

*Branka Zei Pollermann*

QU'EXPRIME LA PROSODIE AFFECTIVE :  
L'ÉTAT DU CORPS OU L'ÉTAT DE L'ESPRIT ?

Proposition d'un modèle unifié  
de l'émotion et de cognition

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 de l'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 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 (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 une 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 coll., 1984 ; Roseman, 1991 ; Scherer, 1984). Il est également acquis que des processus dits « émotionnels » sont nécessaires pour certaines prises 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 soient différents de ceux impliqués dans l'émotion (Lane et coll., 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 LA COGNITION

Notre modèle est basé sur deux approches épistémologiques complémentaires : l'épistémologie génétique (Piaget, 1970) et la théorie sémiotique de la connaissance (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 à 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 un des trois facteurs atteint un niveau critique, c'est-à-dire plus élevé ou plus bas par rapport aux variations habituelles, l'organisme déclenche une réaction psychophysiologique émotionnelle. Cet état de l'organisme peut devenir 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 liées à ses humeurs, ses attitudes générales, sa personnalité (s'étendant sur une période plus longue) et celles liées aux états de courte durée tels que ses émotions et ses stratégies communicatives. Les variations acoustiques se manifestent à trois niveaux :

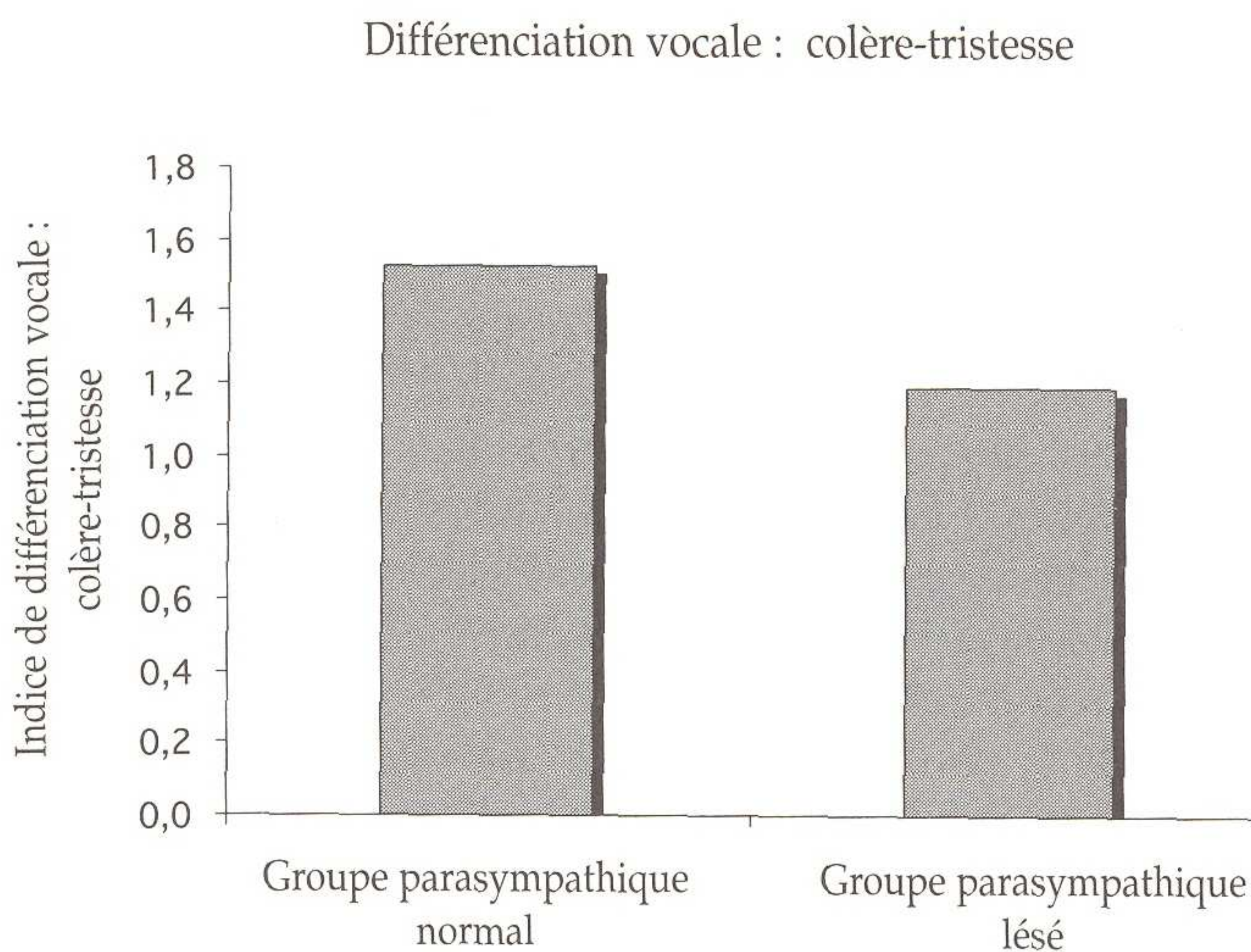
- suprasegmental : la hauteur moyenne de la voix, les contours mélodiques, l'énergie acoustique, le débit ;
- segmental : la précision articulatoire, articulation tendue ou relâchée, rapide ou lente ;
- intrasegmental : la qualité de la voix.



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, de peur ou de 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 à 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 articulatoire 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, de tristesse et de colère ont été induites au travers du récit de leurs propres expériences

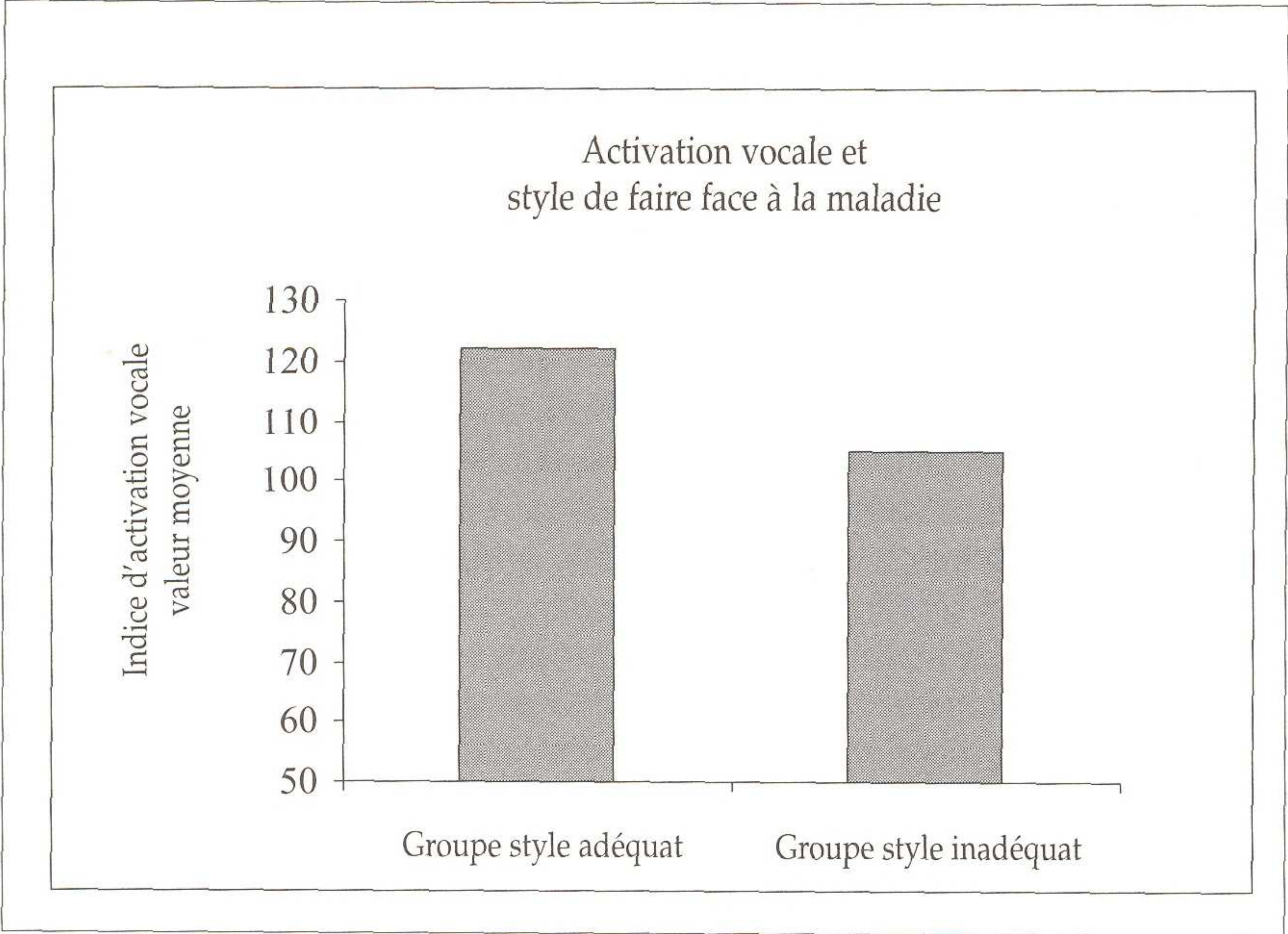




émotionnelles. Les analyses acoustiques des voix des patients nous ont permis de calculer un indice d'activation vocale (ActVoc). Il représente le cumul de valeurs normalisées de cinq paramètres vocaux : la F0 moyenne, l'écart de la F0, le coefficient de variation de la F0, l'écart de l'énergie voisée 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 ( $\Delta$  CT) a été significativement plus faible ( $F = 40.96$  ;  $P = .000$ ). Nous en 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 (Zei Pollermann et Bernhardt, 2006), 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 dans ce but (Morris et coll., 1985). Un style *adéquat* correspond à une attitude combative et confiante (Obrist, 1981) pour 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. Les valeurs normalisées des cinq paramètres vocaux énumérés ci-dessus ont été cumulées pour constituer l'indice d'activation vocale (ActVoc). L'analyse de la variance (ANOVA) a montré que l'indice ActVoc 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 dans la voix au niveau de l'indice d'activation vocale.

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

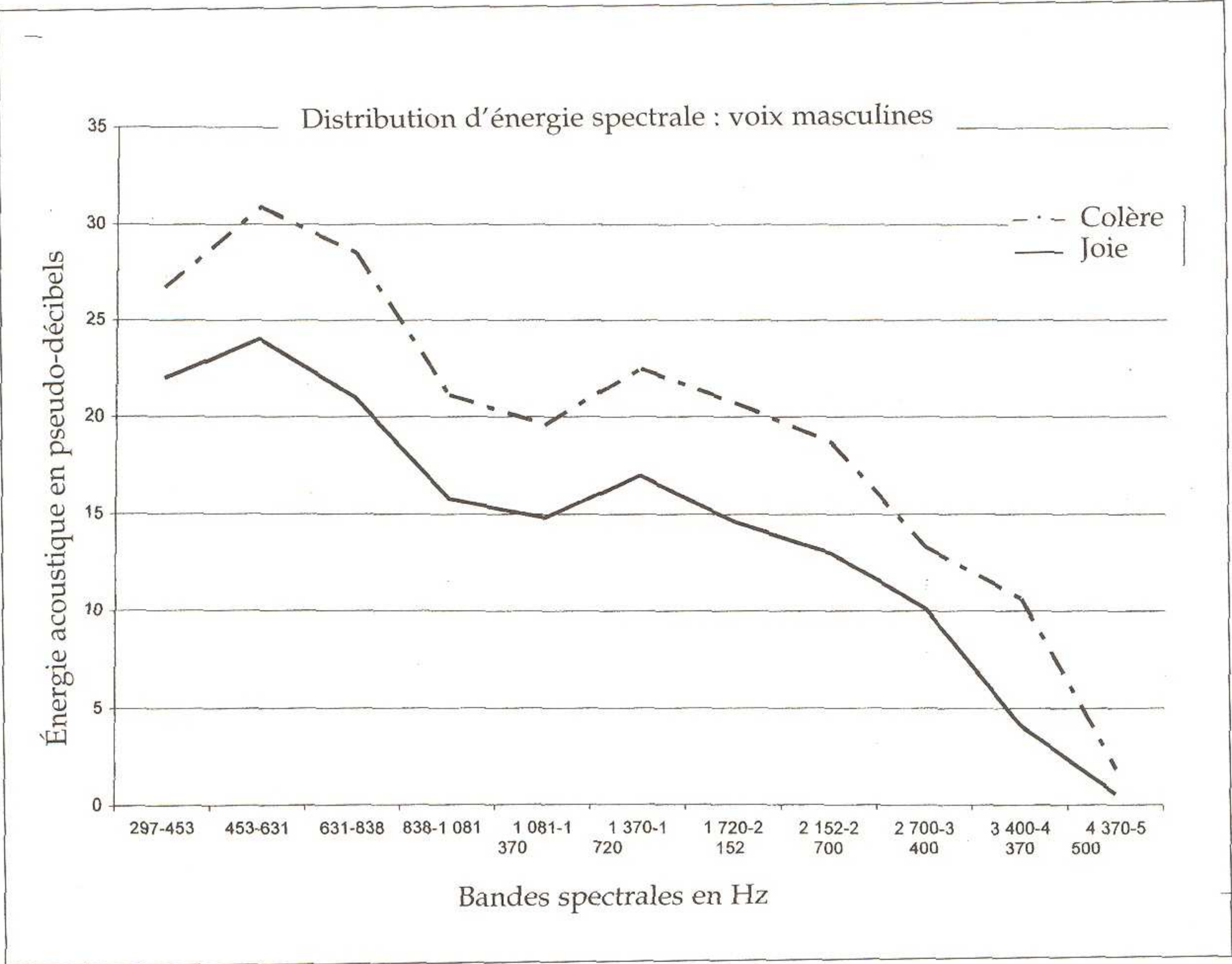
Pour cerner les indices vocaux de la valence hédonique, nous avons comparé le spectre moyen<sup>1</sup> de la colère (émotion à forte activation et valence hédonique négative) avec celui de la joie (émotion à forte activation et valence hédonique positive). Le spectre a été découpé en bandes de largeur de 1.5 Bark. Les résultats des analyses statistiques (T tests appareillé) 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 différenciation significative s'étend de 300 Hz – 3.400 Hz ;  $P < 0.01$ , alors que chez les femmes ( $N = 20$ ), elle est de 840 Hz – 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.

Au niveau socioculturel la voix reflète les processus de la régulation de l'interaction. Scherer (Scherer, 1980) propose que la voix reflète l'influence simultanée de deux forces : une déterminée par l'état psychobiologique 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 symptôme de l'état psychophysiologique du locuteur ? Les considère-t-il

---

1. Aussi connu sous le terme de *long term average spectrum* (LTAS).





comme des indices intentionnels non arbitraires mais conventionnalisés tels que les symboles, ou 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 avoir que le statut d'indice spontané ou de symbole, 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 psychophysiologique 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 des critères pour une solution de 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-affectifs 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 <i>Push-Pull</i>	Statut sémiologique des indices vocaux	Durée
Émotions Par ex. : joie, colère, tristesse, peur	Zone critique	<i>push</i> dominant	Indice spontané Symptôme	Courte à moyenne
Humeurs Par ex. : gai, irritable, déprimé, inquiet	Zone habituelle à critique	<i>push</i> dominant ou mixte	Indice spontané ou faussement spontané	Moyenne
Attitudes personnelles Par ex. : préférer, haïr, aimer, apprécier	Zone habituelle	<i>push</i> dominant ou mixte	Symptôme et/ou Symbole	Moyenne à longue
Attitudes interpersonnelles Par ex. : distant, chaleureux, méprisant	Zone habituelle	<i>pull</i> dominant	Symptôme et/ou Symbole	Courte à moyenne
Traits de personnalité Par ex. : nerveux, timide, anxieux, extraverti	Zone habituelle	<i>push</i> dominant	Symptôme	Longue