

PERCEPTION(S) DE LA VOIX CHANTÉE

par Michèle CASTELLENGO

L'exposé est articulé selon trois parties. Dans la première nous rappelons les particularités anatomiques et neuro-physiologique du système auditif humain qui conditionnent la perception auditive. Dans la deuxième partie nous exposons les bases d'une approche cognitive «écologique»¹ d'auditeurs communiquant au sein d'une société culturelle donnée. A la lumière de ces résultats et tout en nous fondant sur l'analyse acoustique, nous proposons des éléments d'interprétation de la diversité des écoutes de la voix humaine et de certaines techniques vocales.

1 - Psychophysique et neurosciences

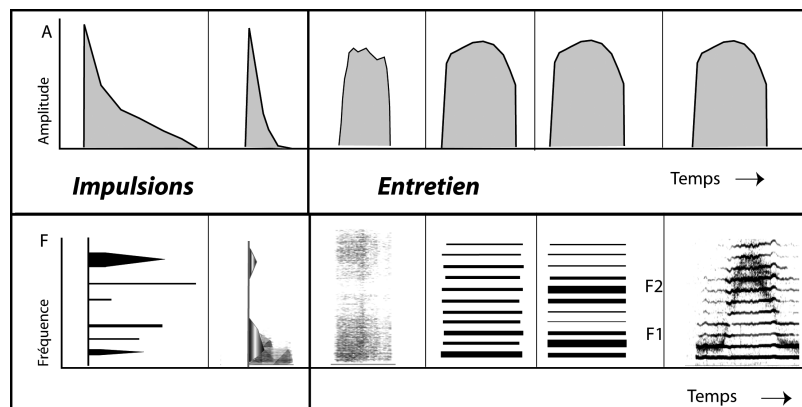
De l'examen des trois étapes de la perception sonore : réception (oreille), transmission (voies acoustiques) et traitement des données sensorielles (centres supérieurs), on tire des données psychoacoustiques fondamentales :

- **Efficacité et discrimination auditive.** Zone de plus *grande sensibilité de l'oreille* située de part et d'autre de 3kHz: *formant du chanteur* et de l'acteur et zone d'efficacité spectrale pour produire un crescendo d'intensité. Zone de meilleure *discrimination sensorielle* : celle où les seuils différentiels de fréquence et d'intensité sont les plus fins.
- **Codage de la «sensation de hauteur» sonore.** *Codage tonotopique* associé à la sensation de hauteur spectrale (détection des formants vocaliques) et *codage temporel* associé à la sensation de hauteur tonale (fréquence fondamentale des sons périodiques et leurs limites : 30 Hz vers le grave, 4000 Hz vers l'aigu). Application à la perception des voix graves : sensation de rugosité, «fry», hauteur indistincte.
- Sensibilité différente aux **modulations** d'*amplitude* (4Hz) et de *fréquence* (6Hz); cf vibrato
- Transition : audition d'un chant «Kargiraa»², diversité des écoutes.

2 - Percevoir c'est catégoriser, faire «sens».

2.1 Qu'est-ce ? Typologie acoustique : la voix parmi les autres sons.

La catégorisation libre. Test avec 30 sons reconnaissables, très brefs (700 ms), issus de l'environnement quotidien. Groupements réalisés sur la base d'une typologie acoustique des sources sonores :



- Liens typologie acoustique et catégorisation perceptive,
- Critères temporels primant sur les critères spectraux,
- Variabilité temporelle et signature spectrale (formants) de la voix humaine.

¹ fondée sur les stratégies d'un être vivant confronté à son environnement

² (CD Smithsonian Folkways Record; CD SF 40017; page 1)

2-2 - Catégorisation et perception sonore.

Principes de la catégorisation perceptive selon E. Rosch et D. Dubois (en particulier notion de prototype). Applications à la perception sonore et plus particulièrement à la voix humaine.

2-3 - Catégorisations de la voix humaine : Qui? Quoi? Comment?

- Situation de communication sonore (réseau de relations sociales, culturelles)
- Connaissances acquises, motivations personnelles : Elaboration et remaniement des catégories selon les connaissances acquises (mémorisées) et les motivations personnelles (du moment).
- Notion de *prototype* (propre à l'auditeur), qui sert de référence à l'évaluation.
- Nature et niveau de l'expertise (chanteur d'opéra, de variété, phoniatre, acousticien...). L'expertise est à double tranchant. Perception plus discriminante mais effet de «filtre», d'occultations.

2-4 - Un schéma de l'écoute humaine en situation réelle

La perception opère dans les deux sens : du bas vers le haut (de l'oreille vers les centres supérieurs) et du haut vers le bas (par sélection des éléments qui nous intéressent, dans le cadre d'une stratégie perceptive anticipatrice qui fait sens pour le sujet).

- Psychophysique : (Bottom-up). Mesures avec stimuli paramétrés. Données fondamentales de la sensation mais, travail avec sons « normalisés »; extraction des paramètres physiques. Objectif : caractériser un auditeur «normal».
- Cognition «écologique», «personnalisée» : (Top-Down/Bottom-up). Tests d'écoute avec sons réels complexes, variables, en contexte situé : percevoir c'est donner «sens» au son. Prise en compte des spécificités des sujets (connaissances, motivations etc) et de l'inscription dans un temps, dans un lieu (la vie). Etude des stratégies de traitement des données acoustiques mises en oeuvre pour «faire sens».

3 - Quelques exemples de perception de la voix chantée

3-1 - Ambiguïtés de l'écoute d'une voix

L'auditeur peut prêter attention à l'aspect sémantique (parole), à l'intonation, (mélodie) aux mouvements spectraux (formants vocaliques), aux indices affectifs... etc

3-2 Formants vocaliques et harmoniques du son glottique

- Importance perceptive des deux premiers formants vocaliques F1 et F2.
- Catégorisation perceptive des voyelles : prise en compte des variables (sexe, âge, tessiture, chant). En voix chantée les formants vocaliques interfèrent de façon complexe avec le son fondamental harmonique produit par les cordes vocales : ex «formant-tuning» et chant diphonique.
- «Formant-tuning». L'oreille discrimine aisément les cinq premiers harmoniques³ de sorte que le «formant-tuning» (généralement H1, H2 ou H) 3est perçu indépendamment de la reconnaissance de la voyelle : il contribue à la sonorité globale de la voix.
- «Chant diphonique» : technique de «formant-tuning» poussée à l'extrême dans laquelle la largeur formantique est considérablement réduite. Techniques vocales de mélodie «harmonique».
 - Technique 1, la mélodie est produite par le contrôle de F1 ; ex Kargiraa⁴
 - Technique 2, la mélodie est produite par le contrôle de F2 (600 à 2000 Hz; ex Xhomii).
 - Troisième 3 : les chanteurs accordent F1 et F2 à l'octave (voyelles antérieures)

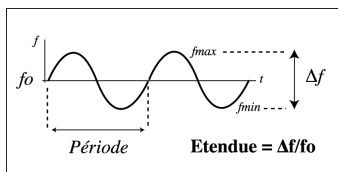
³ ils sont séparés par des intervalles plus larges que ceux des «filtres» cochélares

⁴ Remarque : sur le site du CREM Tran Quang Hai utilise une dénomination qui fait référence à la technique vocale et non à l'acoustique. Le chant dit à double cavité (kargiraa) correspond au diphonique de F1 et le chant à simple cavité au diphonique de F2.

- Rôle du deuxième formant vocalique dans la perception de la hauteur. La zone de fréquence de F2, très discriminante pour la reconnaissance des voyelles, est une *zone d'écoute* privilégiée par les auditeurs humains (cf langues sifflée).
- Le modèle harmonique. Combinaison du modèle harmonique et de la zone d'écoute expliquant des paradoxes de la sensation de hauteur : la «quarte» des grosses cloches; la «quintina» sarde.

3 - 2 Variations de la fréquence fondamentale de la voix humaine

- Instabilités de fréquence d'un son tenu «droit» : caractéristiques de l'émission par un être vivant. Le jitter est 3 à 4 fois plus grand que celui d'un son instrumental (violin, flûte à bec).
- Le **vibrato** : indice cognitif d'une voix chantée lyrique (cf le Teremin, l'illusion d'une voix chantée alors même qu'aucune voyelle n'est synthétisée)



La hauteur perçue d'un son vibré dépend de la *Rapidité* du vibrato vocal : 6 à 8 oscillations par seconde de sa *largeur* (variable, de 1/2 ton, à une tierce mineure) ce qui est beaucoup plus que le vibrato habituel des instruments (violin, flûte, saxophone).

Expériences de perception de sons vibrés lors des transitions et dans le cas où la durée est très brève (notes piquées).

- Le **trille vocal**. Lorsque la largeur du vibrato dépasse 2 tons, la perception change assez brusquement. Ce n'est plus la hauteur moyenne, f_0 qui est perçue, mais deux hauteurs alternées, voisines des fréquences f_{mx} et f_{min} . cf site Castellengo (2002)

3 - 3 - A propos de la qualité vocale

Voici quelques idées sur le travail mené au LAM de 2003 à 2007. Thème abordé par Maeva Garnier (Bib 2003) sous la direction de Nathalie Henrich. Recherche pluridisciplinaire qui s'inscrit dans le cadre de la conception cognitive de la perception présentée dans cet article. Point de départ = pratique des chanteurs et des pédagogues. Les principales étapes sont :

- des chanteurs interprètent la même phrase musicale avec des qualités vocales différentes. Enregistrement du son et de l'EKG.
- des professeurs de chant écoutent les enregistrements et commentent de façon libre ce qu'ils entendent.
- Les verbalisations sont transcrites et étudiées selon les méthodes d'analyse du discours développées en psycholinguistique par Danièle Dubois.
- Les résultats permettent de définir les différents champs cognitifs d'appréhension de la qualité vocale (cf schéma) et les principales catégories de qualité vocale.
- Travail d'écoute des extraits regroupés par catégories pour sélectionner les données acoustiques à examiner.
- Analyses acoustiques (zone du formant du chanteur, vibrato, pente spectrale etc).
- Interprétation des données acoustiques en fonction des principales catégories de qualité vocale.

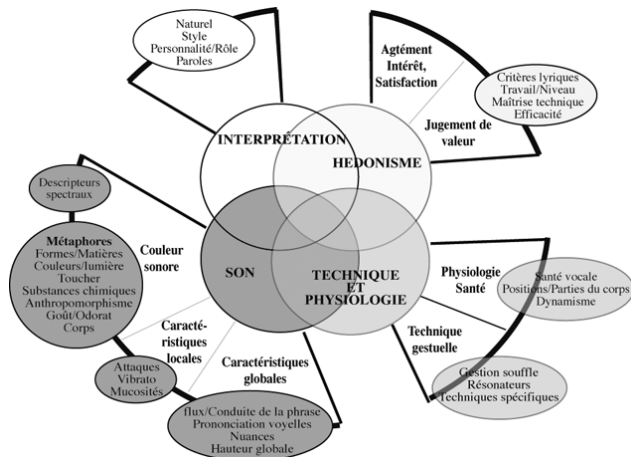
La méthode a été reprise avec une base de données enregistrée spécifiquement pour l'étude des qualités vocales. Voir l'article JIMS de 2007. Parallèlement à cette recherche Nathalie Henrich a réuni un groupe pluridisciplinaire (chanteurs, phoniâtres, orthophonistes, acousticiens) se donnant pour objectif de définir un vocabulaire consensuel ayant trait aux qualités vocales du chant lyrique, associé à une base sonore d'exemples sélectionnés en commun. cf publi Henrich 2007, Tallin.

Ces exemples témoignent de la complexité des études de «perception de la qualité vocale». Il s'agit d'études qui impliquent un fort investissement de la part des participants. Lorsque l'effort peut être soutenu sur une longue durée, un groupe d'experts se constitue, groupe au sein duquel l'élabore un vocabulaire commun entre chercheurs de disciplines différentes et le partage des connaissances se développe. Les participants s'enrichissent mutuellement de résultats dont la pertinence reste sous le contrôle des acteurs de la recherche.

3 - 4 - Ce qui n'a pas été abordé à Giens

Le diagramme ci-dessous donne une idée du «cahier des charges» d'un chanteur tenu de chanter «juste», sur une grande étendue en fréquence, avec un beau timbre, en respectant l'articulation du texte, avec une puissance suffisante pour que la voix porte, tout en communiquant une émotion artistique attendue. Plusieurs de ces objectifs imposent des choix entre des contraintes contradictoires entre elles (exemple : portée, timbre, articulation, tessiture...), choix qui varient selon le type de voix (en particulier homme, femme). Il faudrait aussi aborder :

- Comment sont perçues les différences de qualité entre les deux mécanismes laryngés?
- Comment interagissent les différentes qualités vocales avec la tessiture : à spectre similaire la



perception varie grandement entre le grave, le médium et l'aigu ce qui retentit sur l'appréciation des voyelles et sur celle de la sonorité globale de la voix.

- Quel est l'effet de la tessiture sur l'intelligibilité du texte.
- Comment le vibrato interfère avec l'appréciation des qualités vocales? et plusieurs autres questions qui seront autant de sujets passionnants pour les jeunes chercheurs.

4 - Conclusions

1 - La voix humaine est la plus étonnante et la plus complexe des productions sonores

- par la **richesse de sa structure acoustique**, variable en hauteur et en spectre. C'est le seul «instrument» de timbre spectral.
- par la **diversité des modalités d'appréhension** qu'elle suscite.
- Son étude est **indissociable** de celle du réseau de **communications entre humains** dans lequel la voix prend sens pour un auditeur.

2 - Pour un auditeur humain une voix est l'expression d'une personne.

- Impact émotionnel d'une voix qui «interpelle» (cf musique electro-acoustique)
- Interprétation perceptif à divers niveaux : comportemental (du locuteur), sémantique (langue), esthétique, toujours par référence à des prototypes personnels mémorisés (pas de norme absolue)

3 - Prise en compte de deux approches complémentaires de la perception humaine

- Approche psychophysique (et neurosciences). Etudes par paramètres⁵, point de vue généraliste (le sujet = récepteur moyen). Fournit des résultats *fondamentaux* (données physiologiques) mais s'en tenir aux conditions de validité (stimuli de laboratoire).
- Approche «phénoménologique, philosophique, cognitive». Le point de départ en est un auditeur inséré dans un contexte de vie, parce que «percevoir c'est faire sens» (théorie de la forme, catégorisation). Conception globale : tout interagit (cf intensité). Démarche *incontournable* pour comprendre les jugements de qualité (musique, voix chantée). Ce doit être le cadre idéal des échanges entre chanteurs et chercheurs.

4 - La voix chantée est une «production» culturelle qui se définit au sein d'un réseau d'échanges entre humains.

L'étude acoustique des qualités techniques et esthétiques d'une voix : «belle voix, voix émouvante» est relative à la culture considérée. Pas d'universaux de la voix chantée!

⁵ séduisant et rassurant : c'est de la science...

Références bibliographiques

Partie 1

- Collectif - Botte, M. C., Canévet, G., Demany, L., Sorin, C. (Ed.). (1989). *Psychoacoustique et perception auditive*. Paris: INSERM / SFA / CNET - Ed. Médicales Internationales.
- Castellengo, M. (1994). La perception auditive des sons musicaux. In A. Zenatti (Ed.), *Psychologie de la musique* (pp. 55-86). Paris: Presses Universitaires de France.
- Fitch, W. T., & Fritz, J. B. (2006). Rhesus macaques spontaneously perceive formants in conspecific vocalizations. *J.A.S.A.*, 120(4), 2132-2141.
- Leipp, E. (1984). *Acoustique et musique* (4ème ed.). Paris: Masson.
- Moore, B. C. J. (2004). *An introduction to the psychology of hearing* (5ème éd. ed.). London: Elsevier, Academic Press.
- Patterson, R. D. Page personnelle. from <http://www.neuroscience.cam.ac.uk/directory/profile.php?rdp1>
- Peretz, I. (2010). Page personnelle from http://www.brams.org/members-fr/peretzi_fr
- Pillot, C. (2004). *L'efficacité vocale dans le chant lyrique : Aspects physiologique, cognitif, acoustique et perceptif*, Université Paris 3, Paris.
- Stevens, S. S., & Davis, H. (1983). *Hearing - Its Psychology and Physiology*: Acoustical Society of America / American Institute of Physics.
- Sundberg, J. (2001). Level and center frequency of the singer's formant. *J. of Voice*, 15(2), 176-186.

Partie 2

- Dubois, D. (1997). *Catégorisation et cognition*. Paris: Kimé.
- Guyot, F., Castellengo, M., & Fabre, B. (1997). Etude de la catégorisation d'un corpus de bruits. In D. Dubois (Ed.), *Catégorisation, représentation et systèmes symboliques* (pp. 45-58). Paris: Kimé.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Paris: Gallimard.
- Rosch, E. (1976). Classifications d'objets du monde réel : origines et représentations dans la cognition. *Bull. de Psychologie*(N° spécial : La mémoire sémantique), 242-250.
- Vanderveer, N. J. (1979). *Ecological Acoustics : Human Perception of environmental sounds* (Vol. Thèse, Faculty of Cornell University). Ann Arbor: University Microfilms N° 8004002.

Partie 3

- Bailly, L., Henrich, N., Pelorson, X., & Gilbert, J. (2008). *Vocal folds and ventricular bands in interaction : comparison between 'in vivo' measurements and theoretical predictions*. Paper presented at the Acoustics'08, Paris.
- Castellengo, M. (1993). *Fusion or separation : from vibrato to vocal trill*. Paper presented at the S.M.A.C., Stockholm.
- Castellengo, M. (2002). Les sources acoustiques. In D. Mercier (Ed.), *Le livre des techniques du son* (Vol. 1, pp. 45-80). Paris: Dunod.
- Castellengo, M. (2002). The human voice and its registers : the value of interdisciplinary collaboration. From http://www.med.rug.nl/pas/Conf_contrib/Castellengo/Castellengo_bio_touch.htm
- Castellengo, M., D'Alessandro, C., & Richard, G. (1989). *Study of vocal pitch vibrato perception using synthesis*. Paper presented at the 13th International Congress on Acoustics, Belgrade.
- Castellengo, M., Lortat-Jacob, B., & Léothaud, G. (2001). *Pitch perception : five voices with four sardinian singers*. Paper presented at the ISMA, Perugia.
- D'Alessandro, C., & Castellengo, M. (1994). The pitch of short-duration vibrato tones. *J.A.S.A.*, 95(3), 1617-1630.
- Fant, G. (1970). *Acoustic Theory of Speech Production with Calculations based on X-Ray Studies of Russian Articulations* (2ème ed. Vol. 2). La Hague: Mouton.
- Garnier, M. (2003). *Approche de la qualité vocale dans le chant lyrique : perception, verbalisation et corrélats acoustiques*. Mémoire de DEA ATIAM, Université Pierre et Marie Curie, Paris.
- Garnier, M., Henrich, N., Castellengo, M., Sotiropoulos, D., & Dubois, D. (2007). Characterisation of Voice Quality in Western Lyrical Singing : from Teacher's Judgements to Acoustic Descriptions. *Journal of Interdisciplinary Music Studies (JIMS)*, 1(2), 62-91.
- Hai, T. Q., & Zemp, H. (2010). Analyses acoustique et physiologique des techniques vocales (Publication.: http://www.crem-cnrs.fr/realisations_multimedia/animations/diphonique/hai1.html)
- Henrich, N., Bezard, P., Expert, R., Garnier, M., Guerin, C., Pillot, C., et al. (2007). *Perception and verbalisation of voice quality in Western lyrical singing : contribution of a multidisciplinary research group*. Paper presented at the 3rd C.I.M. (Conference on Interdisciplinary Musicology), Tallin (Estonia).
- Schonle, J., & Horan, K. (1980). The pitch of vibrato tones. *J.A.S.A.*, 67, 246-252.
- Tamba, A., & Castellengo, M. (1969). La musique du théâtre Nô Japonais. *Bulletin du GAM (Groupe d'Acoustique Musicale)*, N°39.
- Tehardt, E. (1982). Pitch of complex signals according to virtual-pitch theory : tests, examples, and predictions. *J.A.S.A.*, 71, 671-678.
- Terhardt, E. (2007). Page personnelle. from <http://www.mmk.e-technik.tu-muenchen.de/persons/ter.html>
- Vanderveer, N. J. (1979). *Ecological Acoustics : Human Perception of environmental sounds* (Vol. Thèse, Faculty of Cornell University). Ann Arbor: University Microfilms N° 8004002.