le domaine de places n'est pas isomorphe.

Considérons pour commencer les cas d'intermodalité. Ici, si le nombre de dimensions est différent, alors des E-classificateurs se rattachant à des modalités sensorielles différentes ne peuvent plus prendre leurs arguments dans un même domaine de places — car un domaine de places ne peut pas être à la fois unidimensionnel et bidimensionnel. Donc, l'intermodalité figurale est impossible dans certains cas. Par exemple, si le goût est unidimensionnel et la vue bidimensionnelle, alors une même figure ne peut pas être à la fois visuelle et gustative.

En ce qui concerne la synesthésie, on dira ceci : si le nombre de dimensions est différent, alors il ne peut pas y avoir isomorphie des domaines de places ; donc, la synesthésie figurale est impossible dans certains cas. Par exemple, si le goût est unidimensionnel et la vue bidimensionnelle, alors une figure visuelle ne peut pas être transposée en une figure gustative.

La question de la n -dimensionnalité des modalités sensorielles, il est vrai, est particulièrement difficile et discutable. Une première hypothèse possible est celle selon laquelle [a] toutes les modalités sensorielles ont nativement la même n -dimensionnalité. Certains auteurs, par exemple Helmholtz, Lotze et Wundt, ont ainsi défendu l'idée que toutes les modalités sensorielles étaient nativement unidimensionnelles. D'autres, comme Kant, ont soutenu qu'elles étaient nativement tridimensionnelles. La seconde option est de poser que [b] les modalités sensorielles présentent des n -dimensionnalités différentes. Maints auteurs ont soutenu que certaines modalités sensorielles — au moins la vision — étaient nativement pluridimensionnelles et d'autres nativement unidimensionnelles (Kries 1904 ; Boernstein 1955 ; P.F. Strawson 1959 : 64–66). Il peut sembler à première vue (mais ce point devra être rectifié) que

¹ Naturellement, l'identité des n -dimensionnalités — ici le fait que la transposabilité d'une figure dans une autre exige que leurs E-classificateurs soient exprimables par des prédicats de même arité — n'est pas une condition suffisante pour l'isomorphie des domaines de places. On doit y ajouter les relations de position (voir *supra*) avec leurs propriétés formelles comme, dans le cas de la succession temporelle, la transitivité (si t suit u et u suit v , alors t suit v), l'antiréflexivité (t ne se suit pas lui-même) et l'asymétrie (si t suit u , alors u ne suit pas t). D'un domaine de places U unidimensionnel mais défini par une relation non transitive, n'est-il pas naturel de dire qu'il n'est pas isomorphe au continuum temporel T , au sens où une T -figure ne peut être transposée en une U -figure ?

si l'on choisit l'option [a], alors l'isomorphie des domaines de places peut toujours être garantie : la multimodalité figurale est dans tous les cas possible sous le rapport du domaine des places. Si on préfère [b], alors la multimodalité figurale est impossible dans certains cas, à savoir dans tous les cas où le nombre de dimensions (1, 2 ou 3) est différent.

L'option retenue par les gestaltistes de l'école de Berlin est clairement l'option [a] : toutes les modalités sensorielles partagent nativement un unique espace primitivement tridimensionnel. « L'espace perceptuel est *un*, clame Koffka (1936 : 303), et il peut être rempli avec des objets de différentes modalités sensorielles. » De même Köhler :

On peut avoir l'expérience des choses et de leurs propriétés par le toucher plutôt que visuellement. Les choses sont également ressenties comme chaudes ou froides ; elles dégagent une odeur, sont lourdes et émettent des sons. Toutes ces expériences sont localisées dans un unique espace perceptuel, que ce soit avec précision ou simplement de manière vague. (Köhler 1947 : 214)

Ainsi, la localisation des sons dans l'espace tridimensionnel a été abondamment étudiée et discutée, particulièrement par Wertheimer et Hornbostel après qu'ils eurent mis au point, durant la première guerre mondiale, un appareil acoustique destiné à mesurer la distance d'un sous-marin ou d'un feu d'artillerie adverse (Hornbostel & Wertheimer 1920 ; Allers & Schmiedek 1925 ; Wallach 1938 ; Ash 1995 : 187–190). Mais ce point est également généralisable aux sens dits proximaux (odorat, goût et toucher). Pour reprendre un exemple déjà cité, le mouvement d'un objet qui pique la main par intermittence n'est généralement pas plus localisé sur la peau bidimensionnelle qu'une couleur ne l'est sur la rétine : l'expérience tactile nous le donne plutôt comme un mouvement localisé dans l'espace tridimensionnel (Köhler 1947 : 100–101).

Les motifs de ce choix apparaissent clairement dans la réfutation par Koffka de l'argument de Berkeley en faveur de la bidimensionnalité de l'espace visuel (Koffka 1936 : 115–116). Berkeley, commente-t-il, arguait que, comme la rétine est bidimensionnelle, l'espace de la perception visuelle doit l'être aussi. Ainsi, la distance entre un point *a* et l'œil projette sur la rétine un unique point *b* : or le point *b* « demeure invariablement le même » quelle que soit la distance *a*—*b*. Et naturellement, ce raisonnement peut être généralisé à toutes les modalités sensorielles. Alors qu'un triangle projette sur la rétine une figure bidimensionnelle, un son produit dans le tympan une figure unidimensionnelle, etc. Selon Koffka, l'argument de Berkeley repose sur deux fausses présuppositions. D'abord, il « inclut l'hypothèse de constance en supposant que nous pouvons investiguer la totalité de l'espace perceptuel en examinant

ses points individuels séparément, un par un ». Ensuite, Berkeley présuppose une corrélation entre la n -dimensionnalité du stimulus et celle du phénomène perceptuel : si l'espace rétinien est bidimensionnel, alors l'espace phénoménal de la vision doit l'être aussi. Seulement, cette présupposition est fausse. Il est vrai que, dans certains cas, deux points inégalement distants de b peuvent en apparaître également distants. Sirius, par exemple, ne nous apparaît généralement ni plus ni moins distante que la planète Vénus. Mais cela ne prouve pas du tout que la distance subjective qui me sépare des deux astres est nulle, comme elle le serait si l'espace vu était bidimensionnel. La n -dimensionnalité du stimulus n'est pas nécessairement identique à celle du phénomène qu'il produit¹.

Pourtant, la manière dont nous avons formulé le problème prête à confusion. Car ce n'est pas sans risque de confusion que Grelling et Oppenheim identifiaient, dans l'article cité, le domaine de places de la mélodie au « continuum temporel ». Nos analyses précédentes n'avaient pas trait, du moins directement, à la n -dimensionnalité des modalités sensorielles, mais à celle, constitutive de figures, des domaines de places d'E-classificateurs. L'hypothèse que le sens auditif est nativement tridimensionnel n'exclut pas le moins du monde qu'une figure mélodique, comme le supposent Grelling et Oppenheim, soit unidimensionnelle — pas plus d'ailleurs, pour prendre un autre exemple, que la tridimensionnalité native du sens visuel n'empêche un triangle d'être bidimensionnel. Assurément, les deux figures sont perçues dans un espace tridimensionnel. Je vois un triangle dessiné sur le papier sous mes yeux ; j'entends fredonner le thème du chat dans la rue, à une certaine distance. Cependant, la tridimensionnalité n'est évidemment pas *constitutive* des figures *triangle* et *thème du chat*, lesquelles sont invariantes si je modifie leur localisation dans l'espace tridimensionnel — par exemple si je m'en approche ou m'en éloigne².

Il semble finalement que notre discussion de la n -dimensionnalité de l'espace et du temps ne soit pas d'une grande aide pour notre problème, suggérant seulement la modeste hypothèse suivante : si le domaine de places

¹ Le souci de construire pour chaque modalité sensorielle des structures n -dimensionnelles irréductibles à leur correspondant physique se retrouve par exemple dans le système d'intervalles musicaux de Samojloff, qui observe qu'« il serait intéressant de clarifier de quelle manière on pourrait ordonner adéquatement en un système les sensations d'autres organes sensoriels, par exemple celles des organes olfactifs et gustatifs » (Samojloff 1923 : 239).

² Naturellement, rien n'empêche que certaines figures musicales soient constitutivement tridimensionnelles. Voir (Dyck 2021), lequel, entre autres, cite l'exemple de *Fidelio* de Beethoven, dont la partition prescrit des trompettes éloignées (hors scène).

d'une figure est défini sur une relation à n_1 dimensions et celui de sa transposition sur une relation à n_2 dimensions, alors $n_1 = n_2$ et $0 < n_i \leq 3$. Contrairement à celle mentionnée plus haut selon laquelle $n = 1$, cette hypothèse ne permet pas d'affirmer que, n étant constant, la synesthésie serait toujours possible sous le rapport de la n -dimensionnalité.

Ce dernier résultat suscite pourtant quelques remarques plus significatives.

Selon Grelling et Oppenheim, comme on l'a vu plus haut, l'identité des E-classificateurs n'est pas suffisante pour la transposition : il y faut encore l'isomorphie des domaines de places. Mais le problème est que la distinction entre les classificateurs et leurs domaines de places n'est peut-être pas aussi claire que ne le suggèrent Grelling et Oppenheim. Considérons à nouveau la partition écrite du thème du chat et sa version sonore, en nous limitant, pour simplifier, à l'E-classificateur HAUTEUR. Une particularité de la partition est qu'elle représente par des relations spatiales non seulement les relations temporelles de la version sonore, mais aussi les intervalles mélodiques et harmoniques.

D'une part, on parle de « hauteur » sonore et l'E-classificateur HAUTEUR est représenté dans la partition par la relation spatiale « ... au-dessus de... » ou « ... en-dessous de... ». De même, nous disons que deux notes sont plus ou moins « proches » selon que leur intervalle est plus ou moins grand, et cette relation de proximité est représentée dans la partition par des relations de proximité au sens spatial.

D'autre part, l'antériorité d'une note dans le temps est représentée dans la partition par la relation spatiale « ... à gauche de... ». Assurément, « ... à gauche de... » n'est jamais employé pour dénoter une relation temporelle. Cependant, au lieu de « ... à gauche de... », on pourrait également dire, en anglais, « *before* », qui a un double sens spatial et temporel. Ce qui pourrait suggérer qu'espace et temps, quoique non isomorphes, ont en commun quelque trait structurel, qui transparaît dans le fait que la relation temporelle « ... antérieur à... » comme la relation spatiale « ... à gauche de... » sont exprimés par des prédicats à deux places. Le passé est « derrière » nous et l'avenir « devant », etc. C'est ce que remarquait très bien Köhler en 1947 :

Le temps expérimenté a certaines caractéristiques en commun avec l'espace expérimenté, en particulier avec la dimension spatiale qu'on indique par les mots « devant » et « derrière ». Partout et dans toutes les langues, les mots se référant à des relations dans cette dimension sont employés comme des termes pour des relations temporelles. En anglais, nous pouvons avoir quelque chose « *before us* » ou « *behind us* » en un sens à la fois spatial et temporel ; nous

regardons « *forward* » dans l'espace comme dans le temps ; et la mort s'approche dans le temps tout comme quelqu'un s'approche de nous dans l'espace. Du point de vue de l'isomorphisme, on s'attendrait à ce qu'il y ait une similitude de correspondance (*a corresponding kinship*) entre le corrélat physiologique du temporel et celui de cette dimension spatiale particulière. En tout cas, des « points » temporels forment des groupes temporels tout comme des points donnés simultanément tendent à former des groupes dans l'espace. Cela vaut pour l'audition et le toucher non moins que pour la vision. (Köhler 1947 : 150–151)¹.

Comprises en un certain sens, ces remarques pourraient conduire à ne conserver que l'arité des classificateurs au détriment de la n -dimensionnalité des modalités sensorielles. Plus précisément, il s'agirait de déplacer le problème de la transposition synesthésique de la n -dimensionnalité du domaine de places à l'arité des classificateurs, lesquelles poseraient désormais des problèmes dans une large mesure distincts et indépendants. Cela aurait pour effet de rendre non pertinent le fait que la version sonore et la partition n'ont pas la même n -dimensionnalité du point de vue des modalités sensorielles. Dans le cas qui nous occupe, le temps devrait être inclus dans la liste des E-classificateurs. Exactement comme la hauteur sonore, il serait par exemple conçu comme une fonction qui, à chaque donnée δ , associe une valeur $t \in \mathbb{R}$.

Une conséquence de cela est qu'il serait envisageable de voir dans la partition une transposition synesthésique de la version sonore². Tout dépendra des classificateurs : le thème du chat devra pouvoir être décrit exhaustivement au moyen de classificateurs transmodaux (audiovisuels). Plus précisément, si certains P-classificateurs sont identiques et si les mots « au-dessus de... », « ... *before*... », etc., ne sont pas équivoques (par exemple métaphoriques)³, c'est-à-dire dénotent les mêmes classificateurs transmodaux comme le suggèrent les remarques ci-dessus, alors la partition du thème du chat remplirait les trois critères de Grelling et Oppenheim pour être une transposition de sa version sonore : [1] le domaine du classificateur binaire ESPACE_PLAN de la

¹ Cette constatation peut être étendue à une multitude d'autres caractères. « Des mots comme "homogène" et "continu", disait par exemple Köhler, se réfèrent pareillement, bien entendu, à des attributs de surfaces dans l'espace ou d'étendues dans le temps » (Köhler 1947: 174–175).

² En un sens analogue, Grelling & Oppenheim (1988a) eux-mêmes évoquent le plan d'une maison : « Le plan de la maison (comme complexe) est sa *Gestalt* pour une transposition par changement d'échelle de mesure ; car le plan est un invariant pour cette transposition. »

³ Sur cette condition, voir (Hornbostel 1925 : 190).

partition est isomorphe au domaine de la combinaison (binaire) des deux classificateurs unaires TEMPS et HAUTEUR de la version sonore ; [2] les E-classificateurs sont identiques ; [3] certains P-classificateurs sont identiques.

Cette dernière conception présente des avantages. D’abord, elle rend mieux compte du fait qu’un carré rouge, par exemple, est une figure bidimensionnelle alors même que — comme on peut tout au moins le supposer — l’espace visuel est tridimensionnel. Ensuite, comme je le préciserai en conclusion, elle permet de mieux délimiter le problème de la figuralité comprise au sens des deux critères d’Ehrenfels. La question de savoir si la partition est ou non une transposition de la version sonore, par exemple, pourra être tenue pour une question oiseuse, dont la réponse dépend simplement du choix des classificateurs. Ce point suggère que les critères d’Ehrenfels — sursommativité et transposabilité — sont peut-être insuffisants et qu’il convient d’en envisager d’autres, particulièrement des critères fondés sur la distinction figure-fond et les phénomènes de saillance ou d’importance à l’intérieur de systèmes dynamiques (Seron 2022).

Remarques conclusives

Dans le sillage de la quasi-analyse carnapienne, la clef de voûte de l’article de Grelling et Oppenheim est l’idée que la méréologie des figures n’est pas *réelle*, mais qu’elle est essentiellement une affaire de descriptions, c’est-à-dire de concepts. En d’autres termes, aucune des parties propres de la figure n’est une partie réelle. Toutes sont des « parties logiques » ou, pour user d’une expression brentanienne, des « parties distinctionnelles ». Se demander de quelles parties se compose une figure mélodique, c’est un peu comme se demander si le citron se compose de deux moitiés ou de trois tiers de citron : il ne pourrait sortir grand-chose d’une controverse entre moitiéistes et tiersistes. Une grande part des malentendus au sujet de la *Gestalttheorie*, y compris chez certains de ses défenseurs, viennent du fait qu’on n’a pas vu que, tout au moins au niveau non expérimental, elle parlait moins des phénomènes figuraux que de notre discours sur les phénomènes figuraux. Oppenheim et Grelling ont été attentifs à cette distinction. Ces auteurs ne défendent pas, par exemple, la thèse *psychologique* absurde que l’expérience d’une figure se ferait au moyen de classificateurs, mais plutôt la thèse *philosophique* suivante : le concept de figure utilisé en psychologie ne doit pas être défini méréologiquement. Cette thèse n’a rien à voir avec ce qu’est en réalité l’expérience figurale, qui est, au mieux, une question pour les psychologues.

Sans doute ces éléments nous incitent-ils à franchir un pas de plus. D'une part, ils se ramènent à ceci : [1] la méréologie n'est pas réelle (toutes les parties propres sont des parties logiques) ; [2] le point de vue figural — la description en termes de totalités plutôt qu'en termes d'associations d'éléments — est meilleur en psychologie. D'autre part, nous affirmons que [2] *ne dérive pas de [1]*. Ce point est évident, si l'on songe que rien n'empêche de choisir des classificateurs méréologiques. Par exemple, un groupe de trois carrés contigus pourra être décrit au moyen du classificateur NOMBRE_DE_CARRÉS, dont la valeur sera 3. De tels classificateurs ne dénoteront pas des parties réelles, mais des parties logiques analogues aux moitiés et aux tiers de citron. Nous obtenons alors une description non totalisante, méréologique, mais dépourvue de parties propres réelles et en ce sens, pourrait-on dire, « quasi-méréologique ». Bref, une fois admis que toute partition est affaire de classificateurs, c'est-à-dire de description et de langage, la distinction entre approches totalisante et méréologique devient elle-même une question intralinguistique. Cette idée a été clairement énoncée par Moritz Schlick, quand il reprochait à Köhler de préconiser indûment l'approche gestaltiste en physique :

La question de savoir si quelque chose est une totalité (*Ganzheit*) est analogue à la question de savoir s'il existe réellement une géométrie euclidienne. Les deux questions, comme telles, sont vides de sens. (...) Une seule question pourvue de sens reste ouverte : est-il plus pratique de décrire le monde de façon euclidienne ou de façon non euclidienne ? (...) Dans l'ensemble, la description en termes de totalité se révèle plus pratique en psychologie, la description sommatique plus pratique en physique. Les deux sont possibles absolument parlant. Köhler croit avoir démontré qu'il y a aussi des figures en physique. Sa tentative prouve seulement que le point de vue totalisant est également possible. *Mais il n'y a aucune cloison réelle entre somme et figure. Il n'y a que différentes façons de parler, différentes formes sous lesquelles on présente les choses.* (Schlick 2020 : 316, je souligne ; cf. Schlick 1935)

Quelle conclusion tirer en ce qui concerne la multimodalité figurale ?

D'abord, la conception de Grelling et Oppenheim — et plus encore son dérivé esquissé à la fin de la section précédente — semble sinon élargir significativement le concept de figure, du moins en énoncer des conditions nécessaires mais non suffisantes. Comme je l'ai suggéré plus haut, on peut se demander si cette conception capture vraiment la compréhension *psychologique* de ce qu'est une figure et s'il ne convient pas d'adjoindre aux deux critères d'Ehrenfels d'autres conditions de nature psychologique, spécialement en ce

qui concerne la distinction figure-fond¹. Il n'est pas sûr, par exemple, qu'il soit psychologiquement pertinent de considérer qu'une même figure mélodique s'instancie dans sa partition et sa version sonore individuelles. Quoi qu'il en soit, si l'on s'en tient à la dernière conception exposée plus haut, il semblerait que la synesthésie ne pût dans aucun cas être exclue *a priori*. Nous dirons alors qu'elle dépend seulement de l'existence, dans notre lexique, de classificateurs transmodaux (dont les domaines sont supposés isomorphes). L'intermodalité figurale, quant à elle, serait possible dans tous les cas, même si les domaines des E-classificateurs ne sont pas isomorphes. Ce résultat peut être vu comme trivial, puisqu'il ne fait que répéter la contrainte d'isomorphie.

Ensuite, c'est du même coup les différences entre modalités sensorielles et le problème de la multimodalité figurale qui passent au second plan. Comprise au sens indiqué plus haut, l'idée que toutes les parties sont des parties logiques a pour effet d'estomper considérablement d'une part les différences entre modalités sensorielles, d'autre part la différence entre la figure et l'expérience totale². L'expérience individuelle est correctement décrite comme une expérience totale et multimodale, où la présence de figures (au sens défini par Grelling et Oppenheim) et de modalités sensorielles dépend des classificateurs. Comme l'affirmait Michael Tye :

En réalité il n'y a pas d'entités telles que des expériences purement visuelles, des expériences purement auditives ou des expériences purement olfactives dans la conscience normale de tous les jours. (...) Ce sont des fictions dues à l'imagination des philosophes et des psychologues. (...) Il y a simplement une expérience multimodale, descriptible de manière plus ou moins riche. (Tye 2003 : 28)

Les distinctions entre modalités sensorielles dépendent des classificateurs et donc du langage ; la partition modale de l'expérience — par exemple en deux perceptions partielles visuelle et auditive — est purement *logique*. Naturellement, cette idée n'a aucune implication psychologique. (La question n'est absolument pas de savoir si les différences entre les modalités sensorielles existent également pour les êtres dépourvus de langage. Un être sans langage comme un hareng est assurément capable de distinguer entre deux données individuelles dont l'une est visuelle et l'autre auditive. C'est-à-dire qu'il fait

¹ Secondairement, la question est de savoir si la description de Grelling et Oppenheim est suffisante pour distinguer entre figures ehrenfelsiennes et « systèmes d'effectuation » (voir *supra*).

² La seconde distinction correspond à celle de Köhler entre figures fortes et faibles (Köhler 1920 : 114–172 ; Lewin 1926 : 302).

l'expérience d'un *contraste* entre l'une et l'autre. Mais fait-il la différence entre les *types* « visuel » et « auditif » ? Il n'y a qu'un moyen de le savoir : il faut lui poser la question. Si le hareng ne répond pas, nous ne serons pas plus avancés. Mais s'il répond, il sera doté de langage et nous n'aurons pas la réponse non plus. La question, en conséquence, est telle qu'il est *a priori* impossible de lui apporter une réponse. Conformément au sain principe wittgensteinien, nous dirons que cette question est vide de sens et n'a pas à être posée.) En revanche, à la suite de Schlick, nous dirons que le choix des classificateurs comme indirectement l'introduction de figures multimodales sont motivés par la psychologie — à savoir par ses résultats antérieurs, ses finalités, sa méthodologie (expérimentation, introspection), etc.

Bibliographie

- Allers, R., & Schmiedek, O. (1925), « Über die Wahrnehmung der Schallrichtung, II. Mitteilung », *Psychologische Forschung*, 6, p. 92–112.
- Ash, M. (1995), *Gestalt Psychology in German Culture, 1890-1967. Holism and the Quest for Objectivity*, Cambridge MA, Cambridge University Press.
- Boernstein, W.S. (1955), « Classification of the human senses », *Yale Journal of Biology and Medicine*, vol. 28, p. 215.
- Brentano, F. (1979). *Untersuchungen zur Sinnespsychologie*, Hamburg, Meiner.
- Buytendijk, F.J.J. (1970), « Some aspects of touch », *Journal of Phenomenological Psychology*, vol. 1, 1, p. 99–122.
- Carnap, R. (1929), *Abriss der Logistik. Mit besonderer Berücksichtigung der Relationstheorie und ihrer Anwendungen*, Wien: Springer-Verlag.
- Carnap, R. (1998), *Der logische Aufbau der Welt*, Hamburg, Meiner.
- Chrudzinski, A. (2013), « Gestalt, equivalency, and functional dependency : Kurt Grelling's formal ontology », dans N. Milkov, V. Peckhaus (éds.), *The Berlin Group and the Philosophy of Logical Empiricism*, Dordrecht, Springer, p. 245–261.
- Dyck, J. (2021), « Spatial music », *European Journal of Philosophy*, févr. 2021, p. 1–14.
- Ehrenfels, C. von (1890), « Über Gestaltqualitäten », *Vierteljahrsschrift für wissenschaftliche Philosophie*, 14, 3, p. 249–292.
- Grelling, K., & Oppenheim, P. (1938). « Der Gestaltbegriff im Lichte der neuen Logik », *Erkenntnis*, 7 (1937/1938), p. 211–225. Repris dans *Die Berliner Gruppe. Texte zum logischen Empirismus*, Hamburg, Meiner, 2015, p. 377–394.
- Grelling, K., & Oppenheim, P. (1988a), « Supplementary remarks on the concept of Gestalt », dans B. Smith (éd.), *Foundations of Gestalt Theory*, München Wien, Philosophia Verlag, p. 206–209.

- Grelling, K., & Oppenheim, P. (1988b), « Logical analysis of “Gestalt” as “functional whole” », dans B. Smith (éd.), *Foundations of Gestalt Theory*, op. cit., p. 210–216. Réédité dans *Gestalt Theory*, 21/1 (1999), p. 49–54.
- Helson, H. (1933), « The fundamental propositions of Gestalt psychology », *Psychological Review*, 40/1, p. 13–32.
- Hornbostel, E.M. von (1925), « Die Einheit der Sinne », *Melos, Zeitschrift für Musik*, V, 6, p. 290–297.
- Hornbostel, E.M. von, & Wertheimer, M. (1920), « Über die Wahrnehmung der Schallrichtung », *Sitzungsberichte der preussischen Akademie der Wissenschaften*, 20, p. 388–396.
- Husserl, E. (1970). *Philosophie der Arithmetik. Mit ergänzenden Texten (1890–1901)*, éd. par L. Eley, *Husserliana*, vol. 12, Den Haag, Nijhoff.
- Husserl, E. (1984). *Logische Untersuchungen, Husserliana*, vol. 19/1, éd. par U. Panzer, The Hague Boston Lancaster, Kluwer, 1984.
- James, W. (1950), *The Principles of Psychology*, New York, Dover, vol. 2.
- Katz, D. (1925), *Der Aufbau der Tastwelt*, Leipzig, Johann Ambrosius Barth, 1925.
- Koffka, K. (1936), *Principles of Gestalt Psychology*, London, Kegan Paul, Trench, Trubner & Co.
- Köhler, W. (1920), *Die physischen Gestalten in Ruhe und im stationären Zustand. Eine naturphilosophische Untersuchung*, Braunschweig, Vieweg & Sohn.
- Köhler, W. (1925), *Gestaltprobleme und Anfänge einer Gestalttheorie*, München, Bergmann.
- Köhler, W. (1947). *Gestalt Psychology. An Introduction to New Concepts in Modern Psychology*, New York, Liveright.
- Kries, J. von (1904), « Zur Psychologie der Sinne », dans W. Nagel (éd.), *Handbuch der Physiologie des Menschen*, vol. 3/1, Braunschweig, Vieweg und Sohn, p. 16–29.
- Leclercq, B. (2020), « Foundational mereology as a logical tool for descriptive psychology », dans D. Fisette, G. Fréchette, H. Janoušek (éds.), *Franz Brentano's Philosophy after One Hundred Years. From History of Philosophy to Reism*, Springer, p. 125–147.
- Marty, A. (1908). *Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie*, vol. 1, Halle, Niemeyer.
- Lewin, K. (1926), « Vorbemerkungen über die psychischen Kräfte und Energien und über die Struktur der Seele », *Psychologische Forschung*, 7, p. 294–329.
- Lewin, K. (1936), *Principles of Topological Psychology*, New York London, McGraw-Hill.
- Locke, J. (1997). *An Essay Concerning Human Understanding*, London, Penguin Books.
- Meinong, A. (1894), « Beiträge zur Theorie der psychischen Analyse », *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 6, p. 340–385 ; Schluß, *ibid.*, p. 417–455.
- Meinong, A. (1903), « Bemerkungen über den Farbenkörper und das Mischungsgesetz », *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 33, p. 1–80.

- Rescher, N., & Oppenheim, P. (1955), « Logical analysis of Gestalt concepts », *The British Journal for the Philosophy of Science*, 6/22, p. 89–106.
- Schlick, M. (1935), « Über den Begriff der Ganzheit », *Erkenntnis*, 5, p. 52–55.
- Schlick, M. (2020), « Gestaltpsychologie », dans D. Fisette, G. Fréchette, F. Stadler (éds.), *Franz Brentano and Austrian Philosophy*, Cham, Springer, p. 309–326.
- Samojloff, A. (1923), « Die Anordnung der musikalischen Intervalle im Raume », *Psychologische Forschung*, 3, p. 231–240.
- Scheerer, M. (1931), *Die Lehre von der Gestalt, ihre Methode und ihr psychologischer Gegenstand*, Berlin, De Gruyter.
- Seron, D. (2022), « Décire n'est pas tout : Kurt Lewin sur l'émotion », *Philosophie*, 152, p. 36–47.
- Spearman, C. (1926), « Two defects in the theory of "Gestalt" », *Proceedings and Papers of the 8th International Congress of Psychology at Groningen*, Groningen, Noordhoff, p. 190–197.
- Spearman, C. (1937a), *Psychology Down the Ages*, vol. 1, London, MacMillan.
- Spearman, C. (1937b), « The confusion that is Gestalt-psychology », *The American Journal of Psychology*, vol. 50, 1/4, p. 369–383.
- Strawson, P.F. (1959), *Individuals. An Essay in Descriptive Metaphysics*, London New York, Routledge.
- Stern, D. (1998), *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*, London, Karnac Books.
- Straus, E. (1978). *Vom Sinn der Sinne. Ein Beitrag zur Grundlegung der Psychologie*, Berlin Heidelberg New York, Springer.
- Trémault, É. (2020), *Structure et sensation. Une critique de la psychologie de la forme*, Paris, Vrin.
- Tye, M. (2003), *Consciousness and Persons. Unity and Identity*, Cambridge MA, MIT Press.
- Wallach, H. (1938), « Über die Wahrnehmung der Schallrichtung », *Psychologische Forschung*, 22, p. 238–266.