

Université de Genève
Faculté de psychologie et des sciences de l'éducation

LA VOIX COMME INDICE AFFECTIF

PAR K. SCHERER et B. ZEI

La parole, en tant qu'instrument vocal d'interaction sociale, comporte non seulement les éléments linguistiques tels que les phonèmes mais aussi des traits porteurs d'information non linguistique. Ces traits, comme par exemple la qualité de la voix ou la prosodie, apparaissent soit spontanément en caractérisant l'état physiologique concomitant à l'état affectif du sujet parlant, soit de façon intentionnelle en caractérisant les attitudes du locuteur ainsi que ses stratégies d'interaction sociale. Ces dernières incluent également les traits que l'émetteur produit dans le but de les faire paraître comme spontanés.

Une analyse acoustique de la voix permet alors de repérer, d'une part, les traits vocaux relatifs aux états affectifs et physiologiques du locuteur et, d'autre part, ses attitudes ou ses stratégies d'interaction sociale.

ESQUISSE DES PROCESSUS PHYSIOLOGIQUES SOUS-JACENTS À LA PRODUCTION DE LA PAROLE

La production de la parole peut être décrite en termes d'une coordination spécifique des éléments mobiles des systèmes respiratoire, phonatoire et articuloire. Chacun de ces systèmes a une fonction spécifique dans l'expression vocale et détermine certains traits acoustiques de la parole:

— Les structures du système respiratoire (en particulier les muscles intercostaux externes, internes et externes obliques, le diaphragme, le muscle droit de l'abdomen et le grand dorsal) contrôlent la pression sous-glottique et règlent le débit de l'air lors de la phonation. Les différents «patterns» respiratoires influencent les variations dans l'intensité et dans la hauteur de la voix ainsi que la continuité et/ou la discontinuité de la parole.

— Les structures du système phonatoire (avec le larynx et les muscles thyro-aryténoïdiens) mettent en vibration le flux d'air passant dans le canal vocal. Leur activité se manifeste acoustiquement sous forme du ton laryngien. Ce ton est un son complexe composé d'une série de sons appelés harmoniques. Le plus grave, et en principe le plus intense, des harmoniques est dit fondamental (F0). Dans la parole, c'est le fondamental qui détermine l'impression de la hauteur mélodique de la voix. Normalement la hauteur du fondamental se situe entre environ 80 et 500 Hz, et varie en fonction de l'âge et du sexe du

locuteur aussi bien qu'en fonction de son intonation ou son état d'éveil affectif. La fréquence du ton laryngien et sa complexité harmonique contribuent également à l'impression de la qualité (timbre) de la voix humaine et varient selon l'état des cordes vocales (leur épaisseur ou tension par exemple) ainsi que d'après le tonus musculaire général.

— Les structures du système articuloire (avec la langue, le pharynx, le voile du palais, les lèvres et le maxillaire inférieur) modifient la forme et le volume des cavités ainsi que la rigidité et/ou la souplesse de leurs parois. Les corrélatifs acoustiques de ces modifications se présentent sous forme de variations spectrales qui sont à la base de la différenciation des éléments phonétiques. Cependant, ces variations déterminent aussi d'autres «qualités» vocales, en particulier le timbre de la voix.

Or, le fonctionnement de ces trois systèmes est influencé par les différents états dans lesquels se trouve l'organisme. Une activation ergotaxique, par exemple, ou un état d'alerte physiologique aussi bien qu'un état affectif particulier peuvent modifier le tonus musculaire et de par ce fait influencer le fonctionnement des trois systèmes mentionnés. Ces derniers à leur tour vont alors modifier certains paramètres acoustiques de la voix (pour le résumé des principaux paramètres acoustiques de la voix humaine, se référer à la liste 1, en annexe).

APPLICATION DE L'ANALYSE DE LA VOIX EN PSYCHIATRIE

Nombre d'études ont signalé que les états psychopathologiques s'accompagnent de caractéristiques vocales très spécifiques (Chevrie-Muller, 1971; Darby, 1981; Fonagy, 1983; Menahem, 1983; Scherer, 1987). Certains auteurs (Moses, 1954; Ostwald, 1963; Eldred & Price, 1958; Bannister, 1972; Roessler & Lester, 1976) rapportent des données témoignant d'un rabaissement de la F0 chez les dépressifs. D'autres études (Tolkmit, Helfrich, Standke, Scherer, 1982), traitant de l'évolution de l'état des patients, signalent que la F0 des schizophrènes et des dépressifs baisse après une période d'hospitalisation et de psychothérapie. Ceci impliquerait une hausse de la F0 pendant la dépression. La figure 1 montre les valeurs de la F0 relevées avant et après la thérapie chez 32 femmes (dont 17 dépressives et 15 schizophrènes) suivies au

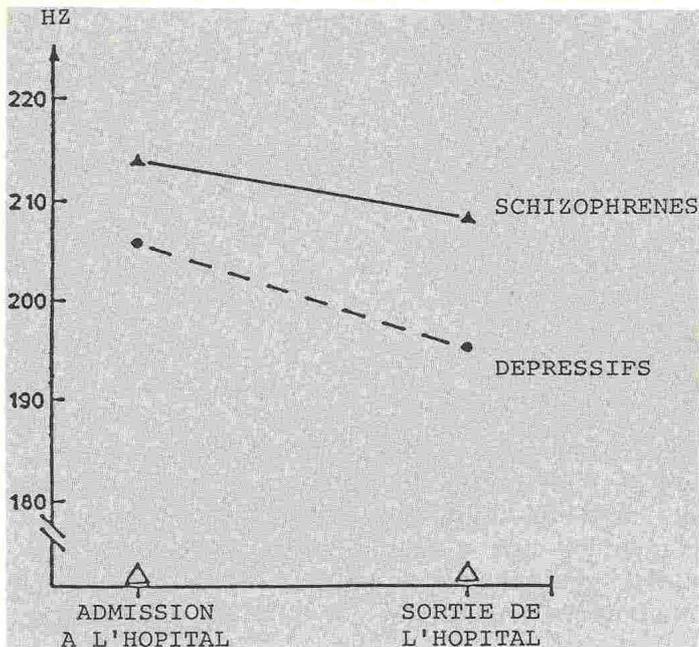


Fig. 1. — Valeurs moyennes de la F0 relevées avant et après la thérapie.

Langley-Porter Neuropsychiatric Institute à San Francisco. Etant donné que la F0 élevée peut provenir d'une hausse de la tension des cordes vocales et que cette dernière à son tour peut être due à une augmentation générale du tonus musculaire, on pourrait conclure qu'au terme du traitement hospitalier, la voix des patients était plus basse parce que la thérapie a eu comme effet une décontraction musculaire générale.

Laver (1980) et Scherer (1979) suggèrent que, pour des raisons physiologiques, une augmentation du tonus de la musculature du système phonatoire entraîne nécessairement une hausse de la fréquence minimum du fondamental (F0 minimum — pour une définition voir la liste en annexe).

Les données concordant avec cette interprétation ont été obtenues récemment dans une étude interdisciplinaire portant sur la dépression, et effectuée à l'Institut Max Planck de psychiatrie (Munich). Les résultats de cette étude mettent en évidence que la baisse de la F0 minimum constitue un des indices sûrs de l'amélioration de l'état du patient (Klos, Ellgring, Scherer, 1987). La figure 2 montre d'une part les valeurs de la F0 minimum et,

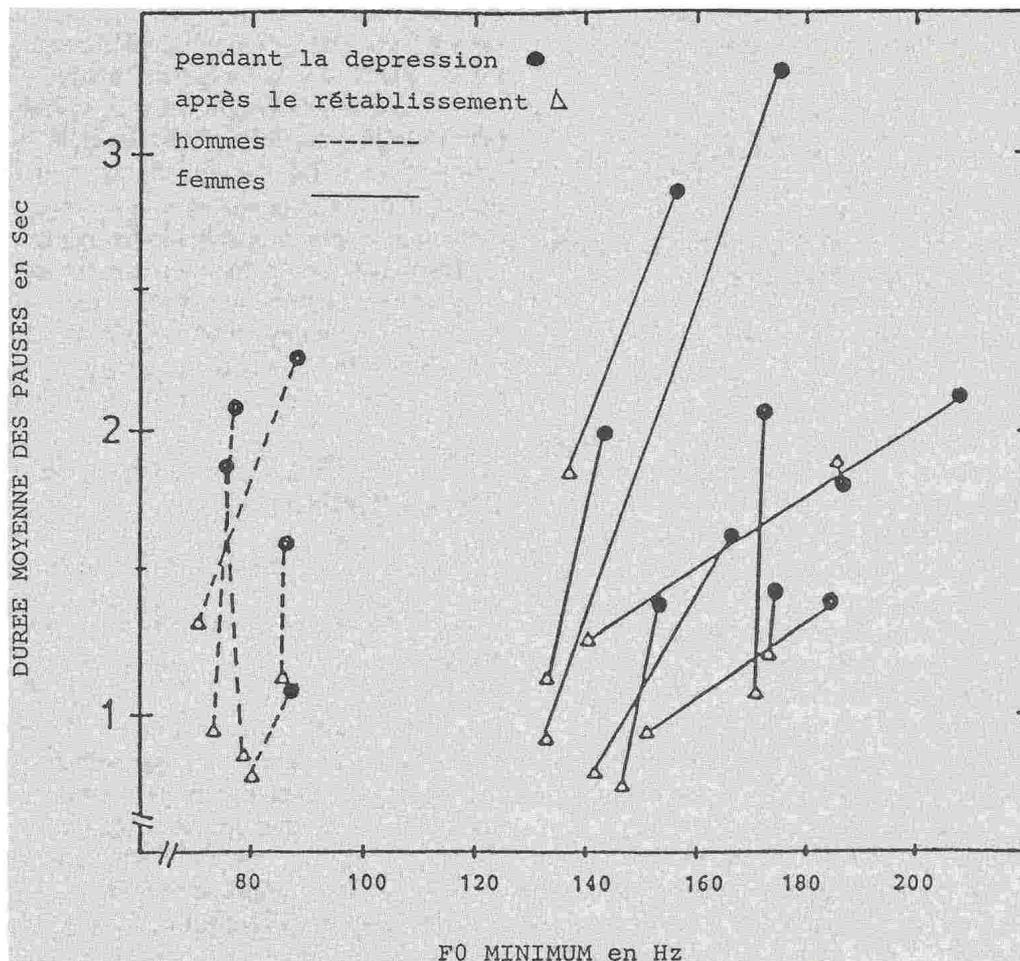


Fig. 2. — Valeurs moyennes de la F0 minimum et de la durée des pauses, relevées pendant la dépression et après le rétablissement.

d'autre part, celles indiquant la durée moyenne des pauses à l'intérieur des énoncés chez 16 patients (11 femmes et 5 hommes) pendant leur dépression et après leur rétablissement.

D'autres études (Newman & Mather, 1938; Zuberbier, 1957; Ostwald, 1964; Hargreaves & Starkweather, 1965; Bannister, 1972) signalent que l'éventail du fondamental (pour une définition voir la liste en annexe) et sa variabilité sont assez réduits chez les dépressifs. Cela pourrait être mis en relation avec les observations cliniques qui précisent que la voix des dépressifs est monotone. L'impression de monotonie est imputable surtout à une stéréotypie des courbes intonatives mais aussi aux aspects plats des contours mélodiques.

Les résultats les plus concordants concernent les effets de la dépression sur l'intensité de la voix. Ainsi a-t-il été trouvé que, en général, les dépressifs ont une voix faible en intensité (Moses, 1954; Zuberbier, 1957; Eldred & Price, 1958; Whitman & Flicker, 1966).

Quant à la qualité de la voix (déterminée par la répartition d'énergie acoustique dans le spectre), on a observé chez des dépressifs (Tolkmitt et al. 1982) que l'énergie acoustique de la zone spectrale de 240-440 Hz augmente après la thérapie, tandis que celle de la zone de 440-920 Hz décroît. Ces résultats se prêtent à l'interprétation suivante: pendant les phases aiguës de la dépression, le tonus musculaire peut être très élevé (Goldstein, 1965; Malmö, 1975) et entraîner des constrictions du pharynx qui, à leur tour, créent les conditions de phonation favorisant les résonances des harmoniques au-dessus de 500 Hz. Par contre, une diminution d'anxiété entraînerait une diminution générale du tonus musculaire avec les changements vocaux correspondant.

La précision de l'articulation de la parole semble être également liée au tonus musculaire. Certaines études (Zuberbier, 1957; Tolkmitt et al., 1982) indiquent que l'articulation des dépressifs souffre d'un manque de précision. Ce phénomène peut provenir soit d'une hypotension, soit d'une hypertension des organes articulatoires, entraînant des problèmes de coordination des mouvements articulatoires. Tolkmitt et al. (1982) ont observé, chez les dépressifs, qu'à la suite d'un traitement psychiatrique, l'articulation de certaines voyelles était plus précise. Cette précision a pu être repérée au niveau acoustique sous forme de changements de valeurs «formantielles» (F1, F2, F3) de certaines voyelles.

D'autres paramètres étroitement liés à l'état affectif du locuteur sont: la vitesse d'élocution, la longueur des pauses et l'accent tonique. L'étude citée plus haut (Klos et al.) a montré que la vitesse d'élocution chez des patients dépressifs s'accélère en moyenne de 23% lors de leur rétablissement. Les auteurs rapportent également un raccourcissement des pauses (de 59% pour femmes et 56% pour hommes). Ce fait est dû à une augmentation des capacités cognitives nécessaires pour la production de la parole (par exemple: la planification syntaxique et les choix lexicaux). Ainsi, dans la figure 2, qui représente les résultats pour chacun des patients observés, on note que la baisse de la FO minimum et le raccourcissement des pauses accompagnent les améliorations des états affectifs chez tous les patients.

Bien que de tels résultats soient encourageants, il faut aussi signaler que tous les résultats obtenus dans ce domaine ne sont pas concordants. Ceci est probablement dû au manque d'homogénéité des symptômes ou des degrés comparables de la gravité du trouble. Chez les

Tableau I. — Prédications des types de la voix correspondant aux différents troubles affectifs.

| | <i>Configuration pharyngienne</i> | <i>Tonus du conduit vocal</i> | <i>Registre</i> |
|---------------------------------------|-----------------------------------|-------------------------------|-------------------------|
| Euphorie | ample | moyennement tendu | légèrement de poitrine |
| Manie | ample | extrêmement tendu | extrêmement de poitrine |
| Ahédonie | très resserré | légèrement tendu | neutre |
| Sentiment d'impuissance | resserré | relâché | de tête |
| Troubles d'anxiété | resserré | moyennement à fortement tendu | de tête |
| Peur agitée | resserré | extrêmement tendu | très de tête |
| Frustration, insatisfaction chronique | resserré | moyennement tendu | moyennement de poitrine |
| Indifférence, apathie | neutre | légèrement relâché | neutre |
| Honte, culpabilité | resserré | légèrement tendu | de tête |

patients dépressifs en phase maniaque, par exemple, on peut s'attendre à ce que la F0 moyenne et l'étendue de la F0 soient élevées, tandis que chez ceux en phase dépressive ou retardés la F0 moyenne et l'étendue de la F0 seront réduites.

Tout en disposant de nombreuses données témoignant d'une forte influence de la dépression sur la voix, il a été rarement possible de fournir une interprétation cohérente des données expérimentales faute de cadre théorique qui le permettrait. Le récent modèle dit «modèle des processus composant l'émotion» (component-process-model) (Scherer, 1981, a, b, c; 1984, a, b) pourrait servir de cadre conceptuel pour une interprétation cohérente des données expérimentales. Ce modèle permet une différenciation des états affectifs d'après plusieurs critères relevant de composantes physiologiques et perceptivo-cognitives de l'organisme réagissant aux stimulus d'origine endogène et/ou exogène. Chacun des états serait le produit d'une évaluation de la situation donnée, effectuée par le sujet de plusieurs points de vue tels que, par exemple: son agrément, sa pertinence à l'égard des buts et/ou des besoins, et la capacité du sujet de la maîtriser. Les différents troubles affectifs peuvent alors être expliqués en termes des mauvais fonctionnements des processus d'évaluation. Par exemple, une estimation trop basse de la capacité de maîtrise d'une situation peut entraîner chez le sujet l'anxiété ou la dépression. Ce modèle permet également d'étudier (et de prévoir) les caractéristiques vocales relatives aux différents états émotionnels. Le tableau I

présente l'effet hypothétique d'un certain nombre de ces états sur trois dimensions centrales de la production de la voix, à savoir: la configuration pharyngienne, le tonus du conduit vocal et le registre de la voix.

Le tableau II présente en termes acoustiques les prédictions des changements vocaux relatifs aux différents troubles affectifs (pour une discussion plus détaillée de ces hypothèses, voir Scherer, 1987).

LES EFFETS MÉDICAMENTEUX SUR LA VOIX

Grâce aux liens qui existent entre l'affectivité, la physiologie et la production de la voix, il est également possible d'étudier les effets des agents psychopharmacologiques sur la voix (Feldstein et Weingaertner, 1981). La figure 3 montre les résultats d'une étude menée par Helfrich, Standke, Scherer (1984) concernant les changements de la hauteur de la voix survenus après l'administration de deux médicaments antidépresseurs et du placebo.

L'analyse de la voix se présente donc comme un outil entièrement inoffensif et facile à utiliser pour le contrôle des effets médicamenteux.

CONCLUSION

Les exemples donnés plus haut montrent que les déterminants psychophysiologiques de la production de la voix, mis en évidence par l'analyse acoustique, peuvent jouer

Tableau II. — Prédications relatives aux changements des paramètres acoustiques de la voix correspondant aux différents troubles affectifs.

| | Euphorie | Manie | Anhédonie | Sentiment d'impuissance | Troubles d'anxiété | Peur agitée | Frustration, insatisfaction chronique | Indifférence, apathie | Honte, culpabilité |
|------------------------|----------|-------|-----------|-------------------------|--------------------|-------------|---------------------------------------|-----------------------|--------------------|
| Perturbation F0 | | | | + | + | + | | | |
| Moyenne de la F0 | + | | | + - | + | ++ | + - | - | + |
| Eventail de la F0 | + | ++ | - | - | + | ++ | - | | |
| Variabilité en F0 | + | ++ | - | - | + | ++ | - | | |
| Contour de la F0 | + | + | | - | + | ++ | - | | + |
| Moyenne du F1 | - | - | + | + | + | + | + | + | + |
| Moyenne du F2 | | | - | - | - | - | - | - | - |
| Domaine du F1 | + - | + - | - | + - | - - | - - | - - | - | - |
| Précision des formants | | + | | - | + | + | + | | + |
| Intensité moyenne | + | ++ | | - - | + | + | + | + - | |
| Eventail de l'int. | + | + | | - | | + | + | | |
| Variabilité de l'int. | + | + | | - | | + | | | |
| Eventail fréquentiel | + | ++ | | + | + | ++ | + | + | |
| Energie de haute fr. | + - | + - | + | + - | + | ++ | + | + - | + |
| Bruit spectral | | | | + | | | | | |
| Débit de la parole | + | + | - | - | + | ++ | | | |
| Temps de transition | - | - | | + | - | - | | | |

+ = augmentation, - = diminution. Symboles répétés indiquent un accroissement du degré prévu du changement. Deux symboles opposés désignent le cas où les types des voix antécédentes exercent des influences opposées.

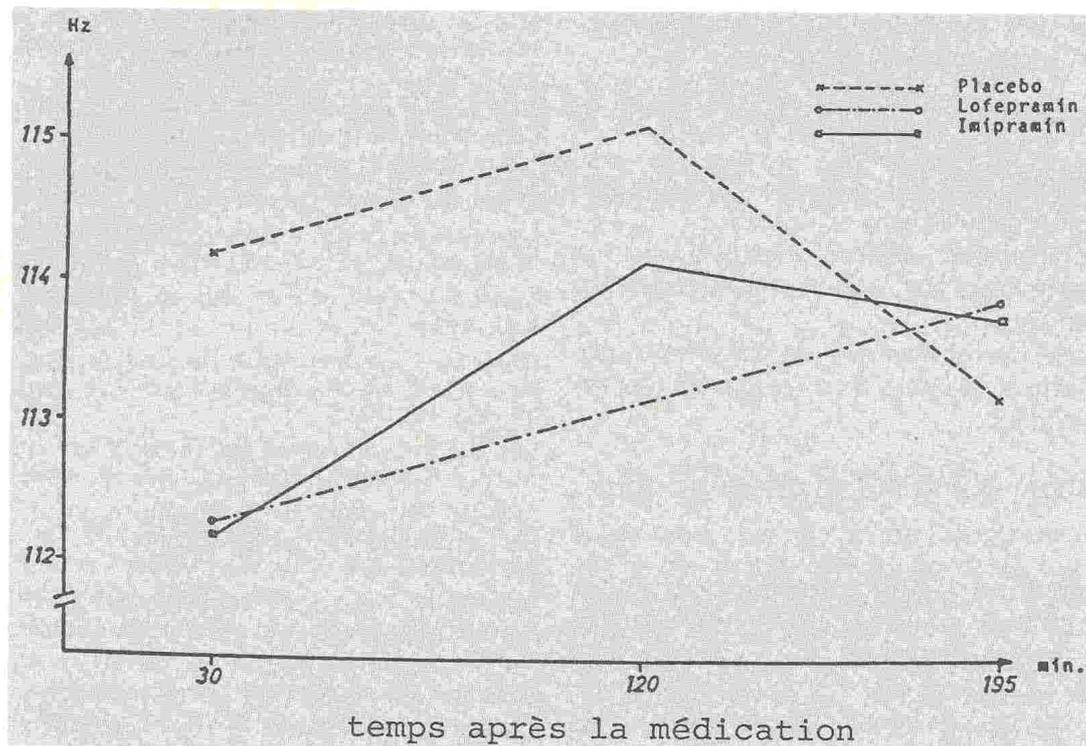


Fig. 3. — Valeurs moyennes de la F0 relevées au moment de la prise du médicament et quelque temps après.

un rôle important dans la compréhension et l'interprétation des données cliniques portant sur les troubles affectifs. Ainsi rend-elle possible d'une part une meilleure différenciation et un meilleur diagnostic des troubles affectifs, et d'autre part un meilleur contrôle de l'évolution de la maladie. L'analyse de la voix peut aussi donner des indications concernant l'efficacité de la thérapie, comme elle peut servir au dépistage précoce des effets secondaires des médicaments (par exemple les troubles moteurs liés à l'absorption de tricycliques). Pour mieux comprendre la nature de la relation qui existe entre l'affectivité et la vocalisation, l'étude clinique des troubles émotionnels devrait faire partie de l'étude de la psychobiologie de l'émotion. Une telle approche exigerait alors des études basées sur diverses observations cliniques dans un cadre interprétatif interdisciplinaire appliqué aux processus affectifs (voir Scherer, 1987). En concluant, il faut signaler que la technologie de l'analyse de la voix est largement accessible grâce à la disponibilité des logiciels appropriés fonctionnant sur les micro-ordinateurs. Elle est aussi facile à effectuer, que ce soit dans des cliniques ou dans des cabinets privés, par un personnel paramédical formé pour ce but.

ANNEXE LISTE DES PRINCIPAUX PARAMÈTRES ACOUSTIQUES DE LA VOIX

I. Traits relatifs à la fréquence fondamentale (F0)

- Perturbations de la F0: légères variations dans les pulsions glottiques produites dans l'intervalle d'une seconde.

- Moyenne de la F0: fréquence moyenne des vibrations des muscles thyro-arythénodiens (dits cordes vocales) dans un énoncé.
- F0 minimum: valeur moyenne des 5% des valeurs minimum des F0 réalisées dans les syllabes atones correspondant au niveau physiologique de la phonation c'est-à-dire le niveau phonologiquement non marqué.
- Etendue de la F0: la différence entre la valeur la plus haute et la valeur la plus basse du fondamental dans un énoncé.
- Variabilité de la F0: mesure de la dispersion des valeurs du fondamental (l'écart-type par exemple).
- Courbe de la F0: variations de la hauteur du son fondamental à l'intérieur d'un énoncé.

II. Traits relatifs au spectre acoustique

- Moyenne du F1: valeur moyenne du premier formant (F1 correspondant à la plus basse concentration d'énergie acoustique dans le spectre) calculée à la base de plusieurs réalisations d'une même voyelle dans un énoncé.
- Moyenne du F2: valeur moyenne du deuxième formant (la deuxième concentration d'énergie acoustique, partant du F0) calculée à la base de plusieurs réalisations d'une même voyelle dans un énoncé.
- Domaine du F1: largeur de la plage d'harmoniques renforcés par résonance et correspondant au premier formant.
- Précision des formants: degré de concordance entre les valeurs formantielles réalisées et celles prévues

selon un modèle standardisé du système phonologique de la langue en question.

III. Traits d'intensité

- Intensité moyenne: valeur moyenne d'énergie des ondes sonores réalisées dans un énoncé.
- Etendue de l'intensité: la différence entre la valeur la plus haute et la valeur la plus basse en intensité mesurées dans un énoncé.
- Variabilité de l'intensité: mesure de la dispersion des valeurs d'intensité réalisées dans un énoncé (l'écart-type par exemple).

IV. Traits relatifs à la distribution d'énergie acoustique

- Eventail fréquentiel: la différence entre la valeur de la F0 et celle du point spectral le plus élevé présentant une énergie acoustique vocale.
- Energie de haute fréquence: proportion d'énergie située dans la région de hautes fréquences (par exemple au-dessus de 1 KHz).

BIBLIOGRAPHIE

- Bannister M.L.: *An instrumental and judgemental analysis of voice samples from psychiatrically hospitalized and nonhospitalized adolescents*. Unpublished Ph.D. thesis, University of Kansas, 1972.
- Chevrie-Muller C., Dodart F., Seguir-Dermier N., Salmon D.: *Etude des paramètres acoustiques de la parole au cours de la schizophrénie de l'adolescent*. *Folia Phoniatica*, 23: 401-428, 1971.
- Darby J.K.: *Speech and voice studies in psychiatric populations*. In: J.K. Darby (Ed.): *Speech evaluation in psychiatry*, pp. 253-284. Grune & Stratton, New York, 1981.
- Eldred S.H., Price D.B.: *A linguistic evaluation of feeling states in psychotherapy*. *Psychiatry*, 21: 11-121, 1958.
- Feldstein S., Weingaertner H.: *Speech and psychopharmacology*. In: J.K. Darby (Ed.): *Speech evaluation in psychiatry*, pp. 379-396. Grune et Stratton, New York, 1981.
- Fonagy I.: *La vive voix*. Payot, Paris, 1983.
- Goldstein I.B.: *The relationship of muscle tension and autonomic activity to psychiatric disorders*. *Psychosomatic Medicine*, 27: 39-52, 1965.
- Hargreaves W.A., Starkweather J.A., Blacker K.H.: *Voice quality in depression*. *Journal of Abnormal Psychology*, 70: 218-220, 1965.
- Helfrich H., Standke R., Scherer K.: *Vocal indicators of psychoactive drug effects*. *Speech Communication*, 3: 245-252, 1984.
- Klos Th., Ellgring H., Scherer K. (submitted for publication): *Vocal changes in depression*. Munich: Max-Planck-Institut für Psychiatrie.
- Laver J.: *The phonetic description of voice quality*. Cambridge: Cambridge University Press, 1980.
- Malmö R.B.: *On emotions, needs, and our archaic brain*. Holt, Rinehart & Winston, New York, 1975.
- Menahem R.: *La voix et la communication des affects*. *L'Année Psychologique*, 83: 537-560, 1984.
- Moses P.J.: *The voice of neurosis*. Grune and Stratton. New York, 1954.
- Newman S., Mather V.G.: *Analysis of spoken language of patients with affective disorders*. *American Journal of Psychiatry*, 94: 913-942, 1938.
- Ostwald P.F.: *Soundmaking: The acoustic communication of emotion*. Springfield: Charles C. Thomas, 1963.
- Ostwald P.F.: *Acoustic manifestations of emotional disturbance*. *Disorders of Communication — Research Publications*, 42: 450-465, 1964.
- Roessler R., Lester J.W.: *Voice predicts affect during psychotherapy*. *Journal of Nervous and Mental Disease*, 163: 166-176, 1976.
- Scherer K.R.: *Nonlinguistic vocal indicators of emotion and psychopathology*. In C.E. Izard (Ed.): *Emotions in personality and psychopathology*, pp. 493-529. Plenum, New York, 1979.
- Scherer K.R.: *Vocal indicators of stress*. In J. Darby (Ed.): *Speech evaluation in psychiatry*, pp. 171-187. Grune & Stratton, New York, 1981a.
- Scherer K.R.: *Speech and emotional states*. In J. Darby (Ed.): *Speech evaluation in psychiatry*, pp. 189-220. Grune & Stratton, New York, 1981b.
- Scherer K.R.: *Wieder die Vernachlässigung der Emotion in der Psychologie*. In W. Michaelis (Ed.): *Bericht über den 32. Kongress der Deutschen Gesellschaft für Psychologie in Zürich*, 1980, pp. 304-317. Hogrefe, Göttingen, 1981c.
- Scherer K.R.: *On the nature and function of emotion: A component process approach*. In K.R. Scherer & P. Ekman (Eds.): *Approaches to emotion*, pp. 293-318. Hillsdale N.J., Erlbaum, 1984a.
- Scherer K.R.: *Emotion as a multicomponent process: A model and some cross-cultural data*. In P. Shaver & L. Wheeler (Eds.): *Review of personality and social psychology* (Vol. 5, pp. 37-63). Beverly Hills, CA: Sage, 1984b.
- Scherer K.: *Vocal assessment of affective disorders*. In J.D. Maser (Ed.): *Depression and expressive behaviour*, pp. 57-83. Hillsdale N.J., Erlbaum, 1987.
- Tolkmitt F., Helfrich H., Standke R., Scherer K.R.: *Vocal indicators of psychiatric treatment effects in depressives and schizophrenics*. *Journal of Communication Disorders*, 15: 209-222, 1982.
- Whitman E.N., Flicker D.J.: *A potential new measurement of emotional state: A preliminary report*. *Newark Beth-Israel Hospital*, 17: 167-172, 1966.
- Zuberbier E.: *Zur Schreib- und Sprechmotorik der Depressiven*. *Zeitschrift für Psychotherapie und Medizinische Psychologie*, 7: 239-249, 1957.

Adresse:

Faculté de psychologie et des sciences de l'éducation, Université de Genève, 1211 Genève 4.