



Physiologie de la voix et paramètres de la phonation

Nathalie Henrich
ICP-CNRS Grenoble

L'appareil phonatoire humain se décompose en trois niveaux :

- n le niveau sous-glottique, qui comprend le diaphragme, les poumons et la trachée,
- n le niveau glottique, qui comprend le larynx dans lequel s'insère les cordes vocales,
- n le niveau supra-glottique, qui comprend la partie haute du larynx, le pharynx, les cavités buccale et nasale.

Le rôle de chaque niveau sur la production d'un son peut se modéliser de façon physique ou mathématique.

Le niveau sous-glottique s'apparente à une soufflerie, et il permet de régler la pression en entrée du système phonatoire (pression sous-glottique P_s).

Le niveau glottique intervient dans la production des sons voisés, où il joue le rôle d'un excitateur acoustique. Le débit d'air qui traverse la glotte au cours du temps est modulé par la vibration des cordes vocales, ce qui génère une onde acoustique qui se propage dans le conduit vocal et qui est rayonnée par les lèvres. Les paramètres qui définissent cette onde acoustique sont principalement la période fondamentale d'un cycle glottique, la durée de l'impulsion glottique relativement à la période (quotient ouvert), l'asymétrie de l'impulsion glottique (coefficient d'asymétrie), son amplitude, et la force du contact à la fermeture (fermeture abrupte ou molle). On peut également y associer les sources de bruit (souffle, raucité).

Au niveau supra-glottique, le conduit vocal joue principalement le rôle d'un articulateur. Il permet la production des voyelles et des consonnes. Dans le cas des voyelles, il fait office de résonateur, et permet de sélectionner les bandes de fréquence à renforcer par ajustement des fréquences et largeurs de bande des résonances acoustiques. Ces renforcements d'énergie visibles sur le spectre d'un signal de parole sont appelés les formants. La position des deux premiers formants définit entièrement l'espace vocalique. Dans le cas des consonnes, des sources aéroacoustiques de bruit sont générées selon la position des constriction. Le conduit vocal est modélisé par sa fonction d'aire, qui varie en fonction de la position des articulateurs. Ses propriétés acoustiques s'en déduisent par calcul de la fonction de transfert entre la pression acoustique à la glotte et la pression acoustique aux lèvres. Le rayonnement aux lèvres est modélisé par un filtre passe-haut.

Les paramètres articulatoires et leurs conséquences acoustiques seront décrits en détail par J. Vaissière. Nous ne les détaillerons donc pas dans cette présentation. Nous nous centrerons sur les paramètres de source glottique.

La qualité vocale est fortement dépendante des caractéristiques de la source glottique. Celle-ci se modélise physiquement par un ensemble de systèmes masse-ressort. Le signal acoustique résultant peut être modélisé par des fonctions mathématiques. Les nombreux modèles mathématiques de signaux que l'on trouve dans la littérature décrivent la source glottique



dans le domaine temporel, à l'aide d'un nombre restreint de paramètres (Doval & d'Alessandro, 1999). La variation temporelle de ces paramètres a une incidence directe sur le spectre du signal de source glottique, et plus globalement sur le son produit (Henrich, 2001). Le quotient ouvert et le coefficient d'asymétrie définissent la position d'un renforcement spectral en basse fréquence, le « formant glottique ». Ils interviennent dans la tension de la voix (qualité relâchée ou pressée d'une voix). Néanmoins seules des variations importantes de ces paramètres (au-delà de 10%) sont réellement perceptibles (Henrich et al., 2003). La force du contact à la fermeture influence la pente spectrale en haute fréquence et elle a donc une grande influence perceptive. La composante bruitée est modélisée par les variations temporelles fines en amplitude et en fréquence (shimmer et jitter), ainsi que par l'ajout d'un bruit (souffle).

Quelles sont les mesures possibles des paramètres de source glottique ? La fréquence fondamentale, ainsi que les paramètres de bruit (jitter, shimmer, bruit additif) peuvent se mesurer directement sur le signal audio en sortie des lèvres. La fréquence fondamentale, le jitter et le quotient ouvert peuvent se mesurer indirectement par électroglottographie, une méthode de mesure non-invasive du contact entre les cordes vocales (Henrich et al., 2004). La mesure des autres paramètres de source nécessite l'estimation préalable du signal de source glottique, par des techniques de filtrage inverse.

B. Doval, C. d'Alessandro, 1999. The spectrum of glottal flow models. Notes et Documents LIMSI 99-07, 22p. (<http://www.limsi.fr/Individu/cda/NDL9907.ps>)

Henrich N. (2001) [Etude de la source glottique en voix parlée et chantée : modélisation et estimation, mesures acoustiques et électroglottographiques, perception](#) PhD thesis, Université Paris 6, 2001.

Henrich N., Sundin G., Ambroise D., d'Alessandro C., Castellengo M. and Doval B. (2003) [Just noticeable differences of open quotient and asymmetry coefficient in singing voice](#), Journal of Voice, vol. 17 (4), pp. 481-494.

Henrich N. , d'Alessandro C. , Castellengo M. and Doval B. (2004) [On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation](#), J. Acoust. Soc. Amer., Vol. 115 (3), pp. 1321-1332.