



HAL
open science

Relations entre évaluation acoustique et perceptive de la qualité de la voix : une approche dynamique

Nicolas Audibert, Claire Pillot-Loiseau, Lise Crevier-Buchman, Sophie Fagniard, Bernard Harmegnies, Kathy Huet, Amelia Pettrossi, Myriam Piccaluga, Reina Remman, Virginie Roland, et al.

► To cite this version:

Nicolas Audibert, Claire Pillot-Loiseau, Lise Crevier-Buchman, Sophie Fagniard, Bernard Harmegnies, et al.. Relations entre évaluation acoustique et perceptive de la qualité de la voix : une approche dynamique. 9èmes Journées de phonétique clinique JPC 2023, Mathieu Balaguer, Lila Gravelier, Julien Pinquier, Jun 2023, Toulouse, France. pp.85-88. hal-04266870

HAL Id: hal-04266870

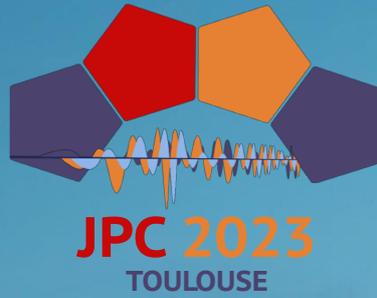
<https://hal.science/hal-04266870v1>

Submitted on 2 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain



9^{ème} Journées de Phonétique Clinique

"Prendre la mesure de la parole"



UNIVERSITÉ TOULOUSE - JEAN JAURÈS
15-17 JUIN 2023

Conception graphique: © Benoît Colas, UT2J/CPRS-DAR. Photographie: © DR.



Relations entre évaluation acoustique et perceptive de la qualité de la voix : une approche dynamique

Nicolas AUDIBERT¹
Claire PILLOT-LOISEAU¹
Lise CREVIER-BUCHMAN^{1,3}
Sophie FAGNIART²
Bernard HARMEGNIES²
Kathy HUET²
Amelia PETTIROSSI^{1,4}
Myriam PICCALUGA²
Reina REMMAN¹
Virginie ROLAND²
Lola TERNY²
Clémence VERHAEGEN²
Véronique DELVAUX²

¹Laboratoire de Phonétique et phonologie UMR 7018, CNRS/Université Sorbonne
Nouvelle, Paris, France.

²Service de Métrologie et Sciences du Langage,
Institut de Recherche en Sciences et Technologies du Langage & Fond National de La
Recherche Scientifique,
Université de Mons, Belgique

³Hôpital Foch : Service de Laryngologie Phoniatre,
92150 Suresnes, France

⁴Imagerie & Cerveau,
UMR1253 INSERM/Université de Tours,
37000 Tours, France

La notion de qualité vocale (QV) est complexe et multiparamétrique. En contexte clinique, on recommande typiquement une évaluation fondée sur une diversité d'outils/indicateurs (Kreiman & Gerratt, 2010 ; Barsties & De Bodt, 2015 ; Brockmann-Bauser & De Paula Soares, 2023), entre autres : évaluation perceptive par le clinicien à l'aide d'échelles (GRBASI, CAPE-V), auto-évaluation par le patient (VHI), indicateurs acoustiques simples (jitter, shimmer, HNR, CPPs, etc.) ou composites (DSI, AVQI), dont l'interprétation est modulée par l'évaluation de la fonction pneumo-phonatoire et l'imagerie des plis vocaux. Deux problématiques non résolues traversent l'histoire de la recherche en évaluation de la QV : (i) la question des relations, non linéaires et non transparentes, entre paramètres acoustiques et jugements perceptifs ; (ii) le lien entre évaluation de la QV et l'échantillon évalué, /a/ tenus vs. échantillons de parole, plus largement la question de la dynamique temporelle de la QV dans la parole.

Dans cette étude exploratoire, nous adressons ces deux questions de façon conjointe en analysant dans quelle mesure un ensemble de paramètres acoustiques dynamiques permettent de prédire les jugements perceptifs sur les mêmes échantillons de parole. Les productions analysées sont issues d'enregistrements effectués par Delvaux et Pillot-Loiseau (2020) à l'aide de MonPaGe (Laganaro *et al.*, 2021), comprenant la phrase « Mélanie vend du lilas » produite par 39 femmes (21-85 ans) et 40 hommes (20-88 ans) normophoniques issus de la région parisienne. Neuf orthophonistes expérimentées ont jugé les phrases à l'aide de l'échelle GRBASI (distribution des grades : *Figure 1*).

Ces jugements perceptifs sont ici exploités et mis en relation avec les analyses acoustiques suivantes : le rapport harmonique sur bruit, le CPPs (Maryn *et al.*, 2009) et le rapport d'énergie entre les bandes 0-1kHz et 1-4kHz, dont la pertinence perceptive a été démontrée (Kitzing, 1986), ont été mesurés sur trois points de chaque segment de la phrase pour 39 femmes et 40 hommes. Des modèles GAM par variable et sexe ont permis de modéliser l'évolution des mesures dans la phrase, en corrigeant l'influence de F0 sur ces mesures.

On constate que le pouvoir prédictif des indices acoustiques par rapport aux altérations perçues évolue au cours du temps (*Figure 2*), c'est-à-dire en fonction du contenu segmental et/ou de la position prosodique, différemment entre hommes vs. femmes (même après correction pour la différence de F0) et de l'indice concerné. Ces résultats seront discutés à la conférence en lien avec la littérature pertinente (par exemple Pouchoulin *et al.*, 2008).

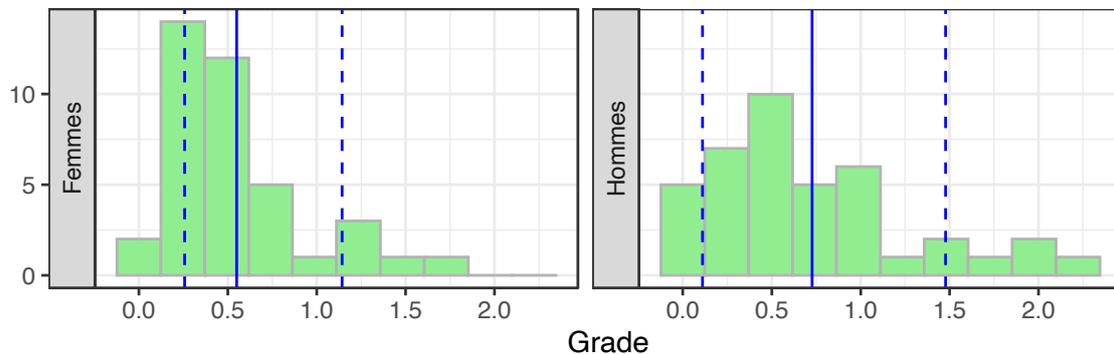


Figure 1. Distribution des grades moyens attribués par les 9 orthophonistes expérimentées aux 39 femmes et 40 hommes pris en compte dans l'analyse. Les lignes verticales bleues indiquent la valeur médiane (trait plein) et les quantiles à 10% et 90% (pointillés).

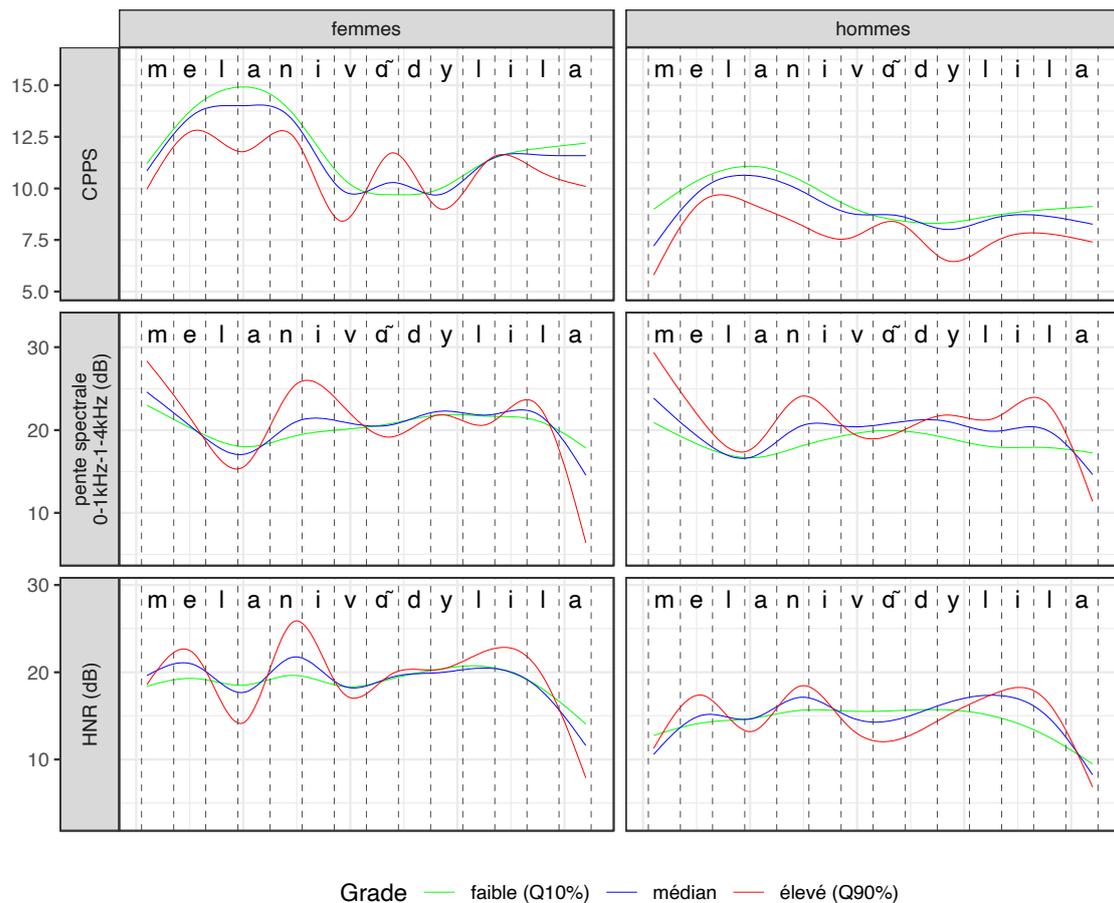


Figure 2. Évolution dans la phrase « Mélanie vend du lilas » de chacune des trois mesures retenues en fonction du grade moyen, obtenue par une modélisation GAM tenant compte des variations de F0, séparément pour hommes et femmes. Les trajectoires sont représentées pour le 1^{er} décile (F : 0,14 ; H : 0,11), la médiane (F : 0,43 ; H : 0,56) et le 9^e décile (F : 1,14, H : 1,44) des valeurs de grade. La couleur de fond rose pâle indique une différence significative uniquement entre grade faible et élevé de cette population sur la portion de signal correspondante, le rose plus foncé une différence significative également entre grade faible et médian.

Références bibliographiques

- Barsties, B., & De Bodt, M. (2015). Assessment of voice quality : Current state-of-the-art. *Auris Nasus Larynx*, 42(3), 183-188. <https://doi.org/10.1016/j.anl.2014.11.001>
- Brockmann-Bausser, M., & De Paula Soares, M. F. (2023). Do We Get What We Need from Clinical Acoustic Voice Measurements ? *Applied sciences*, 13(2), 941. <https://doi.org/10.3390/app13020941>
- Delvaux, V., & Pillot-Loiseau, C. (2020). Perceptual Judgment of Voice Quality in Nondysphonic French Speakers : Effect of Task-, Speaker- and Listener-Related Variables. *Journal of Voice*, 34(5), 682-693. <https://doi.org/10.1016/j.jvoice.2019.02.013>

- Kitzing, P. (1986). LTAS criteria pertinent to the measurement of voice quality. *Journal of phonetics*, 14(3-4), 477-482.
- Kreiman, J., & Gerratt, B. R. (2010). Perceptual Assessment of Voice Quality : Past, Present, and Future. *Perspectives on voice and voice disorders*, 20(2), 62-67. <https://doi.org/10.1044/vvd20.2.62>
- Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fournet, M., Chiuvé, S. C., Lopez, U., Trouville, R., Ménard, L., Burkhard, P. R., Assal, F., & Delvaux, V. (2021). Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French : the MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*, 35(11), 1060-1075. <https://doi.org/10.1080/02699206.2020.1865460>
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *Journal of the Acoustical Society of America*, 126(5), 2619-2634. <https://doi.org/10.1121/1.3224706>
- Pouchoulin, G., Fredouille C., Bonastre JF, Ghio A, Giovanni A. (2008). Analyse Phonétique dans le Domaine Fréquentiel pour la Classification des Voix Dysphoniques. *Journées d'Étude sur la Parole (JEP 2008)*, Avignon.

Remerciements

Ce travail a été soutenu par le Partenariat Hubert Curien (PHC) TOURNESOL n°46241WH et par le Laboratoire d'Excellence (LabEx) *Empirical Foundations of Linguistics* (EFL) n°ANR-10-LABX-0083. Il contribue à l'IdEx Université de Paris (ANR-18-IDEX-0001).