



HAL
open science

Analyse des différents stimuli auditifs : musique, langage et bruit. Etude comparative

Christophe d'Alessandro

► **To cite this version:**

Christophe d'Alessandro. Analyse des différents stimuli auditifs : musique, langage et bruit. Etude comparative. Bernard Lechevalier; Hervé Platel; Francis Eustache. Le cerveau musicien. Neuropsychologie et psychologie cognitive de la perception musicale - Chapitre 2, De Boeck Supérieur, pp.33-46, 2006, Questions de personne, 9782804162801. 10.3917/dbu.leche.2006.01.0033 . hal-01712523

HAL Id: hal-01712523

<https://hal.science/hal-01712523>

Submitted on 5 Dec 2021

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.

Analyse des différents stimuli auditifs : musique, langage et bruit. Étude comparative.

Christophe d'Alessandro
LIMSI-CNRS
BP 133 - F91403 Orsay
Christophe.D'Alessandro@limsi.fr

Introduction

L'exploration fonctionnelle de la perception des sons indique une activation d'aires corticales différentes pour des stimuli auditifs de catégories différentes : parole, musique et bruit. Il faudrait donc comprendre dans quelle mesure cette différenciation serait intrinsèque, c'est à dire déjà présente dans les propriétés acoustiques du signal, et dans quelle mesure elle résulterait de processus cognitifs, c'est à dire serait construite activement par le sujet. Tous les sons auxquels l'être humain est exposé quotidiennement ne sont pas porteurs de sens. Et le sens de ceux qui en portent peut-être lui-même de complexité variable, depuis le signal d'avertissement jusqu'au concept complexe. Une classification simpliste des sons en termes de « parole », « musique » et « bruits » n'est pas sans poser de nombreux problèmes. Ainsi avant de passer à la comparaison des aspects physiques des stimuli auditifs, il est nécessaire d'aborder le sonore dans sa dimension fonctionnelle. Traditionnellement, l'aspect fonctionnel des sons de la parole est l'objet d'étude de la phonologie. C'est donc sous le chapitre de la phonologie que nous discuterons des différences fonctionnelles entre parole musique et bruit. L'aspect physique, acoustique en l'occurrence, des sons sera ensuite abordé sous le chapitre de la phonétique, puisque traditionnellement la phonétique est la science des sons de la langue (ici étendu aussi au langage des sons qu'est la musique), sous l'aspect de la production, de la perception, et d'une façon générale sous l'aspect acoustique.

Aspects phonologiques

Parole, musique et bruit

Avant d'envisager les différences entre parole, musique et bruit, il faut d'abord tenter de définir ces 3 mots.

- Parole. Il s'agit de la manifestation sonore du langage. En tant que telle, la parole est un objet linguistique, c'est à dire déterminé par la langue. Son but est avant tout communicatif, la langue doit permettre de représenter et d'échanger des concepts, du sens. La pensée doit clairement pouvoir utiliser l'outil linguistique. La langue est un système sémiotique, c'est à dire un système de signifiants et de signifiés, qui possède une double articulation. La première articulation est celle des mots, unités minimales ou atomiques du sens. Deux mots vont permettre de distinguer deux nuances de sens. La seconde articulation est celle des phonèmes, unités minimales ou atomiques du son. Deux phonèmes vont permettre de distinguer deux mots (par exemple [pas] et [bas]). Au-delà du contenu conceptuel (mots), la prosodie (mélodie, rythme) de la parole précise et nuance le discours.
- Musique. La musique est l'art de combiner les sons. Il s'agit donc encore d'un système sémiotique, système de signifiants que l'on combine pour constituer un signifié. Contrairement à la parole, système à double articulation ou le signifiant (le son) est lui-même organisé en système (les phonèmes), les signifiants de la musique sont a priori quelconques. Berlioz écrivait, il y a plus d'un siècle et demi que « tout corps sonore mis

en œuvre par le compositeur est un instrument de musique (Traité d'instrumentation et d'orchestration) ». Ainsi, en toute généralité la musique est un système sémiotique pour lequel tout objet sonore peut servir de signifiant. Même s'il n'est pas douteux qu'il y ait une sorte de signifié, celui-ci n'est pas de nature conceptuelle ou linguistique, mais de nature esthétique. Le but est en effet un but expressif, expression de la « pensée musicale ». Le système sémiotique n'a qu'une seule articulation, comme le discute Nattiez (1975, pp. 194-217), dans le cadre de la musique classique, il s'agit de l'articulation du texte musical en « notes », un mot qui fait référence explicite à l'écriture, la notation. D'une façon générale, la musique est un discours que l'on peut décomposer en unités minimales, même s'il ne s'agit pas toujours de notes, au sens d'une représentation exacte des valeurs mélodiques, rythmiques, ou de l'intensité. Certains ont tenté de trouver aussi une double articulation en musique (par exemple Levi-Strauss, 1963), mais il semble plus justifié de ne distinguer qu'une seule articulation, sans chercher de "mots" musicaux, mais directement des "phrases" à partir des "notes".

- Bruits. La meilleure définition du ou plutôt des bruits est peut-être la plus négative. Les bruits sont les sons qui ne ressortissent ni de la catégorie de la parole, ni de celle de la musique. Se sont donc les sons de l'environnement, naturels ou humains, qui n'ont pas de but communicatif ou expressif. Même si on considère les bruits qui ont un sens (bruits d'avertissement, signaux), il n'y a pas de système sémiotique, car pas de combinaison entre signifiant pour créer de nouveaux signifiés : à un signifiant correspond de façon biunivoque un signifié, il y a *signal* et non *signe*.

Compétence et attention

Ainsi, il devient clair que la nature d'un son, et a fortiori sont contenu acoustique, ne suffisent pas, à eux seuls, pour aborder sa perception. Le sujet est partie prenante et active du processus de perception, c'est un processus cognitif. Pour la parole, en plus du signal sonore il faut nécessairement une connaissance de la langue, donc au minimum des phonèmes et des mots, afin de communiquer des concepts. Pour la musique, en plus du signal sonore il faut nécessairement un langage musical, en général sous forme d'éléments repérables par une théorie (notes, échelles, modes), afin de former des phrases et d'exprimer des idées musicales. Dans le cas des bruits, le signal sonore fonctionne vraiment comme un signal : il indique la provenance ou la nature du bruit, mais sans communication ni expression véritable, il n'y a pas de niveaux intermédiaires entre son et interprétation du son. Lorsque les bruits deviennent complexes et porteur de sens, même de façon réduite, on parlera volontiers de « musique » ou de « langage » : ainsi du mécanicien qui écoute la « musique » du moteur de camion pour en détecter la panne, ou de l'ornithologue qui repère les oiseaux à leur « langage » (encore que dans ce dernier cas il y puisse y avoir effectivement une organisation du son dans un but communicatif).

Puisque le son ne peut être parole que par la médiation de la langue, les sons d'une langue étrangère ne sont pas parole pour le sujet qui ne la comprend pas, ils ne sont que bruit. Sans intention et compétence linguistique, il n'y a pas parole. Un premier élément qui transforme le son en parole est la compétence linguistique. De plus, le sujet récepteur de la parole doit y prêter attention pour qu'elle fonctionne comme parole : le voisin qui bavarde produit un bruit pour le sujet qui n'est pas concerné, même (surtout) s'il possède la compétence linguistique. Un second élément qui transforme le son en parole est l'attention de l'auditeur.

La musique n'est musique que si d'une part elle est intelligible pour le sujet, et si le sujet veut écouter ce son particulier comme musique. Pour certain, le chant du vent ou du silence sera musique, et non bruit. Pour d'autres, une musique savamment élaborée, restera inintelligible et ne sera que bruit. Ainsi, comme pour la parole ce n'est pas le signal qui détermine sa perception, mais bien une relation entre le signal et le sujet : à la fois compétence et attention.

Enfin les bruits ne sont tels qu'en tant que sons dépourvus de contenu intelligible, ou avec un contenu informatif très pauvre.

Donc, a priori, le projet de caractériser les différents stimuli auditifs sur la base de propriétés acoustiques seules est voué à l'échec : il faudrait considérer le stimulus lui-même, et aussi le sujet impliqué dans la perception et la cognition de ce stimulus. Si on étudie la parole, c'est celle de la compétence linguistique du sujet, si on étudie la musique, c'est celle dont le langage est connu du sujet, même s'il n'est pas un musicien "actif", capable de jouer ou d'aborder théoriquement cette musique. Dans les deux cas, il faut en plus de la compétence, que le sujet coopère avec le signal en lui prêtant attention. Signal acoustique, compétence linguistique ou musicale, attention, voici les conditions que l'on va supposer réunies par la suite.

Modalités d'écoute

Ce phénomène attentionnel dans l'écoute des sons peut se subdiviser encore, entre une écoute analytique et une écoute holistique, ou globale. Dans un même extrait musical, on peut se concentrer sur les hauteurs et le rythme (par exemple prendre une dictée musicale) ou bien sur l'interprétation, ou bien sur le timbre (c'est l'écoute par exemple d'un orgue par le facteur d'orgue), ou bien écouter globalement "la musique". De même l'écoute de la parole peut se focaliser sur l'orthographe (dictée), sur le sens (analyse d'un discours), voire sur la qualité de voix de la personne qui parle.

Pour un même stimulus sonore, l'auditeur peut à volonté, après un entraînement éventuel utiliser différents modes d'écoute. Par exemple, pour une phrase de parole, l'auditeur en mode « parole » peut en noter les mots et les phonèmes, ou bien l'intonation, sous forme de ton hauts et bas, d'accents. En mode « musique », il peut en noter les hauteurs, sur chaque syllabe, en termes d'intervalles musicaux, ainsi que le rythme, sur une grille métrique. C'est un exercice qui demande un certain entraînement, tout comme la dictée musicale. Dans ce cas, ce n'est plus le sens de la phrase ou ses nuances qui sont perçues, mais la mélodie au sens musical du terme.

La modalité d'écoute, contrairement à la compétence, et dans une certaine mesure à l'attention, est sous le contrôle conscient du sujet percevant. Cela implique qu'en fonction de directives et d'un acte conscient, le même stimulus doit pouvoir susciter des tâches cognitives différentes, et sans doute mobiliser des aires corticales différentes.

En plus des modalités d'écoute, il existe des stimuli auditifs composés intrinsèquement de plusieurs modalités sonores (parole, musique, bruit). Par exemple le chant peut (s'il est intelligible) combiner parole et musique. Suivant les situations, c'est la musique qui dominera (dans le contexte par exemple du chant lyrique) aux dépens de la parole, ou bien la parole qui dominera (dans le contexte par exemple de la psalmodie liturgique) aux dépens de la musique. Enfin, la parole et la musique ne sont jamais (ou presque) entendues sans que des bruits n'entrent dans la composition du stimulus auditif. Dans ce cas, il est probable que la perception sera elle-même composée (on peut à volonté porter son attention sur la parole, ou sur le bruit). Dans le chant, intelligible et en présence de bruit, les trois modalités sont bien sur mélangées.

Phonèmes

Du point de vue fonctionnel, il est classique de considérer un double niveau d'articulation pour la description des langues, celui des phonèmes, ou plus petites unités distinctives de son, et celui des monèmes, ou plus petites unités distinctives de sens.

Ainsi chaque langue retient un ensemble de sons, parmi ceux que pourrait produire l'appareil vocal. Il est intéressant de considérer la nature des sons que la langue utilise. Les phonèmes sont des objets linguistiques : l'inventaire des phonèmes d'une langue s'appuie sur des critères

distinctifs. Le phonème est la plus petite unité sonore qui, substituée à une autre, change le contenu linguistique d'un mot. Par exemple changer le premier son de /p/ de "peau" (/po/) en /b/ aboutit à un mot différent : "beau" (/bo/). On distingue donc les phonèmes /p/ et /b/. Il est remarquable que dans toutes les langues, les systèmes de phonèmes soient toujours bâtis sur des distinctions du contenu spectral des sons : voyelles, consonne correspondent à des variations de timbre. L'aspect mélodique et rythmique de la parole n'intervient que de façon marginale dans les systèmes phonologique. Ces aspects interviennent dans certaines langues seulement, et échappent en majorité à la notation par l'écriture : ils sont du domaine de l'interprétation, de la pragmatique plus que du système. On peut reconstituer les phonèmes des langues disparues, mais pas leurs mélodies ou leurs rythmes.

Solfège

D'après Jacobson (cité par Nattiez, 1975 p 198), aux structures phonologiques de la parole correspondent les gammes musicales : "une intervention de la culture dans la nature, un artifice qui impose des règles logiques au continuum sonore". L'aspect fonctionnel de la musique, n'a pas fait pas l'objet d'une attention théorique identique pour toutes les musiques. Pour les musiques classiques, qui ont une théorie bien établie (musique occidentale, musique indienne, musique arabe etc.), apparaît un système musical, avec souvent une notation et donc un solfège.

Contrairement au cas de la parole, le solfège fait presque entièrement abstraction du timbre instrumental, pour se concentrer sur les aspects mélodiques et rythmiques. On note les hauteurs mélodiques, les valeurs rythmiques, on discute des modes et échelles mélodiques, des rythmes simples et composés.

Ainsi la musique est en premier lieu du domaine de la mélodie et du rythme, et marginalement de celui du timbre. C'est à dire que le timbre peut échapper totalement à la notation.

Aspects phonétiques

Les quatre principaux attributs du signal sonore, dans la dimension perceptive, sont la hauteur mélodique, la durée, l'intensité et le timbre. De ces attributs, la hauteur, la durée et l'intensité correspondent à des stades primaires de la communication sonore : ce sont les premières dimensions perçues et produites dans l'ontogenèse, et sans doute dans la communication animale. Domaine de la pensée non conceptuelle, ces dimensions du sonore sont aussi celles qui véhiculent les contenus les plus subtils, contenus esthétiques de la musique, et contenus pragmatiques, finesses de sens et d'attitudes de la parole, d'une façon générale la dimension du jeu avec le sonore.

Les variations significatives, les contrastes de timbre, sont réservés aux aspects conceptuels de la structuration sonore, l'activité linguistique.

Parmi les trois types de stimuli, la parole est le plus structuré, celui dont le système sémiotique est le plus complexe, celui qui peut exprimer les concepts les plus précis avec le plus de finesse et de contraste. C'est aussi celui qui est le plus contraint du point de vue sonore : l'appareil vocal (ou ses avatars électroniques) est le seul instrument en mesure de produire systématiquement de grandes variations spectrales, de la parole. Par contre, tout signal acoustique peut a priori devenir musique. Il y a certes des instruments musicaux très élaborés (mais aucun aussi élaboré que ne l'est la voix humaine), mais la musique est étonnement robuste au changement instrumental. Contrairement à la parole, ce n'est pas un dispositif unique qui est en mesure de produire de la musique, puisque c'est l'organisation des sons qui définit la musique. Enfin, le bruit est la classe de stimuli sonore la plus générale, puisque toutes les sources de son peuvent produire un ou du bruit.

L'appareil vocal

Les propriétés acoustiques de l'appareil vocal, instrument de la parole, sont exploitées par le langage pour donner leur substance aux phonèmes. Le but linguistique est de produire quelques dizaines d'objets sonores contrastés, aptes à être organisés en système. C'est par des variations de position de l'appareil vocal, par des gestes, que ces sons vont être produits. On peut décomposer l'appareil phonatoire en trois sous-ensembles fonctionnels :

1. Les poumons et la trachée-artère : La source d'énergie nécessaire pour produire de la parole réside dans les muscles abdominaux et thoraciques. Les poumons comprimés par cette musculature, agissent de façon analogue à un soufflet et fournissent un débit d'air.
2. Ce débit d'air est modulé par les cordes vocales, bandes musculaires dans le larynx, entre la trachée-artère et la cavité pharyngée. La glotte est l'ouverture entre les cordes vocales. Les cordes vocales peuvent produire des impulsions isolées et des sons périodiques sur une étendue d'environ 3 octaves.
3. Le débit d'air modulé par les cordes vocales est modifié (filtré) par conduit vocal. Le conduit vocal est une cavité à géométrie variable, qui se subdivise en pharynx, cavité orale, cavité nasale. La conformation du conduit vocal dépend des organes articulateurs : mâchoires, lèvres, langue.

La production des sons par l'appareil vocal peut être décrite par un modèle source/filtre. Dans ce modèle, le son est produit par le filtrage fréquentiel d'une source sonore. La source est associée à la phonation, production initiale du son, et le filtre à l'articulation, c'est à dire à la position du conduit vocal. Le conduit vocal assure une fonction de filtrage acoustique des signaux de source, le filtre mis en jeu évoluant avec la conformation des articulateurs. La cavité formée par le conduit vocal possède des résonances, les formants, et des antirésonances ou antiformants, lors de l'utilisation de la dérivation nasale. La position des organes articulateurs peut évoluer rapidement, et changer ainsi la géométrie du conduit vocal, et donc ses propriétés acoustiques. En résumé, les résonances acoustiques du conduit vocal correspondent aux formants spectraux de la parole. Au contraire la partie source contrôle la fréquence fondamentale, la durée des sons et leurs intensités.

Les sons de la parole

Ces variations permettent de construire le système de phonèmes. La première subdivision des phonèmes est liée au mode d'excitation et à la stabilité du conduit vocal : c'est la séparation voyelles/consonnes. Les voyelles correspondent à une vibration quasi-périodique des cordes vocales et à une position stable du conduit vocal.

On distingue les différentes voyelles grâce à l'emplacement des premiers formants spectraux. Les deux premiers formants suffisent à définir les voyelles du français. Si la dérivation nasale est ouverte (grâce à l'abaissement du voile du palais), les voyelles sont nasalisées. Les voyelles nasales se distinguent acoustiquement des voyelles orales par les anti-résonances introduites par les fosses nasales.

Les consonnes se répartissent en plusieurs classes : semi-voyelles, fricatives (sourdes et sonores), plosives (sourdes et sonores), nasales, liquides. Les semi-voyelles des consonnes produites par une évolution rapide du conduit vocal entre deux voyelles. Les fricatives sont produites par une constriction sans obstruction du conduit vocal. L'excitation est un bruit de friction, le conduit vocal est ainsi perturbé en amont et en aval de la constriction, en première approximation seule la partie aval contribue à la "coloration" du signal de source. Les cordes vocales peuvent entrer en vibration en même temps que le bruit de friction, la fricative étant alors voisée, ou laisser passer l'air sans émettre de son, la fricative étant alors sourde. Les plosives résultent d'une obstruction complète du conduit vocal, puis de son relâchement (ou explosion, d'où le terme de plosives). S'il y a vibration des cordes vocales pendant l'occlusion, l'occlusive (ou plosive) est voisée, sinon elle est sourde. Les consonnes nasales sont produites

par vibration des cordes vocales, occlusion dans le conduit oral, et abaissement du voile du palais, qui entraîne l'apparition d'anti-résonances dues aux fosses nasales. L'occlusion de la cavité buccale étant complète, le son est rayonné par les narines. Les liquides, assez semblables aux semi-voyelles, résultent d'une excitation voisée et de rapides mouvements articulatoires, principalement de la langue.

Prosodie

En quelque sorte superposée au flot de phonèmes, la prosodie est souvent présentée comme la musique de la parole. En effet elle ne sert pas (du moins en Français) à construire les mots. Elle ponctue, rythme et module le discours en distribuant accentuation et intonation. L'intonation est la mélodie, variation de hauteur fondamentale dans le temps. En parole elle assume une fonction distinctive (l'intonation distingue un énoncé interrogatif d'un énoncé affirmatif, par exemple). L'accent, qui met en valeur des segments de parole de durée supérieure au phonème (syllabe, groupe de syllabes, mots), assume des fonctions contrastive, syntaxique et expressive. Les paramètres acoustiques mis en jeu dans la prosodie sont des variations de la fréquence de vibration des cordes vocales, des variations d'intensité, et des variations de durée phonémiques ce qui donne le rythme d'élocution ;

En plus des variations significatives d'intonation, produites volontairement, une micro-mélodie est générée par l'enchaînement des phonèmes, indépendamment de toute action volontaire par le simple jeu de la réaction du conduit vocal sur la source laryngée dont la périodicité est ainsi légèrement modifiée.

Cette micro-prosodie perturbe les courbes de fréquence fondamentale, qui dans la parole ne présentent pas de paliers stables. Ainsi la prosodie de la parole n'évoque pas de « notes » musicales. Dans la parole chantée, au contraire, le chanteur contrôle sa fréquence fondamentale pour produire des notes clairement identifiables. Alors, la prosodie n'est plus de type « parole », mais de type « chant », est la parole devient musique.

Les sons musicaux

Alors que la parole est un jeu sur les timbres, le jeu musical va porter essentiellement sur les dimensions de hauteur, durée et intensité, qui sont les seules notées dans une partition. Le timbre est défini, en quelque sorte par défaut, par l'instrument : timbre du violon ou de la clarinette, ou par l'instrumentation (la combinaison des instruments).

Bien sur l'invariance de timbre des instruments musicaux est relative. Cependant, pour que l'attention porte surtout sur les aspects mélodiques, il faut une variation relativement pauvre de la dimension du timbre, pauvre en regard de la parole.

Les notes, éléments atomiques du son musical sont avant tout des éléments stables. Alors que dans la parole, l'information est portée surtout par les sons transitoires, les consonnes (certaines langues peuvent ne noter que les consonnes dans l'écriture), dans la musique il y a recherche de la stabilité (le chant par exemple privilégiera les voyelles, sous forme de vocalises). Fonagy (1983, p. 310) définit ainsi une mesure de "musicalité" de la voix, comme la régularité, la stabilité de la courbe mélodique sur une syllabe. La voix humaine est perçue comme du chant au delà d'un certain degré de cette "musicalité".

La production de beaucoup de sons instrumentaux, tout comme la production de la voix, montre des sources sonores relativement localisées, qui excitent des corps résonants. Par exemple on peut citer l'anche d'une clarinette et le corps de l'instrument, la corde du violon et la caisse etc. Cette relative indépendance de la source d'énergie sonore et de sa transformation permet d'appliquer le modèle source/filtre aussi aux instruments.

Le filtre résonateur est souvent une cavité acoustique (instruments à vents), ou un dispositif mécanico-acoustique (caisse de résonance des instruments à corde). Sa fonction est de transformer le signal de source. Certains instruments possèdent, comme la voix, des formants.

Cependant, contrairement à la parole, le filtre (donc les formants) des instruments musicaux ne change pas rapidement dans le temps. Si on imagine par exemple une clarinette, pour une note donnée, le filtre (le corps de l'instrument) ne change pas. Si on change de note, on change la longueur vibrante, donc surtout la fréquence fondamentale, et beaucoup moins le timbre typique de l'instrument. Il existe cependant des instruments pour lesquels le jeu musical est obtenu par variation du spectre, comme la guimbarde, mais dans ce cas les variations portent sur un seul formant. Cela correspond à des possibilités de variations spectrales qui sont très limitées comparé à la parole. Un instrument de musique ne parle pas.

Les bruits

Comme nous l'avons vu, la définition du "bruit" ne se trouve pas que dans les propriétés du signal, mais plutôt dans son interprétation. En traitement du signal, la notion de bruit est liée à celle de signal aléatoire. Un bruit est considéré comme la réalisation d'un processus aléatoire. Mais la parole, par exemple, utilise certains phonèmes qui sont des physiquement des bruits (fricatives sourdes par exemple), tout en étant des sons de parole réalisés, perçus et interprétés comme tels. De même certains sons musicaux sont physiquement des bruits, comme les percussions sans hauteur fixe.

Conclusion

Des considérations précédentes ont peut tirer quelques conclusions sur les différences entre parole, musique, et bruit, au niveau fonctionnel, au niveau du signal acoustique, et sans doute au niveau des stimuli transmis dans le nerf auditif :

1. La musique instrumentale utilise un timbre homogène, alors que la parole n'est qu'un changement continu du timbre. La parole est une « mélodie de timbre ». Ce qui porte le sens linguistique est le contraste spectral entre les unités de parole. Au contraire, dans l'instrument de musique on cherche un timbre égal et régulier sur toute la tessiture de l'instrument.
2. Du point de vue de la hauteur mélodique, le jeu musical instrumental exploite un système de notes stables. Le système musical classique est basé sur les hauteurs mélodiques. Les hauteurs ont donc tendance à rester constantes sur une note, et à évoluer par pallier. Au contraire en parole le système mélodique est défini de manière approximative quant aux hauteurs, avec un système de 4 ou 5 tons peu stables.
3. Du point de vue rythmique la musique exploite généralement une métrique contrôlée, alors que la parole exploite une métrique assez pauvre et irrégulière.
4. Un son ne peut être de la parole sans la compétence linguistique de l'auditeur. Donc le signal acoustique est une condition nécessaire mais non suffisante pour que le son soit parole. De la parole inintelligible est un bruit.
5. De même, un son ne peut pas être de la musique sans la compétence musicale de l'auditeur, bien que cette compétence soit beaucoup moins spécialisée que dans le cas de la parole. Une musique inintelligible est un bruit.
6. En plus de la compétence, l'attention de l'auditeur est une condition nécessaire à l'interprétation d'un signal donné en termes de parole, ou de musique. Sans attention, parole ou musique sont des bruits.
7. Pour un même stimulus acoustique, l'auditeur peut utiliser des modalités d'écoute tout à fait différentes. Ainsi dans la parole, on peut à volonté se concentrer sur les mots, le sens, sur la mélodie, sur la couleur des voyelles, sur le souffle etc. De même pour le son musical, où l'écoute peut certainement mettre en jeu des processus cognitifs très différents.
8. Tout objet sonore peut donner naissance à de la musique, si l'auditeur peut y reconnaître une forme d'organisation mélodique, rythmique, ou même dans le domaine des timbres.

9. Les bruits n'ont en principe pas d'organisation en termes de systèmes pour les hauteurs, le rythme ou le timbre. Cependant tout objet sonore peut devenir un bruit s'il est en conflit avec la compétence, l'attention ou les attentes du sujet.

Références :

Les références qui suivent sont citées sans ordre particulier dans chaque catégorie, et ne sont pas toutes référencées dans le texte. Cependant, elles peuvent aider le lecteur peu familier avec l'un ou l'autre des domaines à aborder les aspects acoustiques et fonctionnels de la parole ou de la musique.

Comparaisons entre parole et musique

- Jean-Jacques Nattiez, 1975. Fondements d'une sémiologie de la musique. Collection 10-18, série "Esthétique". No 1017, Paris.
- Nicolas Ruwet, 1972. Langage, musique, poésie. Le seuil, Paris.
- Ivan Fonagy, 1983. La vive Voix. Essais de psycho-phonétique. Payot, Paris.
- Ivan Fonagy, 1981. Emotion, Voice and Music. In Research aspects on singing, Publications of the Royal Swedish Academy of Music, 51-79.
- Roman Jakobson, 1971. Musicologie et linguistique. Musique en Jeu, n°5, 57-59, Edition du Seuil, Paris (traduction française de Musikwissenschaft un Linguistik, 1932)
- Claude Lévy-Strauss, 1964. Le Cru et le Cuit. Plon, Paris.
- Georges P. Springer 1971. Le langage et la musique : parallèles et divergences. Musique en Jeu, n°5, 31-44, Edition du Seuil, Paris (traduction française de Language and Music : Parallel and Divergencies, 1956)

Sur la parole

- Henri Méloni (coordinateur) (ouvrage collectif), 1996 Fondements et perspectives en traitement automatique de la parole}, Universités Francophones, AUPELF-UREF, Paris.
- René Boite & Murat Kunt, , 1987. Traitement de la parole. Dunod, Paris.
- Calliope (ouvrage collectif), 1989. La parole et son traitement automatique. Masson, Paris.
- James L. Flanagan, 1972. Speech analysis, synthesis and perception. Springer-Verlag, Berlin.
- Jean-Sylvain Liénard, 1977. Les processus de la communication parlée. Masson, Paris.
- William. J. Hardcastle, 1976. Physiology of speech production, Academic Press, Londres.
- Gunnar Fant, 1970. Acoustic theory of speech production. Mouton, La Hague-Paris.
- Fernand Carton, 1974. Introduction à la phonétique du Français. Bordas, Paris.
- Jean-Louis Duchet, 1981. La phonologie ; Col. "Que sais-je?", Presse Universitaires de France, Paris.
- André Martinet, 1956. La description phonologique. Librairies Droz, Genève, et Minard, Paris.
- Bertil Malmberg, 1954. La phonétique. Col. "Que sais-je?", Presses Universitaires de France, Paris.
- Bertil Malmberg, 1974. Manuel de phonétique générale. Editions Picard, Paris.
- Roman Jakobson, 1963. Essais de linguistique générale. Les éditions de minuit, Paris.
- John Lyons, 1968. Introduction to theoretical linguistics. Cambridge University Press.
- Ferdinand de Saussure, 1915. Cours de linguistique générale. Payot, Paris, 1972.

Sur la musique

- Marie-Claire Botte, Georges Canévet, Laurent Demany, Christel Sorin, 1988. Psychoacoustique et perception auditive. Edition INSERM, Paris.
- Arthur Benade Fundamental of Musical Acoustics, Dover publications, New York, 1976, 1990.
- Pierre Schaeffer, Traité des objets musicaux, Edition du Seuil, Paris, 1966, 1977.
- Emile Leipp, Acoustique et musique, Masson, Paris, 1977.
- Jürgen Meyer, Acoustics and the performance of music, Verlag das Musikinstrument, Frankfurt/Main, 1978.
- John Pierce, Le son musical, Pour la science, Belin, 1983.
- William Strong, G.R. Plilnik, Music, speech, audio, Soundprint, Provo, UT, 1992.
- Denis Mercier (éditeur), Le livre des techniques du son, Eyrolle, 1987-1993 (3 tomes)
- Curtis Road, L'Audionumérique, Dunod, Paris, 1998 (traduction de Computer Music Tutorial, MIT Press.)
- Arlette Zenatti (sous la direction de) 1994. Psychologie de la musique. PUF, Paris.