

Evaluation de la qualité vocalique à partir de l'analyse du signal modulé : validation de construit

Anna MARCZYK¹
Benjamin O'BRIEN¹
Alain GHIO¹
Virginie WOISARD²
Pascale TREMBLAY³

¹Laboratoire Parole et Langage,
UMR 7309 CNRS,
Université Aix-Marseille

²URI Octogone-Lordat, Toulouse, France

³Centre de recherche CERVO, Québec, Canada

L'intelligibilité de la parole est portée par des modulations d'amplitude évoluant en fonction du temps et de la fréquence (Rosen, 1992) où les fluctuations à cadences rapides (8-16 Hz) et lentes (1-2 Hz et 4-8 Hz) véhiculent respectivement l'information segmentale et suprasegmentale (Elliott & Theunissen, 2009). L'analyse du signal modulé a été récemment appliquée avec succès à l'évaluation de la parole déviante : dysarthrique (Liss, 2010) et apraxique (Basilakos et al., 2017) et consistait essentiellement à quantifier le degré global d'atténuation des modulations à différentes cadences associée à la réduction de l'intelligibilité (cf. la Figure 1). L'objectif de notre présentation est triple et vise à (i) identifier des fluctuations spectrotemporelles porteuses d'information relative à la qualité vocalique ; (ii) proposer un indice de qualité vocalique basé sur ces fluctuations, et (iii) tester la validité de cet indice sur trois corpus différents de la parole atypique.

Pour cela, nous avons extrait et analysé les productions vocaliques des passages lus « La bise et el soleil » et « La chèvre de monsieur Seguin », préalablement étiquetées et segmentées, dans 3 corpus de parole : (1) la parole vieillissante chez les chanteurs *versus* non-chanteurs (Marczyk et al., soumis) ; (2) la parole dans la maladie de Parkinson (Ghio et al., 2012), et (3) la parole déviante des patients présentant des cancers de la cavité buccale et du pharynx (Woisard et

al., 2020). Le premier de ces corpus a été utilisé pour établir la validité de construit à partir des corrélations positives et négatives entre les mesures classiques de la qualité vocaliques (basées sur la relation entre F1 et F2) et les modulations spectrotemporelles. L'indice résultant de ces analyses était ensuite testé sur les deux derniers corpus dans le but de déterminer son efficacité pour différencier les groupes (pathologique *versus* contrôle) et les effets du traitement (L-Dopa). Les résultats montrent que la qualité vocalique est associée à une atténuation des modulations à cadences lentes (1-2Hz) et une amplification simultanée des modulations à cadences plus rapides (8-16 Hz, 16-32 Hz). L'indice basé sur ce transfert d'énergie permet de différencier les groupes (clinique du groupe contrôle) et les effets du traitement (ON/OFF L-Dopa).

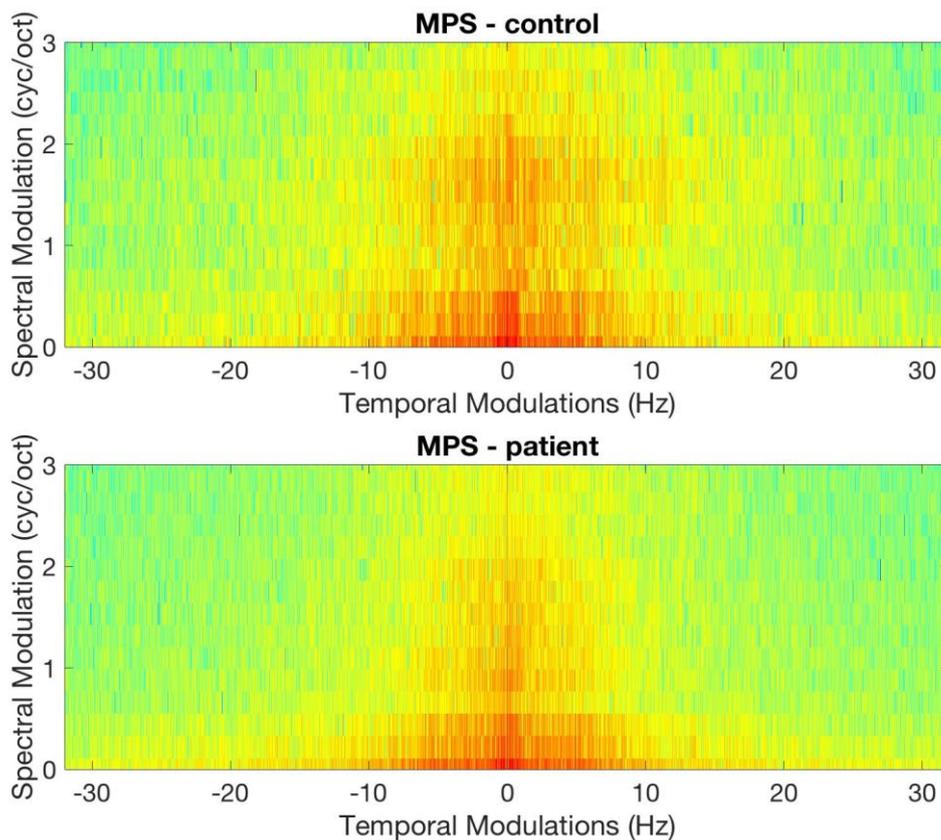


Figure 1. Représentation en modulations spectrotemporelles d'un passage lu par un locuteur contrôle (en haut) et un patient (Parkinson, en bas). Globalement, on observe une atténuation d'énergie (en rouge) à travers les modulations temporelles chez le patient.

Références bibliographiques

- BALAGUER, M., BOISGUERIN, A., GALTIER, A., GAILLARD, N., PUECH, M., & WOISARD, V. (2019). Assessment of impairment of intelligibility and of speech signal after oral cavity and oropharynx cancer. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 136(5), 347–352.
- BASILAKOS, A., YOURGANOV, G., DEN OUDEN, D.-B., FOGERTY, D., RORDEN, C., FEENAUGHTY, L., & FRIDRIKSSON, J. (2017). A Multivariate Analytic Approach to the Differential Diagnosis of Apraxia of Speech. *Journal of Speech, Language, and Hearing Research*, 60, 3378–3392.
- ELLIOTT, T. M., & THEUNISSEN, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Computational Biology*, 5(3).
- GHIO, A., POUCHOUIN, G., TESTON, B., PINTO, S., FREDOUILLE, C., et al. (2012). How to manage sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers?. *Speech Communication*, Elsevier : North-Holland, 2012, 54 (5), pp.664-679.
- LISS, J. M. (2010). Discriminating dysarthria type from envelope modulation spectra. *Journal of Speech, Language, and Hearing Research*, 53(5), 1246–1255.
- MARCZYK, A., BELLEY, E., SAVARD, C., ROY, J.-P., VAILLANCOURT, J., TREMBLAY, P. (soumis) Can amateur singing help maintain articulation: Insights from vowel analysis.
- ROSEN, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 336(1278), 367–373.
- WOISARD, V., ASTÉSANO, C., BALAGUER, M., FARINAS, J., FREDOUILLE, C., et al. (2020). C2SI corpus: a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, Springer Verlag.