

Sous la direction de  
Marie-France Castarède  
et Gabrielle Konopczynski

# Au commencement était la voix



*la vie de l'enfant*

 érès

*Branka Zei Pollermann\**

QU'EXPRIME LA PROSODIE AFFECTIVE:  
L'ÉTAT DU CORPS OU L'ÉTAT D'ESPRIT ?  
Proposition d'un modèle unifié  
de l'émotion et de cognition

\*Hôpitaux Universitaires de Genève, Clinique d'accueil, d'urgences et de  
liaison psychiatriques

Pour comprendre le lien entre la prosodie et l'état affectif du sujet il convient d'abord de dresser les critères de distinction entre les réactions émotionnelles et non émotionnelles.

Dès l'aube des premiers propos sur les émotions, deux thèses sur l'origine des émotions s'affrontent : l'origine corporelle (états organiques) versus l'origine purement mentale (états cognitifs). En effet la relation entre les phénomènes cognitifs et organiques est au centre d'intérêt des chercheurs. Les phénomènes émotionnels ont souvent été définis comme un ensemble de réactions psychophysiologiques qui incluent les cognitions, les réactions viscérales et immunologiques, les expressions verbales et non verbales ainsi que les dispositions comportementales et les sentiments subjectifs. Or cette « définition » descriptive ne nous permet pas de distinguer entre les réactions émotionnelles et non émotionnelles. Nous sommes actuellement en présence d'une grande diversité de théories et de définitions très larges et variées avec peu ou pas de consensus sur les critères par lesquels on pourrait distinguer entre une réaction émotionnelle et une réaction non émotionnelle (Ekman et Davidson, 1994). Nous estimons que la théorie tridimensionnelle de Russel et Mehrabian (Russel et Mehrabian, 1977) présente un intérêt particulier pour l'étude de la relation émotion-cognition. Leur théorie – soutenue par des résultats expérimentaux - affirme qu'à tout moment, l'état psychophysiologique de l'individu peut être défini par trois facteurs : valence hédonique ou agrément (agréable vs désagréable), éveil ou activation physiologique (augmentée vs diminuée) et dominance (le pouvoir fort ou faible d'agir sur le stimulus et/ou sur ses conséquences). Ainsi, la tristesse par exemple, serait caractérisée par une valence hédonique fortement négative (désagréable), une activation faible, et une potence encore plus faible. Par contre, la joie serait une émotion à forte valence positive (agréable) accompagnée d'une activation moyenne à forte et une potence moyenne à forte. Toutefois les auteurs ne donnent pas de critères

pour différencier les états émotionnels des états non émotionnels. À ce propos, Scherer (1993) propose un critère unique : la synchronisation temporaire des réactions de plusieurs sous-systèmes du fonctionnement de l'organisme. Cette synchronisation caractériserait la réponse émotionnelle alors qu'elle serait absente dans la réponse non émotionnelle.

Au niveau de l'expression vocale des émotions on devrait donc pouvoir distinguer entre une voix produite dans le cadre de cette synchronisation et celle produite en dehors de la synchronisation des réactions. Cependant, nous pensons que toute réponse adaptative de l'organisme, exige une coordination et/ou synchronisation de ses réactions psychophysiologiques.

Quant à la relation entre l'émotion et la cognition, il est actuellement admis qu'une évaluation cognitive du stimulus (consciente ou inconsciente) précède toute réaction émotionnelle (Lazarus et al, 1984 ; Roseman, 1991 ; Scherer, 1984). Il est également acquis que des processus dits « émotionnels » sont nécessaires pour certains types de prise de décisions considérées comme « rationnelles » (Damasio, 1994) et que du point de vue neurophysiologique il n'existe pas de preuves que les processus neurologiques sous-jacents à la cognition sont différents de ceux impliqués dans l'émotion (Lane et al, 2000).

Nous estimons qu'une approche épistémologique pourrait à la fois clarifier la relation entre la cognition et l'émotion, fournir les critères de définition et mieux cerner les indices vocaux des états affectifs.

## MODÈLE UNIFIÉ DE L'ÉMOTION ET DE COGNITION

Notre modèle est basé sur deux approches épistémologiques complémentaires : l'épistémologie génétique (Piaget, 1970) et la théorie sémiotique de la connaissance de (Prieto, 1975). En voici les idées de base. La connaissance se construit au travers des conduites d'interaction entre le sujet et l'environnement. Cette interaction exige un traitement cognitif des données liées à l'objet et à l'action propre. La connaissance qui en résulte est construite par rapport au sens que le sujet lui attribue (Prieto, 1975). De ce fait elle est porteuse de valeurs subjectives (Piaget, 1954 / 1981). Selon notre modèle, les valeurs de base attribuées aux objets de connaissance (y compris le corps propre et le processus interactif lui-même) sont : la valence hédonique ; l'activation psychophysiologique (effective ou estimée comme nécessaire pour faire face au stimulus et/ou ses conséquences) ; la dominance. Il va de soi que l'attribution de valeurs est influencée par le contexte spécifique. Nous proposons que ces trois types de valeurs constituent les trois dimensions de l'espace affectif propre à toute conduite interactive. De ce fait les aspects cognitifs et affectifs des conduites ne sont qu'artificiellement séparables (Piaget, 1954). Sur quel critère différencier les réactions émotionnelles des réactions non émotionnelles ?

Inspirés des travaux dans le domaine de la physiologie du comportement (Duffy, 1941) et plus récemment dans le domaine des composants neurologiques des réactions émotionnelles (Chapman et Nakamura, 2001) nous proposons le critère suivant pour distinguer entre les réactions émotionnelles et non émotionnelles : lorsque la valeur d'au moins une des trois facteurs atteint un niveau critique c.-à-d. plus élevé ou plus bas par rapport aux variations habituelles, l'organisme déclenche une réaction psychophysique émotionnelle. cognitivement dominant et reconnu par le sujet comme une expérience émotionnelle. La frontière entre une réponse émotionnelle et non émotionnelle se situe dans l'espace délimité par les valeurs critiques des trois dimensions de l'espace cognitivo-affectif. Si par exemple, le sujet attribue à son interaction avec le stimulus une très forte valence hédonique, la réponse de l'organisme sera reconnue comme émotionnelle.

## LA VOIX COMME REFLET DE L'ÉTAT COGNITIVO-AFFECTIF

L'interaction langagière, comme toute autre conduite interactive, reflète les trois dimensions de l'espace cognitivo-affectif. Aussi, la vive voix porte-t-elle des empreintes de la configuration de l'espace cognitivo-affectif du locuteur. Deux types d'empreintes sont à distinguer : celles liés à ses humeurs, ses attitudes générales, sa personnalité (s'étendant sur une période plus longue) et celles liés aux états de courte durée tels que ses émotions et ses stratégies de communication. Les variations acoustiques se manifestent à trois niveaux :

- suprasegmental: la hauteur moyenne de la voix, les contours mélodiques, l'énergie acoustique, le débit ;
- segmental: la précision articulatoire, articulation tendue ou relâchée, rapide ou lente,
- intra segmental: la qualité de la voix.

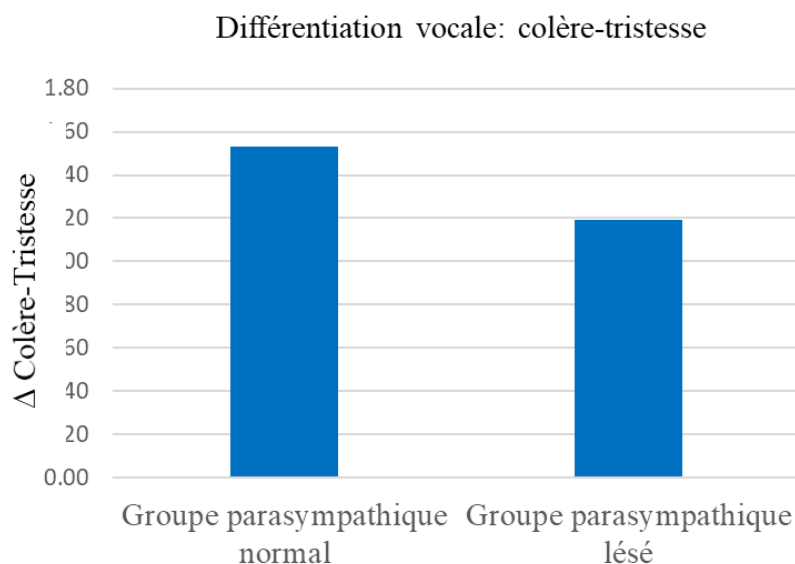
La voix comme symptôme de la dimension « activation » : exemple d'émoussement de l'expression émotionnelle dû aux lésions neurologiques

La relation entre le système nerveux autonome (SNA) et la parole a été modélisée par Williams et Stevens (1972). Selon ces auteurs l'activation accrue de la branche sympathique du SNA apparaît lors de l'expérience de colère, peur ou joie intense. Cette activation provoque une augmentation du rythme cardiaque, et de la tension artérielle et une plus forte distribution du sang dans la musculature striée. Ces changements physiologiques entraînent des modifications de la profondeur et de la fréquence respiratoire, aussi bien qu'une diminution des sécrétions salivaires. L'activation accrue du système

parasympathique conduit à la diminution de la fréquence cardiaque, à la réduction de la pression sanguine et l'éloignement du sang de la musculature

striée lors de l'expérience du chagrin. Ces effets physiologiques influencent le fonctionnement des appareils respiratoire, phonatoire et articulaire en provoquant de multiples changements acoustiques.

Nous avons étudié les changements vocaux au niveau suprasegmental chez 39 patients diabétiques dont certains souffraient de lésions du système nerveux autonome. Les émotions de joie, tristesse et colère ont été induites au travers le récit de leurs propres expériences émotionnelles. Les analyses acoustiques des voix des patients nous ont permis de calculer un indice d'activation vocale (ActVoc). Il représentait le cumul de valeurs normalisées de cinq paramètres vocaux : la F0 moyenne, l'écart de la F0, le coefficient de variation de la F0, l'écart de l'énergie voisée exprimé en dB et le débit (exprimé en nombre de syllabes prononcées par seconde). Un indice de différenciation colère tristesse ( $\Delta$  CT) a été calculé par soustraction de l'ActVoc pour la tristesse de l'ActVoc



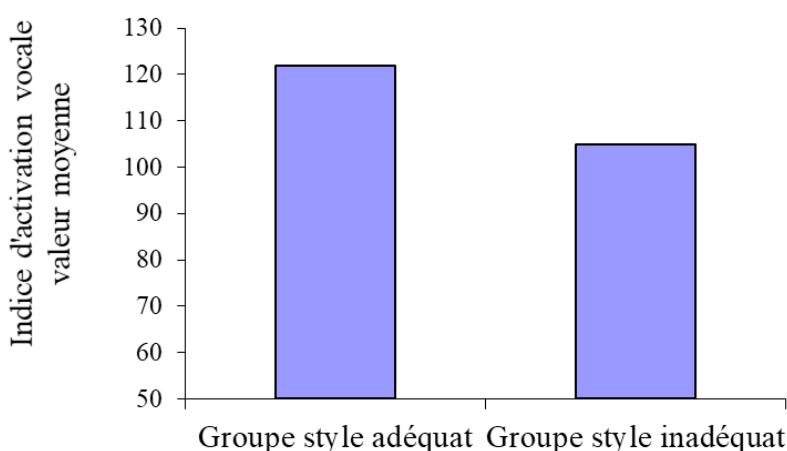
pour la colère. Les résultats de l'analyse de la variance (ANOVA) ont montré que dans le groupe présentant des lésions du système parasympathique l'indice de différenciation ( $\Delta$  Colère-Tristesse) a été significativement plus faible ( $F = 40.96$  ;  $P = .000$ ). Nous concluons que la dimension « activation » se manifeste dans la voix comme symptôme de l'état neurophysiologique du sujet parlant.

#### LA VOIX COMME SYMPTÔME DE LA DIMENSION «DOMINANCE» : EXEMPLE DU STYLE DE FAIRE FACE AU CANCER

Dans une étude pilote (N=10) sur les styles de faire face à la maladie (Bernhardt et Zei 2004) nous avons analysé les voix des patients cancéreux parlant de leur maladie. Leurs styles de faire face ont été indépendamment évalués comme

*adéquats* ou *inadéquats* selon des critères spécifiquement élaborés pour ce but (Morris, et al. 1985). Un style *adéquat* correspondait à une attitude combattive et confiante (Obrist, 1981) de pouvoir gérer la maladie. Des échantillons des voix des patients ont été pris à plusieurs moments de leur récit: le moment de forte activation émotionnelle, le moment de faible activation émotionnelle et le moment où ils parlaient de leur façon de faire face à la maladie.

Activation vocale et  
le style de faire face à la maladie



Les valeurs normalisées de cinq paramètres vocaux (ci-dessus énumérés) ont été cumulées pour constituer l'indice d'activation vocale (ActVoc). L'analyse de la variance (ANOVA) a montré que l'indice cet indice calculé pour l'échantillon prélevé au moment de forte activation émotionnelle était plus élevé dans le groupe ayant un *style actif* évalué comme *adéquat* que dans le groupe dont le style était évalué comme *passif* et *inadéquat* ( $F = 6.136$  ;  $P = 0.38$ ). Nous estimons que la dimension « dominance » s'est reflétée au niveau de la voix dans l'indice d'activation vocale

#### LA VOIX COMME INDICE DE LA DIMENSION «VALENCE HÉDONIQUE»

Pour cerner les indices vocaux de la valence hédonique nous avons comparé le spectre moyen (*long term average spectrum*) de la colère (émotion à forte activation et valence hédonique négative) avec celui de la joie (émotion à forte activation et valence hédonique positive). Le spectre a été découpé en bandes de largeur de 1.5 Bark. Les résultats des analyses statistiques (T tests appariés) ont montré que dans certaines bandes du spectre, l'énergie acoustique de la colère est plus élevée que dans la joie. Chez les hommes (N= 20) la bande

fréquentielle de différentiation significative s'étend de 300 - 3 400 Hz ;  $P < 0.01$ , alors que chez les femmes (N= 20) elle est de 840 - 3 400 Hz ;  $P < 0.05$  (Zei et Archinard 2001). Aussi nous concluons que la valence hédonique se reflète dans les caractéristiques spectrales de la voix (Fig. 3).

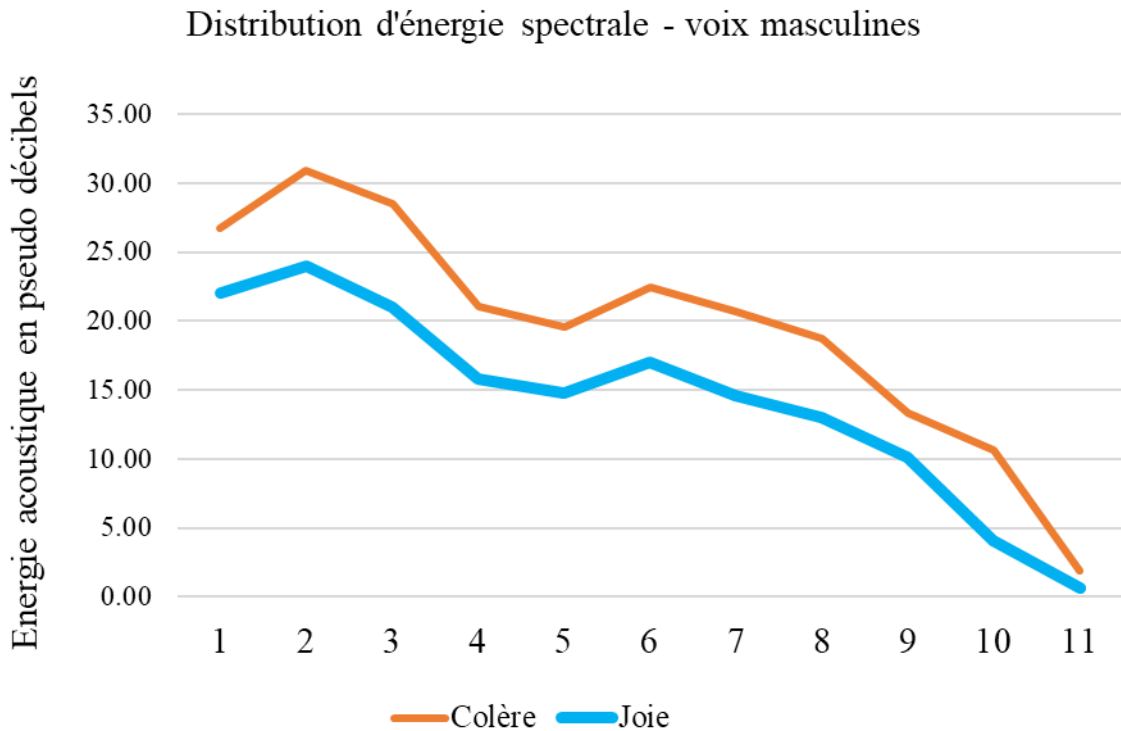


Fig. 3 : Bandes en 1.5 Bark : **1** = 297-453 Hz; **2** = 453-631 Hz; **3** = 631-838 Hz ; **4** = 838-1081 Hz; **5** = 1081-1370 Hz; **6** = 1370-1720 Hz; **7** = 1720-2152 Hz; **8** = 2700-3400 Hz; **9** = 3400-4370 Hz; **10** = 4370-5500 Hz.

Au niveau socioculturel la voix reflète les processus de la régulation de l'interaction. Scherer (1980) propose que la voix reflète l'influence simultanée de deux forces : une déterminée par l'état psycho biologique du locuteur – appelée

la force « push » et l'autre dirigée par des contraintes culturelles et situationnelles – appelée la force « pull ». Aussi convient-il de déterminer le statut sémiologique des indices vocaux de l'affectivité. Quel statut le récepteur attribue-t-il aux variations prosodiques ? Les considère-t-il comme indices spontanés, une sorte de symptômes de l'état psychophysiologique du locuteur ? Les considère-t-il comme des indices intentionnels non arbitraires mais conventionnalisés tels que les symboles, où encore comme des signes proprement dits (Prieto 1975) ? Nous estimons que, du fait que le lien entre l'état cognitivo-affectif et la voix n'est pas arbitraire mais causalement motivé, la prosodie ne peut qu'avoir que le statut d'indice spontané, mais jamais celui de signe.

Pour conclure, le modèle unifié de la cognition et de l'émotion (Zei, 2002) propose que les opérations cognitives impliquées dans toute conduite, se déroulent dans un espace déterminé par trois facteurs : valence hédonique, activation psychophysologique et dominance. De ce fait la prosodie affective exprime l'état cognitivo-affectif dominant au moment de l'énonciation. Nous estimons que la recherche en prosodie peut bénéficier du modèle unifié, car il offre la continuité entre des états traditionnellement considérés comme cognitifs (doute, certitude), en passant par des états émotionnellement « colorés » (amabilité, hostilité) jusqu'aux états pleinement émotionnels (peur, colère, joie, tristesse). Aussi la voix serait, à tout moment, le reflet d'une configuration de l'espace cognitivo affectif du sujet parlant. Le tableau ci-après résume les caractéristiques des états cognitivo-affectif reflétés dans la voix.

Caractéristiques des états cognitivo affectifs reflétés dans la voix.

Type d'état cognitivo – affectif	État des facteurs Valence, Activation, Dominance	Relation “Push”-“Pull”	Statut sémiologique des indices vocaux	Durée
Émotions Par ex.: joie, colère, tristesse, peur, honte,	Zone critique	“push” dominant	Indice spontané - Symptôme	Courte à moyenne
Humeurs Par ex.: gai, irritable, déprimé, inquiet.	Zone habituelle à critique	“push” dominant ou mixte	Indice spontané – ou faussement spontané	Moyenne
Attitudes personnelles Par ex.: préférer, haïr, aimer, apprécier	Zone habituelle	“push” dominant ou mixte	Symptôme et/ou Symbole	Moyenne à longue
Attitudes interpersonnelles Par ex.: distant, chaleureux, méprisant	Zone habituelle	“pull” dominant	Symptôme et/ou Symbole	Courte à moyenne
Traits de personnalité Par ex.: nerveux, timide, anxieux, extraverti	Zone habituelle	“push” dominant	Symptôme	Longue



## BIBLIOGRAPHIE

- Bernhard, J. et Zei Pollermann, B. 2004. "Emotional expression as an indicator of mental adjustment: Voice analysis in patients with early breast cancer". *Submitted*
- Chapman, C.R. et Nakamura, 2001. The affective dimension of pain: mechanisms and implications. In *Emotions, Qualia, and Consciousness*, A. Kaszniak (ed.). Singapore, London, World Scientific. p. 181-210
- Duffy, E., 1941. An Explanation of "emotional" phenomena without the use of the concept of emotion. *The Journal of General Psychology*, 25, 283-293.
- Damasio, A.R. 1994. *Descartes' Error. Emotion, Reason and the Human Brain*. New York: G.P. Putnam's Sons.
- Ekman, P. et Davidson, R.J. 1994 (eds). *The nature of emotion. Fundamental questions*. New York - Oxford: Oxford University Press.
- Lane, R. D.; Nadel, L.; Allen, John J.B.; Kaszniak A.W. 2000. "The Study of Emotion from the Perspective of Cognitive Neuroscience". In *Cognitive Neuroscience of Emotion*. R.D. Lane and Nadel, L. (eds). New York, Oxford, Oxford University Press, p. 3-12.
- Lazarus, R.S.; Coyne, J.C.; Folkman, S. 1984. Cognition, emotion and motivation. The doctoring of Humpty-Dumpty In *Approaches to emotion*, K.R. Scherer & P. Ekman (eds.). Hillsdale, N.J.: Lawrence Erlbaum Associates Inc., 221-237
- Russel, J. and Mehrabian, A. 1977. "Evidence for a Three-Factor Theory of Emotions" *Journal of research in personality*. 11, p. 273-294
- Morris, T.; Blake, S.; Buckley, M. 1985. "Development of a method for rating cognitive responses to a diagnosis of cancer". *Soc. Sci. Med.* 20, p. 795-802
- Obrist, P.A. 1981. *Cardiovascular psychophysiology: a perspective*. New York, Plenum.
- Piaget, J. 1954 / 1981. *Intelligence and Affectivity: their Relationship During Child Development*. Palo Alto: Annual Reviews.
- Piaget, J. 1970. *Epistémologie génétique*. Paris: Presses Universitaires de France.
- Prieto L.J. 1975. *Pertinence et pratique*. Paris, Editions de Minuit.
- Roseman, I.J. 1991. "Appraisal determinants of discrete emotions", *Cognition and emotion*. 5, p.161-200
- Scherer, U.; Helfrich, H.; Scherer, K.R., 1980. Internal push or external pull? Determinants of paralinguistic behaviour. In *Language: social psychological perspectives*, H. Giles; P. Robinson and P. Smith (eds.). Oxford - New York: Pergamon, 279-282.
- Scherer, K.R. 1984. Emotion as a multicomponent process: A model and some cross cultural data. In *Review of personality and social psychology*, P. Shaver (ed.). Vol. 5. Beverly Hills, CA: Sage, p. 37-63
- Scherer, K.R. 1993. Neuroscience Projections to Current Debates in Emotion Psychology. *Cognition and Emotion*, 7, (1), p. 1-41.
- Williams, C. E., et Stevens, K.N. 1972. "Emotion and Speech: Some Acoustical Correlates". *The Journal of the Acoustical Society of America*. 52, p. 1238-1250
- Zei Pollermann B. (2002). "A Place for Prosody in a Unified Model of Cognition and Emotion". In : *Speech Prosody 2002*, Université de Aix-en-Provence, pp 17-22.