



HAL
open science

Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

Etienne Sicard, Anne Menin-Sicard

► **To cite this version:**

Etienne Sicard, Anne Menin-Sicard. Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique. 2021. hal-03177645

HAL Id: hal-03177645

<https://hal.science/hal-03177645>

Preprint submitted on 23 Mar 2021

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.

Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

Etienne Sicard (1,2), Anne Menin-Sicard (2)

INSA-GEI, 135 Avenue de Rangueil, 31077 Toulouse (1)

Laboratoire LURCO www.lurco.unadreo.org (2)

Résumé : Nous abordons dans cet article l'analyse acoustique de la prosodie, en se focalisant notamment sur les notions de fréquence fondamentale, d'intonation, d'étendue vocale et de phonétogramme. Nous tentons de dégager des valeurs de référence pour chacun de ces paramètres acoustiques, pour hommes, femmes et enfants. Une sélection de résultats et analyses statistiques puisées dans la littérature scientifique en archives ouvertes, de mémoires d'orthophonie et de projets de recherche est ensuite proposée, en lien avec l'altération de la prosodie selon différentes pathologies. Nous illustrons les différents concepts à l'aide des logiciels VOCALAB et DIADOLAB, outils cliniques de prise en charge des troubles de la voix et de la parole.

Mots clé : orthophonie, évaluation de la parole, prosodie, dysprosodie, intonation, étendue vocale, VOCALAB, DIADOLAB, PRAAT, Fondamental, FUM, phonétogramme, SRP, speech range profile, altération de la voix, altération de la parole, pathologies de la voix, temps réel, rééducation orthophonique.

Introduction

La prosodie est considérée comme la musique de la parole. Elle favorise l'intelligibilité du discours, permet d'attirer l'attention sur des mots importants et permet de transmettre des émotions ainsi que nuances d'intention. D'après le Dictionnaire d'Orthophonie [Brin-Henry 2018], la prosodie est « l'ensemble des faits supra-segmentaux (intonation, accentuation, rythme, mélodie, tons) qui accompagnent, structurent la parole et qui se superposent aux phonèmes (aspect segmental) ». Nous pouvons donc considérer la prosodie comme l'une des couches d'une structure hiérarchisée où la série de phonèmes constitue le niveau de base, les syllabes le niveau intermédiaire et l'information prosodique le niveau supérieur. Le mot «*prosodie*» vient du grec ancien, où il était utilisé pour une «chanson chantée avec de la musique instrumentale». Plus tard, le mot a été utilisé pour la modulation de la voix humaine en lisant la poésie à haute voix [Nootboom 1997]. Notre approche de la prosodie dans cet article se restreint à des propriétés spécifiques de la parole, au-delà de la simple séquence des phonèmes et des syllabes. Nous nous intéressons en particulier à la modulation de la hauteur et de l'intensité de la voix ainsi que l'allongement ou le rétrécissement de la durée des syllabes.

La dysprosodie

La parole dysprosodique se retrouve dans une large variété de pathologies de la voix (lésions des cordes vocales, atteinte du nerf laryngé supérieur, etc.) ainsi que de troubles neurologiques et psychiatriques, tels que l'autisme, la schizophrénie, les lésions corticales, la maladie de Parkinson, l'ataxie cérébelleuse, la maladie d'Alzheimer, etc.

<i>Indice principal</i>	<i>Lésion typique associée</i>
<i>Hauteur, intensité, débit</i>	Brain stem, Cerebellum, thalamus, striatum, cortex
<i>Contrôle moteur</i>	Striatum, thalamus, cortex
<i>Syndromes comportementaux (apathie, irritabilité, impulsivité)</i>	Striatum, cortex
<i>Apraxie prosodique</i>	Cortex

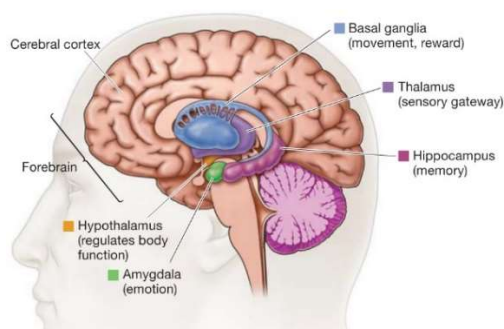


Table 1 : Synthèse d'une approche neurocomportementale de l'étude de la dysprosodie [Sidits 2003]

Un lien entre les altérations prosodiques et les zones du cerveau mise en jeu a été établi par [Sidits 2003], dont nous donnons ci-dessus la table de synthèse. Dans un premier chapitre, nous abordons l'analyse acoustique de la prosodie, en se focalisant notamment sur les notions de fréquence fondamentale, d'intonation et d'étendue vocale, avant de présenter le phonétogramme. Dans une deuxième partie, nous résumons quelques résultats et analyses statistiques puisées dans la littérature scientifique en archives ouvertes en lien avec l'altération de la prosodie selon différentes pathologies.

Pour illustrer nos propos, nous puisons notamment dans les cas de la base de données de l'ERU 15 du laboratoire LURCO/UNADREO, en archives ouvertes [ERU 15], que nous analysons avec les logiciels VOCALAB [Gerip 2020] et DIADOLAB [Gerip 2020b]. Nous présentons des résultats statistiques en lien avec les mémoires d'orthophonie encadrés, notamment celui de [Barbera & Frenay 2016] [Bardelang 2020] [Meyrieux & Moreau 2018], ainsi que par l'exploitation de corpus de voix d'enfant [Remacle 2017], du corpus C2SI en lien avec le cancer ORL [Sicard 2017] et du corpus Voice4PD sur la Maladie de Parkinson [Sicard 2021].

Analyse acoustique de la prosodie

Trois paramètres fondamentaux se dégagent de l'analyse acoustique de la prosodie [Lacheret 2011][Di Cristo 2013]:

- **La fréquence fondamentale** : l'estimation de la fréquence de vibration des cordes vocales à un instant donné,
- **L'intensité** : l'énergie contenue dans le signal à un instant donné,
- **La durée** : mesure d'un intervalle de temps nécessaire pour émettre un segment sonore donné.

Une fois ces informations extraites du discours, nous pouvons extraire la valeur du fondamental, sa distribution statistique, mais aussi l'étendue vocale ou encore le phonétogramme. Nous illustrerons ces notions dans les prochains paragraphes.

Le fondamental F0

Prenons l'exemple d'une portion de phonème /on/, extrait du mot /bonjour/ (figure 1). On constate une périodicité du son, c'est-à-dire une forme qui se répète de façon régulière au cours du temps, même si de légères variations sont observées entre chaque cycle successif. La période de récurrence du son de la voyelle vaut ici environ 5 ms, calculée entre deux pics du signal, dans le sens horizontal qui représente le temps. En inversant cette valeur, on trouve la fréquence correspondante notée F0, égale ici à 200 Hz environ. Il s'agit de la fréquence fondamentale, tout à fait dans la norme d'une voix de femme en parole spontanée.

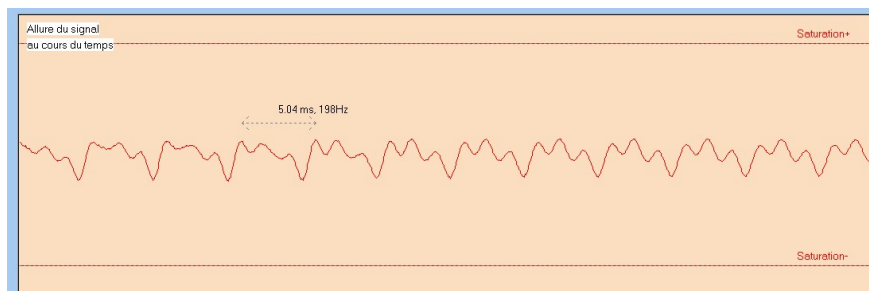


Figure 1 : Allure temporelle d'une portion du phonème /on/ de /Bonjour/ et illustration de la période fondamentale.

La fréquence fondamentale existe dans tous les sons voisés, nous l'appellerons F0 ou encore fréquence fondamentale de la voix et de la parole. Elle correspond au tracé coloré de l'écran Prosodie de DIADOLAB (Fig. 2 en haut), où le temps est en axe horizontal, la fréquence en vertical. Dans PRAAT, elle apparaît selon les mêmes axes en bleu en superposition du spectrogramme, avec une échelle spécifique reportée à droite (Figure 2 en bas).

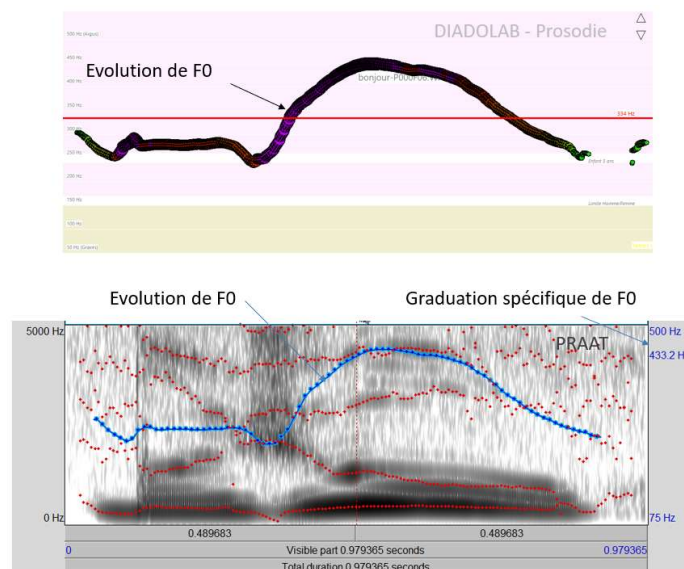


Figure 2 : Détection de la fréquence fondamentale dans DIADOLAB (en haut) et dans PRAAT (en bas), mot /Bonjour/ prononcé par un enfant de 6 ans



Figure 3 : Superposition du contour de F0 et des segments de voyelles et consonnes

Si l'on superpose la segmentation des phonèmes (Figure 3), on constate une variation très importante de F0 sur le phonème /ou/, avec une montée suivie d'une descente, ainsi qu'un allongement significatif (470 ms au lieu de 120 ms en moyenne). La notion d'énergie est symbolisée par la grosseur des ronds, on constate donc aussi sur le /ou/ un maximum d'énergie, ce qui renforce encore l'accent prosodique. L'accent peut donc être une augmentation de hauteur mais aussi d'intensité.

Valeurs de référence

Le fondamental de la voix varie en fonction de l'âge et du genre, comme illustré figure 4 [Vorperian 2016]. Si l'on ne constate que peu de différence selon les voyelles, on note une décroissance de F0 avec l'âge de 4 à 12 ans, une séparation fille/garçon après la mue, vers 12 ans, un F0 autour de 100 Hz pour les hommes à 20 ans, en légère augmentation vers 120 Hz ensuite, et un F0 autour de 220 Hz pour les femmes à 20 ans, qui descend vers 180 Hz ensuite. Concernant la voix adulte et âgée, le mémoire de [Barbera Frenay 2016] montre, sur une centaine de cas, F0 autour de 120 Hz pour les hommes et 200 Hz pour les femmes, avec une tendance à une légère augmentation de F0 pour les hommes et diminution pour les femmes avec l'âge (Figure 5).

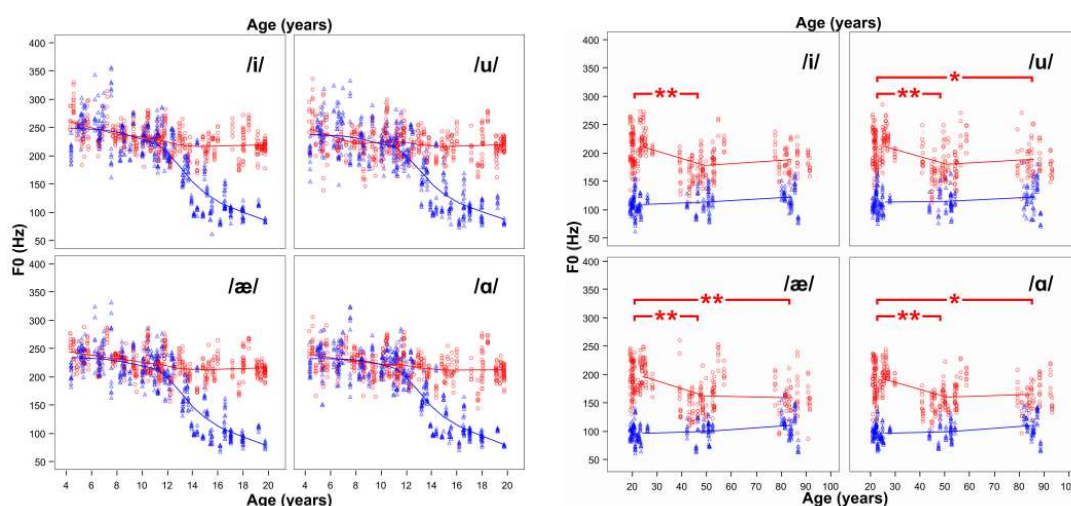


Figure 4 : Evolution de la fréquence fondamentale avec l'âge pour hommes, femmes et enfants, pour une sélection de 4 voyelles [Vorperian 2016]

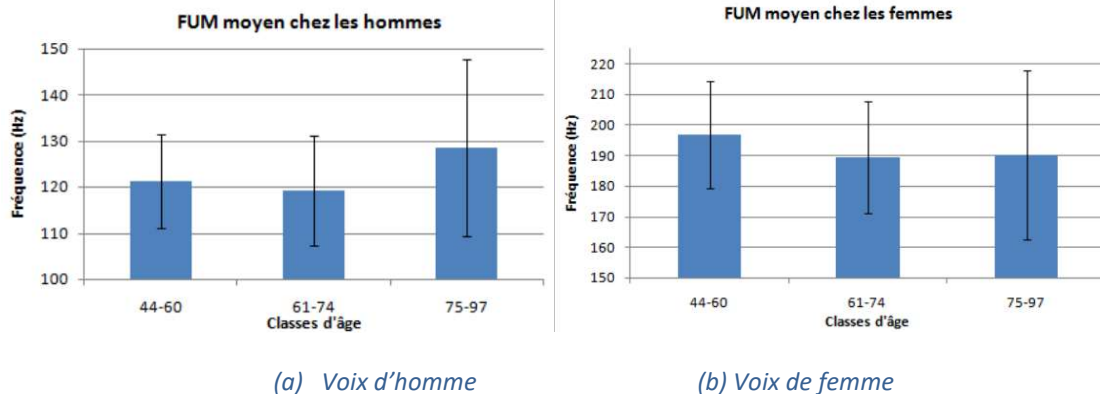


Figure 5 : Evolution du FUM extrait de la voix conversationnelle de 20 hommes et 50 femmes ayant une voix normale, en fonction de l'âge [Barbera & Frenay, 2016]

Genre	FUM	Note au clavier	Plage de variation
Homme	120 Hz	Si1	85 à 140 Hz
Transition H/F	150 Hz	Ré2	140 à 160 Hz
Femme	200 Hz	Sol2	150 à 230 Hz
Enfant	250 Hz	Si2	220 Hz à 350 Hz

Table 2 : Repères standards pour le FUM, le passage de registre et l'étendue vocale pour les hommes, femmes ou enfants

Les valeurs de référence sont reportées à la table 2. Nous considérons une zone de transition entre 140 et 160 Hz, correspondant à la zone d'incertitude entre hommes et femmes. En simplifiant un peu, la voix est considérée comme masculine si F0 est entre 85 et 140 Hz, féminine entre 150 Hz et 230 Hz, puis enfantine au-delà de 250 Hz. Le FUM des enfants de 5 ans est autour de 250 Hz [Sicard 2019], autour de 220 Hz à 8 ans, avant de s'abaisser vers 200 Hz à l'approche de la puberté.

Fondamental usuel moyen

La cumulation des valeurs de F0 peut être dessinée sous forme d'un histogramme, comme reporté à la figure 6 à droite. Chaque point est positionné sur l'axe vertical des fréquences, un peu à la façon d'un boulier. Plus il y a de points de valeur proche, plus les boules s'accumulent et la probabilité (axe X) augmente.

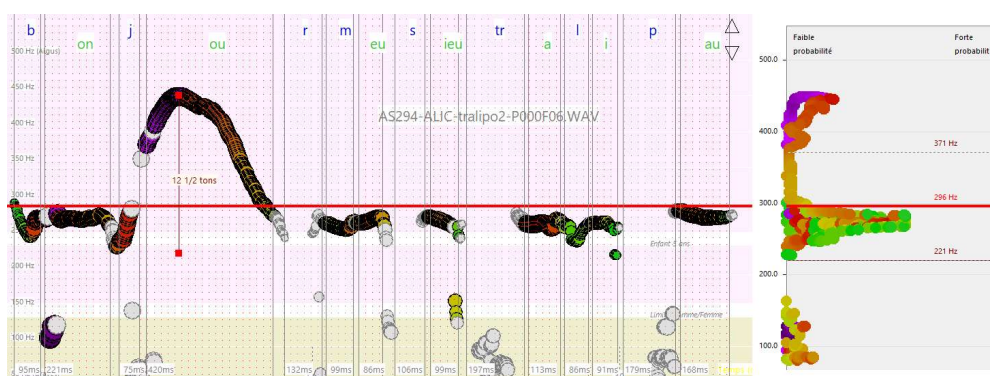


Figure 6 : Histogramme de fréquence fondamentale permettant d'extraire le FUM [Menin-Sicard 2019]

Le fondamental usuel moyen (FUM) est en gros la valeur moyenne des valeurs instantanées plausibles de F0. Par plausible nous entendons des valeurs de F0 ni trop hautes, ni trop basses, ni trop aléatoires d'une valeur à l'autre. Le FUM apparaît comme un trait rouge, ici autour de 300 Hz, ce qui est tout à fait conforme pour un enfant de 6 ans.

Le FUM est assez dépendant de la tâche demandée et du protocole d'enregistrement. Selon qu'il s'agit d'une voyelle isolée, de voix conversationnelle, de texte lu, déclamé, d'une tâche d'énumération ou de description, les valeurs peuvent varier significativement. En outre, le patient peut tenter de s'ajuster à une hauteur correspondant au modèle, même si elle ne correspond pas à son registre de confort. Par exemple, si le modèle est celui d'une voix de femme, un homme pourrait augmenter la hauteur de sa voix.

Nous recommandons de s'en tenir à un protocole précis, avec des instructions et des conditions matérielles et logicielles identiques d'un enregistrement à l'autre. Dans le protocole d'analyse de la voix de VOCALAB [Sicard 2013] [Menin-Sicard 2016], l'extraction du fondamental usuel moyen se fait en temps réel ou en temps différé, sur la base d'une présentation en parole spontanée (« Bonjour je m'appelle...j'ai ...ans et je viens vous voir car »). L'histogramme apparaît alors sous forme d'une représentation statistique des valeurs de F0 selon les notes d'un clavier. La plage de variation de F0 pour le cas AS004 (Figure 7) est de 95 à 250 Hz. VOCALAB évalue le FUM à 155 Hz, qui correspond à la moyenne des F0 les plus fréquents.

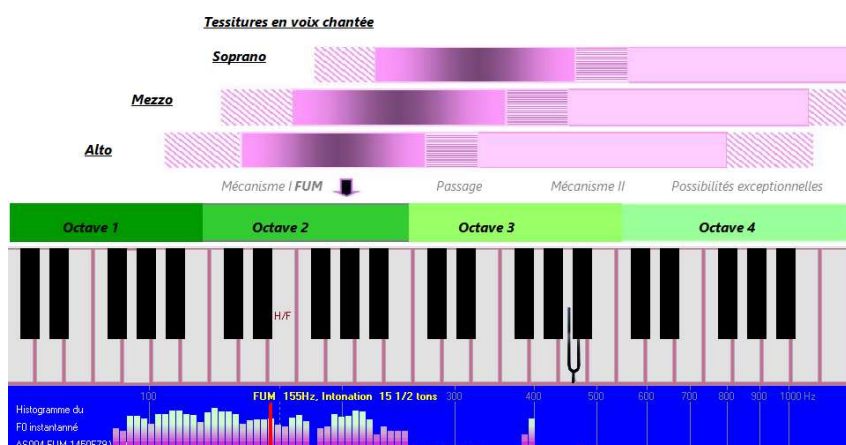


Figure 7 : Histogramme de fréquence fondamentale sur la parole spontanée, cas AS004, permettant d'extraire le FUM [Menin-Sicard 2016]

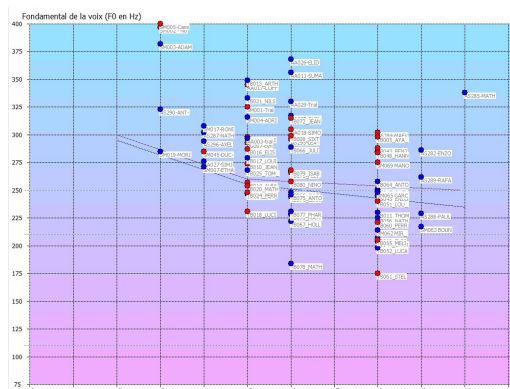


Figure 8 : Evolution du FUM extrait d'une phrase courte /Bonjour Monsieur Tralipau/, parole normale, enfants de 3 à 9 ans montrant la décroissance du FUM. Garçons en bleu, filles en rouge.

Sicard, Menin-Sicard, Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

En reportant le FUM calculé sur chaque production de phrase pour 70 enfants normo phoniques en fonction de l'âge, on constate une décroissance du FUM avec l'âge, et une nette séparation des genres dès 6 ans. On remarque aussi, toujours dans le cas d'un phrase courte telle que /Bonjour Monsieur Tralipau/, une légère élévation du FUM par rapport aux analyses sur des voyelles isolées. Ainsi, à 5 ans, le FUM est autour de 280 Hz, un peu au-dessus de 265 Hz mesuré sur des voyelles [Sicard 2017].

Intonation

Plus la base de l'histogramme présenté du F0 est large, plus la voix peut être considérée comme expressive. A l'inverse, plus la base de l'histogramme est étroite, plus la voix est monotone. Une intonation diversifiée est un indicateur positif. Cette valeur, évaluée par VOCALAB en nombre de demi-tons, renseigne sur les performances vocales, mais aussi sur une possible pathologie du patient. Une intonation limitée peut être signe d'une lésion des cordes vocales, une immobilité liée à la lésion du nerf laryngé supérieur, ou encore une pathologie neurodégénérative. Dans l'exemple de la figure 9, un homme de 75 ans ayant été opéré suite à un cancer ORL (Corpus C2SI) parle avec une voix monotone, la plage principale de variation de F0 étant estimée à seulement 5 ½ tons [Sicard 2017].

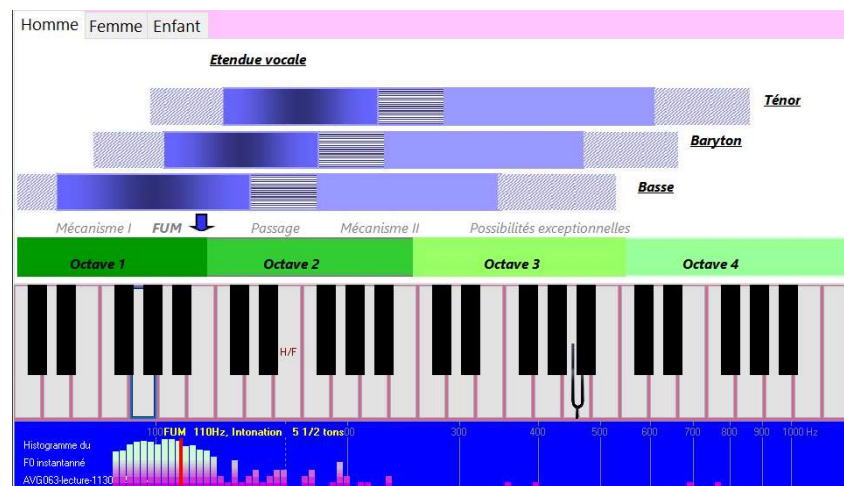


Figure 9 : Histogramme de fréquence fondamentale sur la parole spontanée (fréquence X en log), cas AVG063 du corpus C2SI, avec une intelligibilité faible et une intonation peu variée [Sicard 2017]

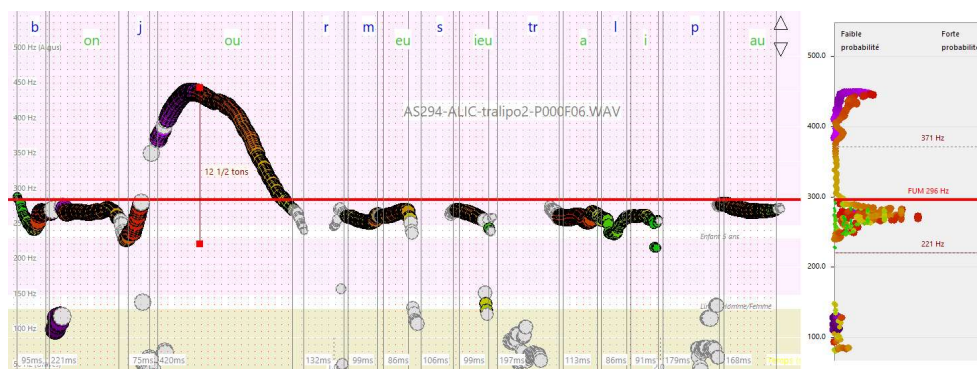


Figure 10 : Tracé de F0 au cours du temps et histogramme de F0 sur la phrase /Bonjour Monsieur Tralipau/, cas AS294 fille de 6 ans, accent conforme sur le /ou/, variation de l'ordre de 12 ½ tons. Fréquence Y en linéaire.

Si l'on prend l'exemple de la phrase « Bonjour Monsieur Tralipau » [Menin-Sicard 2019] de la figure 10 (Cas AS294, fille de 6 ans), on note sur l'histogramme de droite que le fondamental est principalement entre 200 et 300 Hz, mais avec un accent tonique vers 430 Hz, situé sur le /ou/. L'accent est matérialisé par un rond rouge. Dans cette représentation, la fréquence est en vertical, et en échelle linéaire, contrairement à la figure 10 où la fréquence est horizontale et en échelle logarithmique. L'écart-type est estimé à 70 Hz environ. Nous préférons la notion d'intonation en ½ tons pour s'affranchir de valeurs repères qui seraient fonction du genre. L'intonation est évaluée ici à 12 ½ tons, en sélectionnant la hauteur commune la plus basse et la hauteur la plus haute de l'accent sur /on/. Elle correspond à l'excursion fréquentielle de F0, dans sa zone confortable, en éliminant les valeurs extrêmes isolées.

Un exemple d'intonation réduite est donné ci-dessous (fille de 5 ans, cas SM022). L'histogramme est très resserré autour du FUM avec seulement 3 ½ tons de variation. On note sur la figure 11 que l'accent tonique est décalé sur /Monsieur/ et le discours est hésitant, avec une pause détectée et un débit lent.

L'analyse d'autres phrases telles que /Est-ce que tu connais Glochin l'écureuil ?/ [Bardelang 2020] permet aussi de constater des disparités importantes de contour prosodique. Dans le cas TB049 de la figure 12, nous constatons une faible variation de F0 (3 ½ tons en moyenne), très peu d'accent en début et en fin de phrase, celle-ci n'étant alors plus perçue comme interrogative. La même constatation est effectuée pour le cas TB059.



Figure 11 : Tracé de F0 au cours du temps et histogramme de F0 sur la phrase /Bonjour Monsieur Tralipau/, cas SM022 fille de 6 ans, accent non conforme, variation très faible de F0, pause inappropriée

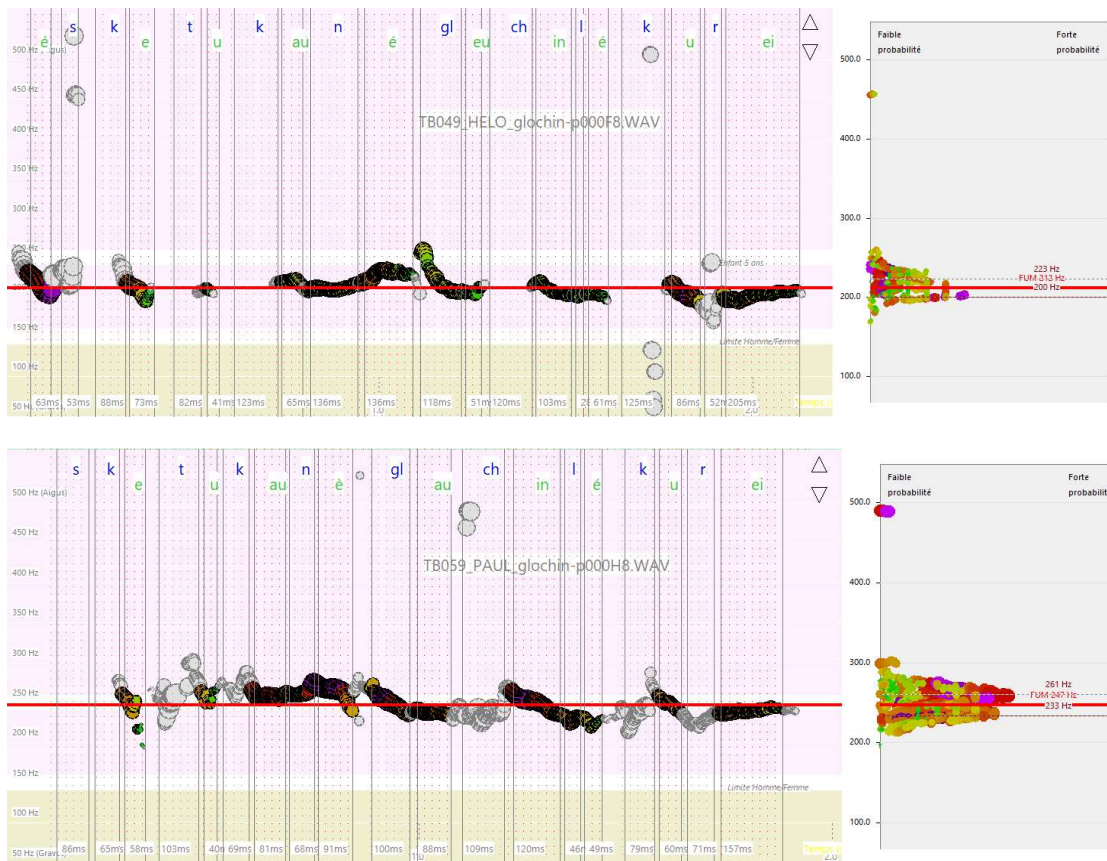


Figure 12 : Tracé de F0 au cours du temps et histogramme de F0 sur la phrase /Est-ce que tu connais Glochin l'écureuil/, cas TB049 fille de 8 ans et TB059 garçon de 8 ans, très variation faible de F0 avec un histogramme resserré autour du FUM

Valeurs de référence

Il existe relativement peu de publications précisant des valeurs de référence concernant l'intonation, à l'exception notable de [Pedersen 2015] (Figure 13), depuis laquelle nous pouvons extrapoler des valeurs de variation de F0 normales, altérées et pathologiques selon l'âge, pour de la parole spontanée.

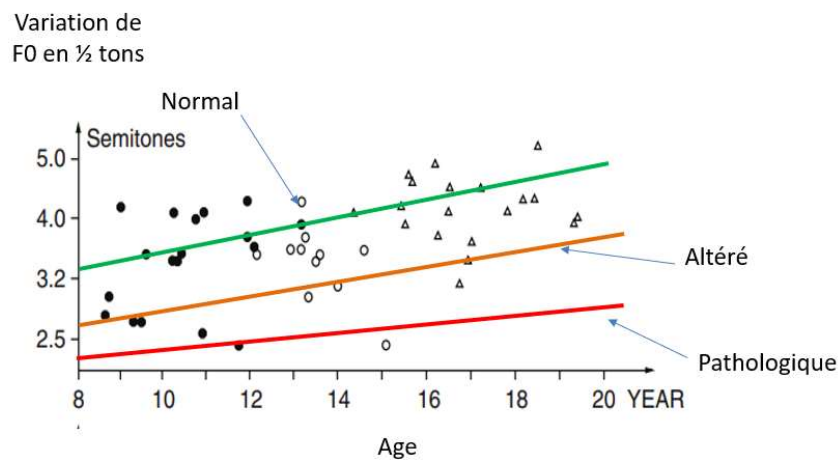


Figure 13 : Variation de F0 en parole continue [Pedersen 2015]

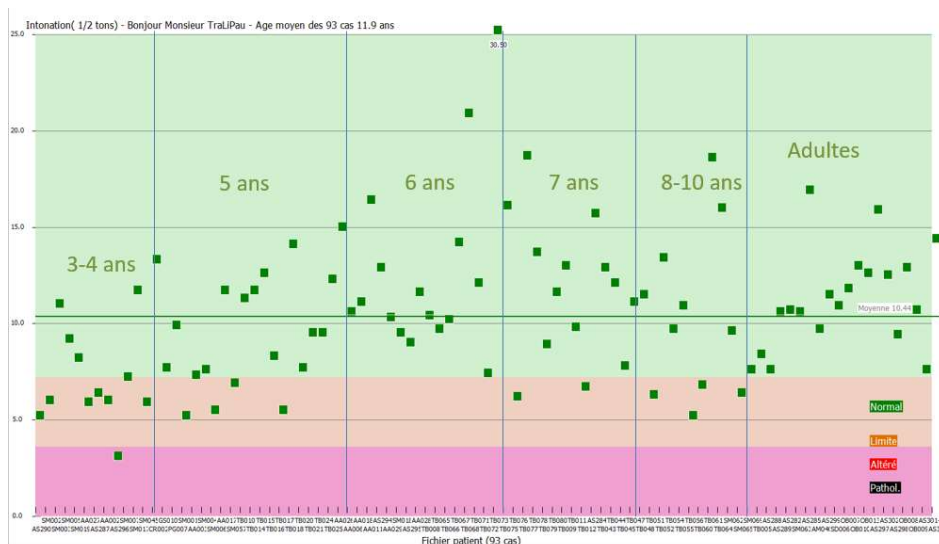


Figure 14 : Variation de F0 sur une phrase courte, analyse de 93 cas de parole normale

Sur une phrase courte telle que /*Bonjour Monsieur Tralipau*/ proposé dans le cadre du bilan de la parole [Menin-Sicard 2019], la prééminence de l’accent tonique sur le /ou/ entraîne une variabilité de F0 beaucoup plus importante que celle mesurée en parole continue par [Pedersen 2015]. Si l’on cumule une centaine d’enregistrements de parole normale de tous âges, nous obtenons la courbe de la figure 14. La moyenne est autour de 10 ½ tons tous âges confondus, soit le double de [Pedersen 2015].

Sur la phrase complémentaire /*Est-ce que tu connais Glochin l’écreuil ?*/ [Bardelang 2020], l’intonation est globalement un peu plus basse (moyenne vers 8,5 ½ tons) mais plus homogène, comme l’indique la figure 15. Elle ne concerne ici que les enfants. Dans DIADOLAB [Gerip 2020], nous considérons les repères de variation de F0 reportés dans la table 3.

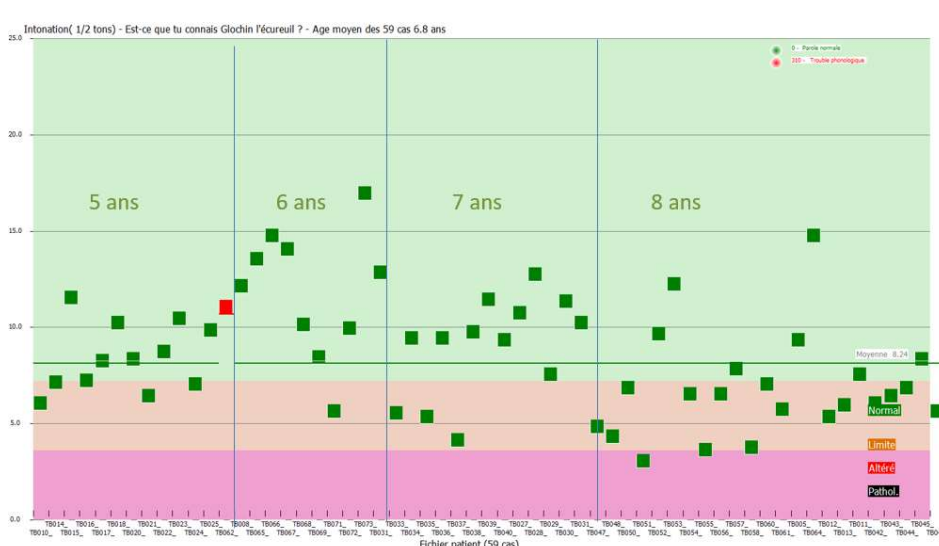


Figure 15 : Variation de F0 sur une 2ème phrase courte, analyse de 59 cas de parole normale

Age	Intonation normale (1/2 tons)	Altérée (1/2 tons)	Pathologique (1/2 tons)
3-4	5	4	3
5-6	6	5	4
7-12	7	6	4
Adulte	9	7	5

Table 3 : Repères proposés dans DIADOLAB pour la variation de F0 selon l'âge

Intonation et état émotionnel

L'impact de l'émotion sur l'intonation d'une phrase courte est considérable. Si l'on prend par exemple le groupe de 11 témoins de l'étude de [Samuel 1998], nous obtenons des valeurs allant de un ½ ton (tristesse) à 6 ½ tons (colère) pour la même phrase */Je pars samedi matin/* (Table 4). On comprend mieux l'importance d'un modèle pré-enregistré, qui dans le cas de */Bonjour Monsieur Tralipau/* est à la fois affirmatif et positif, et se traduit par la fourchette haute des valeurs d'intonation.

Age	Affirmation	Interrogation	Joie	Tristesse	Colère
Intonation (1/2 tons)	2,5	3,5	5	1	6

Table 4: Variation de l'intonation selon le sens de la phrase [Samuel 1998]

Accent tonique

Dans le cadre du protocole d'analyse de la parole [Menin-Sicard 2019], nous évaluons la pertinence de l'accentuation d'une phrase type. Parmi les schémas intonatifs proposés par [Delattre 1966] reportés figure 16, la phrase */Bonjour Monsieur Tralipau/* inclut principalement un motif 2-4_ « Implication », tandis que la phrase */Est-ce que tu connais Glochin l'écureuil/* comporte une interrogation (4-1) suivi d'une question (2-4+).

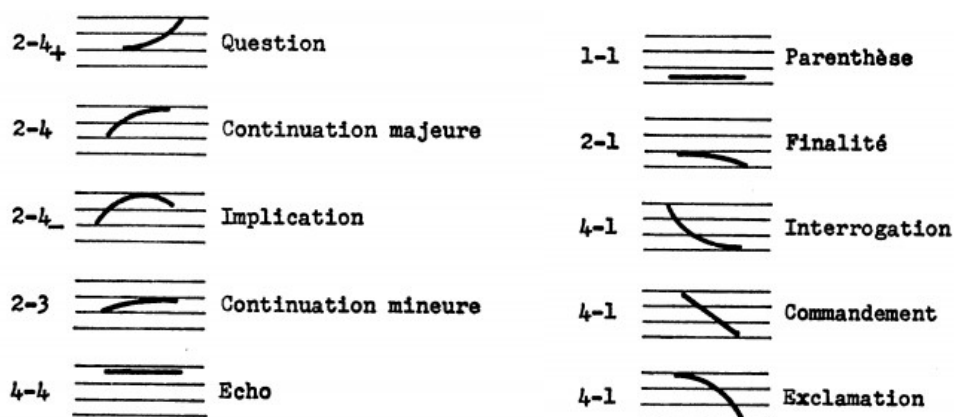


Figure 16 : Les dix intonations françaises les plus fréquentes et leur représentation schématique à l'aide de quatre niveaux de hauteur selon [Delattre 1966]

Formalisme des contours prosodiques

La conception ancienne de Delattre a laissé place à des descriptions plus simples des contours, relativement standardisée (figures 17 et 18), dont se sont inspirés des logiciels d'annotation prosodique automatisés tels que Prosogram [Mertens 2004] ou Prosobox [Goldman 2020]. Une démarche proposée par [Hirst 2005] sous le code INTSINT (*International Transcription System For Intonation*) et illustrée dans [Di Cristo 2013] consiste à repérer les valeurs de hauteur absolues qui délimitent l'espace tonal des locuteurs (M-Mid, T-Top, B-Bottom), puis de définir cinq contours fondamentaux (H-Higher, U-Upstepped, L-Lower, D-Down-stepped, S-Same).

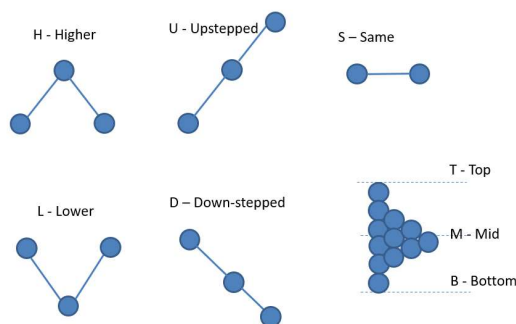


Figure 17 : Illustration des contours intonatifs proposés par [Hirst 2005]

De son côté, [Mertens 2012] a proposé dans son outil Prosogram un codage légèrement différent (figure 18). Il utilise aussi trois niveaux (rebaptisés Low, Medium, High) mais fait une distinction grâce aux minuscules et majuscules entre les mouvements de hauteur, selon qu'ils sont réduits ou amples. Le codage, sur une ou deux lettres est aligné sur les syllabes.

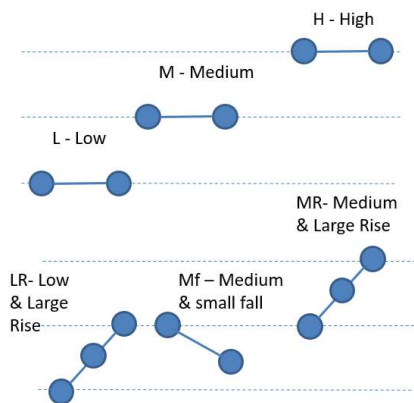


Figure 18 : Illustration des contours intonatifs proposés par [Mertens 2012]

Détection automatisée de l'accent

Dans le cadre du bilan de la parole, nous avons simplement cherché à positionner l'accent tonique majeur en lien avec la phrase, afin de pré-positionner une évaluation de la pertinence de l'accent. Pour ce faire, nous combinons les 4 notions fondamentales de la prosodie, c'est-à-dire la hauteur, l'intensité et la durée des phonèmes, afin d'identifier le ou les accents principaux dans la phrase. La courbe d'accentuation qui apparaît en gris dans la figure 19, superposée à la trajectoire de F0, comporte un maximum qui est symbolisé par un rond rouge. La présence de ce maximum dans le segment /ou/ est conforme. En revanche, un accent inapproprié sur /Monsieur/ est détecté dans le cas AA002 (Figure 20).

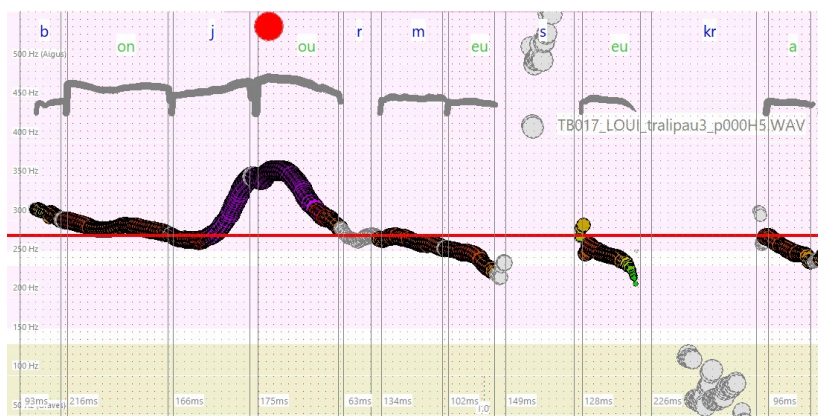


Figure 19 : Détection de l'accent tonique dans DIADOLAB, symbolisé par un point rouge, correspondant à un maximum de la courbe d'accentuation, reportée en gris au-dessus de la courbe de F0, cas TB017, garçon de 5 ans, parole normale

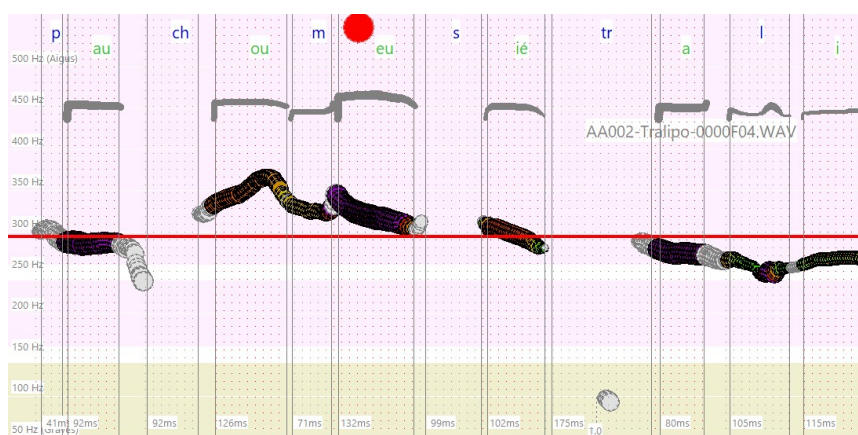


Figure 20 : Détection de l'accent tonique sur /Monsieur/ au lieu de /Bonjour/, cas AA002

Etendue Vocale

L'extraction de F0 sur une sirène permet de déterminer l'étendue vocale. Un écran spécifique a été développé dans VOCALAB [Gerip 2020] pour cette analyse (Figure 21). L'étendue vocale caractérise les performances du patient dans toutes les hauteurs qu'il est capable de produire sans forcer. Nous distinguons en général la zone de mécanisme I (graves, appelée *Modale*), la zone de passage entre les mécanismes I et II, et la zone de mécanisme II (aigus, appelée *Falsetto*).

La transition entre les deux mécanismes est appelée « passage » et s'accompagne souvent d'une modification de la structure acoustique (rupture, décrochage, appauvrissement harmonique). Dans certains cas, tels que AS022 illustré figure 22, la patiente (Femme de 56 ans, nodules des cordes vocales) est incapable de produire certaines notes dans la zone de transition 400-500 Hz et passe en force d'un mécanisme à l'autre. Son étendue vocale est de 24 ½ tons, un peu en dessous de la norme (30 ½ tons). Dans le cas AS031 (Figure 23), le patient (Homme de 77 ans, paralysie récurrentielle unilatérale du nerf laryngé) se révèle incapable de passer en registre II, a une étendue limitée à 12 ½ tons (norme à 30) et fonctionne par paliers plutôt que par montée et descente continue.

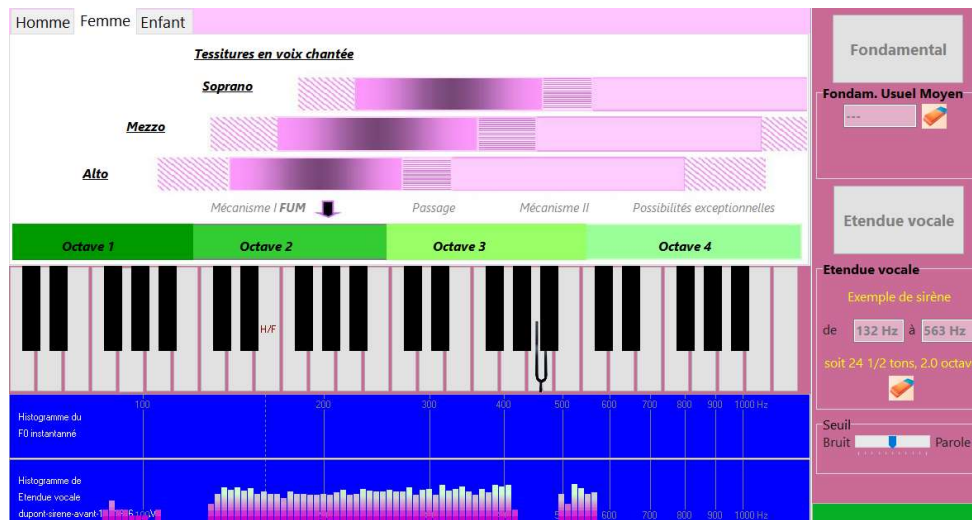


Figure 21 : mesure de l'étendue vocale sur une sirène avec VOCALAB

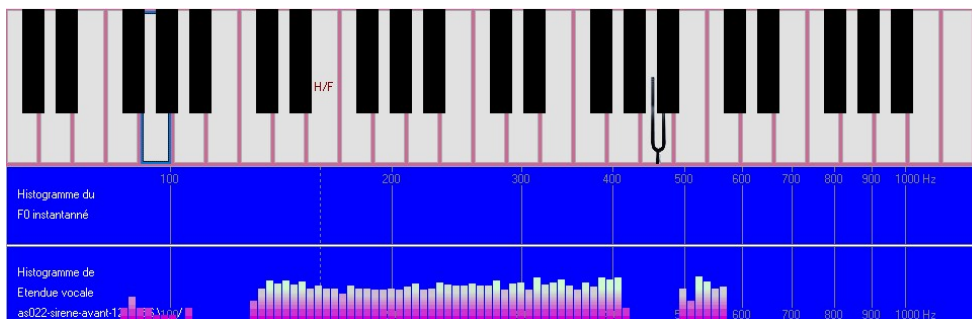


Figure 22 : Exemple d'étendue vocale avec un trou lié à une difficulté de passage de mécanisme (AS022)

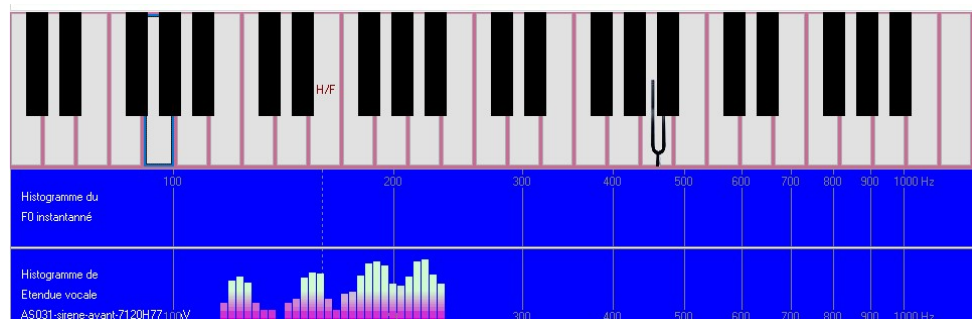


Figure 23 : Exemple d'étendue vocale limitée et irrégulière (AS031)

Valeurs de référence

Dans l'analyse de littérature produite par [Wuyts 2003], le nombre de demi-tons dans le registre I pour les adultes varie de 23 (voix non entraînées) à 26 ½ tons (chanteurs). En cumulant registre I (*modal*) et registre II (*false*), la valeur monte à 30 voire 35 ½ tons dans le cas de chanteurs ou d'enseignants. Pour les enfants de 8-11 ans, l'étendue vocale moyenne est de 27 ½ tons. Les valeurs citées correspondent donc à une plage de l'ordre de 2 à 3 octaves (24 à 36 ½ tons). Nous reproduisons en table 4 les valeurs utilisées dans VOCALAB et DIADOLAB.

Genre	FUM (Hz)	Passage de registre (Hz)	Etendue vocale type	Etendue usuelle en ½ tons	Etendue maximale en ½ tons
Homme	110	220	De 85 à 600 Hz	24	30
			Basse : 75 à 350 Hz		
			Baryton : 100 à 500 Hz		
Femme	200	400	De 150 à 800 Hz	24	30
			Alto : 120 à 780 Hz		
			Mezzo : 140 à 900 Hz		
Enfant 6 ans	250	450	De 160 Hz à 700 Hz	20	25

Table 5: Repères standards pour le FUM, le passage de registre et l'étendue vocale pour les hommes, femmes ou enfants intégrés dans VOCALAB [GERIP 2020]

Dans [Wuyts 2003], les enfants avec nodules ont une étendue de 16 ½ tons (6-7 ans) et 21 ½ tons (8-11 ans), ce qui permet de placer la limite d'altération autour de 25 ½ tons et de pathologie vers 18 ½ tons (Figure 24). Nous utilisons pour base les valeurs proposées par [Pedersen 2005] qui concernent 48 garçons et 47 filles enregistrés dans le cadre d'une chorale. Il s'agit donc d'enfants entraînés qui ont une étendue vocale plus élevée que les enfants sans pratique vocale, comme souligné par de nombreux auteurs.

L'étendue vocale doit pouvoir couvrir au minimum 2 octaves afin d'assurer une expressivité correcte à la parole et garantir une utilisation en voix impliquée. L'étendue vocale considérée comme pathologique est limitée à 1 octave, comme illustré figure 24.

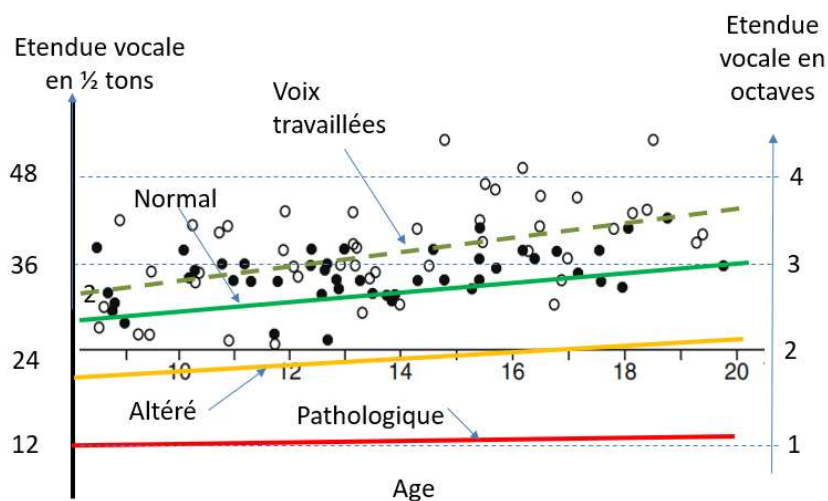


Figure 24 : Valeurs de référence de l'étendue vocale d'après [Pedersen 2005], et proposition de niveau altéré et pathologique

Le phonétogramme

Le phonétogramme, encore appelé *Speech Range Profile (SRP)* en Anglais, est une sorte de « patate vocale », qui caractérise les possibilités du patient en termes d'étendue vocale et de puissance sonore. Dans cette représentation, on considère la fréquence en X en échelle logarithmique (comme un clavier de piano) et l'amplitude en Y en décibel (une autre forme d'échelle logarithmique). Ce mode de représentation est le plus proche de la perception humaine des fréquences et des amplitudes. Deux exemples de phonétogrammes sont donnés figure 25 [Pedersen 2008] [Pabon 2011], correspondant à des voix de femme.

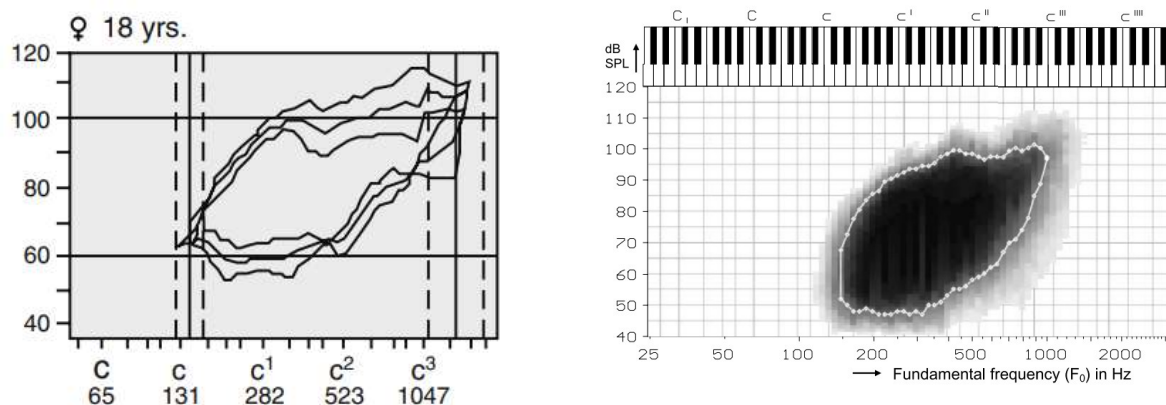


Figure 25: Exemples de phonétogrammes [Pedersen 2008] [Pabon 2011].

Le protocole usuel est de faire prononcer une voyelle /a/ forte et faible, en explorant note par note la dynamique fréquentielle. Le protocole de bilan de la phonation de [Sicard 2013] se focalise plutôt sur 2 sirènes, sur un /ou/, l'une à faible intensité, l'autre à forte intensité. Cette dernière méthode a l'avantage d'être plus rapide et donne les contours du phonétogramme, sans toutefois explorer la zone centrale.

Si l'on superpose d'autres enregistrements tels que la voix conversationnelle, la voix d'appel ou un texte déclamé, on comble petit à petit l'espace amplitude/fréquence, ce qui permet de dessiner l'aire dynamique vocale (Figure 26). Concernant le cas AS042 [Sicard 2021b], l'aire dynamique vocale est peu étendue, l'aire dynamique est significativement en dessous de la norme pour une femme.

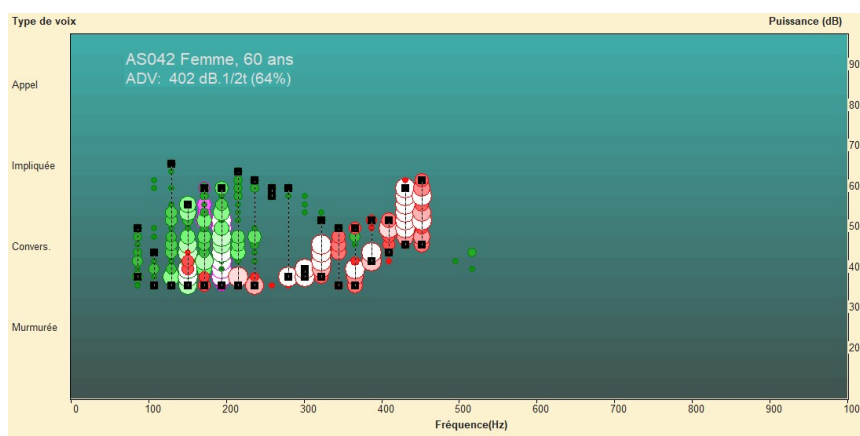


Figure 26 : Phonétogramme et évaluation de l'aire dynamique vocale, cas AS042 [Sicard 2021b]

Valeurs de référence

En simplifiant un peu, un phonétogramme typique d'un adulte se dessine sur 3 octaves (axe des X), soit $3 \times 12 \frac{1}{2}$ tons, avec une différence entre puissance haute et basse de 30 dB en moyenne (axe des Y), qui tend vers 0 aux extrémités basses et hautes de l'étendue vocale. La différence maximale se situe au niveau de la hauteur de confort vocal et atteint entre 30 et 40 dB. On note une certaine différence homme, femme, enfant, en termes de valeurs de fréquence, d'amplitude et donc d'aire vocalique.

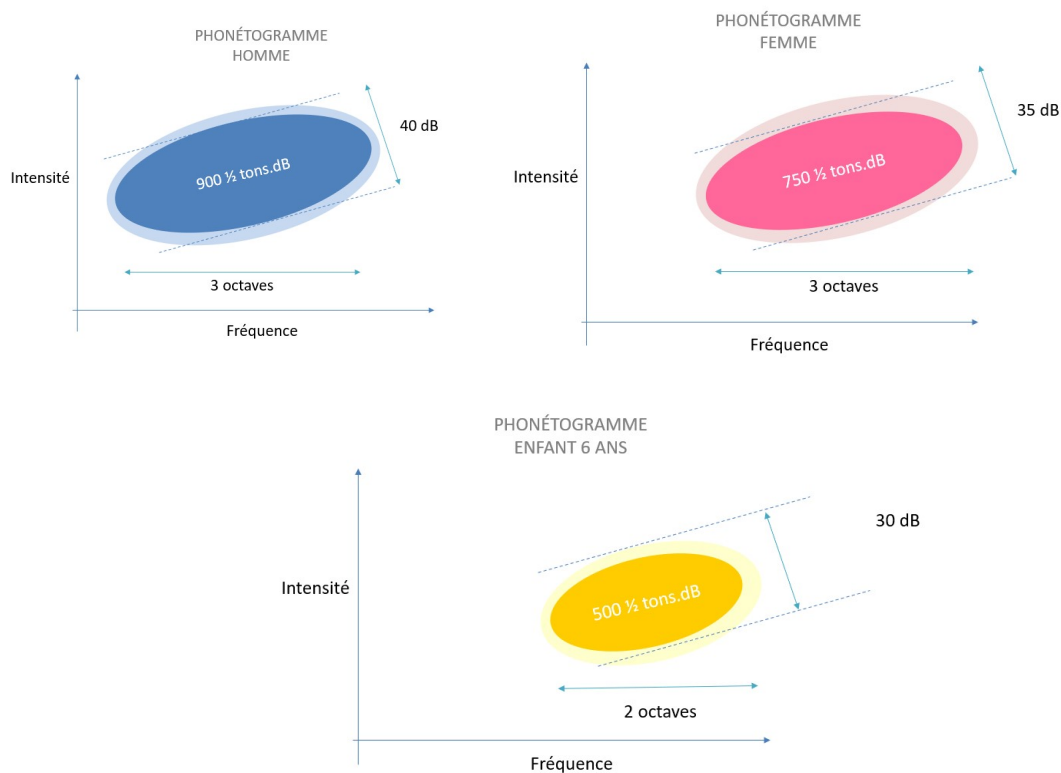


Figure 27: Phonétogramme typique homme, femme et enfant

La surface occupée par le phonétogramme dans l'espace fréquence/dB, encore appelée Aire Dynamique Vocale (ADV) est exprimée en $\frac{1}{2}$ tons.dB. Avant l'évaluation informatisée des phonétogrammes, le calcul de l'aire s'effectuait en cm^2 qui correspond à 32 $\frac{1}{2}$ tons.dB [Pedersen 2008].

La surface du spectrogramme évolue sensiblement avec l'âge. A 8 ans, l'aire dynamique vocale est autour de 500, 750 à la puberté, puis 900 à l'âge adulte, d'après les valeurs proposées dans la littérature [Ikeda 1999] [Pedersen 2003] [Wuyts 2002] (Figure 28). On différencie les aires dynamiques hommes et femmes du fait du décalage en fréquence (vers le haut pour les femmes) et une légère réduction de l'ADV (750) par rapport aux hommes (900).

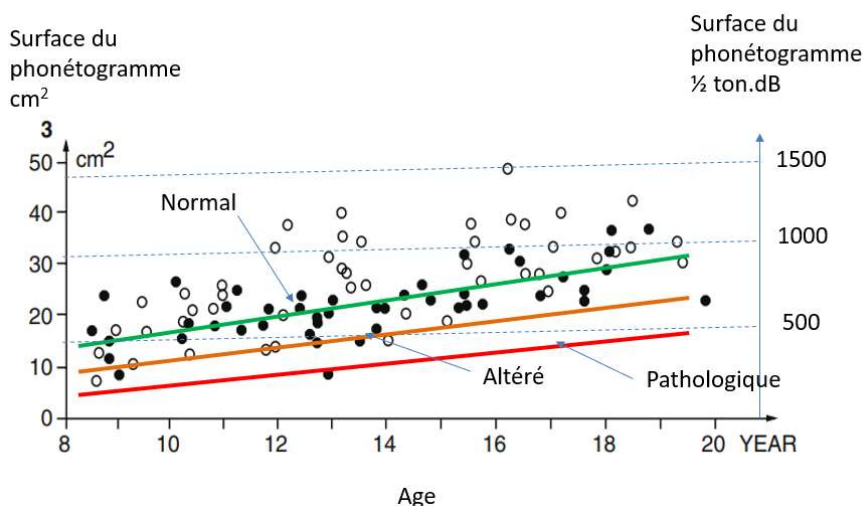


Figure 28 : proposition de valeur normale, altérée et pathologique pour la surface du spectrogramme, sur la base de [Pedersen 2008].

Prosodie altérée

Nous décrivons ici quelques caractéristiques de voix & parole altérées sur la base de publications scientifiques et de groupes de cas significatifs, selon une sélection de pathologies dont certaines sont très communes en clinique orthophonique (dysphonie, nodules), et d'autres ont fait l'objet de programmes de recherches spécifiques auxquels nous avons pu prendre part (C2SI en lien avec les cancers ORL, Voice4PD en lien avec la Maladie de Parkinson). Nous relevons pour chaque type de pathologie des valeurs repères en lien avec le temps phonatoire, l'étendue vocale, l'intonation ou l'aire dynamique vocale, que nous comparons aux normes attendues.

Atteinte du nerf laryngé supérieur

L'étendue vocale et les contours intonatifs peuvent être altérés dans le cas d'atteinte du nerf laryngé supérieur (NLS) [Lepape 2020]. Un corpus de 30 femmes avec suspicion d'atteinte du NLS ont été comparées à 30 autres sans dommage post-chirurgical, et à 30 femmes témoins appariées en âge. Un abaissement des F0 moyennes et maximales avec une réduction de l'étendue vocale maximale d'une tierce sur un texte lu a été observé pour le groupe NLS (table 6). Les auteurs ont noté que l'expression prosodique d'une question et d'une exclamation avait été impactée, et l'étendue vocale significativement réduite.

Paramètre	Suspicion atteinte NLS	Sans dommage chirurgical	Groupe témoin	Norme	Commentaire
FUM	191	217	238	210-250	Tâche de lecture
Intonation (1/2 tons)	8	9	12	6-9	Question /C'est bien toi ?/
Etendue vocale (1/2 ton)	14	16	20	24-30	Réduction significative

Table 6 : Analyse de l'étendue vocale de femmes avec atteinte possible du nerf laryngé supérieur [Lepape 2020]

Apraxie de la parole

La dysprosodie est considérée comme une caractéristique essentielle de l'apraxie de la parole chez l'enfant. [Ballard 2010] met en évidence la difficulté des enfants avec ce type de trouble dans le contrôle des aspects temporels et des accentuations prosodiques de la production de la parole. Cet auteur a étudié l'efficacité d'un traitement ciblant un meilleur contrôle des durées des syllabes sur la base de logatomes de 3 syllabes en l'accentuation fort-faible (par exemple /BA^tigu/) et faible-fort (par exemple, /ba^Tigu/). Les séances de traitement ont été structurées selon l'approche des principes de l'apprentissage moteur (*Principles of Motor Learning* - PML) exposés dans l'ouvrage [Schmidt 2018].

Le but de cette approche est de faciliter l'apprentissage à long terme des habiletés motrices, sur la base d'une pratique à intensité élevée, l'acquisition de compétences multiples et variées en parallèle, avec un ordre aléatoire des stimuli au cours de la session, ainsi qu'une augmentation progressive de la complexité des tâches au cours des séances. En travaillant les clusters /tr/ ou /st/, le patient acquiert des habiletés motrices par rétroaction aux structures plus simples, ici les consonnes isolées /t/, /r/, /s/.

Trois enfants, âgés de 7 à 10 ans, avec une apraxie de la parole légère à modérée et un développement normal du langage ont participé à un traitement intensif de 3 semaines. Tous les enfants ont amélioré leur capacité à contrôler la durée relative des syllabes ainsi que les accents (*lexical stress*) et les contrastes, comme illustré à la figure 29 qui montre une augmentation au fil des séances du pourcentage de productions correctes.

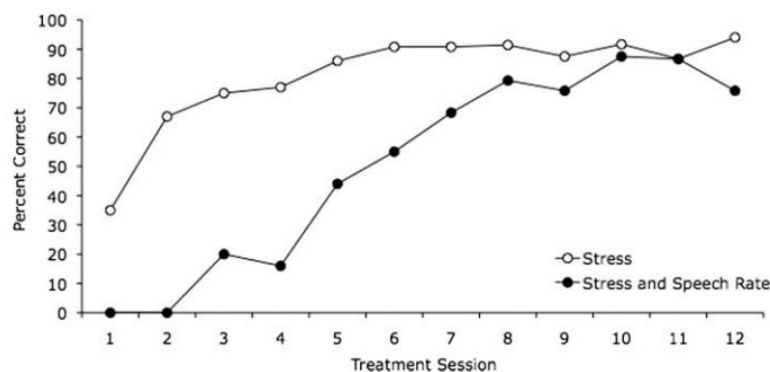


Figure 29 : Pourcentage de production du patient F1 perçues comme correctes soit concernant les accents (cercles vides) soit accent et débit de parole (cercles fermés) au cours des 12 séances de traitement. [Ballard 2010]

Bégaiement

La parole chez les personnes qui bégaiement est considérée comme dysprosodique, principalement du fait d'une plage de variation de la fréquence fondamentale significativement plus petite que les personnes qui ne bégaiement pas. De plus, le rythme de la parole peut être altéré, avec des difficultés à mettre l'accent approprié sur la phrase. Les épisodes de bégaiement sont situés principalement sur des syllabes accentuées.

Dans [Neumann 2018], la parole de 13 hommes adultes avec antécédents de bégaiement a été analysée avant, après et 1 an après un traitement intensif ciblant la fluence. Différents paramètres acoustiques ont été mesurés et comparés à 13 témoins sans troubles. Une série de phrases étaient lues à haute voix avec une prosodie « neutre », « heureuse » et « interrogative ».

Paramètre	Adultes qui bégaiant	Adultes qui bégaiant après traitement	Adultes qui bégaiant 1 an après traitement	Groupe témoin	Norme	Commentaire
FUM	114	137	121	137	110	FUM : formes heureuses et interrogatives plus élevé que la norme
Etendue vocale (1/2 ton)	4	5	4	6	6-9	

Table 7 : Analyse de l'étendue vocale de sujets bégues comparés à des sujets témoins [Lepape 2018]

La parole des patients avant traitement était perçue comme monotone (faible étendue vocale), avec peu d'accents prosodiques ce qui maintient le FUM à une valeur basse, telle que mesurée sur des voyelles tenues [Voropedian 2016]. La thérapie ciblant la fluence a permis de restructurer la prosodie, qui se rapproche de celle des personnes normales (table 7).

Dyspraxie verbale

La dyspraxie verbale (*Developmental verbal dyspraxia – DVD*) est une maladie relativement rare. On la considère aussi sous la dénomination apraxie développementale de la parole (*Developmental apraxia of Speech DAS*). Selon [Charron 2015], trois caractéristiques différencient la dyspraxie verbale des autres atteintes phonologiques ou de la dysarthrie : des erreurs inconstantes affectant les consonnes et les voyelles, des difficultés de coarticulation et des particularités sur le plan de la prosodie. Parmi les marqueurs objectifs, on peut citer un rythme articulatoire lent, des pauses inappropriées, des accentuations inappropriées et des imprécisions dans la planification et la programmation motrice.

Dysphonie

Le but de l'étude de [Wingate 2007] était d'examiner les bénéfices de programmes de renforcement musculaire expiratoire combiné à une thérapie vocale traditionnelle sur 18 utilisateurs professionnels de la voix, la moitié avait des maux de gorge et de fatigue vocale, l'autre moitié présentait des lésions bénignes des cordes vocales. Les séances de thérapie vocale de 45 mn chacune, sous la responsabilité d'une orthophoniste, abordaient notamment l'hygiène vocale, l'hydratation, la minimisation de la parole en situations bruyantes, la réduction des tensions sur les cordes vocales en utilisant des techniques présentées dans les ouvrages de [Boone 2005] et [Stemple 2018]. Le patient avait aussi des exercices de 15 à 20 mn à faire quotidiennement, consignés dans un journal de bord. Les résultats ont montré des améliorations significatives de la plage dynamique vocale, et par conséquent de l'aire dynamique vocale qui rejoint la norme en fin de traitement (Table 8).

Paramètre	Pré-traitement	En cours de traitement	Post traitement	Norme	Commentaire
ADV (1/2 ton.dB)	600	680	920	900 (H), 750 (F)	Dans la norme en fin de traitement

Table 8 : Augmentation de l'aire du phonétogramme au cours de la thérapie [Wingate 2007]

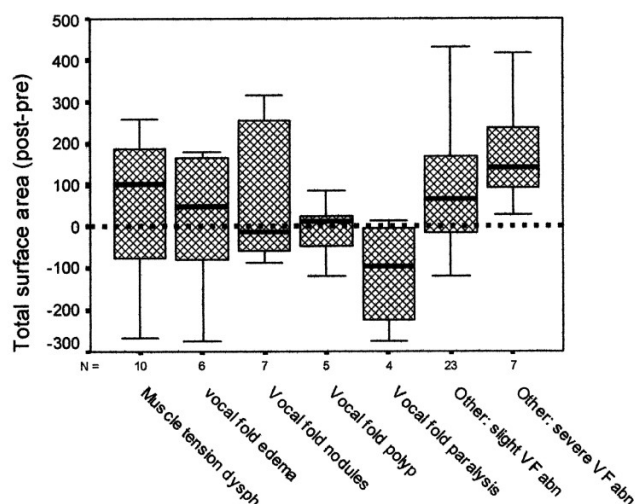


Figure 30 : Modification de l'aire du phonétogramme avant et après thérapie, selon les pathologies [Speyer 2003]

De son côté, [Speyer 2003] a étudié 60 cas de dysphonies de pathologies variées, avant et après prise en charge orthophonique de 2 fois 30 mn/semaine pendant 3 mois. Les éléments généraux de la thérapie vocale comprenaient des conseils d'hygiène de la voix, des exercices généraux d'amélioration du geste vocal et de la parole, ainsi que des exercices plus ciblés selon les besoins des patients et en fonction de leurs possibilités d'expression.

Aucune instruction particulière n'a été donnée pour guider la thérapie. Afin d'éviter un effet lié au thérapeute, pas plus de deux patients ont été pris en charge par un même thérapeute. Les progrès les plus notables concernent les nodules, tandis que la paralysie des cordes vocales montre au contraire une réduction de l'aire dynamique vocale après prise en charge (Figure 30).

Féminisation vocale

L'acquisition d'une voix perçue comme féminine pour les sujets transgenres nécessite d'aborder des subtilités vocales qui rendent cette démarche assimilable à l'apprentissage d'une seconde langue [Blanckaert 2019]. Dans [Morsomme 2016], 9 étapes de thérapie vocale sont décrites : geste vocal sain, soutien respiratoire, élévation de la hauteur tonale, augmentation de la résonance notamment dans le formant F3, courbes montantes et contrastes intonatifs sur la plus large tessiture possible, volume atténué avec souffle, ralentissement du débit, renforcement des points d'articulation, discours au féminin et comportements non verbaux.

Concernant la hauteur tonale, l'objectif visé est d'atteindre la zone autour de 160 Hz, à partir de laquelle la voix commence à être perçue comme voix de femme, et non pas 200-250 Hz, zone commune du FUM chez les femmes.

Les patientes trans sont ainsi amenées au travers d'exercices spécifiques à positionner l'oropharynx vers le haut, afin de raccourcir la longueur du tractus vocal, ce qui a notamment pour effet d'augmenter la valeur des formants et contribue à la perception d'une voix féminine (Table 9).

Paramètre	Avant prise en charge	Après prise en charge	Norme	Commentaire
FUM (Hz)	92	190	110 (H), 200 (F)	Zone de transition 150 Hz
Etendue vocale (1/2 tons)	26	23	30	

Table 9 : Performances de patientes trans avec exercices de féminisation vocale [Blanckaert 2019]

Maladie de Parkinson

La parole de patients atteints de la maladie de Parkinson (MP) est associée à la dysarthrie hypokinétique (HD) qui résulte d'une altération multidimensionnelle de la phonation, de l'articulation et la prosodie. [Skodda 2009] a étudié la modification des paramètres prosodiques pour 27 hommes et 23 femmes MP sur une période de 20 mois en moyenne, qu'il a comparé à 50 témoins appariés en âge. Dans le groupe PD, le débit de parole a diminué du premier au deuxième examen, en particulier chez les patients de sexe masculin. L'intonation a été significativement réduite par rapport au groupe témoin, ce qui témoigne d'une altération caractéristique de la prosodie (Table 10).

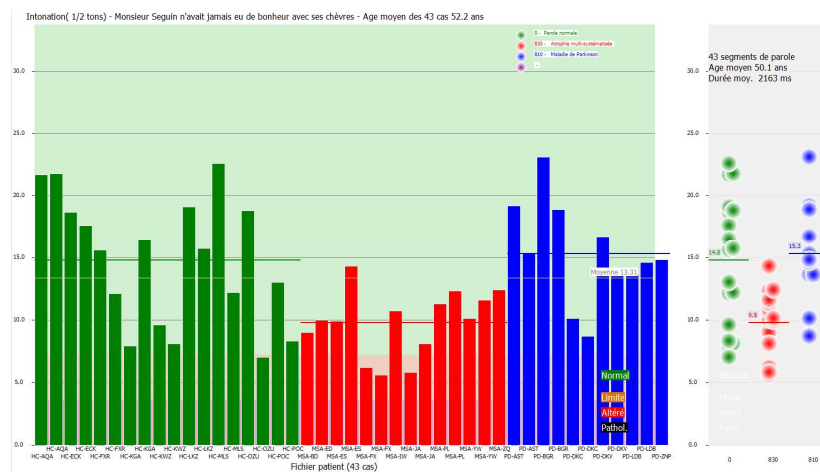
Paramètre	MP	MP+20 mois	Témoin	Norme	Commentaire
Hommes FUM (Hz)	137	133	120	110	Augmentation FUM
Homme Intonation (1/2 tons)	4	4	6	6-9	Discours monotone
Femme FUM (Hz)	192	186	193	200	Décroissance FUM
Femme Intonation (1/2 tons)	5	4	6	6-9	Discours monotone

Table 10 : Performances de patients Parkinsoniens à T0 et T0+20 mois en moyenne, comparés à une groupe témoin [Skodda 2009]

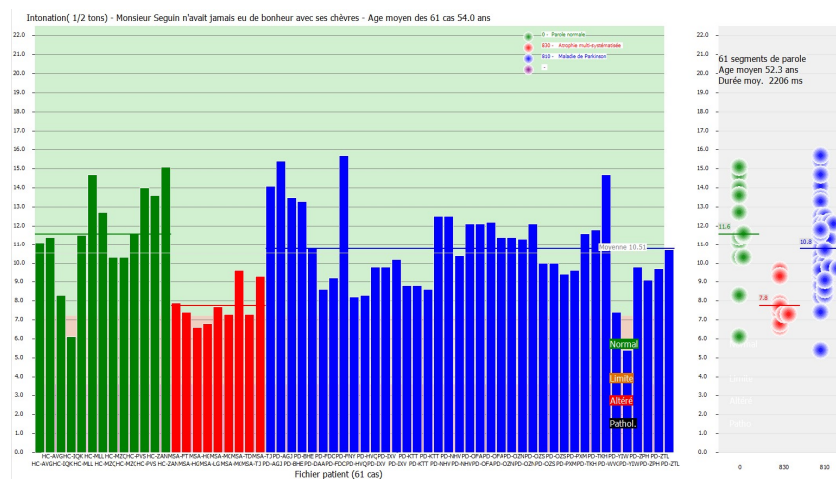
Dans [Skodda 2011] une analyse acoustique a été réalisée sur 169 patients atteints de MP et 64 témoins sains appariés selon l'âge sur la base d'une tâche de lecture de quatre phrases. Aucune différence significative dans le débit n'a été trouvée entre PD patients et témoins, mais les patients MP ont montré une réduction de l'intonation et du temps de pause entre les mots.

De son côté, [Viallet 2002] a étudié l'impact de la L-DOPA sur les paramètres prosodiques, notamment le FUM et l'intonation pour 20 hommes atteints de la MP. Il a montré que la L-DOPA peut restaurer la partie supérieure de la gamme tonale, augmentant ainsi la valeur du FUM qui passe en moyenne de

120 à 130 Hz, ainsi que l'intonation qui passe de 4 à 5 ½ tons, phénomène qui pourrait être attribué à la restauration de l'activité musculaire du larynx.



Groupe femmes, témoins en vert, MSA en rouge, PD en bleu



Groupe hommes, témoins en vert, MSA en rouge, PD en bleu

Figure 31 : Extraction de l'intonation sur le texte lu : réduction significative pour le groupe MSA [Sicard 2021]

Dans son analyse différentielle des syndromes Parkinsoniens, [Hlavnicka 2017] insiste sur des composantes prosodiques spécifiques : moyenne de la durée des segments voisés (*duration of voiced intervals - DVI*) qui augmente avec la sévérité de la dysarthrie, durée des consonnes plosives (*duration of stop consonants - DUS*) ainsi que la dégradation des fricatives non voisées (*decay of unvoiced fricatives - DUF*). Dans cette étude, les performances de patients atteints de la maladie de Parkinson idiomatique (PD), sont comparées à celle de patients souffrant d'Atrophie Multi-systématisée (MSA) ou la paralysie supra nucléaire progressive (PSP).

Dans le cadre du projet Voice4PD [Das 2020], nous avons étudié l'intonation sur une phrase de la fable de Mr Seguin, à l'aide des outils de segmentation et de statistiques de DIADOLAB. Le contraste entre MSA et les autres groupes est assez net, comme l'illustre la figure 31. Il y a bien une caractéristique de discours monotone avec MSA [Sicard 2021], alors que PD a des caractéristiques très proche du groupe témoin.

Nodules chez les enfants

L'étude de [Wuyts 2003] concerne 77 garçons et 41 filles souffrant de dysphonie liée principalement à des nodules, avec quelques cas de polypes ou d'œdèmes. Ces patients sont comparés à 53 garçons et 41 filles témoins. Les âges sont répartis pour moitié à 6-7 ans, pour moitié à 8-11 ans. On constate (table 11) une réduction proche d'une octave de l'étendue vocale et la diminution par 2 de l'aire dynamique vocale chez les garçons.

Paramètre	Patients 6-7 ans	Patients 8-11 ans	Témoins 6-7 ans	Témoins 8-11 ans	Norme	Commentaire
Etendue vocale garçons (1/2 tons)	16	21	23	28	25	Réduction 1 octave
Etendue vocale filles (1/2 tons)	17	22	27	27	25	Réduction 1 octave
ADV garçons (1/2 ton.dB)	340	430	730	950	700	ADV réduite
ADV filles (1/2 ton.dB)	442	670	850	850	650	ADV réduite

Table 11: Comparaison entre enfants avec nodules et enfants témoins concernant l'étendue vocale et l'aire dynamique vocale

Sourds implantés

Dans [Lenormand 2010] est détaillée une étude qui porte sur les enfants sourds implantés, analysant une liste d'altérations regroupées en

- **effets clash** : accents excédentaires /Le Pe Tit/
- **effets rupture** : ruptures inappropriées entre les mots, hésitations /Petit ## Garçon/
- **effets balance** : règles de durée ou d'appariement des groupes syntaxique mal appliquées /Le#petit garçon/

[Desgraves 2016] précise que les sujets implantés utilisent surtout les indices de rythme pour percevoir la prosodie. L'implant leur donne un accès limité à la perception des variations de hauteur à cause d'une dégradation temporelle et spectrale du signal, de par le nombre limité d'électrodes présents dans la cochlée et le couplage imparfait entre les impulsions de stimulation électrique issus de l'implant et les récepteurs du nerf auditif. Les électrodes préformées permettent une stimulation plus précise du nerf auditif car elles sont localisées plus près des cellules nerveuses, ce qui limite le recouvrement des champs électriques entre électrodes adjacentes. Les sujets bénéficiant d'électrodes préformées obtiennent de meilleurs résultats en prosodie que ceux porteurs d'électrodes droites. On peut supposer qu'ils profitent d'indices spectraux de meilleure qualité pour percevoir la prosodie.

Troubles du Spectre Autistique

De manière générale, les analyses acoustiques montrent que les enfants et les adultes augmentent la fréquence F0 et l'intensité pour passer d'une parole neutre à un accent intonatif, avec de manière fréquente un allongement des phonèmes. [Vidou 2018] relève que certaines études montrent que la parole des enfants atteints de troubles du spectre autistique (TSA) sont caractérisées par une intonation exagérée, tandis que d'autres études la qualifient de monotone. On note aussi un emploi excessif de la mise en emphase, un manque de contraste entre la condition neutre et la condition emphase, ainsi que des erreurs dans la position de la syllabe accentuée.

Tumeur du larynx

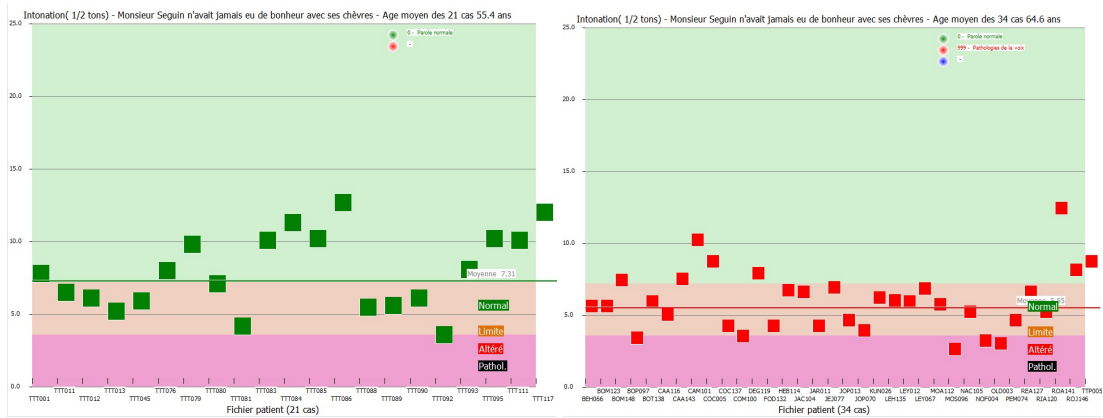
Dans l'étude de [Tuomi 2017], 42 patients diagnostiqués avec une tumeur primitive du larynx et traités avec radiothérapie ont été l'objet d'une étude portant sur l'impact d'une prise en charge orthophonique (PCO) sur les progrès en confort vocal. Les sessions comportaient des exercices de base (relaxation, posture et respiration), l'augmentation de la phonation, le travail sur des mots courts, puis sur de courtes phrases, le travail de l'intonation et de l'accentuation de syllabes, puis des phrases de plus en plus longues et accentuées tout en conservant une articulation détendue. Les dernières séances étaient consacrées à la communication, la gestion des pauses, du contact visuel et du travail de la voix projetée. L'intérêt de l'étude a été de montrer que sans prise en charge, les performances des patients stagnent, tandis qu'elles s'élèvent tout en restant en dessous des normes d'aire dynamique vocale (table 12).

ADV (1/2 ton.dB)	Bilan initial	Bilan à 6 mois	Bilan à 2 mois	Commentaire
Patient avec PCO	320	420	475	Progression
Patient sans PCO	317	316	352	Stagnation
Norme	900 (H), 750 (F)			

Table 12: Evolution de l'aire dynamique vocale de patient ayant été atteints de tumeur du larynx, avec ou sans prise en charge orthophonique [Tuomi 2017]

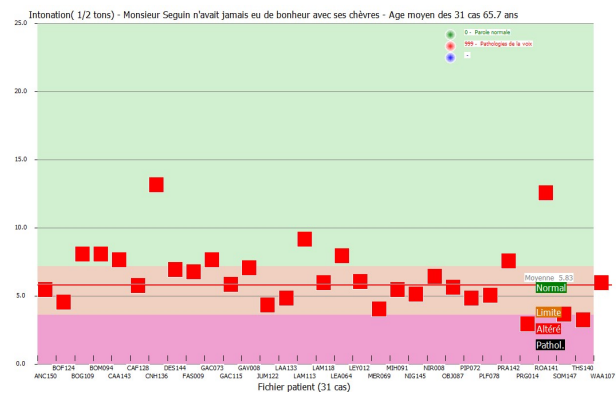
Dans le cadre du projet C2SI [Sicard 2017] nous avons analysé la parole sur un texte lu de 10 secondes environ, tiré de la fable « La chèvre de Monsieur Seguin » d'A. Daudet. Le groupe témoin comportait, au moment de l'étude, 21 cas. Le groupe des 65 patients a été séparé en deux sous-groupes d'après les scores d'intelligibilité fournis par l'équipe médicale : intelligibilité altérée (50 -100/100), comportant 34 patients, et le groupe parole pathologique (0-50/100) de 31 patients. Concernant le groupe témoin, la valeur de l'intonation en ½ tons est un peu plus faible que la norme proposée sur une phrase courte (7.3 au lieu de 9 ½ tons), bien que plus élevée que les valeurs relevées par [Pedersen 2015] (5 ½ tons). La durée de parole analysée est ici plus longue (10 à 15 secondes), avec statistiquement moins d'accents toniques, et le sens du texte est plutôt triste, favorisant peu la prosodie (« n'avait jamais eu de bonheur », « perdait toutes »). Enfin, la moyenne d'âge des témoins est de 55 ans, ce qui peut expliquer une légère réduction de l'intonation.

On note une réduction de 25% environ (2 ½ tons) de la prosodie chez les patients. Dans le cas du groupe pathologique, la valeur moyenne similaire à celle du groupe altéré est un peu surprenante, dans la mesure où l'intelligibilité est très réduite.



a) Groupe contrôle

b) Groupe altéré



c) Groupe pathologique

Figure 32: Intonation du groupe témoin et du groupe patient sur un texte lu « Mr Seguin n'avait jamais eu.. », projet C2S1

Texte lu (phrases Mr Seguin)	Groupe témoin	Groupe patients intelligibilité altérée	Groupe patients intelligibilité pathologique	Commentaire
Intonation (1/2 ton)	7.3	5.5	5.8	Intonation comparable patho/altéré
Norme adulte	9	9	9	Groupe témoin un peu en dessous de la norme

Table 13: Evolution l'intonation sur des témoins et patients ayant été atteints de tumeur du larynx, projet C2S1



Figure 33: Contour hyper-prosodique de /Monsieur Seguin/, cas CNH136, montrant une instabilité en hauteur marquée

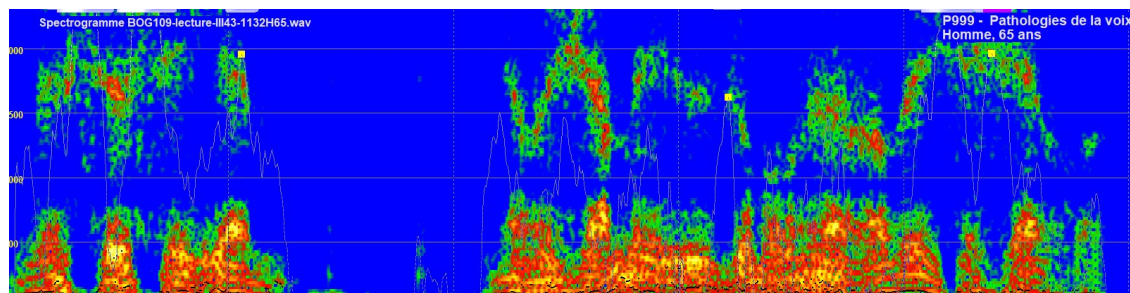


Figure 34: Spectrogramme de la phrase « Monsieur Seguin... » montrant une déstructuration harmonique marquée, avec un décodage de F0 difficile (BOG109)

Certaines valeurs singulièrement élevées sont liées à une instabilité en hauteur marquée (CNH136, femme de 87 ans, figure 33), parfois aussi à des erreurs de détection de F0 du fait d'un rapport bruit/signal très élevé pour certains patients, comme illustré à la figure 32, (BOG109, homme de 65 ans), pour lequel F0 est difficile à détecter et la variation de F0 difficile à évaluer. C'est plus l'appauvrissement harmonique et le bruit qui caractérisent ce type de pathologique, plus qu'une réduction de la variation de F0, la capacité de modifier la hauteur de la voix étant relativement préservée.

Tyroplastie

Les interventions de médialisation de la corde vocale (tyroplastie) ont pour but d'améliorer la fermeture des cordes vocales au cours de la phonation. Ce type d'intervention peut être effectué suite à une paralysie unilatérale des cordes vocales. Dans [Uloza 2005], 32 patients répartis en 17 hommes et 15 femmes sont analysés. Les progrès observés sont significatifs, autant en termes d'étendue vocale, d'aire dynamique vocale (table 14) que d'endurance vocale (temps phonatoire en nette hausse).

Paramètre	Avant intervention	Après intervention	Groupe contrôle	Norme	Commentaire
Etendue vocale hommes (1/2 tons)	13	23	36	30	Nets progrès
Etendue vocale femmes (1/2 tons)	11	20	30	30	Progrès
ADV Hommes (1/2 ton.dB)	180	420	915	900	Progression
ADV Femmes (1/2 ton.dB)	130	325	725	750	Progression
TPM (s) Hommes	8	17	25	20	S'approche de la norme
TPM (s) Femmes	5	13	21	15	S'approche de la norme

Table 14: Evolution de l'aire dynamique vocale, du temps phonatoire maximum et de l'étendue vocale avant et après tyroplastie montrant de très nets progrès

Conclusion

Nous avons abordé dans cet article quelques propriétés spécifiques de la parole, notamment les notions de fréquence fondamentale, d'intonation et d'étendue vocale. Nous avons donné pour chacune de ces notions des exemples de parole normale et altérée. En s'appuyant sur une sélection de publications scientifiques en archives ouvertes, nous avons tenté de séparer les caractéristiques de la parole normale, altérée et pathologique. Nous avons analysé l'altération de la prosodie, principalement sous l'angle de la fréquence, de l'intonation et de l'étendue vocale selon différentes pathologies, notamment la dysphonie et la Maladie de Parkinson. Les illustrations ont été réalisées à l'aide des logiciels VOCALAB et DIADOLAB, dédiés à la prise en charge orthophonique des troubles de la voix et de la parole.

Remerciements

Nous tenons à remercier les orthophonistes contributrices à l'ERU 46 du laboratoire LURCO et les nombreuses orthophonistes qui ont accepté d'enregistrer des témoins et des patients, ainsi que O. Barbera, C. Frenay, C. Grenier, T. Bardelang et M. Lemaitre qui ont enregistré et exploité les données de parole dans le cadre de leur mémoire d'orthophonie. Nous remercions aussi le Dr. V. Woisard et K. Daoudi de nous avoir permis d'analyser de grands corpus dans le cadre des projets ANR C2SI et Voice4PD. Nos remerciements vont aussi à A. Remacle pour l'accès au corpus de voix d'enfants normophoniques.

Nous saluons aussi les efforts de mise en archive ouvertes d'un très grand nombre de résultats scientifiques, en particulier dans le domaine de la parole, ce qui permet une diffusion très large, simple et sans droit d'accès, autorisant de ce fait des analyses comparatives et méta-analyses qui accélèrent ainsi leur mise en application clinique. Enfin, nous remercions l'ensemble des orthophonistes, étudiants et chercheurs qui nous ont accordé leur confiance en utilisant nos méthodologies, approches et outils dans le cadre de leur pratique clinique ou de leurs travaux scientifiques.

A propos des auteurs



Etienne SICARD est professeur à l'INSA Toulouse en électronique, informatique et traitement du signal. Il est directeur de recherches de l'équipe ERU 46 sur la parole pathologique au Laboratoire LURCO et chercheur associé à l'IRIT dans le cadre du projet ANR Voice4PD sur le

diagnostic différentiel des Syndromes Parkinsoniens. Etienne SICARD est co-auteur des logiciels VOCALAB et DIADOLAB pour le bilan et la rééducation orthophonique de la voix et la parole, ainsi que de MICROWIND et IC-EMC en lien avec la conception des circuits intégrés. Il a co-écrit une vingtaine d'ouvrages et plus de 200 publications scientifiques dans le domaine de l'analyse de la voix et parole, des circuits intégrés, et de la compatibilité électromagnétique. Il a été nommé en 2006 *Distinguished Lecturer* de la société savante IEEE.

Site : www.etienne-sicard.fr



Anne MENIN-SICARD est orthophoniste et titulaire d'un master recherche en Sciences du langage de l'université de Grenoble. Elle a exercé en libéral à Toulouse de 1994 à 2015. Anne MENIN-SICARD est co-auteur des logiciels VOCALAB et DIADOLAB.

Dans le cadre de AMS Formation, elle organise et anime des formations aux orthophonistes sur le bilan et la prise en charge de la voix et parole pathologique. Elle a fondé en 2019 Opop, pour la formation en ligne personnalisée à destination des orthophonistes. Chercheuse associée à l'ERU 46 du LURCO, elle développe des outils de mesures objectives et de rééducation de la parole en lien avec le logiciel DIADOLAB. Elle est auteur d'un ouvrage sur l'évaluation objective de la voix (DeBoeck, 2017) ainsi que de nombreuses publications dans le domaine de la voix et de la parole pathologique.

Glossaire

- ADV - Aire Dynamique Vocale
- AMS : Atrophie Multi-Systématisée
- AOS: Apraxia of Speech
- C2SI : Carcinologic Speech Severity Index
- DAS : Developmental apraxia of Speech
- DVD : Developmental verbal dyspraxia
- DUF : decay of unvoiced fricatives
- DUS : duration of stop consonants
- DVI : duration of voiced intervals
- ERU : Equipe de Recherche Unadreo
- F0 : Fondamental de la voix
- FUM : Fondamental usuel moyen
- HC : Healthy control
- HD : dysarthrie hypokinétique
- INTSINT : International Transcription System For Intonation
- LURCO : Laboratoire Unadreo de Recherches Clinique Orthophonique
- MP : Maladie de Parkinson
- MSA : Multiple system atrophy
- NLS : nerf laryngé supérieur
- ORL : oto-rhino-laryngologie
- PCO : prise en charge orthophonique
- PD : Parkinson's Disease
- PML : Principles of Motor Learning
- PRAAT : logiciel d'analyse de la parole « doing phonetics by computer »
- PSP: paralysie supra nucléaire progressive, progressive supranuclear palsy
- SRP - Speech Range Profile
- TSA : troubles du spectre autistique
- UNADREO : Union Nationale pour le Développement de la Recherche et de l'Évaluation en Orthophonie
- VOICE4PD : Differential diagnosis between Parkinson's disease and Multiple System Atrophy using digital speech analysis

Références

[Ballard 2010] Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*.

[Barbera Frenay 2016] Barbera 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*. Université de Lyon.

Sicard, Menin-Sicard, Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

[Bardelang 2020] Bardelang T. (2020), L'apport d'une phrase supplémentaire au protocole de DIADOLAB3. Mémoire d'orthophonie. Université de Lyon. Encadré par A. Menin-Sicard.

[Blanckaert 2018] Blanckaert, E., Pillot-Loiseau, C., & Morsomme, D. (2018). De l'intérêt du Speech Range Profile (SRP) dans le bilan vocal: aspects méthodologiques.

[Blanckaert 2019] Blanckaert, E., Mertens, P., Pillot-Loiseau, C., Didone, V., & Morsomme, D. (2019). L'analyse prosodique: outil d'objectivation de l'efficacité thérapeutique dans le cadre de la féminisation vocale? JPC 2019, Mons, Belgique.

[Boersma 2018] Boersma P. & Weenink D. (2018): Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from <http://www.praat.org/>

[Boone 2005] Boone, D. R., McFarlane, S. C., Von Berg, S. L., & Zraick, R. I. (2005). The voice and voice therapy.

[Brin-Henry 2018] Brin-Henry, F., Courier, C., Lederlé, E., & Masy, V. (2018). Dictionnaire d'orthophonie. Ortho-édition.

[Clerc 2016] Clerc, A., & Regnier, P. Prise en charge orthophonique de la dysarthrie dans la maladie de Parkinson. <https://www.happyneuronpro.com>

[Courtois du Passage 2004] Courtois du Passage, N., & Galloux, A. S. (2004). Bilan orthophonique chez l'enfant atteint d'autisme: aspects formels et pragmatiques du langage. *Neuropsychiatrie de l'enfance et de l'adolescence*, 52(7), 478-489.

[Das 2019] Das, B., Daoudi, K., Klempir, J., & Ruzs, J. (2019, May). Towards disease-specific speech markers for differential diagnosis in Parkinsonism. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 5846-5850). IEEE.

[Delattre 1966] Delattre, P. (1966). Les dix intonations de base du français. *French review*, 1-14.

[Desgraves 2016] Desgraves, C., & Lambert, F. (2016). La perception de la prosodie linguistique chez l'adulte implanté cochléaire post-lingual: intérêt d'un entraînement axé sur la perception de contours mélodiques. <http://thesesante.ups-tlse.fr/1760/>

[Di Cristo 2013] Di Cristo, A. (2013). La prosodie de la parole. De Boeck Supérieur.

[Eichhorn 2018] Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. (2018). Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice*, 32(5), 644-e1.

[Fox 2012] Fox, C., Ebersbach, G., Ramig, L., & Sapir, S. (2012). LSVT LOUD and LSVT BIG: behavioral treatment programs for speech and body movement in Parkinson disease. *Parkinson's disease*, 2012.

[ERU 15] <https://www.unadreo.org> > Les ERU > ERU 15

[Gerip 2020] VOCALAB v4.1 <https://boutiquesante.gerip.com> > Logiciels santé

[Gerip 2020b] DIADOLAB v3.1 <https://boutiquesante.gerip.com> > Logiciels santé

[Goldman 2020] Goldman, J. P., & Simon-Hustinx, A. C. (2020). ProsoBox, a praat plugin for analysing prosody. In *Proceedings of the 10th International Conference on Speech Prosody* (p. 5).

[Gramming 1986] Gramming, P., Gauffin, J., & Sundberg, J. (1986). An attempt to improve the clinical usefulness of phonetograms. *Journal of Phonetics*, 14(3-4), 421-427.

Sicard, Menin-Sicard, Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

[Grenier 2020] Grenier C. (2020), Etude comparative de tests phonologiques existants et du protocole d'analyse objective de la parole avec Diadolab : intérêt et limites d'un test court et des mesures objectives d'intelligibilité et de fluence. Mémoire de Master d'orthophonie. Université de Toulouse. Co-encadrement : S. Michel, E. Sicard.

[Hlavnicka 2017] Hlavnicka, J., et al (2017). Dysprosody Differentiate Between Parkinson's Disease, Progressive Supranuclear Palsy, and Multiple System Atrophy. In INTERSPEECH (pp. 1844-1848).

[Hillenbrand 2009] Hillenbrand, J. M., & Clark, M. J. (2009). The role of f₀ and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics*, 71(5), 1150-1166.

[Hirst 2005] Hirst, D. J. (2005). Form and function in the representation of speech prosody. *Speech Communication*, 46(3-4), 334-347.

[Ikeda 1999] Ikeda, Y., et al. (1999). Quantitative evaluation of the voice range profile in patients with voice disorder. *European archives of oto-rhino-laryngology*, 256(1), S51-S55.

[Kent 2000] Kent, R. D., & Ball, M. J. (Eds.). (2000). *Voice quality measurement*. Singular.

[Lacheret 2011] Lacheret, A. (2011). La prosodie au cœur du verbal. *Rééducation orthophonique*, (246), 87-104. <https://halshs.archives-ouvertes.fr/00636556>

[Lenormand 2010], Lenormand, M. T., & Lacheret, A. (2010). Prosodie chez des enfants implantés cochléaires. <https://halshs.archives-ouvertes.fr/halshs-00636456>

[Le Pape 2020] Le Pape, G., Lazard, D. S., Gatignol, P., Tresallet, C., & Pillot-Loiseau, C. (2020). Voice modulation, self-perception and motor branch of the superior laryngeal nerve. *European Annals of Otorhinolaryngology, Head and Neck Diseases*.

[Lowit-Leuschel 2000] Lowit-Leuschel, A. (2000). Dysprosody. In *Voice quality measurement* (pp. 59-72).

[Menin-Sicard 2013] Menin-Sicard, A. (2013). Approche métacognitive dans le cadre de l'évaluation et la réévaluation de la voix. *Rééducation orthophonique*, 254(254), 121-134.

[Menin-Sicard 2016] Menin-Sicard, A., & Sicard, E. (2016). Evaluation et réhabilitation de la voix: Approche clinique et objective. De Boeck Supérieur.

[Menin-Sicard 2019] Menin-Sicard A., Sicard E. (2019). Méthodologie d'évaluation objective de la phonologie, de la fluence et de la prosodie - Vers un bilan rapide à destination des orthophonistes. *Journées de Phonétique Clinique JPC*, Mai 2019, Mons, Belgique. <https://hal.archives-ouvertes.fr/hal-02127039>

[Mertens 2004] Mertens, P. (2004). Un outil pour la transcription de la prosodie dans les corpus oraux. *Traitement Automatique des langues*, 45(2), 109-130.

[Mertens 2012] Mertens, P. (2012). Transcription of tonal aspects in speech and a system for automatic tonal annotation. In *Advancing Prosodic Transcription*, Stuttgart.

[Meyrieux & Moreau 2018] 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.

[Morsomme 2016] Morsomme, D., & Remacle, A. (2016). Féminiser la voix. De la voix parlée au chant: Bilans, rééducations, pathologies de la voix parlée et chantée.

Sicard, Menin-Sicard, Analyse acoustique de la prosodie dans le cadre de la clinique orthophonique

[Neumann 2018] Neumann, K., Euler, H. A., Kob, M., von Gudenberg, A. W., Giraud, A. L., Weissgerber, T., & Kell, C. A. (2018). Assisted and unassisted recession of functional anomalies associated with dysprosody in adults who stutter. *Journal of fluency disorders*, 55, 120-134.

[Nootboom 1997] Nootboom, S. (1997). The prosody of speech: melody and rhythm. *The handbook of phonetic sciences*, 5, 640-673.

[Pabon 2011] Pabon, P., Ternström, S., & Lamarche, A. (2011). Fourier descriptor analysis and unification of voice range profile contours: method and applications. *Journal of Speech, Language, and Hearing Research*.

[Remacle 2017] 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.

[Samuel 1998] Samuel, C., et al. (1998). Dysprosody after severe closed head injury: an acoustic analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, 64(4), 482-485.

[Schmidt 2018] Schmidt, R. A., Lee, T. D., Winstein, C., Wulf, G., & Zelaznik, H. N. (2018). Motor control and learning: A behavioral emphasis. *Human kinetics*, 6th edition.

[Shriberg 1997] Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research*, 40(2), 273-285.

[Sicard 2017] Sicard, E., Mauclair, J., & Woisard, V. (2017, June). Etude de paramètres acoustiques des voix de patients traités pour un cancer ORL dans le cadre du projet C2SI. In 7èmes Journées de Phonétique Clinique. <https://hal.archives-ouvertes.fr/hal-01510418/>

[Sicard 2019] Sicard, E., Meyrieux, L., Moreau, M., & Remacle, A. (2019). L'analyse acoustique des voix d'enfants de 5 ans: Proposition de valeurs de référence pour les logiciels PRAAT et VOCALAB.

[Sicard 2021] Sicard, E. (2021) Analyse objective de la parole dysarthrique : évaluation d'une sélection d'indices acoustiques. <https://hal.archives-ouvertes.fr/hal-03139503>

[Sicard 2021b] Menin-Sicard, A., Sicard, E. (2021) Etude de cas de voix dans le cadre de la prise en charge orthophonique. Archives ouvertes HAL.

[Sidtis 2003] Sidtis, J. J., & Sidtis, D. V. L. (2003). A neurobehavioral approach to dysprosody. In *Seminars in Speech and Language* (Vol. 24, No. 02, pp. 093-106).

[Skodda 2009] Skodda, S., Rinsche, H., & Schlegel, U. (2009). Progression of dysprosody in Parkinson's disease over time—a longitudinal study. *Movement disorders*, 24(5), 716-722.

[Skodda 2011] Skodda, S., Visser, W., & Schlegel, U. (2011). Gender-related patterns of dysprosody in Parkinson disease and correlation between speech variables and motor symptoms. *Journal of Voice*, 25(1), 76-82.

[Sulter 1995] Sulter, A. M., Schutte, H. K., & Miller, D. G. (1995). Differences in phonetogram features between male and female subjects with and without vocal training. *Journal of voice*, 9(4), 363-377.

[Speedie 1990] Speedie, L. J., Brake, N., Folstein, S. E., Bowers, D., & Heilman, K. M. (1990). Comprehension of prosody in Huntington's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 53(7), 607-610.

[Speyer 2003] Speyer, R., Wieneke, G. H., van Wijck-Warnaar, I., & Dejonckere, P. H. (2003). Effects of voice therapy on the voice range profiles of dysphonic patients. *Journal of Voice*, 17(4), 544-556.

[Stemple 2018] Stemple, J. C., Roy, N., & Klaben, B. K. (2018). *Clinical voice pathology: Theory and management*. Plural Publishing.

[Titze 1992] Titze, I. R. (1992). Acoustic interpretation of the voice range profile (phonetogram). *Journal of Speech, Language, and Hearing Research*, 35(1), 21-34.

[Tuomi 2017] Tuomi, L., Johansson, M., Lindell, E., Folkestad, L., Malmerfors, M., & Finizia, C. (2017). Voice range profile and health-related quality of life measurements following voice rehabilitation after radiotherapy; a randomized controlled study. *Journal of Voice*, 31(1), 115-e9.

[Uloza 2005] Uloza, V., Pribuisiene, R., & Saferis, V. (2005). Multidimensional assessment of functional outcomes of medialization thyroplasty. *European Archives of Oto-Rhino-Laryngology and Head & Neck*, 262(8), 616-621.

[Viallet 2002] Viallet, F., et al. (2002, April). Effects of pharmacological versus electrophysiological treatments on parkinsonian dysprosody. In *Speech Prosody* (pp. 679-682). Laboratoire Parole et Langage.

[Vidou 2018] Vidou, C., Picard-Lauzon, J., Moisoni, V., Trudeau-Fisette, P., Rapin, L., & Ménard, L. (2018). Études des caractéristiques perceptives de l'accent d'emphase produit par des enfants atteints d'un trouble du spectre de l'autisme.

[Vorperian 2016] Vorperian, H. K., Kent, R. D., Austin, D. (2016). Normative Lifespan Database for Vowel Formants: Methodological Considerations and Preliminary Results, *Motor Speech Conference*, 2016

[Wingate 2007] Wingate, J. M., Brown, W. S., Shrivastav, R., Davenport, P., & Sapienza, C. M. (2007). Treatment outcomes for professional voice users. *Journal of voice*, 21(4), 433-449.

[Wuyts 2002] Wuyts, F. L., et al. (2002). Normative voice range profiles of untrained boys and girls. *Journal of Voice*, 16(4), 460-465.

[Wuyts 2003] Wuyts, et al. (2003). Effects of age, sex, and disorder on voice range profile characteristics of 230 children. *Annals of Otology, Rhinology & Laryngology*, 112(6), 540-548.