L'analyse acoustique des voix d'enfants de 5 ans : Proposition de valeurs de référence pour les logiciels PRAAT et VOCALAB

Etienne Sicard, Léa Meyrieux, Mathilde Moreau, Angélique Remacle

▶ **To cite this version:**

Etienne Sicard, Léa Meyrieux, Mathilde Moreau, Angélique Remacle. L'analyse acoustique des voix d'enfants de 5 ans : Proposition de valeurs de référence pour les logiciels PRAAT et VOCALAB. Journées de Phonétique Clinique 2019, May 2019, Mons, Belgique. hal-02434863

HAL Id: hal-02434863
https://hal.science/hal-02434863

Submitted on 10 Jan 2020

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.

L'analyse acoustique des voix d'enfants de 5 ans : Proposition de valeurs de référence pour les logiciels PRAAT et VOCALAB

Etienne SICARD [1]
Léa MEYRIEUX [2]
Mathilde MOREAU [2]
Angélique REMACLE [2,3]

[1] Etienne.sicard@insa-Toulouse.fr
Département Génie Electrique et Informatique,
INSA de Toulouse, France
[2] Angelique.Remacle@uliege.be
Faculté de psychologie et des sciences de l'éducation,
Université catholique de Louvain, Belgique
[3] FNRS & Unité logopédie de la voix,
Université de Liège, Belgique

Résumé : L'objectif de cette étude est d'identifier les points de convergence/divergence entre les paramètres obtenus à l'aide des logiciels PRAAT et VOCALAB sur un corpus de voix de 53 enfants normophoniques de 5 ans, en se focalisant sur la production des voyelles /a/, /i/, /u/ tenues. L'analyse de la fréquence fondamentale (f_0) montre une corrélation très forte des valeurs obtenues avec PRAAT et VOCALAB, une corrélation élevée sur l'analyse du bruit, et des différences importantes entre les indicateurs d'instabilité en fréquence et en amplitude. Cette étude a permis un ajustement des seuils des indicateurs d'altération de la voix de l'enfant de 5 ans dans VOCALAB, et d'extraire des données statistiques sur les voyelles /a/, /i/ et /u/.

Introduction

En clinique, les mesures acoustiques d'une voix sont comparées à des normes correspondant au genre, à l'âge et à la langue de l'individu. De telles normes sont disponibles pour les adultes mais sont plus rares pour les enfants, particulièrement les francophones en-deçà de 6 ans. Pour répondre à ce manque, Remacle et collaborateurs (2019) ont constitué un corpus de voix de 53 enfants normophoniques de 5 ans (26 filles ; 27 garçons), comportant la production des voyelles /a/, /i/, /u/ tenues.

A partir de ce corpus, des valeurs de références pour les paramètres acoustiques ont été établies sous la version 6.0.05 de PRAAT (Remacle et al., 2019) et VOCALAB (Meyrieux & Moreau, 2018).

L'objectif de cette étude est tout d'abord d'identifier les points de convergence et de divergence entre les paramètres obtenus à l'aide de PRAAT et VOCALAB, sur base du corpus décrit ci-dessus. Le 2ème objectif est de fournir aux orthophonistes cliniciennes des repères précis concernant les indicateurs d'altération de la voix de l'enfant.

Analyse du fondamental de la voix

Un certain nombre d'études traitant de la fréquence fondamentale des voix d'enfants (appelée *Pitch* ou f_0) ont été publiées, telles que (Lee, 1999) ou (Hunter, 2012), montrant une fréquence fondamentale autour de 260 Hz vers 5 ans (Figure 1), pouvant s'élever significativement (400 Hz) selon le contexte, notamment à l'école ou en jouant.

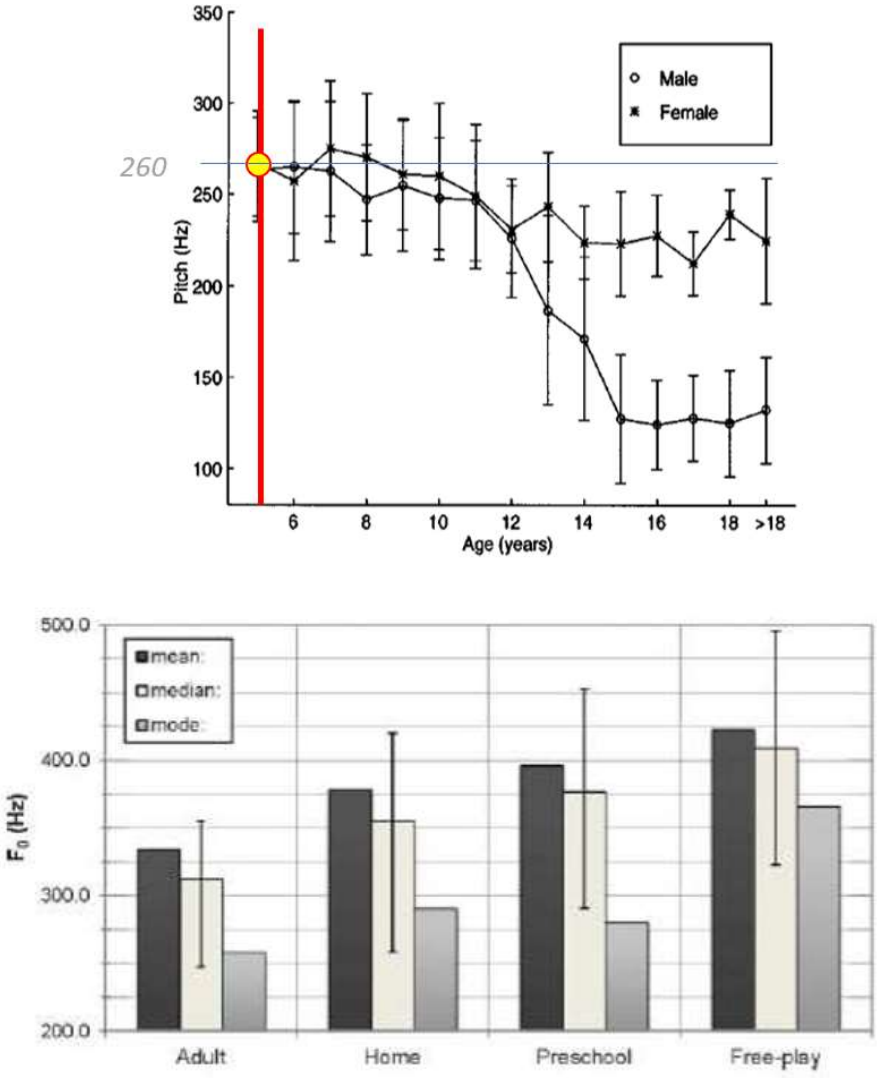

Figure 1 : Analyses de la fréquence fondamentale de la voix de l'enfant (Lee, 1999 en haut), (Hunter, 2012 en bas)

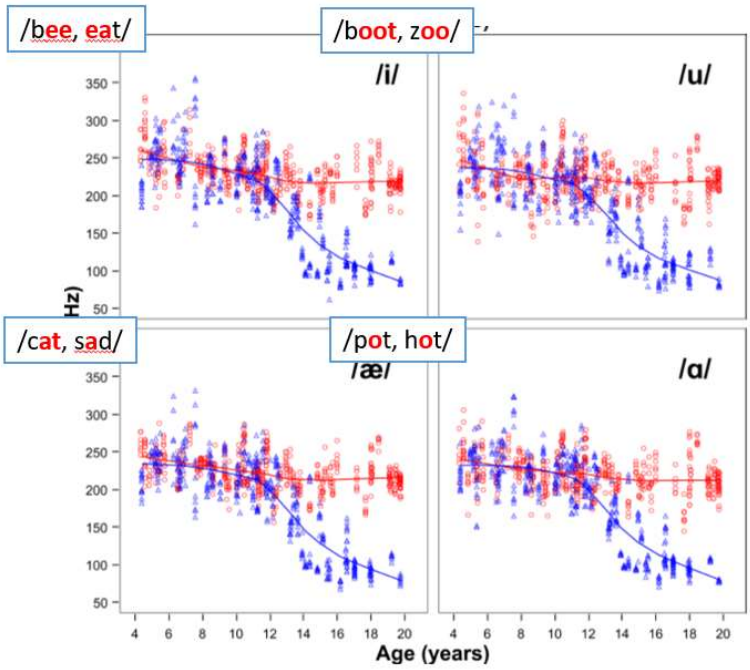

Figure 2 : Différences entre filles et garçons anglophones pour f_0 dans le cas de 4 voyelles distinctes (Vorperian, 2016)

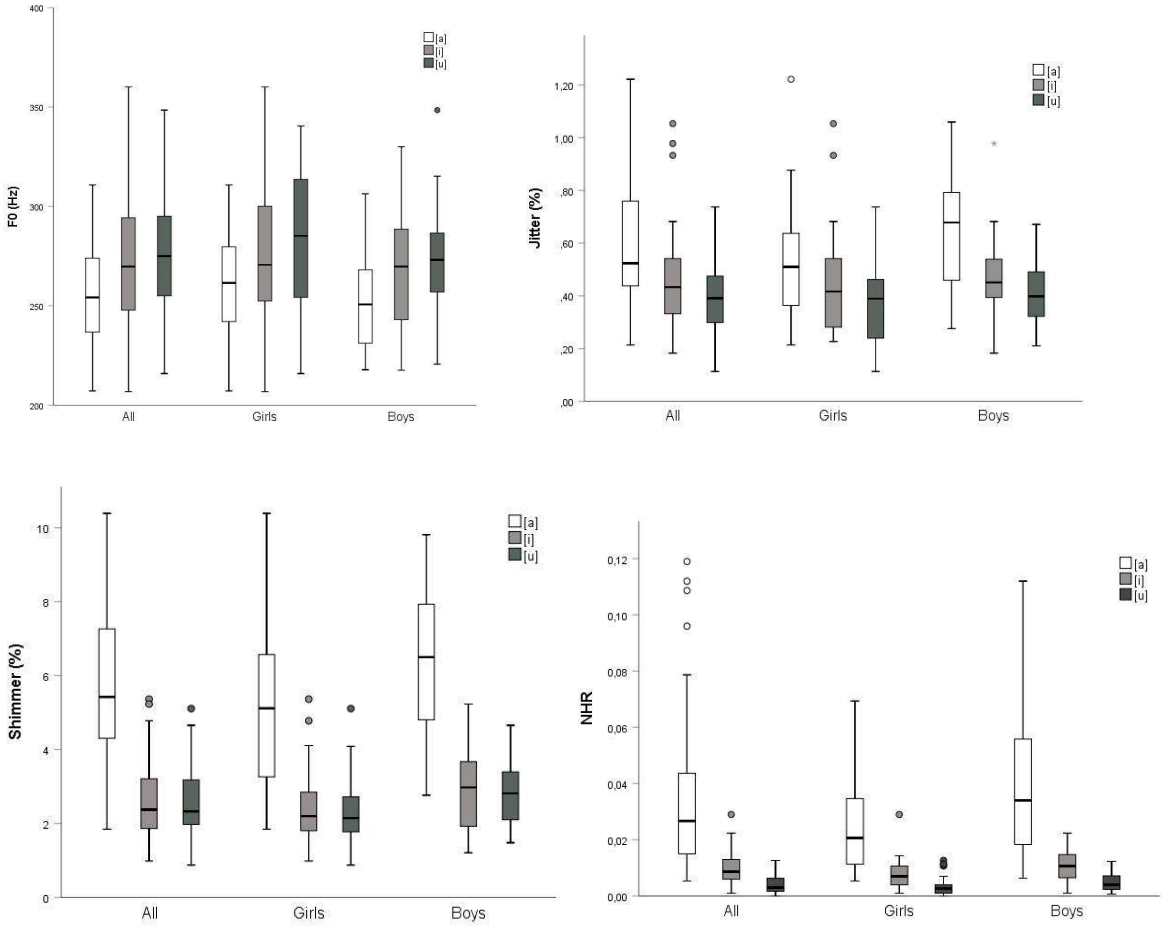

Figure 3 : Extraction avec PRAAT de f_0 , du Jitter, Shimmer et NHR pour 3 voyelles /a:/, /i:/ et /ou:/ prononcées par 50 enfants de 5 ans (Remacle, 2019).

On note des valeurs proches pour filles et garçons jusqu'12 ans environ. L'étude de (Vorperian, 2016) ne montre pas de différence significative entre différentes voyelles (Figure 2), et propose la valeur de 250 Hz à 5 ans.

L'analyse d'un corpus de 50 enfants de 5 ans enregistrés par Genel, Segers et Remacle (Remacle, 2017) a permis d'extraire des valeurs de référence de f_0, autour de 255 Hz pour le /a:/, 271 Hz pour le /i:/ et 277 Hz pour le /ou:/ (Figure 3), ce qui s'approche des valeurs publiées pour des enfants anglophones. Ces calculs ont été effectués avec PRAAT, sur une portion d'une seconde stable au cœur des voyelles.

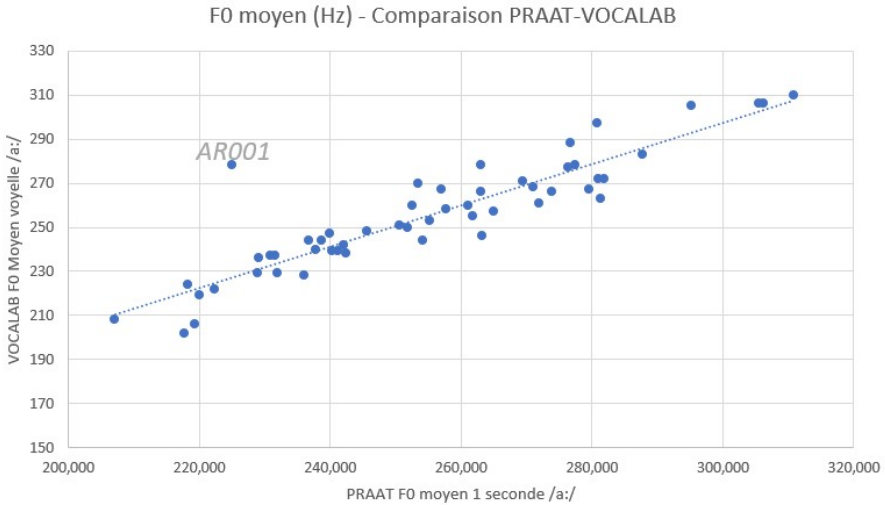

Figure 4 : Corrélation entre f_0 PRAAT et f_0 VOCALAB

La figure 4 présente le résultat de l'extraction de f_0 sur une voyelle /a:/ tenue avec PRAAT (axe horizontal) et VOCALAB (axe vertical). Concernant PRAAT, f_0 a été extrait sur une portion stable d'une seconde au cœur de la voyelle tenue, en évitant les portions instables ou bruitées. Sur VOCALAB, f_0 a été extrait sur l'ensemble de la voyelle.

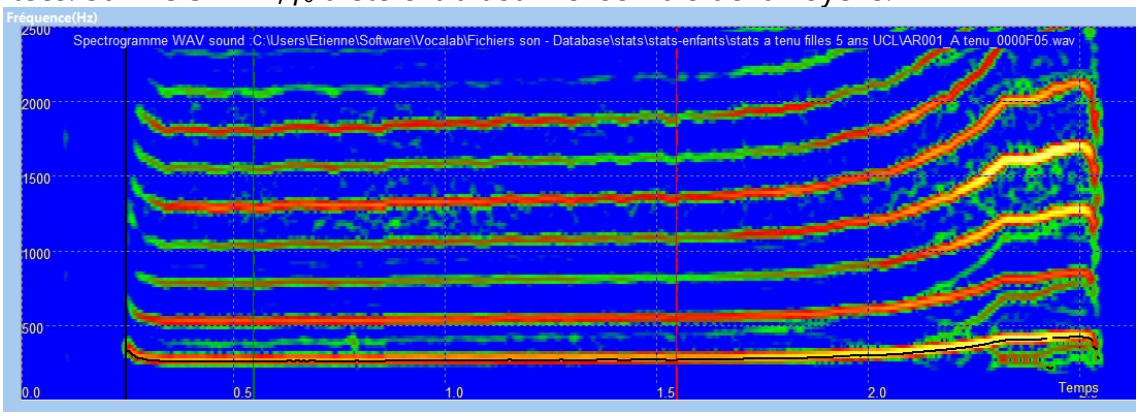

Figure 5 : Spectrogramme du /a:/ du cas AR001

Le cas AR001 qui se distingue dans figure 4 correspond à un /a:/ tenu, qui, malgré la consigne de faire une voyelle tenue sans glissendo, varie en hauteur en fin de production. La valeur extraite par VOCALAB correspond à la moyenne sur l'ensemble de la production du /a:/ (Figure 5). L'analyse du f_0 montre cependant une corrélation de 0.88 (tableau 1) entre des valeurs obtenues avec PRAAT et VOCALAB.

PRAAT Paramètre	PRAAT Moyenne (std)	VOCALAB Paramètre	VOCALAB Moyenne (std)	Corrélation de Spearman
F0 (Hz)	255.12 (24.85)	F0 (Hz)	256.64 (25.30)	r_s=0,888, p=0,01

Tableau 1 : f_0 moyen, déviation et corrélation entre F0 PRAAT et F0 VOCALAB

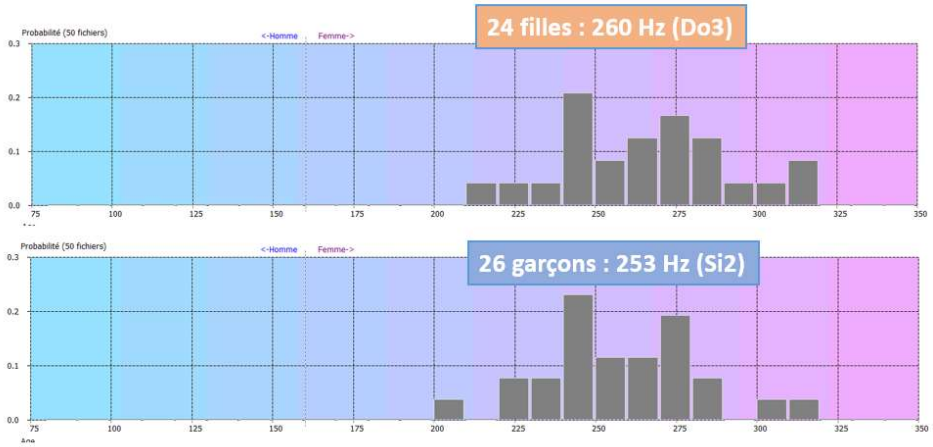

Figure 6 : Histogramme de f_0 moyen du /a:/ tenu pour filles et garçons

Lorsque nous comparons les filles et les garçons, nous n'observons pas de différence significative de f_0 sur /a:/ tenu, seulement de l'ordre de ½ ton sur l'histogramme des valeurs de f_0 de la figure 6, un résultat attendu du fait de l'absence de différences anatomiques du larynx à 5 ans (Remacle et al., 2017).

Instabilité de la hauteur de la voix

Sur PRAAT, le *Jitter* local qui caractérise l'instabilité en hauteur de la voix à court terme a été extrait sur une portion stable d'une seconde au cœur de la voyelle tenue. Sur VOCALAB, le paramètre correspondant appelée « Instabilité en Hauteur » a été extrait sur l'ensemble de la voyelle.

Des différences importantes sont observées entre les indicateurs Jitter -PRAAT et Instabilité en Hauteur – VOCALAB, avec une corrélation particulièrement faible (tableau 2). La figure 7 donne l'exemple de 2 cas où PRAAT considère le Jitter comme faible (0.3 % pour AR005 et 0.6 % pour AR048), tandis que VOCALAB considère l'instabilité en hauteur comme élevée. Ces différences s'expliquent principalement :

- par le calcul court terme de PRAAT (à l'échelle d'un à dix cycles, d'une durée de quelques dizaines de ms),
- par le calcul combinant le court, moyen (100 ms) et long terme (1 sec) de VOCALAB,
- par la portion du signal analysée (1 seconde de signal stable sur PRAAT et la voyelle entière sur VOCALAB).

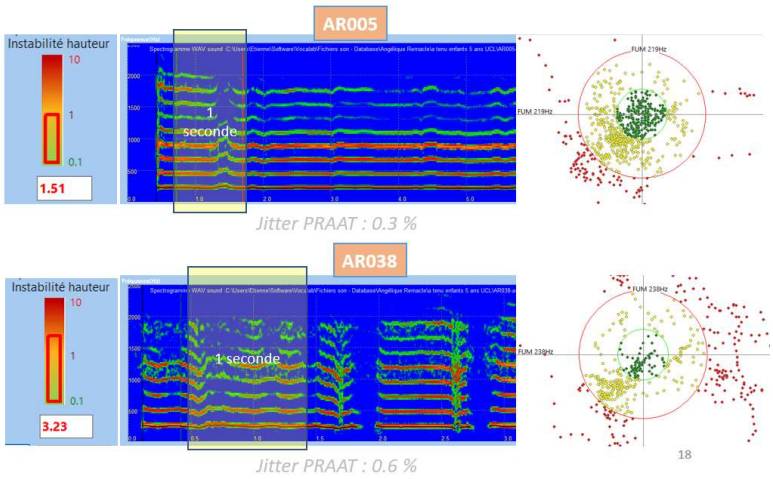

Figure 7 : Jitter-PRAAT et Instabilité en Hauteur – VOCALAB, analyse de 2 cas

Le choix d'une portion stable de signal permet d'éviter les zones de forte instabilité, d'éraillure ou d'interruption de voisement telles que l'on peut les observer notamment sur le cas AR038. Pour une analyse adéquate des indices d'instabilités avec Praat, l'utilisateur doit sélectionner une séquence ininterrompue de cycles. La prise en compte globale de tous les phénomènes au cours de la production de la voyelle ne peut que sanctionner ces indicateurs.

PRAAT PARAMÈTRE	PRAAT MOYENNE (STD)	VOCALAB PARAMÈTRE	VOCALAB MOYENNE (STD)	CORRÉLATION DE SPEARMAN
Jitter Local (%)	0.591 (0.228)	Instabilité Hauteur	1.49 (1.14)	R_s=0,234, P=0,102

Tableau 2 : Jitter PRAAT vs Instabilité en Hauteur VOCALAB, déviation et corrélation

Dû aux différentes portions de signal analysées et aux différents algorithmes utilisés, le tableau 2 illustre aussi des résultats contradictoires en termes d'altération : selon PRAAT, le Jitter est normal (<1%, norme couramment admise), selon VOCALAB, la voix est instable (>1.0, seuil normal/pathologique). La faible corrélation entre Jitter PRAAT et Instabilité en Hauteur VOCALAB a déjà été constatée et discutée dans (Sicard, 2013).

Dans l'illustration de la figure 8, PRAAT ne détecte pas de différence significative dans les périodes cycle à cycle sur le cas AR005, et considère le Jitter comme inférieur à la limite d'altération (1%), tandis que VOCALAB sanctionne l'instabilité en hauteur à cause de phénomènes d'instabilité à moyen et long terme (Indicateur supérieur à 1.0, limite Normal/Altéré).

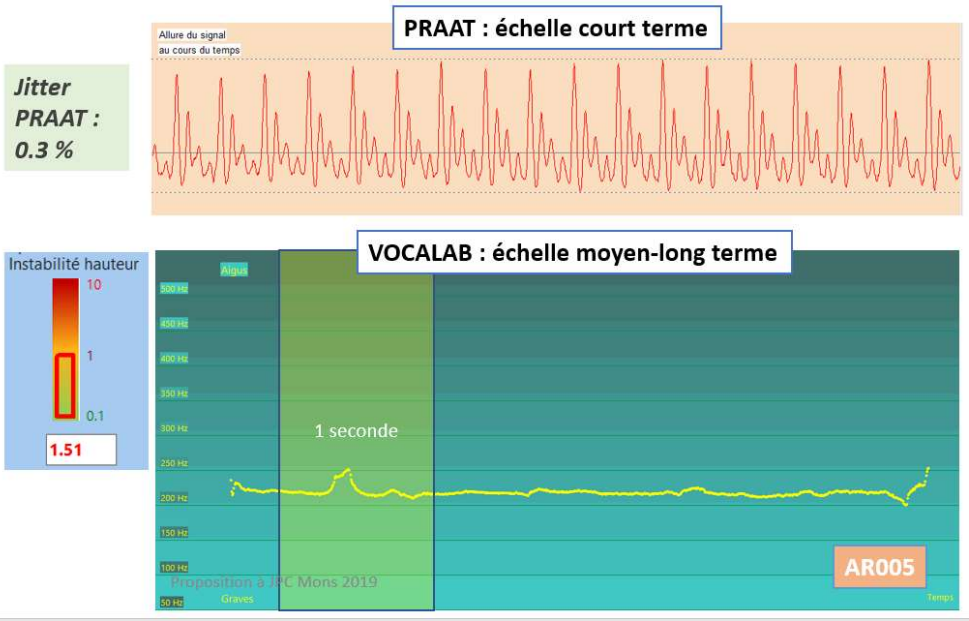

Figure 8 : Jitter-PRAAT et Instabilité en Hauteur – VOCALAB : visualisation à l'échelle cycle-à-cycle et moyen-long terme sur un /a:/ tenu, cas AR005.

Instabilité de l'amplitude de la voix

Sur PRAAT, le *Shimmer local,* qui caractérise l'instabilité en amplitude de la voix à court terme, a été extrait sur une portion stable d'une seconde au cœur de la voyelle tenue. Sur VOCALAB, le paramètre correspondant appelée « Instabilité en Amplitude » a été extrait sur l'ensemble de la voyelle. Là encore, les résultats ne sont pas corrélés (Tableau 3, voyelle /a:/ tenue), mais plus encore, les valeurs sont contradictoires :

- Avec une valeur de 5.8% le *shimmer* PRAAT se situe au-dessus du seuil normal/altéré (autour de 4%)
- Avec une valeur de 0.84, l'instabilité en amplitude VOCALAB se situe au-dessous du seuil normal/altéré (<1.0).

On note une situation inverse à l'instabilité en hauteur, où VOCALAB sanctionnait l'instabilité et PRAAT indiquait un *Jitter* conforme. Ces différences importantes peuvent être liées au mode de calcul (Court-terme pour PRAAT, moyen-long terme pour VOCALAB), à la sensibilité du *Shimmer* au bruit, et au calcul restreint à 1 seconde de signal (attaque et finale exclues) pour PRAAT, et l'ensemble de la production pour VOCALAB.

PRAAT PARAMÈTRE	PRAAT MOYENNE (STD)	VOCALAB PARAMÈTRE	VOCALAB MOYENNE (STD)	CORRÉLATION DE SPEARMAN
Shimmer local (%)	5.824 (2.223)	Instabilité amplitude	0.84 (0.28)	r_s=0,143, p=0,322

Tableau 3 : Shimmer PRAAT vs Instabilité en Amplitude VOCALAB, déviation et corrélation

Bruit/Signal

Nous comparons maintenant le *Noise-to-Harmonics Ratio* (NHR) de PRAAT et le rapport bruit/signal de VOCALAB. Ces deux indicateurs reflètent la quantité de composants apériodiques (bruit) dans la voix. Ils montrent une corrélation relativement élevée pour le /a:/ tenu (Tableau 4). Cependant, le NHR mesuré par PRAAT se situe dans la norme (autour de -25 dB, soit en-dessous de -17 dB considéré comme le seuil normal/pathologique), tandis que VOCALAB considère cet indicateur comme un peu au-dessus de la norme (1.29).

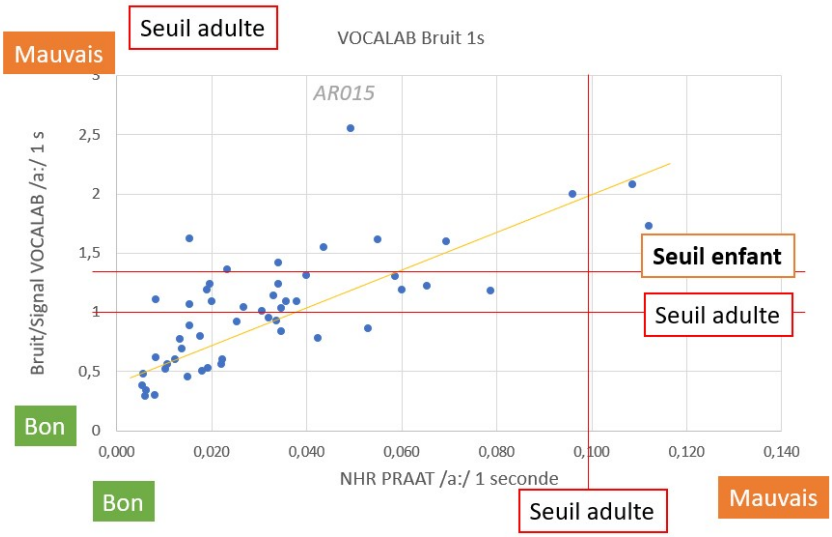

Figure 9 : NHR-PRAAT et Bruit/Harmoniques – VOCALAB : corrélation assez significative avec certains cas singuliers

PRAAT PARAMÈTRE	PRAAT MOYENNE (STD)	VOCALAB PARAMÈTRE	VOCALAB MOYENNE (STD)	CORRÉLATION DE SPEARMAN
NHR	0.034 (0.028)	BRUIT/SIGNAL	1.29 (0.61)	R_S=0,528, P=0,01

Tableau 4 : NHR PRAAT vs Bruit/Signal VOCALAB, déviation et corrélation

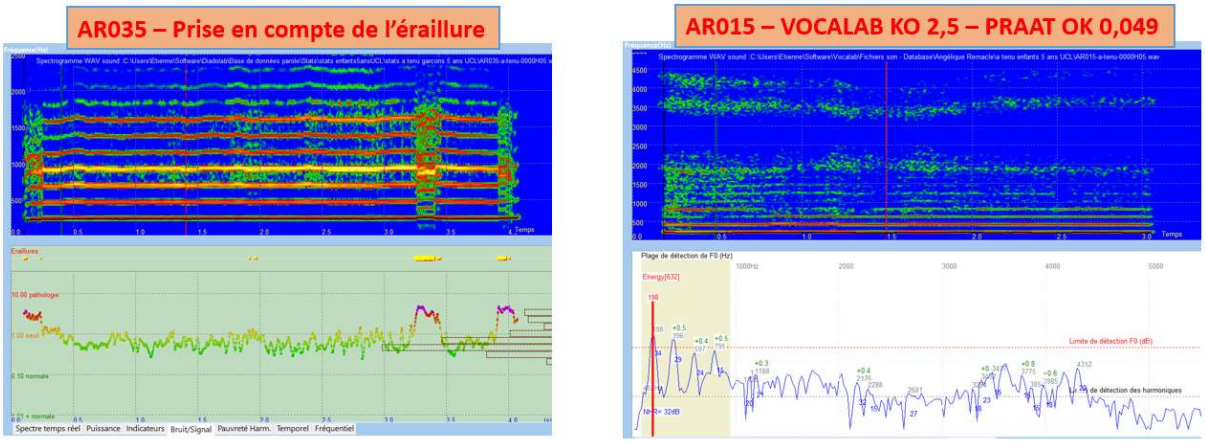

Figure 10 : NHR-PRAAT et Bruit/Harmoniques – VOCALAB : cas AR035 et AR015

Dans la figure 10, nous tentons d'expliciter les différences les plus importantes observées. Dans le cas AR035, le NHR PRAAT est conforme car on a pris soin d'éviter la fin du /a:/ avec une déstructuration harmonique notable correspondant à de l'éraillure. L'indicateur Bruit/Signal porte sur l'ensemble de la production : VOCALAB détecte 3 zones de bruit et d'éraillure, principalement en début et fin de production, et sanctionne l'indicateur correspondant. Dans le cas AR015, PRAAT indique un niveau de bruit faible, ce qui ne se justifie que pour les 3 ou 4 premières harmoniques, les autres étant noyées dans le bruit, comme l'illustre le spectre instantané.

Analyse des voyelles

L'analyse des voyelles /a:/ a été complétée par celle de /i:/ et de /ou:/. La synthèse des résultats est données au tableau 4. La corrélation entre f_0 /i:/, /ou:/ PRAAT-VOCALAB est excellente. On note une corrélation NHR-PRAAT et Bruit/Signal VOCALAB moins nette sur le /i:/, /ou:/ que sur le /a:/. Pour tous les autres indicateurs, les corrélations sont très faibles ou peu significatives.

	PRAAT		VOCALAB		Comparaisons Entre PRAAT et VOCALAB
Voyelle	Paramètre	Moyenne (Std)	Paramètre	Moyenne (Std)	Corrélation De Spearman
/a :/	F0 (Hz)	255.12 (24.85)	F0 (Hz)	256.64 (25.30)	R_s=0,888, P=0,01
/a :/	Nhr	0.034 (0.028)	Bruit/Signal	1.29 (0.61)	R_s=0,528, P=0,01
/a :/	Shimmer Local (%)	5.824 (2.223)	Instabilité Amplitude	0.84 (0.28)	R_s=0,143, P=0,322
/a :/	Jitter Local (%)	0.591 (0.228)	Instabilité Hauteur	1.49 (1.14)	R_s=0,234, P=0,102
/i:/	F0 (Hz)	271.24 (31.86)	F0 (Hz)	271,65 (33,47)	R_s=0,983, P=0,01
/i:/	Nhr	0.009 (0.005)	Bruit/Signal	0,88 (0,62)	R_s=0,247, P=0,083
/i:/	Shimmer Local (%)	2.645 (1.042)	Instabilité Amplitude	0,63 (0,23)	R_s=-0,112, P=0,413
/i:/	Jitter Local (%)	0.457 (0.183)	Instabilité Hauteur	1,79 (1,82)	R_s=0,049, P=0.733
/ou:/	F0 (Hz)	277.35 (31.33)	F0 (Hz)	277,76 (32,20)	R_s=0,967, P=0,01
/ou:/	NHR	0.004 (0.003)	Bruit/Signal	1,27 (0,52)	R_s=0,306, P=0,030
/ou:/	Shimmer Local (%)	2.571 (0.954)	Instabilité Amplitude	0,53 (0,15)	R_s=-0,037, P=0,799
/ou:/	Jitter Local (%)	0.394 (0.141)	Instabilité Hauteur	1,84 (2,2)	R_s=0,298, P=0.035

Tableau 4 : Analyse des 3 voyelles tenues du triangle vocalique, comparaisons PRAAT et VOCALAB

Repères enfant dans VOCALAB

La synthèse des valeurs obtenues pour le /a:/ tenu pour les 5 indicateurs d'altération de la voix de VOCALAB est reportée figure 11 pour les filles et figure 12 pour les garçons. Comme attendu, on note peu de différence entre les 2 sexes. Les indicateurs VOCALAB suggèrent une grande richesse harmonique, du bruit surtout chez les garçons, une certaine instabilité en hauteur et une légère altération de l'attaque (Tableau 5).

Figure 11 : Ensemble des indicateurs d'altération du /a:/ tenu calculés par VOCALAB pour 25 filles de 5 ans

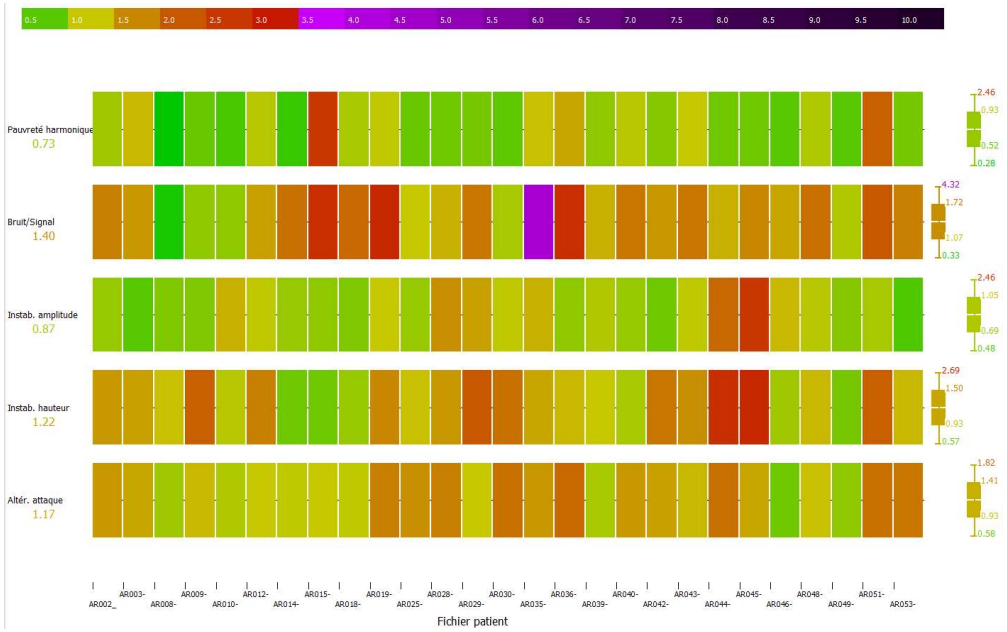

Figure 12 : Ensemble des indicateurs d'altération du /a:/ tenu calculés par VOCALAB pour 27 garçons de 5 ans

GENRE	VOCALAB ALTERATION ATTAQUE	VOCALAB INSTAB. HAUTEUR	VOCALAB INSTAB. AMPLITUDE	VOCALAB BRUIT/SIGNAL	VOCALAB PAUVRETE HARMONIQUE
FILLE	1.14	1.32	0.93	1.16	0.66
GARCON	1.17	1.22	0.87	1.40	0.73
NORME ADULTE	<1.0	<1.0	<1.0	<1.0	<1.0
DIFFERENCE	+	++		++	

Tableau 5 : Valeurs des 5 indicateurs d'altération de la voix de VOCALAB et comparaison à la norme adulte

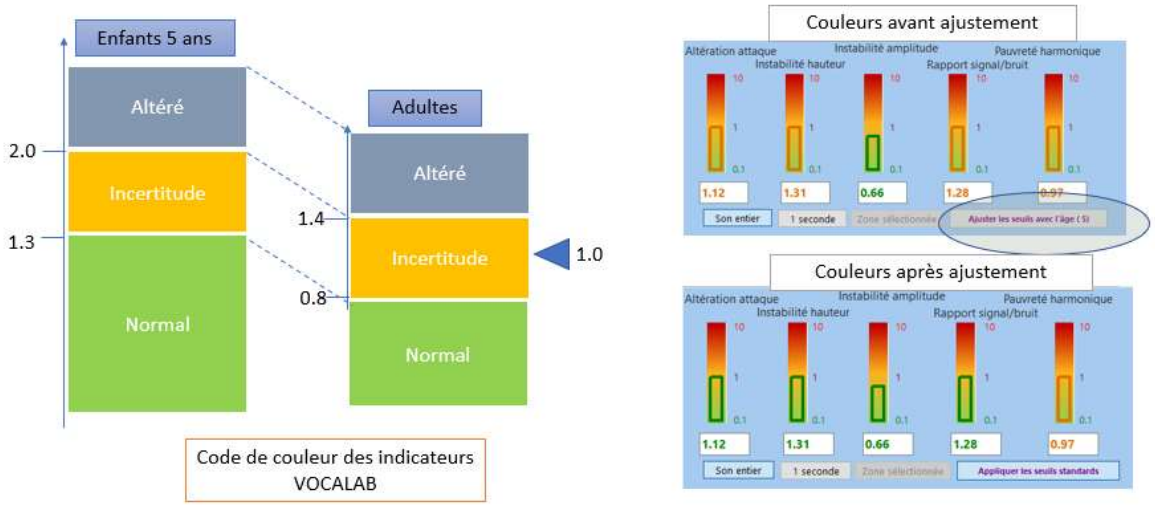

Figure 13. Recoloriage des seuils Normal/Altéré selon l'âge de l'enfant

La synthèse de l'analyse acoustique du /a:/ produit par des enfants de 5 ans (tableau 5) permet d'intégrer des repères d'indicateurs d'altération de la voix de jeunes enfants dans VOCALAB. L'approche proposée, identique à celle des voix âgées (Barbera, 2016), consiste à ne pas modifier les valeurs des indicateurs, mais à proposer un recoloriage des seuils voix normale/altérée (Figure 13) pour certains indicateurs :

- Bruit/Signal : 1.3 = vert au lieu de 0.8
- Instab. Hauteur : 1.3 = vert au lieu de 0.8
- Altération de l'attaque : 1.3 = vert au lieu de 0.8

Les seuils des deux autres indicateurs ne sont pas modifiés :

- Instabilité en amplitude : inchangée
- Pauvreté harmonique : inchangée

L'extrapolation des repères de l'enfant de 5 ans jusqu'à la puberté permet d'adapter les seuils Normal/Altéré et Altéré/pathologique de manière linéaire de 5 ans (cette étude) à 15 ans (voix adulte). Cette approche simpliste justifie la construction d'un corpus d'enfants âgés de 4 à 15 ans pour affiner les repères pour les différents âges intermédiaires.

Conclusion

L'analyse d'un corpus d'enregistrement de voyelles d'une cinquantaine de voix d'enfants de 5 ans a été conduite à l'aide de deux logiciels PRAAT et VOCALAB. Elle a permis d'extraire, d'une part des valeurs de référence pour les deux outils, d'autre part de mettre en évidence des similitudes et des différences dans les résultats obtenus. Des différences significatives entre PRAAT et VOCALAB ont été observées concernant le *Jitter*, le *Shimmer*, tandis que les résultats ont été concordants concernant la fréquence fondamentale et l'analyse du bruit. L'interprétation des différences observées porte principalement sur la différence de mode de calcul : court terme pour PRAAT, moyen/long terme pour VOCALAB.

Bien que les indicateurs de VOCALAB ne soient validés que pour le /a:/, des analyses sur le /i:/ et /ou:/ ont aussi conduites, et différentes tendances similaires ont été observées entre PRAAT et VOCALAB.

Très peu de différences entre filles et garçons ont été observés, ce qui était attendu à cet âge. Une élévation significative de 3 indicateurs sur le /a:/ tenu (l'altération de l'attaque, l'instabilité en hauteur et le bruit/signal) a été constatée dans VOCALAB, tandis que deux indicateurs restaient proche des repères des voix adultes (pauvreté harmonique et instabilité en amplitude). Les seuils normal/altéré des indicateurs d'altération de la voix de jeunes enfants dans VOCALAB ont été modifiés en conséquence, sous forme d'un ajustement de la palette des couleurs, afin de fournir aux orthophonistes des repères précis pour les voix d'enfant.

Références bibliographiques

BARBÉRA, O., FRENAY, C. (2016). Acquisition d'une population adulte tout venant pour la caractérisation de l'évolution de la voix au cours du vieillissement par le logiciel VOCALAB : questionnement des limites entre pathologie et normalité. *Mémoire d'orthophonie, Ecole d'orthophonie de Lyon*. Juin 2016

CARTEI, V., COWLES, W., BANERJEE, R., et al. (2014) Control of voice gender in pre-pubertal children, *British Journal of Developmental Psychology*, VOL. 32, 2014, 100–106.

HUNTER, E. (2012). Impact of four nonclinical speaking environments on a child's fundamental frequency and voice level: A preliminary case study. *Language, Speech, and Hearing Services in Schools*, VOL. 43(3), 253-263.

LEE, S., (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, VOL. 105(3), 1455-1468.

MEYRIEUX, L. & MOREAU, M. (2018) Caractérisation acoustique des productions vocales d'enfants de 5 ans à l'aide du logiciel Vocalab, *Mémoire de Master en logopédie*. Université Catholique de Louvain, Belgique. 2018. Promoteur : REMACLE, A.

REMACLE A., GENEL Y., SEGERS M. (2017), Caractérisation de la voix d'enfants francophones euphoniques âgés de 5 ans, *Congrès de la SFP & Journées Thérapeutiques de la*

voix, Nice, France.

REMACLE, A., GENEL, Y., SEGERS, M., & DE BODT, M. (2019) Vocal characteristics of 5-year-old children: Proposed normative values based on a French-speaking population, *Logopedics, Phoniatrics, Vocology*, VOL. 16, 2019, 1-9.

SICARD, E., & MENIN-SICARD, A. (2013). Implémentation dans VOCALAB d'indicateurs objectifs de la qualité de la voix dans le cadre de l'évaluation de la voix, in *REEDUCATION ORTHOPHONIQUE, Ortho édition*, 2013, 23-27.

VORPERIAN H.K., KENT R.D., AUSTIN D. (2016) "Normative Lifespan Database for Vowel Formants", *Motor Speech Conference*, 2016, USA