

## Voix et images

Roberto Casati

► **To cite this version:**

Roberto Casati. Voix et images. La philosophie du son, Jacqueline Chambon, 1994, 978-2877111096.  
ijn\_00000522

**HAL Id: ijn\_00000522**

**[https://jeannicod.ccsd.cnrs.fr/ijn\\_00000522](https://jeannicod.ccsd.cnrs.fr/ijn_00000522)**

Submitted on 30 Jul 2004

**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.

## 6. DEUX INTERLUDES : LES VOIX ET LES IMAGES SONORES

### 6.1. *Premier interlude : que sont les voix ?*

Daniel Dennett propose une analyse critique du dualisme traditionnel esprit-corps en montrant qu'une partie des raisons qui peuvent motiver un tel dualisme pourraient également justifier un dualisme des voix : il y aurait des voix non matérielles, en plus des voix matérielles. Dennett écrit :

Nous disons «J'entends une voix», «Il a une voix de ténor», «Tu vas te forcer la voix» et «J'ai perdu ma voix». Cela signifie-t-il que la voix est une chose ? Si oui, quelle chose est-elle ? La voix que nous forçons semble être une partie physique du corps aussi peu problématique que le dos ou les yeux que nous forçons, peut-être les cordes vocales ; mais il est certain que l'on n'a pas de cordes vocales de ténor, que l'on ne se réjouit pas d'entendre les cordes vocales de Sutherland, ou encore que l'on ne perd pas ses cordes vocales ; et notre voix, à la différence de nos cordes vocales, peut être transmise par radio à travers les mers et survivre à notre mort sur une bande magnétique. On ne force pas non plus, ni ne reconnaît, ni ne perd aucune vibration [...] (Dennett, 1969 : 8).

Nous proposons ici une solution au problème des voix (qui n'est pas censée s'appliquer au problème du dualisme traditionnel esprit-corps). Il est naturel de concevoir les voix, pas tellement comme des sons ni en général comme des événe-

ments, mais comme les *bases* de dispositions, plus précisément des dispositions d'un sujet à interagir acoustiquement avec le milieu environnant. Un exemple de disposition est la solubilité : la propriété de se dissoudre dans des liquides. Pour qu'un objet soit soluble, point n'est besoin qu'il soit dissout. De même, lorsqu'une voix est attribuée à quelqu'un, ce qui est attribué est une certaine capacité, celle d'interagir acoustiquement avec l'environnement (en somme, de se faire entendre). Cette disposition dépend d'une base non dispositionnelle : un complexe de propriétés de l'organe vocal lui-même. Voyons comment notre proposition rend compte des exemples proposés par Dennett, en suggérant des paraphrases appropriées aux énoncés problématiques :

(a) « Nous forçons notre voix » : *nous modifions une propriété de notre appareil vocal*, parmi celles qui constituent la base de la disposition à interagir acoustiquement avec l'environnement.

(b) « Il a perdu sa voix » : *il a perdu la base en question* (il faut évidemment considérer la base au sens large, incluant aussi ce qui rend possible l'actualisation de la propriété ; dans le cas de certains aphasiques, l'appareil vocal est intact).

(c) « J'ai entendu sa voix » : *j'ai entendu le son de sa voix*. Les défenseurs de la thèse selon laquelle les seules choses que nous entendons sont des sons nieront que les voix sont les bases de dispositions, dans la mesure où les sons ne le sont pas. Mais en fait cette thèse est suspecte. Nous entendons des sons, mais ce ne sont pas les seuls *audibilia*. Nous entendons également des personnes, des voitures, des avalanches. (Certes, nous entendons ces choses en entendant des sons ; cf. sections 11.3 et 12.2) La différence de statut entre les sons et les voix est évidente lorsque nous réalisons que « J'ai entendu le son de cette voix » est acceptable, alors que « J'ai entendu le son de ce son » ne l'est pas. Nous entendons les voix — les

bases — en entendant leurs sons; donc nous entendons des voix *indirectement*. Les sons jouent ici le rôle d'intermédiaires perceptifs dans l'audition des voix.

(d) «J'ai enregistré la voix de la Callas»: *j'ai enregistré le son de la voix de la Callas*. Nous pouvons enregistrer le son de la voix de la Callas; ce faisant, nous enregistrons la voix de la Callas. Que l'on considère aussi «J'ai enregistré les derniers mots de la Callas». Cela veut dire que le son de ces mots a été enregistré, mais aussi que leur contenu a été enregistré.

Un argument supplémentaire, d'ordre un peu plus métaphysique, est le suivant: les bases des dispositions sont préférables aux dispositions tout court comme termes d'une relation perceptive, car les dispositions sont des propriétés de second ordre, et la causalité, qui est au centre de toute relation perceptive, ne peut pas avoir comme terme une propriété de deuxième ordre. Par analogie, la perception de la couleur pourrait être décrite comme la perception de la base d'une certaine propriété d'interagir avec des sujets percevants: ce qui est perçu n'est pas la disposition elle-même, mais sa base (sur la couleur, cf. aussi ch. 11).

## 6.2. *Deuxième interlude: les images sonores.*

Lorsque j'écoute un enregistrement de Maria Callas, est-ce bien le chant de la Callas que j'entends? Ne s'agit-il pas plutôt d'un second événement sonore, qui a une certaine relation avec le chant de la Callas (par exemple, une relation de ressemblance)? Nous proposons de considérer ce cas comme étant exactement parallèle à celui de la perception d'une image. L'événement sonore *a* qui se produit dans ma chambre lorsque j'écoute la radio est une *image sonore* de l'événement *b* qui s'est produit à la Scala le 28 mai 1955, lors de la représentation de *La Traviata*. Le sujet percevant contribue à ce qui fait que *a* est une image de *b*: *a* doit susciter chez lui des capacités de reconnaissance. Il y a aussi une contribution du monde

externe, entre autres un lien causal entre *a* et *b*. L'événement *a* pourrait ne pas susciter de capacités de reconnaissance chez le sujet (parce que, par exemple, l'enregistrement a été totalement effacé; une très mauvaise image de *b* compte à peine comme image tout court); et un événement sonore *c*, indiscernable de *a*, pourrait ne pas être une image de *b* (parce que, par exemple, il a été causé par une manipulation subtile de paramètres auditifs sur un ordinateur qui gère un programme musical). Voir une personne en image est au mieux une forme de perception indirecte : c'est en voyant l'image que je vois la personne. Entendre une chanson à la radio est aussi une forme de perception indirecte : c'est en entendant le son de la radio que j'entends la chanson.

Pour terminer, considérez l'analogie linguistique suivante. Jackendoff (1975) a observé la possibilité de faire référence aux objets représentés dans les images de plusieurs façons :

Dans le tableau, Marie fait deux mètres de hauteur  
Dans le tableau, Marie fait deux centimètres de hauteur

Ces deux phrases ne sont pas forcément contradictoires. La première parle de *Marie dans l'image*; elle est représentée comme faisant deux mètres. La seconde parle de *l'image de Marie*, et cette image fait deux centimètres. Nous suggérons qu'une distinction similaire peut avoir une application dans le cas du son. Ainsi, il peut n'y avoir aucune contradiction entre les deux phrases suivantes :

Au téléphone, Marie parlait doucement  
Au téléphone, Marie parlait fort

Supposons par exemple que la ligne soit perturbée et que le signal soit faible. Dans ce cas, il est possible d'affirmer que l'auditeur a une image acoustique faible d'un événement bruyant, Marie qui hausse la voix.