

## Introduction

Le terme de *timbre* est utilisé pour renvoyer à un certain nombre de propriétés acoustiques des instruments de musique. Si cette notion est souvent traitée dans la littérature ethnomusicologique, elle est définie en quelque sorte par défaut dans la communauté des acousticiens: selon la norme ANSI, le timbre est « ce qui différencie deux sons de même hauteur, de même durée et de même intensité » (American National Standards Institute 1973). C'est sur la base de cette définition qu'ont été menées une bonne partie (en fait la quasi-totalité) des expériences de psychoacoustique relatives au timbre<sup>1</sup>. De fait, selon les communautés scientifiques, le timbre est défini tantôt par rapport aux propriétés acoustiques du son, tantôt par rapport à ses propriétés musicales, en lien étroit avec la spécificité de l'instrument d'un point de vue perceptif. Ce cadre correspond à ce que Castellengo (2002) appelle le *timbre causal*, notion dont la recherche en acoustique musicale n'a pas encore donné de définition claire, même si la catégorisation sur la similitude des sources avait déjà été soulignée par Handel (1989) pour les événements auditifs et par Cadoz (1991) pour la musique en général.

Au cours de nos recherches menées avec des sons produits par le *steeldrum* (Gaillard 2000), nous avons pu observer un comportement singulier des auditeurs concernant la perception temporelle du transitoire d'attaque du son original de cet instrument relativement peu connu, même chez les musiciens. Les auditeurs, incapables de reconnaître le timbre de cet instrument peu conventionnel, ont parfois perçu celui-ci comme n'étant pas un instrument à percussion, et se comportent donc comme les auditeurs des musiques contemporaines « concrètes » ou électroacoustiques, qui déconcertent tout autant par leur étrangeté liée à la difficulté d'identification de la source instrumentale (Castellengo et Dubois, soumis). Les modifications systématiques de certaines particularités acoustiques du transitoire d'attaque du *steeldrum* que nous avons réalisées dans l'expérience présentée ici devraient dès lors nous permettre d'identifier certaines des propriétés du signal physique auxquelles il est possible d'imputer ces effets perceptifs.

# Variabilité temporelle dans la perception des transitoires d'attaque de sons percussifs

Pascal Gaillard

(Laboratoire Jacques-Lordat, Université de Toulouse II)

Michèle Castellengo (LAM-CNRS, Université de Paris VI)

Danièle Dubois (LAM-CNRS, Université de Paris VI)

## Contexte en acoustique musicale

Du point de vue générique de la perception auditive et de la reconnaissance des sources (*timbre causal*), le transitoire d'attaque des sons joue un rôle déterminant. Les transitoires d'attaque des *steeldrums* ont cependant, à cet égard, des caractéristiques originales.

Le *steeldrum* est un instrument de musique né dans les années 1930-1940 dans une île des Petites Antilles, Trinidad. Cette république, indépendante depuis 1962, a su mettre en valeur son patrimoine culturel et musical et faire de cet instrument un emblème.

Cet instrument est fabriqué à partir d'un bidon de pétrole dont on coupe la base à une hauteur directement proportionnelle à la tessiture de l'instrument final: si on veut obtenir une basse, la « jupe » (voir figure 1) sera longue, pour un soprano, elle sera plus courte. Après cette opération, la partie supérieure du bidon est travaillée afin de lui donner une forme globalement concave<sup>2</sup>. Sur cette forme seront délimitées les différentes zones, modelées afin d'émettre des sons possédant des hauteurs déterminées (Alexis, O'Connor et Rossing 1986).

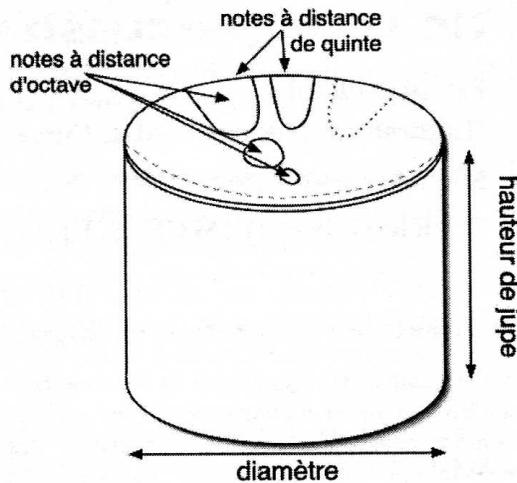
L'acoustique du *steeldrum* a été étudiée par quelques auteurs (Hansen et Rossing 1987; Rossing, Hampton et Boverman 1986) et a fait l'objet d'un regard particulier sur son fonctionnement mécanique en vue d'une modélisation (par exemple, Achong 1996; Achong 1999; Achong et Sinanan-Singh 1997). Sa classification au sein de la famille des instruments de musique pose elle-même quelques difficultés du fait de son mode de fonctionnement origi-

<sup>1</sup> Pour un bilan récent, voir Hajda, Kendall et Carterette, 1997.

<sup>2</sup> Pour une description précise et détaillée de la fabrication des *steeldrums*, voir Gaillard 2000; Kronman 1992.

nal. Selon les situations et les points de vue, il est classé dans la famille des gongs ou des cloches (Rossing 1982), ou dans la famille des jeux de plaques, au même titre qu'un vibraphone (Dournon 1992).

**Figure 1.** Schématisation des différents éléments constituant un steeldrum. La surface supérieure est concave et reçoit l'ensemble des notes. Le fond est ouvert.



Sur le plan acoustique, le son du steeldrum est composé de partiels émis par les différentes zones « accordées » de sorte que les intervalles s'approchent au mieux d'une série harmonique: octave, quinte, double octave. Toutefois, la fabrication artisanale exclut la normalisation. Elle ne permet pas de générer des règles très précises sur la nature du son de steeldrum. Cependant, à la suite des études citées plus haut, nous pouvons décrire le son isolé de chaque note de steeldrum comme un

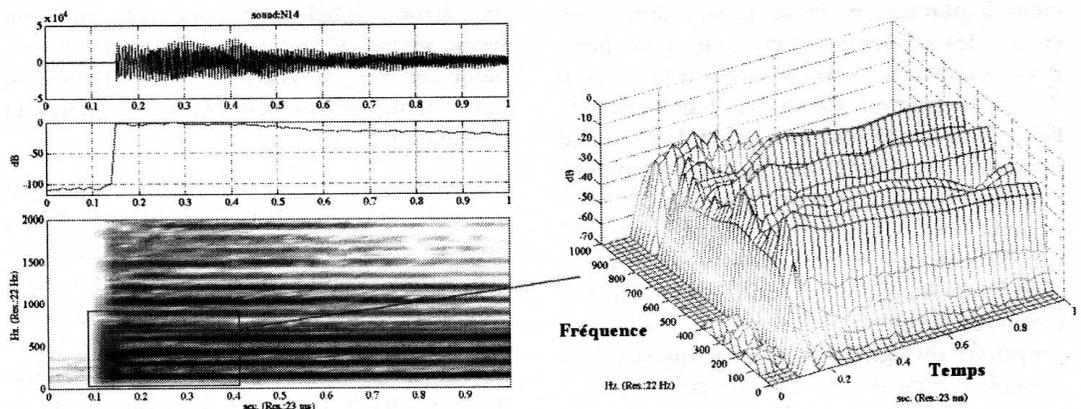
monde sonore complexe, en particulier en ce qui concerne le transitoire d'attaque. C'est précisément cette complexité et les possibilités qu'elle offre en matière de manipulation de sons pour élaborer des stimuli dans le cadre de tests auditifs qui a attiré notre attention.

Un son isolé de steeldrum comporte la plupart du temps les caractéristiques suivantes (voir figure 2):

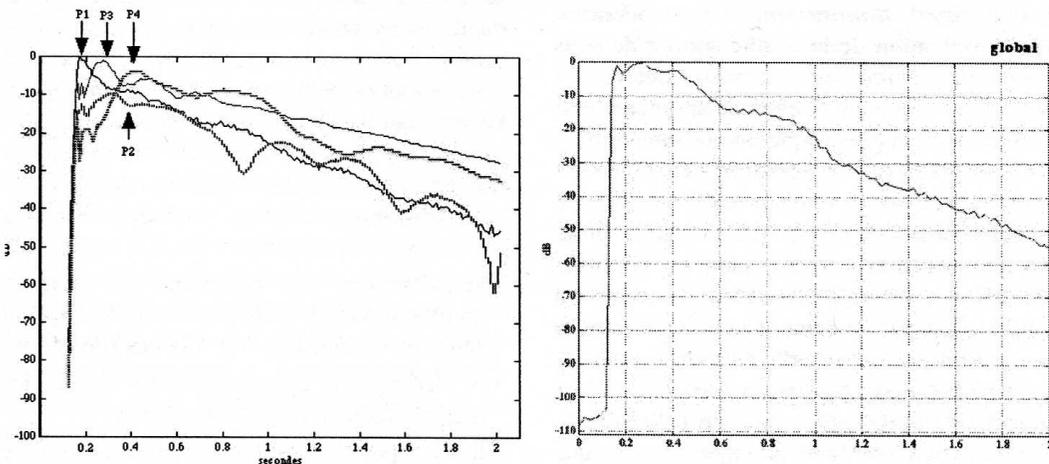
- Quatre partiels (au moins) ont des rapports quasi harmoniques entre eux (plusieurs autres partiels pouvant apparaître avec plus ou moins d'intensité et avec des fréquences de plus en plus éloignées des multiples entiers du premier partiel).
- La décroissance de chacun de ces partiels ne semble pas, *a priori*, en rapport avec la décroissance des autres.
- Plus précisément: l'apparition du maximum de chacun de ces partiels ne se produit pas en même temps (voir figure 3), et ce pour des raisons qui restent assez obscures d'un point de vue physique.

De la dynamique particulière de chaque partiel découle une dynamique globale originale pour un instrument à percussion. En effet, la plupart du temps, le son d'un instrument à percussion possède une décroissance plus ou moins régulière juste après l'impulsion initiale. Rares sont les cas où l'amplitude du son remonte après l'impulsion initiale. Le quatrième partiel est, dans le cas que nous proposons, particulièrement représentatif du retard important qui sépare le début du son et le moment de l'apparition du maximum de ce partiel. L'écart temporel peut atteindre

**Figure 2.** Analyse des quatre premiers partiels d'un son de steeldrum. À gauche, un sonagramme et une courbe des niveaux permettent d'observer l'ensemble des informations acoustiques disponibles. La vue en trois dimensions à droite permet d'évaluer les comportements en intensité en fonction du temps des quatre premiers partiels.



**Figure 3.** Analyse de la dynamique des quatre premiers partiels d'un son de steeldrum. À droite, la globalité du son où on voit le regain d'énergie après le début du son, et à gauche les différents partiels (P1, 2, 3 ou 4) mis en perspective.



plus de 200 ms (comme nous le voyons dans l'exemple de la figure 3). Il semble que des maxima ayant un écart de plus 120 ms sont détectables par un auditeur (Zwicker et Fastl 1999, 275). Par conséquent, les auditeurs sont tout à fait susceptibles de percevoir ces deux événements. Cette caractéristique est propre au steeldrum, et partant, utile pour étudier la perception des transitoires d'attaque.

Les sons produits par cet instrument constituent un ensemble d'objets sonores précieux à étudier. Il convient cependant de s'appuyer sur un cadre théorique dans le domaine de la perception en général et de la perception auditive en particulier afin de mettre au point des tests permettant d'étudier précisément la spécificité des jugements qualitatifs de ces sons musicaux.

### Contexte en psychologie cognitive

Nous nous référerons pour cela à la problématique récente de la catégorisation en psychologie cognitive, principalement développée dans le domaine visuel<sup>3</sup>, mais plus récemment dans d'autres domaines sensoriels<sup>4</sup>, et qui permet, à travers ses concepts et méthodes, d'identifier les critères pertinents de la construction des catégories perceptives. Dans les années 1970, c'est la rencontre de linguistes et de psychologues qui a permis l'émergence d'une représentation du fonctionnement cognitif en rapport avec le langage et plus précisément le lexique, dans le cadre d'une théorie des catégories cognitives et de la catégorisation. Nous retiendrons ici les concepts de prototype et de typicalité qui ont été mis de l'avant pour proposer une représentation des processus cognitifs

mis en œuvre dans la construction des catégories opératoires pour l'adaptation quotidienne (Neisser 1991) ainsi que leur représentation en langue (Dubois 2000).

Le *prototype* est, dans ce contexte, une construction cognitive qui constitue le meilleur représentant de la catégorie à laquelle l'objet perçu va être associé et qui va lui donner sens. Le prototype constitue une connaissance préalable qui interviendra en pilotant les processus descendants (« *top down* » ou « *knowledge driven* ») dans la perception et l'identification de tout nouveau stimulus semblable. La réponse d'identification serait, dans ce cadre, fondée sur une évaluation de la « ressemblance » au prototype, cette distance au prototype étant appelée « typicalité » (Rosch 1978; en français, Dubois 1993). Cette approche a déjà été jugée productive dans différents domaines du sonore, en particulier pour les objets sonores complexes dont l'approche analytique est problématique, étant donnée notre ignorance des paramètres ou propriétés pertinentes et des évaluations holistiques qu'ils suscitent : sons environnementaux, manufacturés, musicaux, vocaux, etc. (Maffiolo 1999; Dubois, Guastavino et Raimbault 2005; Dumortier 2004; Gaillard 2000; Guyot, Castellengo et Fabre 1997; Morange 2005, respectivement).

Cette conception inverse, en quelque sorte, le paradigme classique de la psychophysique qui « calcule » le décalage des réponses « subjectives » à des stimuli décrits « objectivement » par des paramètres physiques. Elle conduit symétriquement à rechercher des corrélations de mesures acoustiques à des structures cognitives repérées dans les organisations catégorielles et dans leurs descriptions verbales.

<sup>3</sup> Pour un bilan en français, voir Dubois 1993; voir également Rosch 1978.

<sup>4</sup> Voir Rouby *et al.* 2002, p. 82-99, en olfaction; Dubois 2000, sur des sons environnementaux (non musicaux).

Appliquée au timbre, cette problématique va permettre de tester l'hypothèse suggérée par Castellengo (Castellengo 2002) selon laquelle le concept de *timbre causal* est opératoire dans l'évaluation de la qualité sonore de sons musicaux: l'auteure suppose en effet que la production d'un son déclenche spontanément et de manière automatique une écoute de type événementiel: quel événement s'est-il produit et où? Comment le son a-t-il été produit (type d'excitation)? Quel type de structure a-t-il été mis en vibration (cordes, plaques, tuyaux)? (Castellengo 1994; 2002) L'identification de la source n'est pas seulement une étape dans la perception du timbre: elle est constitutive de la représentation cognitive qui correspond au concept de timbre. En outre, en l'absence de dénomination simple et partagée (consensuelle) du timbre (à la différence des dimensions de hauteur et d'intensité), c'est le nom de la source sonore qui va permettre de communiquer l'expérience subjective du timbre et de la constituer comme connaissance partagée (Dubois 2000; David 1997; Castellengo et Dubois, soumis; Cheminée 2006).

Les recherches linguistiques menées sur divers matériaux sonores ont en effet montré que le lexique utilisé peut renvoyer à la source, à la matière, au mode de mise en vibration, à l'intention, au système des sons et sa codification, etc., en fonction certes des propriétés du signal acoustique, mais également en fonction des connaissances acquises par les auditeurs lors d'écoutes précédentes et liées à leurs savoirs et savoir-faire (voir par exemple Dubois 2003; Morange 2003; Dubois, Guastavino et Raimbault 2005). Sur le plan acoustique, on a pu montrer que le rôle du transitoire dans le timbre causal est très important, en particulier pour discriminer les deux principales catégories de sons: les sons entretenus et les sons de type impulsif (Hajda, Kendall et Carterette 1997). Dans le cas d'un instrument de musique encore largement méconnu de la plupart des auditeurs européens, comme l'est le *steeldrum*, on fait donc l'hypothèse que la recherche de la cause de la production du son et donc l'identification de l'instrument joue un rôle décisif dans la qualification du timbre.

Nous utiliserons le concept de timbre causal comme représentation catégorielle de la qualification des sons de *steeldrums*, et les méthodes de recherche associées à cette problématique pour mettre en lumière le rôle et la pertinence des transitoires d'attaque dans la reconnaissance du timbre de ces instruments de musique.

## Méthode

La méthode de test utilisée est celle dite de la « catégorisation libre » associée au traitement de données en termes de similitudes représentées par des distances d'arbre (Barthélemy et Guénoche 1988; Sattah et Tversky 1977). Elle permet de repérer à la fois les classes (les catégories) et les critères de catégorisation mis en œuvre par les auditeurs. Ces tests de catégorisation possèdent les caractéristiques suivantes:

- le participant (pas davantage que l'expérimentateur) ne dispose pas d'une description physique exhaustive *a priori* des sons;
- la description linguistique des catégories n'est demandée qu'*a posteriori*;
- il n'y a pas de « bonne » réponse, ni même d'hypothèse sur les critères de classification et donc sur les réponses attendues (les critères de construction *a priori* du matériel fondés dans une description acoustique seule pouvant s'avérer non pertinents pour les participants (voir plus loin, « Stimuli »).

## PROCÉDURE

On présente sur l'écran d'un ordinateur une série d'icônes auxquelles sont associés des sons, à un participant à qui on demande de regrouper par classes les sons selon leur ressemblance, sans lui demander à cette étape d'expliquer les raisons ou critères de ses classements. Chaque participant peut ainsi écouter chaque son autant de fois qu'il le veut et dans l'ordre qu'il le souhaite, sans limite de temps. Les participants écoutent les sons à l'aide d'un casque audio. Il n'y a aucune consigne particulière concernant le nombre et la composition des classes. Les participants sont donc libres de choisir un ou plusieurs traits qui leur semblent pertinents pour former les catégories.

Après la tâche de catégorisation, il leur est demandé de formuler des commentaires sur chacune des catégories qu'ils ont formées. L'analyse de ces commentaires permettra d'expliquer les critères « subjectifs » de catégorisation et, pour certains d'entre eux, de repérer des corrélations avec les mesures physiques des stimuli.

## PARTICIPANTS

Les participants sont au nombre de 24, entre 20 et 25 ans, hommes et femmes, et représentent une population d'un niveau d'études supérieur.

## STIMULI

Le test que nous présentons ici a été conçu à l'issue d'un premier test, portant sur des sons de hauteur et d'intensité variables (Gaillard 2000). Pour cette deuxième série, nous avons retenu treize sons construits à partir d'un son original de steeldrum et de douze sons modifiés par analyse-synthèse à l'aide du logiciel de l'Ircam, AudioSculpt (Gaillard, Castellengo et Legros 2000). La méthode employée pour la construction du corpus de stimuli a été la suivante: chaque partiel d'un son de steeldrum original a d'abord été filtré; nous avons ainsi obtenu onze éléments sonores gardant chacun leur dynamique et leur intensité propre (différente pour chaque composante) et de même durée (6100 ms). Les stimuli ont ensuite été réalisés par synthèse additive des partiels préalablement séparés et dont on avait modifié certaines caractéristiques selon le protocole résumé dans le tableau 1.

**Tableau 1.** Modifications appliquées au son original de steeldrum.

Son	Modifications
1	Mixage des partiels 1, 2, 3 et 7
2	Mise en harmonicité du partiel 5 (+46 Hz)
3	Modification de l'échelle de temps * 2,0
4	Remplacement du partiel 4 par le partiel 4 du son (13), plus 10 dB
5	Atténuation régressive de l'énergie ( <i>FadeIn</i> ) de 121,9 ms à 203,2 ms
6	Atténuation régressive de l'énergie ( <i>FadeIn</i> ) de 121,9 ms à 412,2 ms
7	Accentuation du partiel 4 de 15 dB
8	Mixage des partiels 1, 3, 4 et 7
9	Son entièrement synthétique
10	Son original reconstitué
11	Remplacement du partiel 4 par le partiel 4 du son (13), et retard accentué de 200 ms
12	Mixage des partiels 1, 2, 3, 4 et 7
13	Mixage des partiels 1, 2, 3 et 4

Chacun des 13 sons a une durée égale de 269010 échantillons (6,1 secondes) à une fréquence d'échantillonnage de 44,1 kHz sur 16 bits.

L'objectif n'est pas d'établir un plan expérimental de construction systématique de stimuli, mais d'explorer, à partir des hypothèses sur les propriétés catégorielles, les propriétés pertinentes des éléments typiques et des éléments distants.

Nous avons porté notre attention sur les transitoires d'attaque, car une catégorisation préalable (Gaillard 2000) nous avait conduit à confirmer les résultats déjà observés (voir Grey 1977, par exemple), selon lesquels le transitoire d'attaque (et particulièrement la distance séparant l'apparition du début du son de celle du maximum du partiel 4) est un critère déterminant de catégorisation (en plus de la hauteur et de la composition spectrale, qui servent également de critères de catégorisation). L'hypothèse que nous pouvions formuler à la fin de cette première catégorisation consistait à supposer que les caractéristiques très particulières du transitoire des steeldrums (voir plus haut) sont responsables des « difficultés » à les reconnaître et les classer (dans la catégorie des instruments de percussion, par exemple). En modifiant ces caractéristiques, nous voulions observer les réactions des auditeurs à l'écoute d'un son très complexe, original et difficile à classer.

## Traitement des résultats

Les variables dépendantes sont de deux ordres: les partitions des auditeurs et leurs commentaires sur chacun des groupements qu'ils ont faits. Les premières données permettent de construire deux matrices de similitude. La première matrice permet d'identifier les distances entre participants. Cette première mesure de distance est évaluée à partir de la construction des matrices de chaque participant. À partir des matrices individuelles, nous élaborons la matrice de proximité des participants en estimant la distance entre deux participants (Barthélémy et Guénoche 1988).

Le calcul de la distance ( $d$ ) peut être formalisé de la manière suivante:

$$d_{ij} = \frac{r}{t} \times 100$$

avec  $r$  = nombre de réponses (couples de stimuli à l'intérieur de chaque catégorie) données par les participants et  $t$  = nombre total de paires de stimuli. Par exemple, si deux sujets  $i$  et  $j$  ont mis ensemble le même couple de stimuli, alors il n'y a aucune différence. Nous ne comptabilisons aucun « point ». Si par contre, le participant  $i$  a mis deux stimuli ensemble et que le participant  $j$  ne l'a pas fait, alors nous comptabilisons un point. Le nombre de divergences entre les deux participants  $i$  et  $j$  constitue ainsi la valeur  $r$  de la formule précédente, que nous mettons en relation avec le nombre total de divergences possibles qui est la valeur  $t$ .

Nous constituons alors une matrice des différences avec en abscisses et en ordonnées les numéros des participants. Les participants qui ont des valeurs de différences très fortes par rapport aux autres proposent non pas de mauvaises catégorisations, mais des catégorisations répondant à des stratégies différentes de l'ensemble des autres participants; le second calcul faisant intervenir des valeurs de classement interindividuelles, ces participants « hétérogènes » vis-à-vis de l'ensemble ont été écartés des calculs suivants.

Une fois cette opération faite, le deuxième calcul a pour objectif de mesurer la distance  $D$  entre les stimuli deux à deux (calcul fondamentalement identique au précédent mais en comparant les stimuli et non les participants). Nous pouvons formaliser le calcul de la manière suivante :

$$D_{ij} = 100 \times 1 - \left( \frac{\sum_{n=1}^N d_{ijn}}{N} \right)$$

où  $N$  est le nombre de participants,  $i$  et  $j$  les couples de sons, et  $D$  la mesure de la distance;  $i$  et  $j$  désignent à présent les stimuli, tandis que  $n$  représente le participant en cours de traitement. Les valeurs de  $D$  que nous obtenons sont en pourcentage: plus la valeur de  $D$  pour un couple de stimuli sera grande, plus le couple de stimuli aura été mis dans la même classe (la même catégorie). Pour représenter graphiquement ces valeurs, nous ne mettrons pas en avant une construction graphique paramétrique de type MDS (*Multi Dimensional Scaling*), qui nécessite de faire des hypothèses sur les paramètres à mettre en relation. La représentation en arbre de distances évite cette étape et propose une représentation adaptée à la théorie des prototypes. Ainsi, les nœuds « peuvent s'interpréter comme des entités qui partagent avec un groupe d'objets<sup>5</sup> les caractères communs aux membres de ce groupe » (Barthélemy 1993, 212).

Pour construire les arbres de proximité (visualisation de la distance entre les stimuli), il est nécessaire de calculer les « scores » de chaque stimulus. Le calcul des scores utilise, dans notre cas, l'algorithme de Sattah et Tversky, modifié par Barthélemy et Luong et implémenté dans le programme A.B.C.D. (Barthélemy et Luong 1987; Sattah et Tversky 1977). Les scores sont définis comme la « force de voisinage » de chaque paire de stimuli. Le calcul

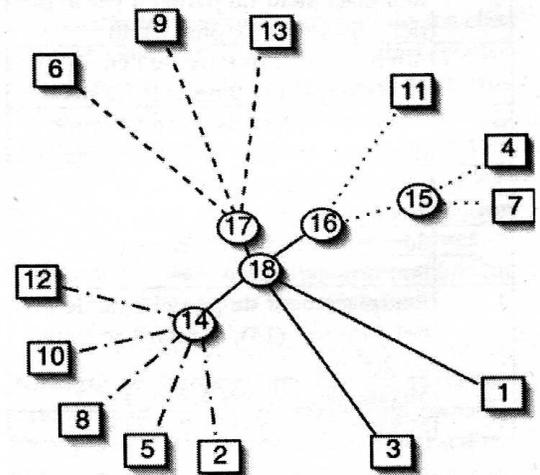
du score cherche également à extraire les regroupements entre stimuli, c'est-à-dire non plus seulement la relation entre deux stimuli, mais également la relation d'un stimulus avec l'ensemble des autres stimuli. La construction des arbres aura pour conséquence de donner une visualisation des groupements où chaque « feuille » (élément terminal d'une arête) représente un stimulus, et où chaque « nœud » (intersection entre deux arêtes) représente les prototypes hypothétiques qui rendent compte de la structure construite à partir des réponses des participants (voir figure 4).

Nous avons construit les arbres de proximité des stimuli. Le critère pertinent de la lecture et de l'interprétation est la longueur des arêtes, et non la relation dans un plan euclidien entre chaque stimulus. Ainsi, la distance entre les sons 4 et 7 (la force de groupement entre ces deux sons) se mesure par la distance de 4 à 15, puis de 15 à 7 (représentée par les lignes en pointillés).

## Résultats et discussion

Voici la représentation en arbre de proximité des stimuli traités dans ce test :

Figure 4. Arbre de proximité des stimuli construit à partir des résultats de la catégorisation des participants.



Le tableau suivant présente les catégories (inférées à partir de l'arbre ci-dessus) effectuées par les participants. En regard de chaque catégorie, nous avons précisé quelques exemples de commentaires formulés pour chacun d'eux, nous permettant ainsi une lecture plus aisée des données.

Du point de vue acoustique, on peut interpréter les regroupements formés par les participants selon au moins deux critères: les

<sup>5</sup> Par objet, on entend ici stimuli.

Tableau 2. Catégories et exemples de commentaires.

Nœud	Son	Exemples de commentaires	Nœud	Son	Exemples de commentaires
17	6	<i>Surtout sourd, plus éteint, attaque pas très nette</i>	14	2	<i>Son de cloche qui va en diminuant</i>
	9			5	
	13	<i>Même note vers le grave</i>		8	<i>Groupe moyen</i>
		<i>On entend moins l'harmonique au-dessus</i>			<i>Même son (!)</i>
		...		...	
16	4	<i>Plus d'aigus que les autres</i>		10	
	7			12	
	11		<i>Les harmoniques aiguës ressortent beaucoup</i>	18	1
		<i>Attaque normale, les harmoniques aiguës sont décalées dans le temps</i>			<i>On entend l'impact</i>
		...		<i>Caverneux comme une bouteille vide (1)</i>	
				3	

caractéristiques spectrales et la vitesse d'émergence du son.

Ces critères sont par ailleurs présents dans les commentaires des participants. En effet, dans un premier temps, une caractérisation de la « hauteur » des sons se retrouve dans les commentaires: « plus d'harmoniques graves », ou « grave, sourd, feutré », ou encore « caverneux comme une bouteille vide », commentaire attribué spécifiquement au son 1 dans lequel il manque le partiel 4. D'autres commentaires font état du caractère « grave » ou « aigu » des catégories que les participants ont formées. Il s'agit bien ici de la hauteur *spectrale* du son, liée aux composantes, et non de la hauteur *tonale* du son, les fréquences des partiels n'ayant pas été modifiées. Les sons qualifiés de « graves » et « sourds » (6, 9 et 13, autour du nœud 17) sont également ceux reconnus comme les moins percussifs.

Les deux paramètres considérés par les participants sont majoritairement le spectre et le transitoire. Le spectre est ici « observé » suivant sa distribution fréquentielle et temporelle. Il s'agit par exemple de la distinction entre les sons dont on a supprimé certains partiels (stimuli 1, mais aussi 8 et 12) et ceux dont on a modifié la distribution temporelle de ces partiels en modifiant de plusieurs manières le partiel 4 (stimuli 4, 7 et 11). Ces deux paramètres permettent d'expliquer les rapprochements:

1. Nœuds 16 et 15: les particularités temporelles du partiel 4; notons les distances d'arbres courtes (4-15, 7-15 et 11-16) attestant d'un rapprochement fort.

2. Nœud 14: richesse spectrale avec encore une fois des distances d'arbres courtes dont on peut inférer de grandes similitudes perçues entre les stimuli.

3. Nœuds 17 et 18: ces deux nœuds peuvent être confondus avec des stimuli présentant des particularités du tout début du transitoire d'attaque (à mettre en opposition avec les nœuds 16 et 15, centrés sur la partie post-transitoire du son, après 200 ms); notons que la diversité de ces sons est représentée par de grandes distances d'arbres (distances 6-17, 9-17, 13-17, 1-18 et 3-18).

La nature de ces deux paramètres va influencer sur la qualification du transitoire qui dépend doublement de la vitesse d'émergence globale du son et de la différence d'énergie entre le minimum et le maximum. Ces deux grands domaines interagissent à travers un paramètre temporel qui définit le moment au-delà duquel le participant considère l'apparition d'une nouvelle information comme ne faisant pas partie du début du son, et en deçà duquel cette même information est intégrée au transitoire. En effet, les événements acoustiques qui surviennent dans le son sont intégrés ou non au transitoire suivant le moment où ils se produisent. Dès lors, deux cas peuvent se présenter:

1. S'ils surviennent assez tôt, les événements acoustiques sont intégrés dans un « objet » acoustique perçu unique et indivisible, le transitoire d'attaque.
2. Si l'événement spectral arrive très tard (de 100 ms ou plus, c'est-à-dire bien au-delà du temps d'intégration de l'oreille)<sup>5</sup>, alors l'information est considérée comme faisant

5 Nous prenons ici la référence des critères d'intelligibilité comme le AI (Indice d'articulation) et le RASTI (Indice de transmission rapide de la parole). Ceux-ci proposent des temps d'intégration de l'oreille allant de 30 à 80 ms.

partie de l'« entretien » du son, engendrant ainsi une grande ambiguïté dans la qualification (percussif ou non) et la reconnaissance du son.

La prise en compte d'un événement acoustique comme faisant partie intégrante du transitoire d'attaque serait donc liée à l'ajustement d'une « fenêtre temporelle » perceptive distincte du temps d'intégration de l'oreille cité plus haut. Une telle fenêtre temporelle permettrait de discriminer ou non deux événements acoustiques en fonction de la consigne donnée au participant ou, dans notre cas, selon la consigne que le participant se donne à lui-même et par conséquent selon la tâche cognitive requise.

Cela nous permet de formuler l'hypothèse de l'existence d'un *seuil variable temporel de prise en compte du transitoire* (voir figure 5).

Cette interprétation s'inscrit dans la théorie de la catégorisation qui postule en particulier le principe d'économie cognitive. La taille de la fenêtre temporelle s'ajuste au niveau de discrimination recherchée, c'est-à-dire en fonction de la finesse d'analyse auditive requise au moment de l'audition d'un son, pour un type d'écoute, pour un auditeur donné. Cela peut avoir pour conséquence une perception du timbre variable selon la diversité des situations d'écoute. En effet, selon la taille de cette fenêtre temporelle, l'identification de la source vibrante, et donc de la causalité du timbre, par exemple, sera plus ou moins facilitée.

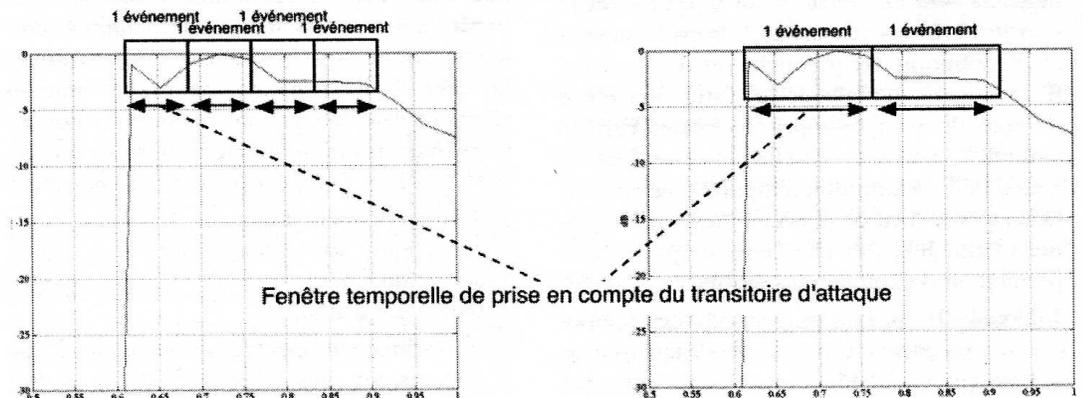
Cette interprétation s'accorde également avec le principe d'efficacité cognitive, qui per-

met une réponse rapide à un stimulus donné par une perception sélective et une mise en relation globale avec les éléments caractéristiques du prototype (comme représentation préalable en mémoire). Ainsi, en « ajustant » la taille de la fenêtre temporelle dont nous supposons l'existence, on peut faire l'hypothèse que les auditeurs sélectionnent et identifient rapidement les seuls indices pertinents nécessaires à une situation donnée et à une catégorisation préalablement formée par leur expérience auditive.

La présence d'un tel *seuil variable* en fonction à la fois des capacités discriminatives et sélectives peut avoir de nombreuses répercussions dans l'évaluation du timbre et de ses qualités. Nous pourrions supposer, par exemple, qu'un musicien, lors de son apprentissage, est amené à « régler » cette fenêtre au plus juste afin de lui permettre de discriminer de façon efficace et rapide soit le type d'un instrument (flûte, violon, etc.) soit la qualité de celui-ci à un moment donné (les différences entre un « bon » et un « mauvais » piano, par exemple). Nous pouvons également supposer que l'autocontrôle auditif effectué par un musicien en train de jouer utilise largement un tel processus, notamment pour l'évaluation de la qualité des détachés.

S'il s'agit de mettre en avant les caractéristiques acoustiques qui permettent d'identifier la source d'un son autant que l'évaluation de la qualité d'un timbre musical (le timbre causal), nous pouvons nous interroger sur la pertinence de chercher *a priori* dans l'information acoustique des paramètres que la perception prendra ou non en compte en fonction de la

**Figure 5.** Illustration de l'hypothèse du seuil variable temporel de prise en compte du transitoire: chaque rectangle en surimposition de cette courbe d'intensité en fonction du temps figure une taille (en temps) de fenêtre, correspondant au seuil temporel de prise en compte du transitoire; à chaque fenêtre correspond un événement acoustique pertinent pour l'analyse auditive du transitoire d'attaque.



situation et des besoins. La mise en lumière de particularités temporelles particulières du son de steeldrum nous a permis de mettre en avant la *variabilité* dans l'approche des auditeurs confrontés à ce type de son. Ce phénomène explique pourquoi certains auditeurs considéraient ce son comme particulièrement percussif, alors que d'autres retenaient plus facilement l'aspect ambigu (presque pas percussif). Cette variabilité dans la perception auditive existe, semble-t-il, également dans d'autres domaines auditifs, comme la phonétique, où nous avons noté que des auditeurs pouvaient avoir des capacités à discriminer dans une tâche auditive des sons d'une langue inconnue qu'ils ne parvenaient pas à entendre dans une tâche sémantique (Magnen, Billières et Gaillard 2005).

Cette variabilité de l'auditeur est peut-être responsable des difficultés que nous avons à définir le timbre à travers des descripteurs toujours plus nombreux et complexes. ◀

## RÉFÉRENCES

- ACHONG, Anthony (1996). « Vibrational Analysis of Mass Loaded Plates and Shallow Shells by the Receptance Method With Application to the Steelpan », *Journal of Sound and Vibration*, vol. 191, n° 2, p. 207-217.
- \_\_\_\_\_ (1999). « Non-Linear Analysis of Compressively/Thermally Stressed Elastic Shell Structures on the Steelpan and Underlying Theory of the Tuning Process », *Journal of Sound and Vibration*, vol. 222, n° 4, p. 597-620.
- \_\_\_\_\_ et K.A. SINANAN-SINGH (1997). « The Steelpan as a System of Non-Linear Mode-Localized Oscillators, Part II: Coupled Sub-Systems, Simulations and Experiments », *Journal of Sound and Vibration*, vol. 203, n° 4, p. 547-561.
- ALEXIS, Clifford, Al O'CONNOR et Thomas D. ROSSING (1986). « Tonal Design and Tuning of Caribbean Steel Drums », *Journal of the Acoustical Society of America*, vol. 80, n° 1, p. 102.
- AMERICAN NATIONAL STANDARDS INSTITUTE (1973). « Timbre », *Psychoacoustical Terminology*, ANSI S3.20-1973, p. 56.
- BARTHÉLEMY, Jean-Pierre (1993). « Similitude, arbres et typicalité », Danièle DUBOIS (dir.), *Sémantique et cognition: Catégories, prototypes, typicalité*, Paris, CNRS, p. 205-224.
- \_\_\_\_\_ et Alain GUÉNOCHE (1988). *Les arbres et les représentations de proximité*, Paris, Masson.
- \_\_\_\_\_ et Nguyen Xuan LUONG (1987). « Sur la topologie d'un arbre phylogénétique: Aspects théoriques, algorithmiques et applications à l'analyse de données textuelles », *Mathématiques et sciences humaines*, n° 100, p. 57-80.
- CADOZ, Claude (1991). « Timbre et causalité », Jean-Baptiste BARRIÈRE (dir.), *Le timbre, métaphore pour la composition*, Paris, IRCAM/Christian Bourgois Éditeur, p. 17-46.
- CASTELLENGO, Michèle (1994). « La perception auditive des sons musicaux », Arlette ZENATTI (dir.), *Psychologie de la musique*, Paris, Presses universitaires de France, p. 55-86.
- \_\_\_\_\_ (2002). « Les sources acoustiques », Denis MERCIER (dir.), *Le livre des techniques du son*, 3<sup>e</sup> édition, Paris, Dunod, vol. 1, p. 48-50.
- \_\_\_\_\_ et Danièle DUBOIS (soumis). « Sémantique des objets sonores: Timbres et musiques », *intellectica*.
- CHEMINÉE, Pascale (2006). « Vous avez dit "Clair"? Le lexique des pianistes entre sens commun et terminologie », *Cahiers du LCPE*, n° 7, « Dénomination, désignation et catégories », p. 49-65.
- DAVID, Sophie (1997). « Représentations sensorielles et marques de la personne: Contrastes entre olfaction et audition », Danièle DUBOIS (dir.), *Catégorisation et cognition: De la perception au discours*, Paris, Kimé, p. 211-242.
- DOURNON, Geneviève (1992). « Organology », Helen MYERS (dir.), *Ethnomusicology, an Introduction*, Londres, MacMillan Press, vol. 1, p. 245-300.
- DUBOIS, Danièle, dir. (1993). *Sémantique et cognition: Catégories, prototypes, typicalité*, Paris, CNRS.
- \_\_\_\_\_ (2000). « Categories as Acts of Meaning: The Case of Categories in Olfaction and Audition », *Cognitive Science Quarterly*, vol. 1, p. 35-68.
- \_\_\_\_\_ (2003). *Perception, Representation and Knowledge: Acoustic Phenomena Between Noise and Sounds*, EAA Symposium on Architectural and Urban Acoustics and Musical Acoustics, Bilbao.

\_\_\_\_\_, Catherine GUASTAVINO et Manon RAIMBAULT (2005). « Les catégories cognitives du bruit urbain : Des discours aux indicateurs physiques », *Acoustique et technique*, n° 39, p. 49-57.

DUMORTIER, Jérôme (2004). « Mise au point d'une méthode d'étude de caractérisation de la qualité sonore des bouches de ventilation », mémoire, Ingénieur CNAM, Conservatoire national des arts et métiers, Toulouse.

GAILLARD, Pascal (2000). « Étude de la perception des transitoires d'attaque des sons de steeldrums : Particularités acoustiques, transformation par synthèse et catégorisation », thèse de doctorat, Musique, acoustique musicale et perception, Université de Toulouse II - Le Mirail, Toulouse.

\_\_\_\_\_, Michèle CASTELLENGO et Claude LEGROS (2000). « Modification de certaines caractéristiques physiques des sons de steeldrums en vue de la réalisation de tests de psychoacoustique », communication, cinquième Congrès français d'acoustique, 3 au 6 septembre 2000, École polytechnique fédérale de Lausanne (Suisse).

GREY, John M. (1977). « Multidimensional perceptual scaling of musical timbres », *Journal of the Acoustical Society of America*, vol. 61, n° 5, p. 1270-1277.

GUYOT, Frédérique (1996). « Étude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : Une approche par la catégorisation », thèse de doctorat, psychoacoustique, Université du Maine, Paris.

\_\_\_\_\_, Michèle CASTELLENGO et Benoît FABRE (1997). « Étude de la catégorisation d'un corpus de bruits domestiques », Danièle DUBOIS (dir.), *Catégorisation et cognition : De la perception au discours*, Paris, Kimé, p. 41-58.

HAJDA, John M., Roger A. KENDALL, Edward C. CARTERETTE et Michael L. HARSHBERGER (1997). « Methodological Issues in Timbre Research », Irène DELIÈGE et John A. SLOBODA (dir.), *Perception and Cognition of Music*, Hove (UK), Psychology Press, p. 253-306.

HANDEL, Stephen (1989). *Listening: An introduction to the Perception of Auditory Events*, Cambridge (MA), MIT Press.

HANSEN, Uwe J. et Thomas D. ROSSING, (1987). « Modal Analysis of a Caribbean Steel Drum », *Journal of the Acoustical Society of America*, vol. 82, n° 1, p. 68.

KRONMAN, Ulf (1992). *Steel Pan Tuning: A Handbook for Steel Pan Making and Tuning*, Stockholm, Musikmuseet.

MAFFIOLO, Valérie (1999). « Caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain », thèse de doctorat, acoustique, Université du Maine, Paris.

MAGNEN, Cynthia, Michel BILLIÈRES et Pascal GAILLARD (2005). « Surdité phonologique et catégorisation : Perception des voyelles françaises par les hispanophones », *Revue PArole*, vol. 33, p. 9-34.

MORANGE, Séverine (2003). « Approches structurale, prosodique, psycho-cognitive de quelques propriétés linguistiques subjectives de la parole : L'exemple de trois Montmartrois âgés », thèse de doctorat, Université de Paris III - Sorbonne nouvelle, Paris.

\_\_\_\_\_, (2005). « Appartenances catégorielles d'un objet complexe, la voix : Approches linguistique et psycho-cognitive », communication au Sensolier, Troisième journée, Ivry - CNRS.

ROSCHE, Elen (1978). « Principles of Categorisation », Eleanor Rosch et Barbara B. LLYOD (dir.), *Cognition and Categorization*, Hillsdale (NJ), Lawrence Erlbaum Associates, p. 27-48.

ROSSING, Thomas D. (1982). « Chimes and Bells », *Percussive Notes*, vol. 19, n° 3, p. 42-57.

\_\_\_\_\_, D. Scott HAMPTON et Joshua BOVERMAN (1986). « Acoustics of Caribbean Steel Drums », *Journal of the Acoustical Society of America*, vol. 80, n° 1, p. 102.

ROUBY, Catherine, Benoist SCHAAL, Danièle DUBOIS, Rémi GERVAIS et A. HOLLEY (2002). *Olfaction, Taste and Cognition*, Cambridge/New York, Cambridge University Press.

SATTAH, Shmuel et Amos TVERSKY (1977). « Additive Similarity Trees », *Psychometrika*, vol. 42, n° 3, p. 19-345.

ZWICKER, Eberhard et Hugo FASTL (1999). *Psycho-acoustics, Facts and Models*, Berlin/New York, Springer-Verlag.