Faculté de Santé
Département Médecine Maïeutique et Paramédical
Centre de Formation Universitaire en Orthophonie

Certificat de Capacité d'Orthophoniste

Grade Master

Mémoire

Validation française du Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

Margaux SHANKS

Sous la direction de :

Madame Sandrine MICHEL, Orthophoniste

Monsieur Timothy POMMÉE, Chercheur postdoctoral et Orthophoniste

Jury composé de :

Monsieur Mathieu BALAGUER

Monsieur Guillaume ROUSSEAU

Juin 2024

DEDICACES ET REMERCIEMENTS

Je souhaite remercier tout d'abord mes deux encadrants, Sandrine Michel et Timothy Pommée, pour leur soutien, leur disponibilité, leur bienveillance et bien sûr pour leurs précieux conseils dont je n'aurais pas pu me passer pour ce travail !

Merci également à Déborah Mbagira pour nos échanges et pour ses conseils, et surtout merci de m'avoir confié le CAPE-Vf !

Je remercie également les membres de mon jury pour l'attention accordée à ce mémoire.

J'adresse tous mes remerciements aux personnes ayant participé à ce mémoire, en commençant par celles qui ont accepté d'être enregistrées. Merci aussi à tous les orthophonistes pour leur implication : qu'ils aient enregistré leurs patients ou évalué des voix, leur rôle a été essentiel dans la poursuite de ce mémoire.

Je tiens à remercier tout particulièrement Ariane Marchika et Guillaume Rousseau, non seulement pour leurs contributions à ce travail mais aussi pour ce qu'ils m'ont apporté dans leur pratique. J'adresse également une mention spéciale à Sarah, Léna, Capucine et Marion qui sauront combien leur aide m'a été précieuse pour ce mémoire.

Plus personnellement, je tiens à remercier mon père et mon frère pour leur légèreté et leur humour quotidiens ainsi que ma mère, qui a toujours su trouver les mots de soutien quand j'en ai eu besoin.

Merci aussi à Félix, source d'apaisement, de compréhension et de réconfort dans les nombreux moments de stress ; merci aussi pour sa patience, ses relectures, ses conseils avisés et surtout sa présence.

Enfin, merci à ma promo dont je suis fière de faire partie ; et merci surtout à ma dream team pour ces 5 années inoubliables qui sont passées beaucoup trop vite. Votre présence et votre soutien ont été indispensables lors de ces études et je suis fière de tout ce que nous avons vécu et accompli ensemble. On a passé tant de bons moments et j'ai hâte d'en vivre encore plus !

TABLE DES MATIÈRES

DEDICACES ET REMERCIEMENTS....1
TABLE DES MATIÈRES....2
TABLE DES ILLUSTRATIONS....5
TABLE DES ABREVIATIONS....6
TABLE DES ANNEXES....6
INTRODUCTION....7
PARTIE THEORIQUE....8
1. La voix....8
1.1. Définition....8
1.2. Anatomophysiologie....8
1.2.1. Soufflerie....8
1.2.2. Vibrateur....9
1.2.3. Résonateurs....9
2. La dysphonie....10
2.1. Définition....10
2.2. Epidémiologie et retentissement....10
2.3. Classification....10
2.3.1. Dysphonie par tension musculaire (DTM)....11
2.3.2. Dysphonie fonctionnelle (psychogène)....11
2.3.3. Dysphonie organique....12
3. Evaluation instrumentale....12
3.1. Examen laryngé....12
3.2. Evaluation acoustique....12
3.2.1. Paramètres vocaux....13
3.2.1.1. Hauteur....13
3.2.1.2. Intensité....14
3.2.1.3. Timbre....14
3.2.1.4. Le phonétogramme....15
3.2.2. Indicateurs pathologiques....15
3.3. Evaluation aérodynamique....16
3.4. Outils disponibles....17
3.5. Intérêts et limites de l'évaluation instrumentale....18
4. Évaluation subjective de la voix....19
4.1. Autoévaluation par le patient....19
4.1.1. VHI et adaptation française....20
4.1.2. Intérêts et limites de l'autoévaluation par le patient....20
4.2. Evaluation par le clinicien....21
4.2.1. GRBAS....21
4.2.2. CAPE-V....22
4.2.2.1. Validité psychométrique du CAPE-V....23
4.2.3. Influence de la langue parlée dans l'évaluation perceptive de la voix....24
4.2.3.1. Adaptation francophone du CAPE-V....24
4.2.4. Intérêts et limites de l'évaluation perceptive....25
PROBLEMATIQUE ET HYPOTHESES....27

I. Problématique 27
II. Hypothèses 27
METHODOLOGIE DE L'ETUDE 28
Objectifs 28
PARTIE 1 : ENREGISTREMENT DES LOCUTEURS 28
1. Populations 28
1.1. Taille, répartition et données démographiques de l'échantillon 29
2. Matériel et méthodes 30
2.1. Documents d'information et de consentement 30
2.2. Protocole d'enregistrement 30
2.3. Distribution du matériel 30
2.4. Déroulement des passations 30
2.4.1. Information et consentement 30
2.4.2. Environnement et outils 31
2.4.3. Enregistrement 31
2.5. Recueil et traitement des données 32
PARTIE 2 : EVALUATION DES VOIX 33
1. Population 33
1.1. Recrutement 33
1.1.1. Pré-recrutement 33
1.1.2. Echantillon final : taille et données démographiques des participants 33
2. Matériel et méthodes 34
2.1. Protocole d'évaluation 34
2.1.1. Ajustement du CAPE-Vf 35
2.2. Essai pilote 35
2.3. Lancement de l'étude 36
3. Analyse statistique 36
3.1. Fiabilité intra-juge 37
3.1.1. CAPE-Vf 37
3.1.2. GRBAS 37
3.2. Fiabilité inter-juges 37
3.2.1. CAPE-Vf 37
3.2.2. GRBAS 38
3.3. Validité de critère 38
3.3.1. Corrélation CAPE-Vf / GRBAS 38
3.3.2. Corrélations CAPE-Vf et GRBAS avec le VHI 38
RÉSULTATS ET ANALYSE 39
1. Statistiques descriptives 39
1.1. CAPE-Vf 39
1.2. GRBAS 40
2. Fiabilité intra-juge 41
2.1. CAPE-Vf 41
2.2. GRBAS 42
3. Fiabilité inter-juges 43
3.1. CAPE-Vf 43
3.2. GRBAS 44
4. Validité de critère 44

4.1. Corrélation CAPE-Vf / GRBAS....44
4.2. Corrélations CAPE-Vf et GRBAS avec le VHI....45
DISCUSSION....46
1. Etude des objectifs et hypothèses....46
1.1. Discussion autour de la première hypothèse : fiabilité du CAPE-Vf....46
1.1.1. Discussion autour de notre hypothèse secondaire : fiabilités du CAPE-Vf et du GRBAS....48
1.2. Discussion autour de la seconde hypothèse : corrélation des scores du CAPE-Vf et du GRBAS....48
1.3. Discussion autour de la troisième hypothèse : corrélation des scores du CAPE-Vf et du GRBAS avec ceux du VHI....49
1.4. Analyse qualitative des évaluations avec le CAPE-Vf....50
2. Limites et biais de notre étude....51
3. Perspectives....52
CONCLUSION....54
BIBLIOGRAPHIE....55
ANNEXES....70
RÉSUMÉ....89
ABSTRACT....90

TABLE DES ILLUSTRATIONS

Figures

Figure 1. Structure du pli vocal selon Giovanni et al. (2014) 9
Figure 2. Evolution de la f_0 moyenne en fonction de l'âge et du sexe (Baken et al., 2000) 13
Figure 3. Lien entre pression sous-glottique et intensité vocale lors d'un son filé d'après Henrich Bernardoni (2012) 14
Figure 4. Exemple de phonétogramme (Pabon, 2004, cité dans Lamesch, 2010) 15
Figure 5. EVA avec étiquetages verbaux telle que dans la version américaine de Kempster et al. (2009) 25
Figure 6. EVA sans étiquetage telle que dans la version francophone de Pommée et al. (2024) 25
Figure 7. Organisation générale du déroulé de l'étude 28
Figure 8. Régions d'exercice des orthophonistes ayant participé à l'étude 34
Figure 9. Répartition des scores en fonction des paramètres vocaux évalués dans le CAPE-V_f) 40
Figure 10. Fréquence d'apparition de chaque score en fonction du paramètre vocal évalué dans le GRBAS 41
Figure 11. EIR incluse dans la version du CAPE-V en Mandarin de Chen et al. (2018) 53

Tableaux

Tableau 1. Données démographiques des locuteurs et pathologies représentées au sein de l'échantillon 29
Tableau 2. Nombre de voix pour chaque grade de sévérité perçue, en fonction du sexe 32
Tableau 3. Répartition des voix dans les quatre listes d'écoute 32
Tableau 4. Composition finale des listes d'écoute 34
Tableau 5. Description des scores de chaque paramètre vocal évalué dans le CAPE-V_f 39
Tableau 6. Tableau de fréquence des scores du GRBAS 40
Tableau 7. Fiabilité intra-juge du CAPE-V_f 42
Tableau 8. Fiabilité intra-juge du GRBAS (Alpha de Krippendorff) 43
Tableau 9. Fiabilité inter-juges du CAPE-V_f 43
Tableau 10. Fiabilité inter-juges du GRBAS (Alpha de Krippendorff) 44
Tableau 11. Corrélations de Spearman entre les scores des paramètres comparables du CAPE-V_f et du GRBAS, pour chaque juge 44
Tableau 12. Corrélations de Spearman entre le score total du VHI, la sévérité globale du CAPE-V_f et le grade du GRBAS 50

TABLE DES ABREVIATIONS

ABI : Acoustic Breathiness Index
AVQI : Acoustic Voice Quality Index
CAPE-V : Consensus Auditory Perceptual Evaluation of Voice
CAPE-V$_f$: version francophone du Consensus Auditory Perceptual Evaluation of Voice
DSI : Dysphonia Severity Index
DTM : dysphonie par tension musculaire
EIQ : Écart interquartile
EIR : échelle à intervalles réguliers
ELS : Société Européenne de Laryngologie
EVA : échelle visuelle analogique
EVA2 : Evaluation Vocale Assistée
f$_0$: fréquence fondamentale
GRBAS : Grade, Roughness, Breathiness, Asthenia, Strain
HNR : harmonics-to-noise ratio
ICC : coefficient de corrélation intraclasse
MDVP : Multidimensional Voice Program
ORL : otorhinolaryngologiste
PIO : pression intra-orale
PSGE : pression sous-glottique estimée
QP : quotient phonatoire
SLA : sclérose latérale amyotrophique
SRP : Speech Range Profile
SVHI : Singing Voice Handicap Index
TMP : temps maximum phonatoire
VHI : Voice Handicap Index
VoiSS : Voice Symptom Scale
VRP : Voice Range Profile
V-RQOL : Voice-Related Quality of Life

TABLE DES ANNEXES

Annexe I : CAPE-V Procedures and Forms (Kempster et al., 2009) 70
Annexe II : CAPE-V$_f$ de Pommée et al. (2024) 74
Annexe III : Protocole d'enregistrement 75
Annexe IV : Version française du VHI de Woisard et al. (2004) 81
Annexe V : Protocole d'évaluation des voix avec le CAPE-V$_f$ 82
Annexe VI : Tableau GRBAS tel que présenté dans notre étude 88

INTRODUCTION

La dysphonie est un trouble touchant la voix et entraînant des manifestations diverses, allant de l'altération des paramètres et/ou des comportements vocaux observables, au ressenti anormal du patient vis-à-vis de sa voix. Sa prévalence est mal documentée en France, bien que ses conséquences aient un impact reconnu sur divers aspects de la vie des personnes touchées (Cohen et al., 2006 ; Deary et al., 2011 ; Scott et al., 1997 ; Titze et al., 1997).

La prise en charge orthophonique de la dysphonie dans toutes ses dimensions repose sur une évaluation adaptée et complète, nécessitant l'intervention de divers professionnels dont les oto-rhino-laryngologistes, les phoniatres et les orthophonistes. L'évaluation doit être, dans le meilleur des cas, multidimensionnelle comme le suggère la Société Européenne de Laryngologie (Dejonckere et al., 2001). Elle comporte idéalement à la fois des méthodes dites « objectives » comme les mesures aérodynamiques et acoustiques de la voix ; et des méthodes dont les résultats sont considérées comme « subjectifs », telles que l'examen direct des cordes vocales, l'auto-évaluation par le patient et l'évaluation perceptive.

A ce jour, l'outil de référence dans le cadre de l'évaluation perceptive de la voix est l'échelle GRBAS d'Hirano (1981). Cependant, il existe un autre outil prometteur adapté et validé en 10 langues : le Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Il est aujourd'hui utilisé dans de nombreux pays par des professionnels experts de la voix. Le CAPE-V est comparable au GRBAS : on retrouve un degré de corrélation élevé entre les évaluations d'un même échantillon vocal avec les deux outils (Karnell et al., 2007). Cependant, diverses études ont montré une meilleure fiabilité inter-juges pour le CAPE-V que pour le GRBAS (Nemr et al. 2012) ainsi que démontré que le CAPE-V serait un outil plus sensible aux paramètres fins des troubles vocaux que le GRBAS (Karnell et al., 2007). De plus, l'évaluation vocale avec le CAPE-V est effectuée selon un protocole standardisé, contrairement au GRBAS.

Dans le cadre de son mémoire de fin d'études, Déborah Mbagira a mené l'adaptation du CAPE-V en langue française, produisant ainsi un outil prêt à l'utilisation mais non validé. Or, pour qu'un outil puisse être utilisé en clinique et en recherche, il est important que son efficacité soit prouvée par le biais de sa validation psychométrique. Ainsi, à travers la validation du CAPE-Vf, ce mémoire s'inscrit dans une démarche de diversification des outils standardisés, valides et fiables à disposition des cliniciens de la voix pour effectuer des évaluations perceptives.

PARTIE THEORIQUE

1. La voix

1.1. Définition

Selon Cornut (2019), la voix est le « support acoustique de la parole ». Elle est créée par le passage de l'air pulmonaire à travers les plis vocaux puis les cavités de résonance, et revêt de nombreux aspects : chuchotement, murmure, cri, chant…

La voix est intimement liée à notre identité : de nombreuses études témoignent du fait que nous émettons des jugements sur le physique, les marqueurs sociaux et l'état psychologique du locuteur sur base de sa voix (Paya Herrero, 2009). Elle est propre à chacun, change et vieillit avec nous (Mueller, 1997). Comme le souligne Cornut (2019), la voix est aussi un important vecteur d'émotions. En effet, certaines activités vocales sont le résultat d'états émotionnels particuliers, comme le sont le rire et les pleurs. Parfois, la voix traduit même une intention au-delà des paroles.

1.2. Anatomophysiologie

L'appareil phonatoire s'organise classiquement selon trois étages (Le Huche & Allali, 2010) : la soufflerie, le vibrateur et les résonateurs.

1.2.1. Soufflerie

La respiration hors phonation s'effectue en deux temps : l'inspiration, suivie de l'expiration. Lors de ce premier temps, les poumons se remplissent d'air : c'est un phénomène actif lors duquel les muscles inspirateurs, c'est-à-dire les muscles élévateurs des côtes, les muscles intercostaux et le diaphragme, sont contractés tandis que les muscles de l'abdomen sont relâchés. L'expiration correspond à la vidange de l'air des poumons. A l'inverse de l'inspiration, c'est un phénomène passif pendant lequel on observe un relâchement des muscles auparavant tendus.

Lors de la prise de parole, on observe une modification de la respiration : le temps d'inspiration est raccourci, tandis que l'expiration s'allonge et devient active. L'air sort des poumons à une pression constante. On peut décrire trois phases successives dans l'activité des muscles respiratoires lors de la phonation (Cornut, 2019). En premier, l'individu prend une inspiration profonde : les muscles inspirateurs se contractent pour empêcher la fermeture de la cage thoracique face à l'importante pression qu'y engendre la prise d'air. Puis, alors que le volume d'air réduit pendant que le sujet parle, la pression de l'air dans les poumons diminue jusqu'à être insuffisante : les muscles inspirateurs cessent leur activité pour laisser place à l'action des muscles expirateurs. Enfin, quand se termine l'émission vocale, le volume d'air pulmonaire et la pression dans la cage thoracique sont minimes : le rôle des muscles abdominaux devient donc essentiel. L'enchaînement de ces mouvements inspiratoires et expiratoires correspond à ce qu'on appelle la coordination pneumophonique (Giovanni, 2021).

Selon le type de voix que souhaite produire le sujet, on relèvera différents débits expiratoires et des variations dans les volumes d'air mobilisés (Holmberg et al., 1988).

1.2.2. Vibrateur

L'air expiré quitte les poumons par la trachée pour venir rejoindre la zone laryngée. Le larynx est un organe de la gorge impliqué dans les fonctions de respiration, de phonation et de déglutition. Il se compose d'un squelette cartilagineux, de muscles et de ligaments. Dans le larynx, on trouve les plis vocaux qui créent la vibration à l'origine de la voix. Au nombre de deux et orientés horizontalement, ils sont composés de plusieurs couches de tissus (Figure 1), qui d'après Hirano (1974) peuvent être sous-divisées en deux structures vibratoires :

- le « corps », constitué du muscle thyro-aryténoïdien appelé « muscle vocal » et du ligament vocal ;
- et la « couverture » qui repose sur le corps, composée de l'épithélium et de la couche superficielle de la lamina propria.

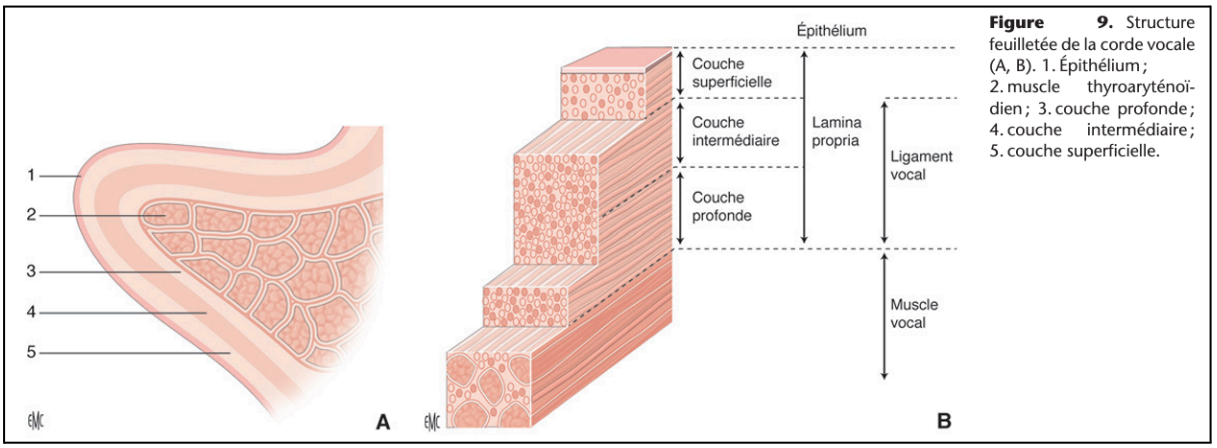


Figure 9. Structure feuilletée de la corde vocale (A, B). 1. Épithélium ; 2. muscle thyroaryténoïdien ; 3. couche profonde ; 4. couche intermédiaire ; 5. couche superficielle.

Figure 1. Structure du pli vocal selon Giovanni et al. (2014)

Les muscles laryngés intrinsèques sont responsables de la mobilisation des plis vocaux. Le crico-thyroïdien et le thyro-aryténoïdien médial permettent leur mise en tension, le crico-aryténoïdien postérieur provoque leur abduction, tandis que le crico-aryténoïdien latéral, le thyro-aryténoïdien et les inter-aryténoïdiens permettent leur adduction (Giovanni & Mattei, 2021).

C'est dans cette position en adduction, couplée à une mise en tension adéquate, que se placent les plis vocaux à l'approche de l'air expiratoire. La pression aérienne sous la glotte augmente jusqu'à ce que l'air s'échappe à travers les plis vocaux. L'ouverture et la fermeture de la glotte à raison de plusieurs centaines de cycles par seconde créent ainsi la vibration à l'origine du son (Lafon & Cornut, 1962).

1.2.3. Résonateurs

D'après Cornut (2019), on appelle résonateurs « les cavités que le son laryngé traverse avant d'arriver à l'air libre ». Ce sont donc l'oropharynx, la cavité buccale, le nasopharynx et les fosses nasales. Ces cavités varient dans leur forme et leur taille car les organes qui les créent sont mobiles. La morphologie des cavités de résonance varie aussi inter-individuellement. Ces variations anatomiques modifient les caractéristiques acoustiques du tractus vocal, produisant ainsi un son vocal propre à chaque individu et à chaque son voisé, dont notamment les voyelles.

Une fois l'air expiratoire passé par les plis vocaux, le son laryngé riche en fréquences passe par les cavités de résonance. Ces dernières, selon leur configuration, renforcent certaines zones fréquentielles avant que le son ne s'échappe par la bouche (Titze, 2001). D'après Titze (1994), les résonateurs exercent également un phénomène rétroactif sur le larynx, qui facilite les vibrations laryngées.

2. La dysphonie

2.1. Définition

Le terme « dysphonie » est employé pour décrire toute pathologie vocale plus ou moins durable ressentie par un individu ou perçue par son entourage (Dejonckere et al., 2001 ; Le Huche & Allali, 2010b). L'atteinte peut concerner la hauteur, l'intensité, le timbre, mais aussi les caractéristiques prosodiques et mélodiques de la voix (Dejonckere et al., 2001). Ces manifestations sont le reflet à la fois de l'état de l'appareil vocal et de l'utilisation de sa voix par un sujet (Le Huche & Allali, 2010a).

2.2. Epidémiologie et retentissement

Bien qu'il existe de nombreuses études sur la prévalence des dysphonies chez les professionnels de la voix, on en retrouve peu sur la prévalence des troubles vocaux dans la population générale. À ce jour, aucune étude n'a été menée sur la prévalence de la dysphonie en France.

Une étude récente de Huston et al. (2023) révèle qu'un américain sur cinq aurait été atteint de dysphonie au cours de sa vie et près d'un répondant sur neuf indique souffrir d'un trouble vocal au moment de l'étude. Les professionnels de la voix, c'est-à-dire des personnes exerçant une profession nécessitant une pratique vocale fréquente et/ou intense, tels que les enseignants, acteurs, chanteurs, etc., sont davantage touchés par ces troubles, du fait des exigences vocales qu'implique leur activité. Aucune disparité des sexes n'est relevée par Huston et al., bien que d'autres études aient rapporté une prévalence et une tendance à consulter plus élevée chez les sujets féminins (Cohen et al., 2012 ; Fritzell, 1996 ; Lyberg-Åhlander et al., 2019 ; Marchese et al., 2022 ; Sheyona & Devadas, 2022).

La dysphonie a un impact sur de nombreux aspects de la vie quotidienne (Cohen et al., 2006) parmi lesquels on peut citer la vie professionnelle (Titze et al., 1997), sociale (Scott et al., 1997) et l'état psychologique (Deary et al., 2011). Ainsi, nous pouvons la considérer comme un enjeu de santé publique (Sheyona & Devadas, 2022).

2.3. Classification

Il existe de nombreuses classifications des troubles vocaux. En effet, aucun consensus quant à la terminologie et aux classifications des troubles vocaux n'est à ce jour établi ; c'est un sujet complexe et controversé - voire problématique (Baker et al., 2007). Face à ce manque, Payten et al. (2022) proposent une classification et une nomenclature reprenant les éléments les plus pertinents et fréquemment rencontrés dans la littérature existante. Les auteurs proposent de classer les dysphonies en trois grandes catégories : les dysphonies par tension musculaire (DTM), les dysphonies fonctionnelles ou « psychogènes » et les dysphonies organiques.

2.3.1. Dysphonie par tension musculaire (DTM)

Ce type de dysphonie est caractérisé par des tensions visibles et palpables au niveau de la musculature intrinsèque et extrinsèque du larynx, ainsi que des mécanismes chroniques de déséquilibre musculaire, le tout dans l'absence de cause neurologique ou structurelle.

Elle peut être **primaire**, avec un mode d'apparition généralement progressif, des signes et symptômes témoignant d'un malmenage ou d'un surmenage vocal prolongé ou en réponse à des changements liés à la santé ou à des exigences sociales, professionnelles ou de performance. Les signes et symptômes sont congruents avec une hyperfonction musculaire (par exemple : tension des muscles laryngés extrinsèques, inconfort musculaire ou douleurs suite à des efforts vocaux…). Sur le plan clinique, on observe une constriction laryngée, supraglottique et/ou une hyperadduction des plis vocaux prolongées lors de la phonation. On observe des caractéristiques acoustiques et/ou perceptives variables, pouvant aller d'une voix normale à forcée ou rauque. Dans la DTM, la phonation et les mouvements laryngés sont améliorés suite à la mise en place de techniques vocales et lors de comportements vocaux non-langagiers.

La DTM peut être **secondaire** à un autre trouble. On retrouve alors une dysphonie ou une aphonie dans laquelle apparaissent les signes et symptômes de la DTM primaire, en présence de lésions des plis vocaux induites par un traumatisme vocal. Parmi les DTM secondaires on retrouve les inflammations ou œdèmes de Reinke induits par traumatisme vocal, les nodules, les ulcères de contact, les polypes…

2.3.2. Dysphonie fonctionnelle (psychogène)

La dysphonie fonctionnelle est caractérisée par une perte du contrôle moteur volontaire et/ou une incapacité à auto-réguler l'initiation, l'inhibition et le maintien de la phonation. Les symptômes ne sont pas imputables à une cause neurologique ou structurelle. On y retrouve les **aphonies ou dysphonies apparaissant de manière soudaine**, après un événement ou une maladie (par exemple une infection des voies respiratoires supérieures, une chirurgie…). Des antécédents de maladie psychogène ou de trouble dissociatif de conversion peuvent être présents. A l'examen laryngé on repère une grande variété de comportements laryngés incompatibles avec les observations relevées lors de l'examen clinique. De manière générale, les symptômes sont incohérents entre eux. Concernant les observations cliniques, des symptômes non-caractéristiques d'un malmenage vocal (cf. DTM) sont présents et on retrouve une impossibilité à obtenir volontairement une amélioration des signes et symptômes. De courts instants d'amélioration fonctionnelle peuvent être observés sur des tâches non-voisées ou des comportements vocaux non-langagiers tels que la toux, le rire ou les pleurs.

Un sous-type de la dysphonie fonctionnelle est la **puberphonie.** Elle apparaît pendant l'adolescence ou à l'âge adulte chez des sujets masculins au développement hormonal normal. Elle est caractérisée par une voix de fausset présente sur une durée prolongée, une instabilité de la hauteur vocale avec des ruptures et/ou des sauts d'octaves, une fermeture glottique incomplète donnant à la voix un aspect soufflé, le tout dans l'absence d'un serrage supraglottique observable.

2.3.3. Dysphonie organique

La dysphonie organique est caractérisée par une atteinte pathologique de la structure et/ou de la mobilité du larynx nécessitant une prise en charge vocale ou une intervention chirurgicale.

La dysphonie organique est **structurelle** quand elle est due à une anomalie structurelle impliquant une partie du larynx. Elle est d'origine congénitale ou développementale, apparaissant sans événement traumatique déclencheur. On y retrouve les lésions laryngées malignes, les anomalies de l'épithélium ou de la lamina propria des plis vocaux, les anomalies vasculaires et les changements du larynx de nature congénitale ou maturationnelle.

On retrouve également des dysphonies organiques de type **inflammatoire**. Elles peuvent être d'origine infectieuse : primaire laryngée (virale ou bactérienne), fongique ou secondaire (infection pulmonaire, rhinosinusite) ; ou non-infectieuse : inflammation du larynx, reflux pharyngo-laryngé, allergie, médicament, d'origine auto-immune ou non-spécifique.

Il existe aussi des dysphonies organiques **neuromusculaires** dues à une atteinte du système nerveux périphérique, comme une pathologie du nerf laryngé supérieur ou récurrent, une myasthénie ou une neuropathie. Elles peuvent aussi être le résultat d'une atteinte du système nerveux central comme dans la sclérose latérale amyotrophique, la maladie de Parkinson ou encore dans les atrophies multisystémiques. Enfin, elles peuvent être issues de troubles du mouvement comme dans le cas de la dysphonie spasmodique et des tremblements dystoniques et vocaux.

Les dysphonies organiques peuvent également être qualifiées de **traumatiques**, catégorie comprenant les diverses séquelles possibles issues d'un événement traumatique tel qu'un traumatisme interne de la muqueuse laryngée, un traumatisme externe de la structure laryngée ou une autre cause qu'elle soit mécanique, chimique ou thermique.

3. <u>Evaluation instrumentale</u>

3.1. Examen laryngé

L'examen du larynx par videolaryngostroboscopie est recommandé par la Société Européenne de Laryngologie (ELS) (Dejonckere et al., 2001). Il est réalisé par un médecin avec un vidéolaryngoscope rigide passé par la bouche du patient, ou bien avec un nasofibroscope. Ce dernier est un tube fin et souple au bout duquel se situe une caméra, que l'on passe par le nez du patient afin d'observer l'aspect et la mobilité du vélum, du pharynx et du larynx au repos et en mouvement (lors de la phonation, de la déglutition…). La stroboscopie, de par sa technique de « flash » lumineux permettant la reconstruction du cycle vibratoire sous la forme d'un « faux ralenti », permet une meilleure visualisation de la vibration des plis vocaux difficilement visible à l'œil nu. La videolaryngostroboscopie offre donc la possibilité d'observer directement l'aspect et la mobilité des plis vocaux, dont la fermeture glottique et la régularité, l'ondulation et la symétrie de la vibration.

Les images recueillies sont enregistrées afin d'être visionnées à posteriori et servir de références lors de futurs examens.

3.2. Evaluation acoustique

3.2.1. Paramètres vocaux

Trois paramètres vocaux sont utilisés pour décrire la voix d'un individu : la hauteur, l'intensité et le timbre.

3.2.1.1. Hauteur

On appelle « hauteur » la sensation auditive d'aigu ou de grave (Ghio, 2021). En termes acoustiques, cela correspond à la fréquence du son, définie comme le nombre de vibrations par seconde ; elle est exprimée en Hertz. Plus la fréquence d'un son est basse, plus il sera perçu comme grave ; plus elle est élevée, plus il sera perçu comme aigu.

La hauteur tonale varie tout au long de la prise de parole autour de ce qu'on appelle la « fréquence fondamentale » ou f_0 (Cornut, 2009). Elle correspond à la fréquence moyenne de vibration des plis vocaux (Whalen & Levitt, 1995). Celle-ci est propre à chaque individu et dépend de nombreux facteurs dont la morphologie des organes de la parole, l'âge (Dehqan et al., 2012), le genre (Skuk & Schweinberger, 2014) et les facteurs linguistiques et culturels liés à l'individu (Traunmüller & Eriksson, 1995 ; Wagner & Braun, 2003). La f_0 moyenne évolue tout au long de la vie chez les hommes comme chez les femmes (Baken et al., 2000) (Figure 2).

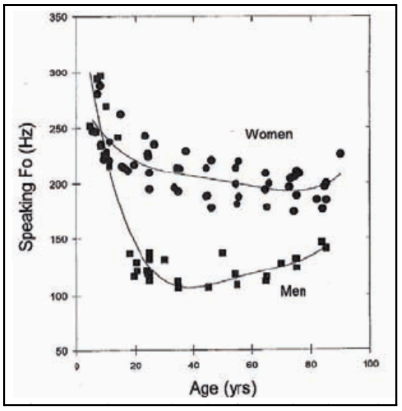


Figure 2. Evolution de la f_0 moyenne en fonction de l'âge et du sexe (Baken et al., 2000)

La totalité des fréquences pouvant être produites par un individu se nomme « étendue vocale ». Au sein de l'étendue on retrouve la tessiture, qui correspond à tous les sons produits aisément et de bonne qualité (Révis & Ravera-Lassalle, 2021). Pour effectuer toute la gamme des sons possibles, il faut basculer entre les différents registres de voix grâce à des changements entre les différents mécanismes laryngés. Ceux-ci sont au nombre de quatre :

- le **mécanisme 0** ou « fry / strohbass » : le plis vocaux sont épais et courts ; la voix est très grave, la f_0 étant située entre 35 et 50 Hz chez les femmes comme chez les hommes (Chen et al., 2002) ;
- le **mécanisme 1** ou « voix de poitrine » : les plis vocaux sont épais et vibrent sur toute leur surface ; c'est le registre le plus fréquemment utilisé dans la voix parlée en occident (Lamesch, 2010) ;
- le **mécanisme 2** ou « voix de fausset » également connu sous le nom « voix de tête » : les plis vocaux s'affinent et s'allongent sous l'action de la bascule du cartilage laryngé ; leur surface de contact diminue, la voix devient plus aiguë (Lamesch, 2010) ;

- le **mécanisme 3** ou « voix de sifflet » : les plis vocaux sont très tendus et allongés (Miller & Schutte, 1993), la voix est très aiguë.

Certaines fréquences font partie de la « zone de recouvrement », ce qui signifie qu'elles peuvent être produites dans deux mécanismes différents. C'est le cas notamment pour certains sons présents dans les mécanismes M1 et M2 (Roubeau et al., 2009).

3.2.1.2. Intensité

L'intensité d'un son dépend de l'amplitude de son signal sonore, qui reflète le niveau de pression acoustique. Plus l'amplitude d'un son est importante, plus on le percevra comme fort ; plus elle est faible, plus on percevra le son comme faible également (Ghio, 2021). On la mesure en décibels (dB).

La voix humaine peut être produite à diverses intensités allant du chuchotement au cri. L'intensité de la voix dépend de la pression sous-glottique à l'entrée de l'appareil vocal, qui elle-même dépend de la force de l'expiration de l'air pulmonaire. Ainsi, comme le souligne Henrich Bernardoni (2012), plus la pression sous-glottique augmente, plus l'intensité vocale en sortie des lèvres augmente (Figure 3).

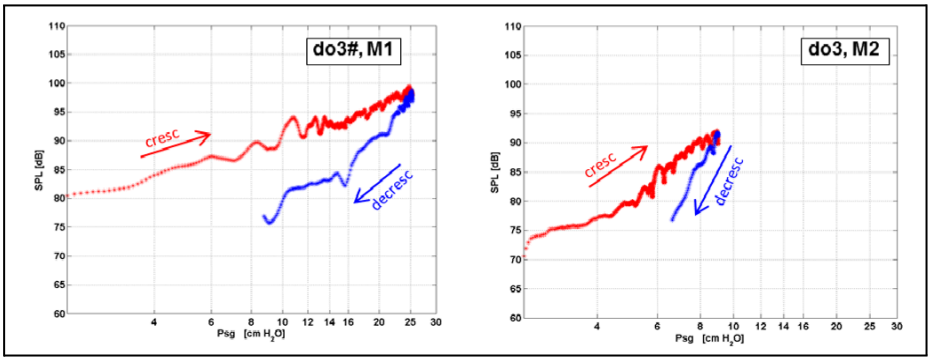


Figure 3. Lien entre pression sous-glottique et intensité vocale lors d'un son filé d'après Henrich Bernardoni (2012)

L'intensité varie tout au long d'une prise de parole, sa valeur oscillant autour d'une « intensité moyenne usuelle » (Cornut, 2019). Selon Gelfer & Young (1997), celle-ci tourne autour de 70 dB pour les hommes et 68 dB pour les femmes. Le locuteur ajuste son intensité vocale en fonction de son environnement, par exemple en l'augmentant lorsqu'il y a du bruit ou quand son interlocuteur est à distance (Morsomme, 2001).

3.2.1.3. Timbre

Le timbre est ce qui donne à la voix son caractère individuel et reconnaissable, sa « couleur ». Il peut même être un signe distinctif pour certains individus, en particulier pour les professionnels de la voix (Cornut, 2019). D'après McAdams & Bregman (1979), le timbre comprend tout ce qui n'appartient pas au domaine de la hauteur tonale ou de l'intensité : c'est donc ce qui permet de distinguer deux sons avec la même fréquence et intensité. Le timbre dépend de propriétés anatomo-fonctionnelles comme la forme des plis vocaux et de leur structure, ainsi que de la forme des résonateurs.

D'un point de vue acoustique, selon Menin-Sicard & Sicard (2016) le timbre correspond à « la répartition des harmoniques et du bruit de la voix ou d'un instrument de musique dans le domaine fréquentiel ». La voix peut être plus ou moins riche en harmoniques. Chez les chanteurs, on retrouve une zone de résonance particulière appelée « formant du chanteur » : les harmoniques présents dans cette zone donnent une sonorité et une richesse singulière à la voix, lui permettant de se distinguer de l'orchestre qui l'accompagne. Une voix peut également être pauvre en harmoniques : on la qualifie alors de « plate », de « terne » ou de « blanche » (Révis & Ravera-Lassalle, 2021).

3.2.1.4. Le phonétogramme

Il existe un test permettant de visualiser la relation entre la hauteur de la voix et la dynamique vocale associée, appelée « champ dynamique ». Pour chaque fréquence, on demande au sujet de produire le son le plus faible possible puis le plus fort possible. Le test peut aussi être effectué sur la base d'une sirène. On obtient donc une représentation graphique permettant de visualiser l'intensité en fonction de la fréquence, comme illustrée dans la figure 4. On l'appelle phonétogramme ou « Voice Range Profile » (VRP). Il existe une version du phonétogramme appelée « Speech Range Profile » (SRP) permettant d'évaluer le champ dynamique sur une tâche plus fonctionnelle de parole. Il semblerait que la sensibilité du SRP soit plus élevée que celle du VRP dans la détection de voix pathologiques (Ma et al., 2007).

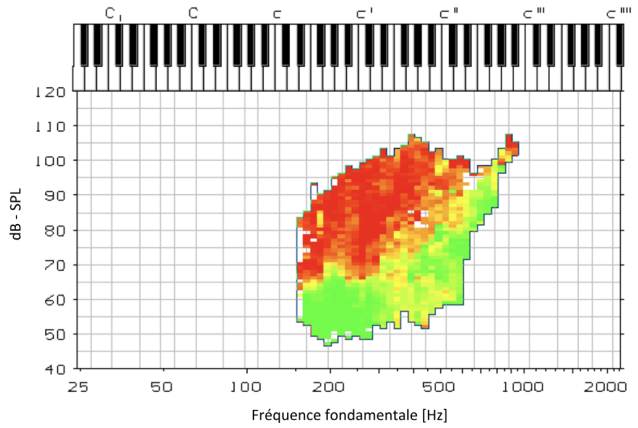


Figure 4. Exemple de phonétogramme (Pabon, 2004, cité dans Lamesch, 2010)

3.2.2. Indicateurs pathologiques

Certaines mesures acoustiques permettent de rendre compte d'une altération vocale : le jitter, le shimmer, le harmonics-to-noise ratio et la f_0 sont les mesures acoustiques les plus utilisées à cet effet lors de l'évaluation vocale (Maryn et al., 2009). Ces deux premières sont intégrées aux recommandations de l'ELS (Dejonckere et al., 2001). Les analyses de ces paramètres sont généralement faites à partir d'un /a/ tenu.

Le **jitter** permet d'évaluer les variations à court terme (c'est-a-dire de cycle en cycle) de la fréquence fondamentale dans le temps. On l'obtient en faisant la moyenne de la différence de f_0 entre deux cycles vibratoires consécutifs (Ghio, 2021). Il permet le calcul d'autres mesures, dont la plus fréquemment utilisée est le « jitter factor » (Giovanni & de Saint-Victor, 2013). Le jitter rend compte de l'instabilité fréquentielle à court terme du son vocal.

Le **shimmer** évalue les variations à court terme de l'amplitude du cycle vibratoire sur une durée déterminée (Ghio, 2021). Il permet donc d'apprécier la stabilité de l'intensité vocale de cycle en cycle. Il est à différencier de l'instabilité de l'intensité à moyen terme, qui correspond au tremblement vocal ou au tremolo.

Les normes pour le jitter et le shimmer varient de manière importante en fonction de la culture, de certains facteurs linguistiques et même de l'outil employé pour les détecter (Demirhan et al., 2016 ; Sicard & Menin-Sicard, 2013), mais elles sont généralement de l'ordre de de 0,5-1% pour le jitter et 3% pour le shimmer (Teixeira et al., 2013), les valeurs en dessous de ce seuil étant considérées comme normales.

Le **harmonics-to-noise ratio** (HNR) est un rapport permettant de comparer les parties harmoniques et non-harmoniques d'un son. Chaque voyelle étant composée de formants (c'est-a-dire de groupes d'harmoniques) spécifiques, la valeur du HNR dépend de celle qui est produite. Il est utilisé pour quantifier le bruit dans un signal sonore : plus le HNR est élevé, plus la voix est riche en harmoniques. Inversement, moins il est élevé, plus la voix semblera faible (Teixeira, 2013), témoignant d'un mauvais accolement des plis vocaux. D'après Boersma (1993), un HNR inférieur à 7 est considéré comme pathologique.

Plusieurs études établissent un lien entre jitter et shimmer, et la raucité et/ou le souffle vocal perçu (Iwata & von Leden, 1970 ; Jones et al., 2001 ; Lieberman, 1963). Cependant, il semblerait que le HNR soit un meilleur indicateur car davantage corrélé au souffle et à la raucité (Ferrand, 2002 ; Krom, 1995 ; Yumoto, Sasaki & Okamura, 1984).

3.3. Evaluation aérodynamique

L'évaluation aérodynamique consiste en la mesure de phénomènes biomécaniques entrant en jeu lors de la phonation, parmi lesquels on peut citer le temps maximum phonatoire, la capacité vitale, le quotient phonatoire, le débit expiratoire et la pression sous-glottique estimée.

Le **temps maximum phonatoire** (TMP) correspond à la durée maximale de phonation après une inspiration profonde, généralement mesurée à partir d'une voyelle. On considère qu'il est le reflet du fonctionnement des plis vocaux (Iwata & von Leden, 1970). Plusieurs essais peuvent être demandés, parmi lesquels le meilleur, c'est-à-dire le plus long, sera conservé (Giovanni & de Saint-Victor, 2013). Le TMP moyen se situe à 25,7 secondes chez les femmes et à 34,6 chez les hommes (Hirano et al., 1968). On considère qu'il est anormal lorsqu'il est inférieur à 10 secondes, et qu'il a un impact sur la vie quotidienne s'il est de moins de 5 secondes (Omori, 2011). Boone (1981) propose de comparer le TMP au temps d'expiration hors phonation en effectuant la mesure du **rapport s/z**. Celui-ci est calculé à partir de la production d'un /s/ et d'un /z/. Selon Boone, ce rapport est de l'ordre de 1 chez les sujets normophoniques et inférieur à 1 chez les sujets dysphoniques, puisque le manque d'efficacité phonatoire entraînerait une diminution du TMP sur le /z/. Cependant, la littérature révèle un manque de consensus quant à la pertinence de cette mesure. Alors que plusieurs études démontrent un lien entre le rapport s/z et la pathologie vocale, d'autres comme celle

de Gelfer & Pazera (2006) trouvent des similitudes entre les résultats de sujets normophoniques et dysphoniques, comme souligné par Gilman (2021).

La **capacité vitale** est le volume d'air maximal mobilisé lors d'un cycle de respiration, mesuré à l'aide d'un spiromètre. Elle se situe entre 1 et 1,5 l chez les adultes (Giovanni, 2021).

Le **quotient phonatoire** (QP) est une mesure de rendement vocal permettant l'estimation de l'air consommé pendant la phonation par unité de temps. On l'obtient en divisant la capacité vitale par le TMP. Le QP moyen est évalué à 145 ml/sec chez les hommes et 137 ml/sec chez les femmes (Hirano et al., 1968) et considéré comme anormal lorsqu'il est inférieur à 69 ml/sec pour les hommes et 78 ml/sec pour les femmes.

Le **débit expiratoire** correspond au débit d'air moyen qui traverse la glotte en une minute. La norme se situe en moyenne à 101 ml/sec pour les hommes et 92 ml/sec chez les femmes. On considère qu'il est pathologique s'il est de moins de 46 ml/sec chez un homme et 43 ml/sec chez une femme (Hirano et al., 1968).

Lors de la phonation, la pression de l'air sous-glottique augmente afin de provoquer la vibration des plis vocaux. Cette pression peut être mesurée grâce à la pression intra-orale (PIO) par le biais d'un capteur de pression dans la bouche. Celui-ci mesure la PIO juste avant l'explosion du [p] lors de répétitions de la syllabe /pi/. On obtient ainsi une mesure, appelée **pression sous-glottique estimée** (PSGE) et exprimée en hPa. Elle est un marqueur de forçage vocal (Hillman et al., 1990). La valeur moyenne normale pour les femmes est de 7,70 hPa, et de 7,46 hPa pour les hommes (Morsomme, 2001).

3.4. Outils disponibles

Il existe une variété d'outils utilisés par des cliniciens de la voix, dont les orthophonistes, lors de l'évaluation vocale en structure hospitalière ou en cabinet libéral.

Les **index de qualité vocale** permettent de quantifier le degré d'un trouble vocal. Cette quantification est effectuée sur la base d'analyses statistiques de certaines combinaisons de paramètres acoustiques, qui sont ensuite comparées à des seuils de référence. Selon les résultats obtenus, la voix évaluée est qualifiée de « saine » ou de « pathologique ». Il existe plusieurs index, dont les plus utilisés sont le Dysphonia Severity Index (DSI) de Wuyts et al. (2000), l'Acoustic Breathiness Index (ABI) de Barsties et al. (2017) et l'Acoustic Voice Quality Index (AVQI) de Maryn et al. (2009), ce dernier ayant été adapté et validé en langue française par Pommée et al. en 2018. Ils peuvent être implémentés dans d'autres logiciels tels que Praat et Audacity, que nous aborderons prochainement.

Divers **logiciels** ont été élaborés pour assister les cliniciens dans la mesure de paramètres acoustiques lors de l'évaluation vocale. Ceux-ci permettent la visualisation, l'analyse et la manipulation de sons enregistrés. Parmi ces logiciels nous pouvons citer Audacity, créé en 2000 par Dannenberg & Mazzoni ; Praat, développé par Boersma & Weenink en 1992 et qui propose une fonctionnalité de « Voice Report » permettant la visualisation de certains paramètres vocaux ; ou

encore le gratuiciel VoxLog qui permet la mesure de l'AVQI et de l'ABI. Cependant, l'ABI n'ayant pas encore été validé en français, les scores-seuils en français n'y sont pas encore intégrés. Conçu spécifiquement pour les orthophonistes, le logiciel Vocalab est un outil polyvalent offrant une alternative plus spécifique à la profession que Praat et Audacity. Il permet notamment d'évaluer cinq indicateurs de perturbation vocale : le fondamental usuel moyen, l'étendue vocale, le TMP, l'aire dynamique vocale et les rapports s/z et a/z (Sicard et al., 2021).

Enfin, d'autres dispositifs tels que l'**Évaluation Vocale Assistée** (EVA2) développée par le Laboratoire de la Parole et du Langage de l'Université Aix-Marseille, ou le **Multidimensional Voice Program** (MDVP) créé par CSL-Kay Elemetrics permettent une évaluation multidimensionnelle des paramètres acoustiques et aérodynamiques de la phonation.

3.5. Intérêts et limites de l'évaluation instrumentale

L'évaluation instrumentale, sous ses diverses formes, présente un avantage majeur : elle permet l'obtention de mesures relativement objectives. L'utilisation d'un protocole standardisé tel que celui proposé par l'ELS (Dejonckere et al., 2001) permet de réduire les biais et d'obtenir une meilleure reproductibilité des mesures effectuées. L'évaluation instrumentale permet également au clinicien d'observer ou de quantifier des phénomènes imperceptibles à l'oreille ou à l'œil nu. Utilisée en complément d'autres mesures subjectives, elle informe le clinicien sur les multiples mécanismes mis en jeu dans la parole et la voix (Mehta & Hillman, 2007).

Cependant, l'évaluation instrumentale présente également ses limites, notamment concernant sa fiabilité qui fait l'objet de diverses critiques. En effet, bien que les résultats soient recueillis suivant des méthodes relativement objectives, le recueil et l'interprétation des résultats reste dépendante de la perception du clinicien qui les effectue (Ghio, 2012). Il existe d'ailleurs un manque de consensus sur les corrélats perceptifs des mesures acoustiques (Brockmann-Bauser et al., 2010) : en effet, le shimmer et le jitter correspondent à différentes caractéristiques vocales (souffle, raucité…) selon l'individu qui les juge. De plus, l'évaluateur effectue certains choix méthodologiques pouvant influencer l'objectivité de la mesure, tels que le choix du stimulus vocal employé ou encore la taille et l'emplacement de la fenêtre d'analyse. Ainsi, il est important que ces paramètres soient contrôlés afin de pouvoir être reproduits lors d'évaluations subséquentes, permettant ainsi une évaluation plus fiable.

La fiabilité des résultats de mesures instrumentales est également contestable du fait de variables indépendantes de l'intervention des cliniciens. Selon Behrman (2005), plusieurs instruments mesurant les mêmes paramètres acoustiques peuvent produire des résultats différents. La notion même de normes et de seuils pathologiques est compliquée à aborder, en particulier pour les mesures acoustiques. Il existe une grande diversité dans les résultats obtenus par les études les ayant investiguées, notamment du fait de variations importantes en termes de protocoles, de matériel et de logiciels d'analyse (Goy et al., 2013). La fiabilité de certains paramètres acoustiques en tant qu'indicateurs pathologiques de la voix est également questionnée. Effectivement, le jitter et le shimmer ont fait l'objet de plusieurs critiques, notamment car leur fiabilité est limitée lorsqu'un signal sonore est instable ou complexe (Giovanni et de Saint-Victor, 2013). D'après Leong et al. (2013), le

jitter ne serait interprétable que s'il est inférieur à 5%. De plus, une étude de Holmberg et al. (2003) révèle que chez des patientes atteintes de nodules et ayant suivi un protocole de rééducation, aucune évolution significative n'est observée au niveau des paramètres acoustiques et aérodynamiques évalués malgré une amélioration de l'aspect des plis vocaux ainsi que de la qualité vocale constatées. Enfin, l'évaluation instrumentale étant effectuée à un instant T ou sur un échantillon de voix limité, celle-ci est peu représentative du trouble dans sa globalité.

L'évaluation instrumentale présente également des limites en termes de faisabilité et de mise en pratique. Tout d'abord, elle requiert un matériel spécifique : microphone et ordinateur pour l'évaluation acoustique, matériel d'exploration médicale pour l'examen laryngé et logiciels et appareils de mesure pour l'évaluation aérodynamique. Ce matériel, en particulier celui nécessaire pour les évaluations aérodynamiques et les examens laryngés, est peu accessible aux professionnels exerçant en cabinet libéral du fait des prix élevés et des contraintes qu'il engendre en termes de place (Behrman, 2005).

Pour que les mesures obtenues soient exploitables, l'emploi de ces instruments nécessite des connaissances techniques spécifiques sur l'utilisation du matériel et l'interprétation des résultats (Giovanni & de Saint-Victor, 2013). En effet, dans son étude sur les méthodes d'évaluation des cliniciens de la voix, Behrman (2005) révèle que 60% des professionnels sondés peinent à appliquer les bonnes méthodes d'analyse acoustique lorsqu'ils sont confrontés à des signaux vocaux irréguliers. Enfin, ces examens peuvent parfois être compliqués à effectuer d'un point de vue pratique. Certains sont relativement invasifs, comme il en est le cas pour l'examen laryngé par nasofibroscope ou vidéolaryngoscope. Le jitter et le shimmer, deux mesures basées sur la détection de f_0, nécessitent un effort de phonation soutenue compliqué à réaliser pour certains patients dysphoniques. Enfin, pour que les mesures soient les plus fiables possibles, elles doivent idéalement être effectuées dans un environnement calme et sans bruit de fond, ce qui n'est pas toujours réalisable.

4. Évaluation subjective de la voix

L'évaluation subjective de la voix peut être menée par le clinicien ou par le patient, dans quel cas on parle d'auto-évaluation. Ces deux approches sont incluses dans le protocole d'évaluation fonctionnelle de la voix pathologique publié par l'ELS (Dejonckere et al., 2001).

4.1. Autoévaluation par le patient

Selon Francis et al. (2017), l'auto-évaluation par le patient correspond à toute mesure de son état de santé issue directement de lui-même, sans interprétation par un tiers. Dans le domaine de la voix, il en existe une multitude permettant d'évaluer divers aspects de la dysphonie et de ses conséquences au quotidien. Parmi celles-ci, les échelles d'auto-évaluation du handicap vocal sont recommandées dans la prise en charge globale de la dysphonie puisqu'elles permettent de comparer l'altération de la qualité vocale perçue par le clinicien à la sévérité du handicap quotidien vécu par l'individu (Dejonckere et al., 2001). Plusieurs échelles permettent de mesurer l'impact de la dysphonie sur la qualité de vie, comme le **Vocal Performance Questionnaire** (Deary et al., 2016), le **Voice Symptom Scale** ou VoiSS (Deary et al., 2003), le **Voice-Related Quality of Life** (V-RQOL)

d'Hogikyan & Sethuraman (1999) et le **Voice Handicap Index** (VHI) de Jacobson et al. (1997), ces deux dernières étant les plus fréquemment utilisées (Reghunathan & Bryson, 2019).

4.1.1. VHI et adaptation française

Le Voice Handicap Index (VHI) a été créé et validé en 1997 par Jacobson et al., dans le but de pallier le manque d'outils évaluant l'impact psychosocial des troubles de la voix. C'est un auto-questionnaire composé de 30 questions, réparties en 3 sous-domaines évaluant l'impact émotionnel, fonctionnel et physique de la dysphonie. Le patient remplit le questionnaire en cochant une case pour chaque item : J pour « jamais » (0 points), PJ pour « presque jamais » (1 point), P pour « parfois » (2 points), PT pour « presque toujours » (3 points) et T pour « toujours » (4 points) en fonction de la fréquence d'apparition de la situation décrite. Le résultat donne un score sur 120. Plus celui-ci est élevé, plus le handicap vocal est considéré comme sévère. Il en existe une version raccourcie, le VHI-10, qui ne comporte que 10 questions (Rosen et al., 2004), ainsi qu'une version destinée aux chanteurs, le SVHI (Cohen et al., 2007).

La passation du VHI présente divers avantages pour l'évaluation et la prise en charge des troubles vocaux. D'après les auteurs, le VHI permet à la fois d'évaluer la perception du patient concernant l'impact de son trouble vocal au quotidien, de mesurer l'efficacité de certaines méthodes de rééducation vocale et d'apprécier l'impact fonctionnel de certaines thérapies ou interventions médicales employées dans le cadre de troubles vocaux.

Le VHI a fait l'objet de multiples adaptations dans d'autres langues, dont le français (Woisard et al., 2004). Tout comme la version originale, l'outil francophone comporte 30 questions évaluant les mêmes 3 sous-domaines. Sa validité a été établie au niveau de la cohérence interne, la fidélité temporelle et la validité de contenu.

4.1.2. Intérêts et limites de l'autoévaluation par le patient

L'auto-évaluation par le patient est un élément essentiel de l'évaluation de la voix dysphonique : elle permet de recueillir la perspective et le vécu du patient, que ceux-ci soient en accord avec les observations du clinicien - ou non. Par conséquent, les résultats de ces évaluations sont essentiels pour la prise de décisions concernant le suivi du patient dysphonique : si la pathologie du patient n'a pas de conséquences sur sa santé et ni sur son vécu, une intervention chirurgicale et/ou orthophonique ne semble pas justifiée (Francis et al., 2017).

Cependant, l'auto-évaluation du patient dans le cadre d'un trouble vocal présente quelques limites, comme le soulignent Francis et al. (2017). Si certains questionnaires s'appuient sur une population choisie, d'autres, tels que le VHI, le VoiSS ou encore le V-RQOL, restent imprécis quant à la population à laquelle ils sont adressés ou auprès de laquelle ils ont été validés ; ils sont donc parfois proposés, à tort, à des individus pour qui ils n'ont pas été conçus. La validité de ces outils peut aussi parfois poser question : ceux-ci sont souvent construits sans participation de patients concernés par les mesures évaluées. Les questions, censées refléter le vécu des patients, sont donc élaborées par des personnes qui n'ont jamais fait l'expérience du trouble. Un autre écueil relevé provient de la difficulté à interpréter les scores, en particulier si certaines questions ont été laissées sans réponse. En effet, la signification des scores n'est pas toujours explicitée par les auteurs (Guyatt et al., 2002).

Enfin, il a été démontré à plusieurs reprises que les résultats d'auto-évaluations par le patient sont faiblement corrélées à ceux de mesures objectives, qu'elles soient acoustiques (Wheeler, Collins, & Sapienza, 2006) ou aérodynamiques (Cheng & Woo, 2010 ; Hsiung, Pai, & Wang, 2002 ; Woisard et al., 2007). La correspondance entre les résultats d'auto-évaluations et ceux de l'évaluation perceptive par le clinicien présente des résultats mitigés selon les études. En effet, le degré de corrélation rapporté varie, allant de modéré (Dejonckere et al., 1996 ; Mateos-Serrano et al., 2021) à plus faible (Maunsell et al., 2022). Toutefois, il n'en reste pas moins important d'envisager l'auto-évaluation, les mesures objectives et l'évaluation perceptive comme complémentaires dans leurs contributions à la prise en charge d'un trouble vocal, dans la mesure où elles permettent d'appréhender le trouble dans ses diverses dimensions.

4.2. Evaluation par le clinicien

L'évaluation perceptive est l'outil le plus fréquemment utilisé dans l'évaluation clinique de la voix (Behrman, 2005). Il existe de nombreux outils utilisés lors de l'évaluation perceptive, parmi lesquels on peut citer : le **Buffalo Voice Profile** (Wilson, 1987), le **Stockholm Voice Evaluation Approach** (Hammarberg, 2000), le **GRBAS** (Hirano, 1981) et le **Consensus Auditory Perceptual Evaluation of Voice** (CAPE-V) de Kempster et al. (2009). Ces deux derniers sont les outils les plus fréquemment utilisés et mieux acceptés (Barsties & De Bodt, 2015).

4.2.1. GRBAS

Le GRBAS (Hirano, 1981) est une échelle d'évaluation perceptive créée par Hirano en 1981. Elle est utilisée mondialement et dans diverses disciplines par des cliniciens et chercheurs pour évaluer la voix (Nemr et al., 2012). Elle permet l'évaluation de 5 paramètres perceptifs de la voix, à savoir :

- « G » ou « *grade* », qui permet d'apprécier la sévérité du trouble dans sa globalité ;
- « R » ou « *roughness* » : la raucité ;
- « A » ou « *asthenicity* » : l'asthénie, c'est-à-dire le manque de puissance de la voix ;
- « B » ou « *breathiness* » : le souffle vocal ;
- « S » ou « *strain* » : l'impression de forçage, d'une hyperfonction de la voix.

Chaque paramètre est côté selon quatre degrés de sévérité : 0 (normal), 1 (altération légère), 2 (altération modérée) et 3 (altération sévère).

L'évaluation perceptive avec le GRBAS peut être effectuée sur la base de tout stimulus vocal. Ce peut être un échantillon de parole spontanée, permettant ainsi une représentation fidèle du trouble vocal ; ou un /a/ tenu, voyelle laxe observable sans effet de coarticulation et qui permet de mettre en lumière d'éventuelles variations du timbre et de la stabilité de la voix (Giovanni & de Saint-Victor, 2013).

Depuis sa création, le GRBAS a été le sujet de diverses études, chacune apportant des modifications ou des critiques quant à sa validité et sa pertinence. En 1996, Dejonckere et al. proposent l'ajout d'un « I » pour décrire l'instabilité du trouble vocal. Les paramètres « G », « R » et « B » semblent être les plus pertinents, certaines études démontrant que l'asthénie et le forçage sont moins fiables (Dejonckere et al., 1993 ; Révis, 2021).

4.2.2. CAPE-V

Le CAPE-V, ou « Consensus Auditory-Perceptual Evaluation of Voice », est un outil standardisé d'évaluation perceptive de la voix (Annexe I). Il fut créé lors d'une conférence de consensus organisée par le « Special Interest Group 3 - Voice and Voice Disorders » de l'ASHA en juin 2002 à Pittsburgh, aux Etats-Unis. Cette conférence rassemblait des professionnels experts de la voix dans le but de pallier le manque d'outils standardisés, basés sur des données probantes et incorporant des recommandations de bonnes pratiques qui sont disponibles dans le cadre de l'évaluation perceptive de la voix (Kempster et al., 2009).

L'objectif principal du CAPE-V est de décrire la sévérité des aspects auditivo-perceptifs d'un trouble de la voix, de façon à ce que cette description puisse être communiquée à d'autres cliniciens. Son objectif secondaire est de contribuer aux hypothèses concernant les mécanismes anatomophysiologiques d'un trouble, dans le but d'évaluer la nécessité d'effectuer des examens complémentaires (Kempster et al., 2009). Il n'est cependant pas destiné à être utilisé comme seul outil d'évaluation pour déterminer l'origine d'un trouble vocal, ni à démontrer une relation de 1:1 avec les résultats d'autres examens vocaux (Kempster et al., 2009).

Dans la version américaine de Kempster et al. (2009) l'évaluation perceptive se fait sur la base d'un protocole standardisé composé de :

- 2 voyelles tenues : /a/ et /i/ de 3 à 5 sec
- 6 phrases, chacune induisant un comportement vocal spécifique :
  - « The blue spot is on the key again. » : production de toutes les voyelles de l'anglais ;
  - « How hard did he hit him? » : production d'une attaque douce avec le /h/ ;
  - « We were away a year ago. » : phrase produite en phonation continue ;
  - « We eat eggs every Easter. » : production d'attaques dures ;
  - « My mama makes lemon muffins. » : présence de sons nasaux ;
  - « Peter will keep at the peak. » : présence de phonèmes occlusifs non voisés.
- un échantillon de parole spontanée d'environ 20 secondes en réponse à une question telle que « Parlez-moi de votre problème de voix » ou « Dites-moi comment fonctionne votre voix »

Il est conseillé aux cliniciens d'enregistrer leur patient dans un environnement calme où il peut être assis.

Une fois que le clinicien a écouté les enregistrements des trois tâches, il évalue 6 caractéristiques de la voix : « overall severity » (sévérité globale), « roughness » (raucité), « breathiness » (souffle), « strain » (forçage), « pitch » (hauteur) et « loudness » (intensité). Pour chaque caractéristique, le clinicien place un curseur indiquant le degré d'atteinte vocale perçu sur une échelle visuelle analogique (EVA) allant de 0 à 100 mm. Il mesure ensuite la distance jusqu'au curseur en partant de la gauche et reporte sa mesure sur l'espace prévu à droite : il obtient donc une note sur 100. Si l'atteinte vocale pour une caractéristique donnée est uniforme au cours des 3 tâches, un seul curseur sera placé. Si elle varie selon la tâche, l'évaluateur place plusieurs curseurs et note un numéro sous chacun (par exemple : #1 sous le curseur correspondant à la tâche de voyelle tenue,

#2(a) pour celui de la première phrase…). Deux EVA vierges sont également mises à disposition pour que le clinicien puisse évaluer d'autres caractéristiques vocales si besoin.

Sous l'EVA sont notées les mentions :

- « MI » : « mildly deviant », indiquant une atteinte légère ;
- « MO » : « moderately deviant », indiquant une atteinte modérée ;
- et « SE » : « severely deviant » qui correspond à une atteinte sévère.

Elles servent de référence au clinicien quant au placement du curseur sur l'EVA.

A droite de chaque EVA on trouve la mention « C/I ». L'évaluateur entoure le terme qui décrit le mieux le paramètre évalué :

- « C » : « consistent », c'est-à-dire constant. La caractéristique vocale cotée est présente à travers toutes les tâches.
- « I » : « intermittent ». La caractéristique vocale est présente uniquement pendant certaines tâches ou à certains moments d'une tâche.

Il est également possible de faire des commentaires sur la résonance : normale / autre (préciser) ainsi que sur d'autres caractéristiques comme la diplophonie, le fry, la voix de fausset, l'asthénie… Cette catégorie est ouverte et laisse au clinicien la liberté d'ajouter les informations qu'il juge pertinentes.

4.2.2.1. Validité psychométrique du CAPE-V

La validité et la fiabilité du CAPE-V ont été établies pour sa version américaine par Karnell et al. en 2007. Les auteurs ont obtenu une bonne correspondance entre les résultats d'évaluations des mêmes voix avec le CAPE-V et le GRBAS, ainsi qu'une fiabilité inter-juge et intra-juge correcte pour le CAPE-V. Cependant, les voix étaient évaluées avec le CAPE-V et le GRBAS au même moment, induisant une potentielle influence d'un outil sur l'autre. Les résultats du GRBAS et du CAPE-V ont également été comparés à ceux de deux auto-évaluations par les patients, le V-RQOL et le VHI : la correspondance entre les deux types de mesures s'est révélée faible. Ces résultats étaient attendus puisque ces deux types d'évaluations mesurent des versants différents de la dysphonie. En effet, les évaluations perceptives reflètent la perception de la voix dysphonique par un clinicien, alors que les auto-évaluations quantifient la perception du patient vis-à-vis des conséquences sociales et émotionnelles du trouble vocal et de leur potentiel impact sur la vie quotidienne.

Zraick et al. (2011) ont répété l'expérience, recrutant cette fois-ci davantage de cliniciens pour évaluer les voix, afin d'obtenir un échantillon de professionnels plus varié. Pour éliminer le biais méthodologique de l'étude de Karnell et al. (2007) entraîné par l'évaluation simultanée des voix avec le GRBAS et le CAPE-V, les échantillons vocaux étaient évalués avec un outil sur une session, puis l'autre sur une deuxième session une semaine plus tard. Les auteurs ont déterminé que les fiabilités inter-juge et intra-juge du CAPE-V étaient plus élevées que pour le GRBAS. De même, une étude de Nemr et al. (2012) rapporte un degré de corrélation élevé entre les évaluations d'un même échantillon vocal avec le CAPE-V et le GRBAS, ainsi qu'une meilleure fiabilité inter-juges pour le CAPE-V que pour le GRBAS.

4.2.3. Influence de la langue parlée dans l'évaluation perceptive de la voix

La perception et la description de la qualité vocale dépendent fortement d'aspects socio-culturels (Dejonckere et al., 2001 ; Yiu et al., 2008), dont la langue. En effet, diverses études démontrent l'impact de la langue maternelle de l'évaluateur dans la perception de la qualité de voix. Une étude de Ghio et al. (2015) révèle des différences significatives dans l'évaluation de la raucité par des locuteurs français et italiens : il semblerait que ces derniers tendent à considérer comme normales des voix perçues comme rauques (« R » dans le GRBAS) par les évaluateurs français. Dans une étude semblable comparant les évaluations perceptives effectuées par des locuteurs japonais et américains, les jugements des critères « A » et « S » (*asthenia* et *strain*) du GRBAS n'étaient pas cohérents entre les deux groupes.

Un changement dans la langue employée par le locuteur peut être source de variations, tel qu'il en est le cas pour la f_0 (Engelbert, 2014 ; Järvinen et al., 2013). Les accents régionaux semblent également avoir un impact sur la perception de la qualité de voix. Ghio et al. (2013) démontrent dans leur étude que l'accent régional n'impacte pas le jugement perceptif de voix dysphoniques de grade modéré à sévère. Cependant, l'accent régional a un impact important sur la perception de voix légèrement dysphoniques à normales : celles-ci sont considérées comme anormales par des évaluateurs non familiers avec l'accent.

4.2.3.1. Adaptation francophone du CAPE-V

Au vu de l'influence de la langue parlée sur l'évaluation perceptive de la voix, le CAPE-V a été adapté dans de nombreuses langues. En 2023, Déborah Mbagira a élaboré une adaptation en langue française du CAPE-V (CAPE-V$_f$) dans le cadre de son mémoire de fin d'études à l'Université de Liège en Belgique (Pommée et al., 2024). L'outil (Annexe II) a été construit à partir d'avis d'experts français et belges recueillis par un processus Delphi en trois tours. Le premier tour consistait en un recueil des données démographiques des participants et de leurs connaissances dans le domaine, ainsi qu'un premier sondage de leur avis sur les phrases en français proposées. Ces phrases avaient été élaborées avec l'aide d'un linguiste francophone, sur base des critères de construction des phrases originales. Suite à ce premier tour, des sélections et ajustements de tâches furent effectués selon les recommandations des experts. Un deuxième tour fut donc lancé, lors duquel le consensus fut atteint pour toutes les nouvelles propositions. La version finale de l'adaptation en langue française du CAPE-V, validée au complet lors du troisième tour du processus Delphi, est donc composée de trois tâches.

La première tâche est une voyelle tenue. La voyelle laxe /a/ a été choisie car elle est relativement neutre sur le plan articulatoire et produite avec un tractus vocal ouvert. Elle permet d'obtenir un aperçu de la phonation sans influence de la coarticulation opérant en parole continue.

La deuxième tâche est la lecture de phrases. Six phrases ont été élaborées, chacune induisant un comportement vocal spécifique.

- « Dites-nous les amis, où allez-vous samedi ? », qui permet d'analyser la production des extrêmes du triangle vocalique (/i/, /u/, /a/).

- « Son chat Fifi saute sur la souris. », permettant l'analyse des transitions entre les phonèmes voisés et les phonèmes non voisés.
- « Elle va aux Iles de la Madeleine » phrase entièrement voisée lors de laquelle on peut observer la présence/absence d'arrêts ou de spasmes et la capacité à maintenir le voisement.
- « Irma, où va Alain aujourd'hui ? A Ottawa. », qui permet d'évaluer les éventuelles attaques glottiques dures.
- « Lundi matin, maman nous a montré les nénuphars. », phrase comportant une grande proportion de phonèmes nasalisés permettant d'observer une éventuelle hyponasalité.
- « Dobi comptait payer les deux pastèques. » qui permet d'évaluer les éventuelles fuites nasales et l'impact des phonèmes occlusifs non voisés sur la phonation.

La dernière tâche consiste en la production de discours spontané en réponse à la phrase « Présentez-vous brièvement (nom, prénom, région, activité principale). » Cette phrase a été modifiée par rapport au protocole original. En effet, les questions « Parlez-moi de votre problème de voix » ou « Dites-moi comment fonctionne votre voix » incluses dans la version américaine ont été écartées au profit d'une question plus neutre, évitant ainsi une influence de la charge émotionnelle sur la production vocale.

Une modification importante a été apportée au système de cotation des paramètres dans le CAPE-Vf. Alors que la version américaine et les adaptations dans d'autres langues incluent les inscriptions « MI » (« mildly deviant » qui signifie « légère déviation »), « MO » (« moderately deviant », c'est-à-dire « déviation modérée ») et SE (« severely deviant » ou « déviation sévère ») sous l'EVA (Figure 5), les créateurs de la version francophone ont opté pour le retrait ces marqueurs avec consensus des experts durant l'étude Delphi (Figure 6), dans l'objectif d'éviter un effet de regroupement de points sous les marqueurs verbaux de sévérité (Pommée et al., 2024).

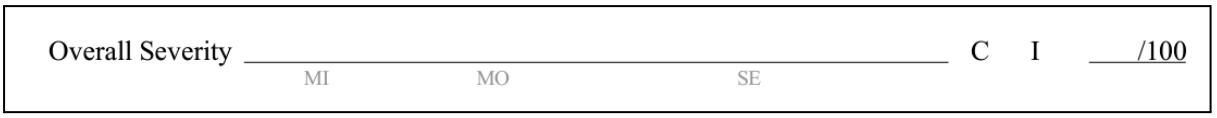


Figure 5. EVA avec étiquetages verbaux telle que dans la version américaine de Kempster et al. (2009)

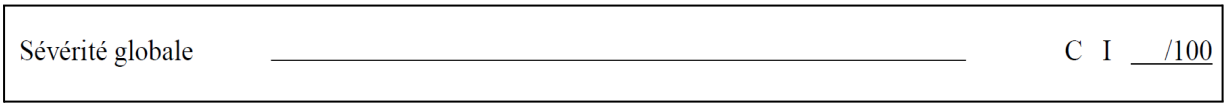


Figure 6. EVA sans étiquetage telle que dans la version francophone de Pommée et al. (2024)

4.2.4. Intérêts et limites de l'évaluation perceptive

D'après Oates (2009) l'évaluation perceptive présente de nombreux avantages. En effet, les termes employés sont intuitifs et compréhensibles par la population générale, permettant un bon transfert d'informations du clinicien au patient. De plus, les outils les plus fréquemment utilisés se ressemblent dans les aspects de la voix qu'ils évaluent, ainsi que par leurs définitions de ces aspects. Enfin, l'évaluation perceptive présente des avantages en termes de coût, temps requis et confort du patient et ne nécessite pas de connaissances techniques pointues pour être mise en pratique.

Cependant, l'évaluation perceptive fait l'objet de diverses critiques, notamment en raison de sa subjectivité. En effet, pour pouvoir être utilisée comme outil en clinique, elle doit répondre à des exigences en termes de fiabilité et de validité (Oates, 2009). Chaque évaluateur se réfère à ses propres références internes de voix normales ; cependant, ces références peuvent varier inter-individuellement et même intra-individuellement (Kreiman et al., 1993). Certaines sources d'erreurs entraînant une diminution de la fiabilité des résultats d'évaluations perceptives ont été identifiées, parmi lesquelles on peut citer : le type de tâche (appariement vs. estimation de gravité du trouble, etc.), le type d'échelle de cotation (EVA, ordinale…), le type d'échantillon vocal évalué, la fatigue de l'évaluateur, la sensibilité perceptive de l'évaluateur ou les caractéristiques vocales qu'il doit évaluer (Oates, 2009).

Plusieurs solutions sont envisagées par Oates (2009) pour limiter l'impact des facteurs influençant la perception du clinicien et ainsi améliorer la fiabilité et la validité de leurs résultats ; nous allons illustrer ces points à l'aide d'une comparaison des outils GRBAS et CAPE-V.

Tout d'abord, il est recommandé d'utiliser un outil construit sur des bases théoriques solides et incluant des caractéristiques vocales pertinentes et bien définies, assurant ainsi une bonne validité de contenu. Les attributs de la voix inclus dans le CAPE-V ont été choisis parmi ceux apparaissant le plus fréquemment dans la littérature internationale ; ils ont ensuite été validés par un groupe d'experts lors de la conférence ASHA. Six caractéristiques vocales sont proposées dans le protocole du CAPE-V, et deux espaces sont prévus pour que le clinicien en décrive d'autres s'il en ressent le besoin. L'évaluation de la voix avec le GRBAS se limite elle à l'appréciation de cinq attributs vocaux, sans proposer au clinicien la possibilité d'évaluer une autre caractéristique vocale saillante.

Comme cité précédemment, le type d'échelle de mesure qu'inclut l'outil peut également influencer la fiabilité des résultats. Le GRBAS propose une échelle ordinale de notation allant de 0 à 3, à la différence du CAPE-V qui emploie une échelle visuelle analogique de 100 mm. Ce choix n'est pas sans conséquences sur la fiabilité des mesures : alors qu'un écart d'un point entre deux évaluations sur une échelle de 4 points comme celle du GRBAS représente une différence de 25%, une différence d'un point sur l'échelle visuelle analogique du CAPE-V n'équivaut qu'à un écart de 1% (Karnell, 2007). L'EVA permet ainsi une évaluation plus sensible des caractéristiques vocales (Grant et al., 1999 ; Hasson & Arnetz, 2005). De plus, l'EVA, en tant qu'échelle continue, permet d'effectuer des statistiques paramétriques alors que les résultats d'échelles ordinales comme le GRBAS doivent être analysés avec des tests non-paramétriques (Kim, 2017).

La présence d'un protocole détaillant les procédures de passation et d'enregistrement, ainsi que les tâches vocales à effectuer est également un atout puisqu'il permet de limiter les biais liés à la méthodologie. Le CAPE-V propose à l'évaluateur des instructions exhaustives, guidant ainsi le clinicien à travers les trois tâches vocales élaborées à partir d'une réflexion commune approfondie de la part des auteurs. Ce n'est pas le cas pour le GRBAS qui n'est accompagné d'aucune consigne et peut s'effectuer sur la base de n'importe quelle production vocale.

Enfin, l'évaluation perceptive de la voix se fait idéalement avec un outil dont la fiabilité et la validité ont été établies. Or, bien que le GRBAS soit l'outil dont l'utilisation est la plus répandue, sa fiabilité est contestée (Ghio et al., 2021). Alors que cela semble essentiel, il n'existe donc à ce jour

aucun outil standardisé d'évaluation perceptive de la voix dont la validité et la fiabilité ont été attestées en langue française.

PROBLEMATIQUE ET HYPOTHESES

I. Problématique

L'évaluation perceptive fait partie intégrante de l'évaluation vocale multidimensionnelle dans le cadre des dysphonies. Malgré son importance, on relève un manque d'outils standardisés, valides et fiables à la disposition des cliniciens français dans ce domaine. Or, les cliniciens de la voix, dont font partie les orthophonistes, ont besoin d'outils sensibles, spécifiques et fiables pour évaluer et suivre l'évolution des troubles vocaux de leurs patients.

Au vu de la supériorité psychométrique du CAPE-V par rapport au GRBAS établie par de précédentes études, nous proposons par notre travail de recherche de répondre à la question : ***L'adaptation française du CAPE-V est-elle un outil fiable et valide pour l'évaluation perceptive de la voix ?***

II. Hypothèses

Première hypothèse : L'adaptation française du CAPE-V est un outil d'évaluation perceptive de la voix montrant une bonne fiabilité intra-juge (ICC ⩾ 0.75) ainsi qu'une bonne fiabilité inter-juge (ICC ⩾ 0.75) pour les différents paramètres.

> *Hypothèse secondaire :* Les fiabilités intra-juge et inter-juge du CAPE-V sont supérieures à celles du GRBAS.

Deuxième hypothèse : L'adaptation française du CAPE-V est un outil d'évaluation perceptive de la voix montrant une bonne validité de critère, mesurée par une forte corrélation avec les différents paramètres (r ⩾ 0.70) de l'échelle GRBAS.

Troisième hypothèse : La corrélation du score de sévérité globale du CAPE-V avec le score total du VHI est plus faible qu'avec le G du GRBAS (r < 0.70), étant donné que ces deux outils évaluent des construits différents du phénomène vocal multidimensionnel.

METHODOLOGIE DE L'ETUDE

Objectifs

L'objectif de notre étude est d'établir la validité de l'adaptation française du CAPE-V (CAPE-Vf). Par souci de clarté, nous présenterons l'étude en deux parties (Figure 7), bien que certaines des étapes de celles-ci aient été menées simultanément. La première phase de l'étude consiste en un recueil d'enregistrements de voix auprès de personnes dysphoniques et normophoniques, suivi d'un tri de ces données en quatre listes. La deuxième phase correspond à la phase d'évaluation des voix, lors de laquelle nous avons envoyé les enregistrements à treize orthophonistes afin qu'ils les évaluent avec le CAPE-Vf ainsi que le GRBAS.

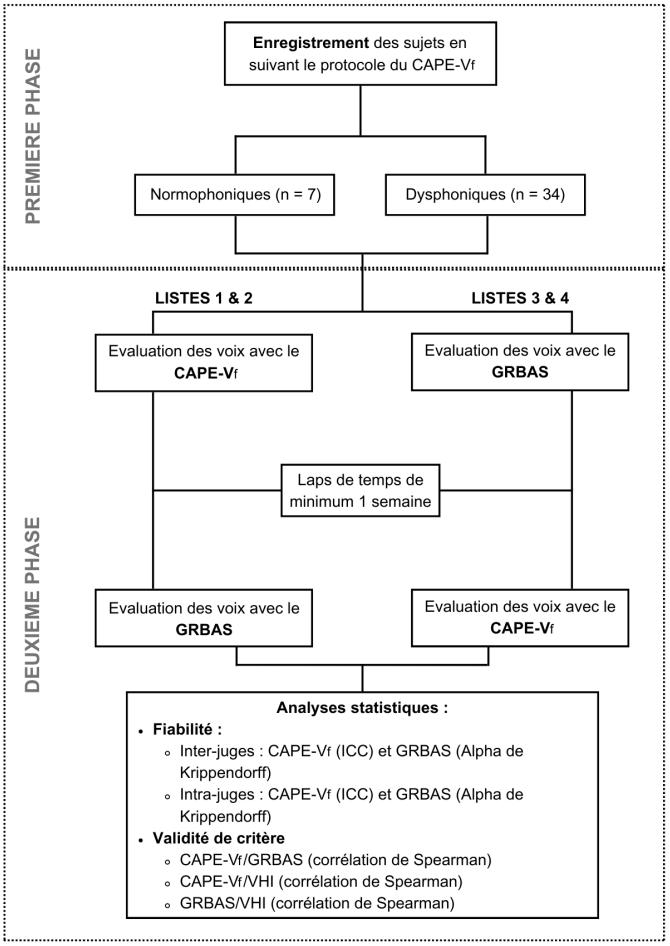


Figure 7. Organisation générale du déroulé de l'étude

PARTIE 1 : ENREGISTREMENT DES LOCUTEURS

1. Populations

Nous avons recruté pour notre étude des sujets dysphoniques ainsi que des sujets normophoniques.

Le recrutement des sujets **dysphoniques** a eu lieu au sein de la patientèle de quatre orthophonistes pratiquant des rééducations vocales en cabinet libéral, dans les régions Occitanie et Bourgogne-Franche-Comté. Nos critères d'inclusion étaient les suivants : patients adultes,

dysphoniques diagnostiqués, francophones natifs. Nous souhaitions obtenir un échantillon de voix représentatif du large champ de pathologies pouvant entraîner une dysphonie. Ainsi, nous n'avons exclu aucune pathologie. Des sujets **normophoniques** ont été recrutés dans notre entourage. La seule condition pour participer était l'absence d'antécédents de troubles vocaux.

1.1. Taille, répartition et données démographiques de l'échantillon

Trente-quatre personnes dysphoniques et sept personnes normophoniques (Tableau 1) ont été enregistrées dans le cadre de notre mémoire. Quinze pathologies différentes ont été recensées.

Tableau 1. Données démographiques des locuteurs et pathologies représentées au sein de l'échantillon

		LOCUTEURS DYSPHONIQUES		LOCUTEURS NORMOPHONIQUES	
Données démographiques	**Sexe**	Homme (n = 13)	Femme (n = 21)	Homme (n = 4)	Femme (n = 3)
	Âge moyen ***Ecart-type***	54.46 *18.68*	56.76 *13.09*	37.25 *15.80*	46.33 *20.31*
	Plage d'âge	23 - 83	30 - 90	25 - 57	23 - 60
Pathologies représentées	**Dysphonie dysfonctionnelle**	3	5	*N/A*	*N/A*
	Paralysie récurrentielle	2	3	*N/A*	*N/A*
	Dysodie	1	3	*N/A*	*N/A*
	Nodules	0	2	*N/A*	*N/A*
	Presbyphonie	1	1	*N/A*	*N/A*
	Vergeture	1	1	*N/A*	*N/A*
	Kyste post-opératoire	0	2	*N/A*	*N/A*
	Polype post-opératoire	1	0	*N/A*	*N/A*
	Cancer de l'oropharynx	1	0	*N/A*	*N/A*
	Chimiothérapie	0	1	*N/A*	*N/A*
	Papillomatose laryngée	0	1	*N/A*	*N/A*
	Dyskinésie laryngée	1	0	*N/A*	*N/A*
	Oedème	1	0	*N/A*	*N/A*
	Mue faussée	1	0	*N/A*	*N/A*
	AVC	0	1	*N/A*	*N/A*
	Pathologie indéterminée	1	0	*N/A*	*N/A*

2. Matériel et méthodes

2.1. Documents d'information et de consentement

Afin d'assurer le respect de la réglementation relative à la protection des données à caractère personnel, tous les participants ont reçu une **note d'information** ainsi que des **formulaires de consentement au traitement des données**. La note d'information présente les objectifs et le déroulé de l'étude, les bénéfices ainsi que les risques associés à la participation et les renseignements sur le traitement des données et les droits du participant sur celles-ci. Le formulaire de consentement au traitement de données à caractère personnel et le formulaire d'autorisation de réutilisation et de traitement de données à caractère personnel permettent de recueillir le consentement écrit du sujet vis-à-vis de sa participation à l'étude et de l'utilisation de ses données à des fins de recherche.

2.2. Protocole d'enregistrement

Afin d'obtenir une procédure d'enregistrement standardisée, nous avons créé un protocole d'enregistrement à destination des cliniciens (Annexe VI). Celui-ci contenait toutes les informations nécessaires pour le recueil des données auprès des participants de l'étude, à savoir :

- le matériel et les conditions d'enregistrement requis,
- la procédure d'information et de recueil de consentement à mener auprès des patients,
- la méthode de pseudonymisation et de transfert des données,
- le détail des tâches vocales que doivent produire les participants.

Le protocole devait être lu par le clinicien dans sa totalité avant le début du recueil de données.

2.3. Distribution du matériel

Les cliniciens ont reçu par voie postale un dossier contenant :

- le protocole d'enregistrement,
- un formulaire de la version française du VHI (Woisard et al., 2004) pour chaque patient (Annexe VII),
- une autorisation de réutilisation et de traitement de données à caractère personnel (2 exemplaires par patient), un formulaire de consentement (2 exemplaires par patient) et une notice d'information pour chaque patient,
- des cartes flash avec les phrases du protocole CAPE-Vf.

Afin que les cliniciens puissent nous transmettre les enregistrements de façon sécurisée, nous leur avons créé un lien d'accès à un espace de dépôt personnel sur FileSender Renater.

2.4. Déroulement des passations

2.4.1. Information et consentement

Avant tout, la notice d'information et les deux documents de consentement ont été présentés aux potentiels participants dysphoniques et normophoniques. Tous les sujets souhaitant participer à l'étude ont signé les documents. Les participants ont bénéficié du délai de réflexion de leur choix avant la signature.

2.4.2. Environnement et outils

Les participants ont été enregistrés dans un environnement calme et sans distraction. Un micro-casque ou micro sur pied était placé à 45 degrés et à une distance de 6 à 10 cm de la bouche. Les enregistrements audio ont été effectués en mode mono, avec une résolution d'au moins 16 bits (idéalement 32 bits) et un taux d'échantillonnage du signal d'au moins 22 kHz (idéalement 44100 Hz). Pour éviter un effet de fatigue ou d'échauffement vocal chez les sujets dysphoniques, l'enregistrement était effectué en début de séance d'orthophonie.

2.4.3. Enregistrement

Afin que les cliniciens se familiarisent avec le processus d'enregistrement ainsi que pour nous permettre de vérifier la qualité des enregistrements, ils devaient effectuer un enregistrement d'essai avec leur propre voix. Celui-ci devait être transféré via FileSender avant de procéder à l'enregistrement des participants afin que nous puissions vérifier sa qualité ainsi que le respect des consignes d'enregistrement. Avant de commencer l'enregistrement, les locuteurs ont rempli un questionnaire VHI.

La première tâche à enregistrer était le /a/ tenu. Avant de commencer, le clinicien dit au participant : « La première tâche consiste à dire le son /a/. Tenez ce /a/ aussi stable que possible, avec votre voix habituelle, jusqu'à ce que je vous demande d'arrêter. Nous allons le faire trois fois. » Le participant effectue donc la tâche 3 fois, pendant 3-5 secondes à chaque fois.

La deuxième tâche du protocole est la lecture de phrases. Au nombre de six, celles-ci étaient présentées sous forme de « carte flash » montrant ainsi une seule phrase cible à la fois. Le participant était chargé, dans la mesure du possible, de les faire défiler. Le clinicien devait veiller à ce que le participant ait terminé de lire la phrase en entier avant qu'il ne regarde la prochaine carte afin que le bruit des cartes qui défilent ne perturbe pas l'enregistrement de la voix. La consigne était la suivante : « Veuillez lire les phrases suivantes une à la fois, comme si vous parliez à quelqu'un dans une vraie conversation. » (L'individu exécute la tâche, lisant chaque phrase, l'une après l'autre.) Si le participant avait des difficultés à lire, le clinicien pouvait lui demander de répéter les phrases après lui, en notant cette particularité.

- Phrase 1 : « Dites-nous les amis, où allez-vous samedi ? »
- Phrase 2 : « Son chat Fifi saute sur la souris. »
- Phrase 3 : « Elle va aux Îles de la Madeleine. »
- Phrase 4 : « Irma, où va Alain aujourd'hui ? À Ottawa. »
- Phrase 5 : « Lundi matin, maman nous a montré les nénuphars. »
- Phrase 6 : « Dobi comptait payer les deux pastèques. »

La dernière tâche à effectuer est la tâche de discours semi-spontané. Le clinicien devait susciter au moins 20 secondes de discours conversationnel en réponse à la question : « Présentez-vous brièvement en tant que Jean/Jeanne Dupont, en indiquant votre région d'origine et votre activité principale. » Le participant était libre d'inventer des éléments s'il préférait ne pas partager d'informations personnelles.

En cas d'erreur nécessitant un ré-enregistrement, les cliniciens étaient autorisés à réitérer l'expérience un maximum de 3 fois.

2.5. Recueil et traitement des données

Les enregistrements, documents de consentement, VHI et données démographiques recueillis auprès des participants nous ont été envoyés par transfert sécurisé via FileSender Renater.

Les enregistrements ont ensuite été écoutés et triés sur l'ordinateur sur Microsoft Excel. Ont été retenus seuls ceux qui satisfaisaient nos exigences en termes de son (qualité sonore de l'enregistrement, absence de bruits perturbateurs en arrière plan…) et de respect du protocole du CAPE-V$_f$. De légères modifications ont été apportées à certains des enregistrements telles que l'augmentation du volume sonore.

Une fois les enregistrements sélectionnés, nous les avons classés en groupes selon la sévérité de l'altération perçue (Tableau 2) :

- NA : « non altérée »
- AL : « altération légère »
- AM : « altération modérée »
- AS : « altération sévère »

Les degrés de sévérité ont été attribués par consensus de nos trois avis.

Tableau 2. Nombre de voix pour chaque grade de sévérité perçue, en fonction du sexe

FEMMES				HOMMES				
NA	AL	AM	AS	NA	AL	AM	AS	*TOTAL*
n = 6	n = 6	n = 6	n = 6	n = 6	n = 3	n = 5	n = 3	n = 41

Enfin, nous avons réparti les voix en quatre listes d'écoute de manière à ce que la répartition des degrés de sévérité et des sexes soit la plus équitable (Tableau 3).

Tableau 3. Répartition des voix dans les quatre listes d'écoute

Locuteurs	Sévérité	n	LISTE 1	LISTE 2	LISTE 3	LISTE 4
Hommes	NA	6	2 + 1 rep	1	1	2 + 1 rep
	AL	3	1 + 1 rep	2 + 1 rep	1 + 1 rep	2 + 1 rep
	AM	5	2 + 1 rep	2 + 1 rep	1 + 1 rep	1 + 1 rep
	AS	3	2 + 1 rep	1 + 1 rep	2 + 1 rep	1 + 1 rep
Femmes	NA	6	1	2 + 1 rep	2 + 1 rep	1
	AL	6	2 + 1 rep	1 + 1 rep	2 + 1 rep	1 + 1 rep
	AM	6	1 + 1 rep	1 + 1 rep	2 + 1 rep	2 + 1 rep
	AS	6	1 + 1 rep	2 + 1 rep	1 + 1 rep	2 + 1 rep
TOTAL		41	12 + 7 rep n = 19	12 + 7 rep n = 19	12 + 7 rep n = 19	12 + 7 rep n = 19

La mention « rep » signifie qu'un des enregistrements de voix est répété au sein de la liste. Ainsi, pour chaque degré de sévérité, un des enregistrements apparaît deux fois. Ceci a permis de recueillir plusieurs évaluations pour un même enregistrement de voix, donnant ainsi la possibilité d'évaluer la fiabilité intra-juge.

Par souci de faisabilité nous avons volontairement limité le nombre d'enregistrements par liste à dix-neuf ; un nombre plus élevé d'enregistrements aurait augmenté significativement le temps d'évaluation pour les juges ainsi que potentiellement entraîné une fatigue et une diminution de l'attention chez ceux-ci.

Lorsque nous avons constitué les listes, nous avons renommé les enregistrements « enregistrement_1 » à « enregistrement_19 ». Nous les avons ordonnés de manière à ce que les enregistrements répétés ne soient pas trop proches, limitant ainsi le risque que les évaluateurs se réfèrent à leur précédente évaluation de la même voix.

PARTIE 2 : EVALUATION DES VOIX

1. Population

1.1. Recrutement

1.1.1. Pré-recrutement

Pour ce qui est de la deuxième partie de notre étude, nous avons commencé par étudier la faisabilité de notre projet. Pour ce faire, nous avons sollicité par mail environ 950 orthophonistes identifiés sur le registre de professionnels formés Ostéovox (Piron, 1998). Nous avons également contacté quelques ORL et phoniatres par le biais de leur adresse mail professionnelle.

Un court questionnaire détaillant les objectifs de notre étude ainsi que le rôle qu'ils auraient s'ils participaient leur a été transmis. Le questionnaire proposait aux cliniciens de renseigner leur contact, leur profession (sélection parmi trois options : orthophoniste, ORL et phoniatre), leur nombre d'années d'expérience dans le domaine de la voix et leur région d'exercice. En remplissant ce questionnaire, les cliniciens acceptaient d'être recontactés au sujet de notre étude une fois certaines modalités déterminées suite à notre essai pilote (cf. partie 2.2) telles que la durée de passation et les modalités de celle-ci. Cinquante-cinq professionnels, dont 54 orthophonistes et 1 phoniatre, ont montré un intérêt pour la participation à notre étude en répondant positivement à notre questionnaire.

1.1.2. Echantillon final : taille et données démographiques des participants

Nos critères d'inclusion pour l'échantillon final des juges étaient les suivants : orthophoniste, phoniatre ou ORL avec au moins un an d'expérience dans le domaine de la voix.

Suite à notre relance, vingt-et-une orthophonistes ont confirmé leur participation à l'étude. Elles ont été réparties dans quatre listes d'écoute (cf. Tableau 3). Les listes 1, 3 et 4 étaient composées de cinq orthophonistes et la liste 2 comprenait six orthophonistes. Cependant, seules treize orthophonistes nous ont envoyé les résultats de leurs évaluations, modifiant ainsi la composition des listes d'écoutes pré-établies (Tableau 4).

Tableau 4. Composition finale des listes d'écoute

LISTE 1	LISTE 2	LISTE 3	LISTE 4
Juges 1 et 2	Juges 3, 4 et 5	Juges 6, 7 et 8	Juges 9, 10, 11, 12 et 13

Ainsi, notre échantillon final de juges était composé de sept orthophonistes ayant entre 1 et 10 ans d'expérience dans le domaine de la voix, deux ayant entre 11 et 20 ans d'expérience et trois ayant entre 21 et 30 ans d'expérience. Une orthophoniste avait plus de 30 ans d'expérience. Les orthophonistes étaient issues de cinq régions françaises (Figure 8).

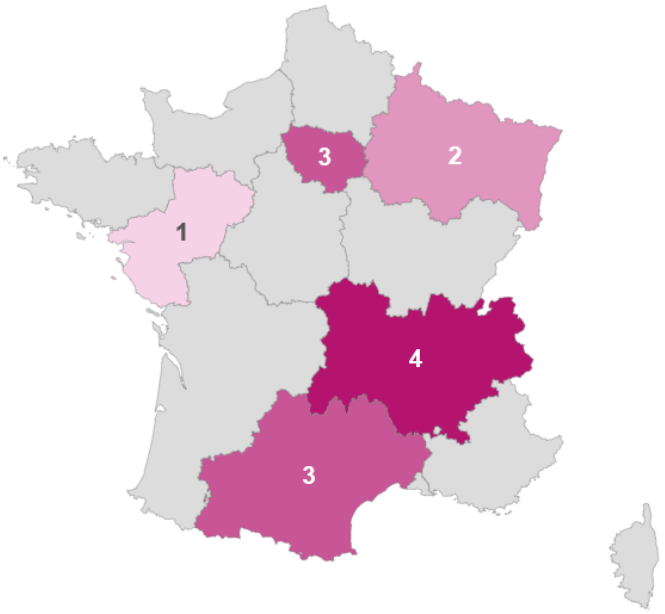


Figure 8. Régions d'exercice des orthophonistes ayant participé à l'étude

2. Matériel et méthodes

2.1. Protocole d'évaluation

Afin que les participants effectuent les évaluations dans des conditions les plus similaires possibles, nous avons élaboré un protocole intitulé « Procédure d'évaluation perceptive de la voix avec l'outil CAPE-V » (Annexe VIII). Deux versions du protocole ont été créées :

- une dans laquelle il est indiqué que l'évaluation des voix doit être effectuée avec le CAPE-Vf en premier, puis le GRBAS ; cette version du protocole a été envoyée aux évaluateurs des listes 1 et 2 ;
- et une où l'ordre des évaluations est inversé (GRBAS puis CAPE-Vf) ; cette version est celle qu'ont reçue les évaluateurs des listes 3 et 4.

Ceci permettait de contrebalancer les éventuels effets de l'utilisation d'un outil sur l'autre, minimisant ainsi les biais potentiels liés à l'ordre dans lequel les participants étaient exposés aux outils.

Le protocole détaillait tout d'abord le matériel envoyé aux évaluateurs, à savoir :

- 19 copies du CAPE-Vf vierges mis à part la mention « Nom » où sont déjà renseignés les noms des enregistrements à évaluer : « enregistrement_1 » à « enregistrement_19 » ;

- un tableau pour renseigner les résultats du GRBAS (Annexe IX) ;
- un lien d'accès FileSender Renater pour télécharger les 19 voix à évaluer.

Il était précisé aux participants qu'ils bénéficiaient d'une période de cinq semaines à dater de la réception du protocole pour effectuer les deux sessions d'évaluations. Un schéma illustrant la temporalité de l'étude était accompagné d'une mention précisant que les deux sessions étaient à espacer d'un laps de temps d'au moins une semaine. Ce délai permettait de limiter le risque d'influence d'un outil sur l'autre.

Enfin, la procédure à mener lors des deux sessions d'évaluation était détaillée. Le protocole s'achevait avec les instructions à suivre une fois que les deux sessions d'évaluations étaient terminées : les évaluateurs étaient invités à nous transmettre les CAPE-Vf et le tableau GRBAS complétés et priés d'effacer les enregistrements de tout support informatique en leur possession.

2.1.1. Ajustement du CAPE-Vf

Pour des raisons de faisabilité, nous avons opté pour une modalité exclusivement digitale. Le CAPE-Vf a donc été envoyé en format pdf aux évaluateurs et ceux-ci devaient le compléter sur leur ordinateur. Nous avons effectué quelques modifications pour adapter le CAPE-Vf à cette modalité : nous avons enlevé la mention « Mesurer la distance (en millimètres) entre l'extrémité gauche de l'échelle et le point de repère pour obtenir un score sur 100. » ainsi que l'espace prévu pour noter le score. En effet, il aurait été compliqué pour les évaluateurs de modifier la taille du pdf CAPE-Vf sur leur écran d'ordinateur de manière à ce que l'échelle visuelle analogique (EVA) fasse exactement 100mm. Ainsi, nous avons effectué les mesures de l'EVA nous-même après réception des évaluations par les juges.

Consigne de notation : Placer un point de repère sur la droite de 100mm pour déterminer le niveau d'atteinte.
- Extrémité gauche = absence d'altération
- Extrémité droite = altération maximale

Mesurer la distance (en millimètres) entre l'extrémité gauche de l'échelle et le point de repère pour obtenir un score sur 100.

Sévérité globale _____ C I /100

Raucité _____ C I /100

Consignes dans la version du CAPE-V de D. Mbagira

Consigne de notation : Placer un point de repère sur la droite de 100mm pour déterminer le niveau d'atteinte.
- Extrémité gauche = absence d'altération
- Extrémité droite = altération maximale

Sévérité globale _____ C I

Raucité _____ C I

Consignes telles que présentées dans notre protocole

2.2. Essai pilote

Avant le lancement de l'étude, notre protocole a été testé lors d'un essai pilote effectué par quatre étudiantes en orthophonie issues du centre de formation de Toulouse. L'objectif de cet essai était de recueillir divers avis quant à la qualité et la clarté des instructions fournies par le protocole ainsi que de déterminer un temps de passation moyen pour chaque outil (CAPE-Vf et GRBAS). Le protocole avait été légèrement modifié afin de répondre aux objectifs de l'essai pilote ; certaines parties, notamment sur la temporalité de l'étude et le transfert des résultats des évaluations étaient donc grisées afin que les participants puissent tout de même les lire et éventuellement faire des commentaires.

Ainsi, les participants devaient noter :

- le temps d'évaluation des échantillons vocaux avec le CAPE-Vf (à compter du début de l'écoute du premier échantillon, jusqu'à la fin de l'évaluation du dernier) ;
- le temps d'évaluation des échantillons vocaux avec le GRBAS (idem) ;
- tous commentaires relatifs au protocole et à la procédure d'évaluation, y compris les parties grisées.

Deux des étudiantes ont effectué les évaluations des 19 voix avec le CAPE-Vf et deux autres étudiantes ont évalué les voix avec le GRBAS.

Ainsi, les temps de passation ont été estimés à 45 min pour le CAPE-Vf et 25 min pour le GRBAS. Nous avons également effectué quelques modifications de notre protocole suite aux retours des testeurs, à savoir :

- mentionné la possibilité que certaines des voix pouvaient être répétées au sein de la liste d'écoute, étant donné que certaines étudiantes avaient exprimé leur surprise à cet égard ;
- transformé la phrase « Il est important que vous écoutiez toutes les tâches vocales avant de remplir le formulaire CAPE-V. » en « Il est important que vous écoutiez l'enregistrement sélectionné en entier avant de remplir le formulaire CAPE-V correspondant. » En effet, cette première formulation portait à confusion, certaines des participantes ayant pensé qu'il s'agissait d'écouter tous les enregistrements avant de débuter les évaluations.

2.3. Lancement de l'étude

Nous avons transmis aux évaluateurs notre protocole final ainsi que les enregistrements vocaux anonymisés par le biais d'un lien FileSender Renater sécurisé. Une fois les évaluations terminées, les évaluateurs nous ont renvoyé le tableau du GRBAS et les formulaires CAPE-Vf complétés ainsi que divers commentaires sur la passation des sessions d'évaluation ou sur le CAPE-Vf lui-même.

3. Analyse statistique

Les données ont été saisies sur ordinateur, sur le logiciel Microsoft Excel. Elles ont ensuite été traitées avec le logiciel Jamovi (version 2.5 ; The jamovi project, 2024).

Notons que le CAPE-Vf offre la possibilité d'évaluer les diverses tâches du protocole séparément. Par souci de faisabilité, nous avons choisi de faire la moyenne entre les scores lorsque

les tâches vocales étaient évaluées séparément. En effet, il n'a pas été possible de comparer plusieurs évaluations d'un paramètre pour une même voix, étant donné que tous les juges n'ont pas évalué les mêmes paramètres séparément.

3.1. Fiabilité intra-juge

3.1.1. CAPE-Vf

La fiabilité intra-juge nous permet de déterminer si deux évaluations d'une même voix avec le CAPE-Vf, effectuées par un même évaluateur à deux instants différents, sont cohérentes entre elles. Nos données étant de type continu, nous avons utilisé le test de **coefficient de corrélation intraclasse (ICC)**. Nous avons choisi un modèle à deux facteurs à effets aléatoires, dans lequel nos juges et nos enregistrements vocaux sont traités comme des échantillons aléatoires issus de leurs populations respectives. Le type d'ICC a été défini comme « évaluateur unique » (single) parce que l'usage prévu du CAPE-V est qu'une voix sera évaluée par un seul clinicien et non pas en faisant la moyenne de plusieurs évaluateurs. Nous avons choisi comme définition « cohérence » (consistency) plutôt qu'« accord absolu » (agreement) puisque nous ne cherchons pas un accord strict entre les scores des évaluations, mais plutôt à ce que les valeurs soient cohérentes entre elles.

Nous avons calculé l'ICC de chaque juge pour tous les paramètres vocaux inclus dans le CAPE-Vf, illustrant ainsi pour chacune le niveau d'accord entre ses deux évaluations d'une même voix. Nous avons donc obtenu pour chaque juge un ICC pour la sévérité globale, la raucité, le souffle, le serrage, l'intensité et la hauteur. Selon Koo & Li (2016), un ICC <0.5 témoigne d'une faible fiabilité ; un ICC supérieur à 0.5 et allant jusqu'à 0.75 démontre une fiabilité modérée ; un ICC supérieur à 0.75 jusqu'à 0.9 montre une bonne fiabilité ; et un ICC supérieur à 0.90 représente une excellent fiabilité.

Nous avons également calculé la médiane des différences absolues pour chaque paramètre et pour chaque juge. Cette donnée a permis de donner un aperçu de l'étendue des différences entre les évaluations répétées, sur les échelles de 0 à 100 du CAPE-Vf.

3.1.2. GRBAS

De même que pour le CAPE-Vf, nous souhaitions voir si un évaluateur évalue la même voix avec le GRBAS à deux instants différents de manière cohérente. Nous avons ainsi calculé **l'alpha de Krippendorff,** adapté aux données ordinales, pour chaque paramètre du GRBAS et pour chaque juge. D'après Marzi et al. (2024), un alpha inférieur à 0.67 représente un accord faible ; un alpha situé entre 0.67 et 0.79 témoigne d'un accord modéré ; un alpha supérieur ou égal à 0.80 montre un accord satisfaisant et un alpha égal à 1 signifie qu'il y a un accord parfait.

3.2. Fiabilité inter-juges

3.2.1. CAPE-Vf

Nous souhaitions déterminer si les évaluations des mêmes voix effectuées avec le CAPE-Vf mais par différents évaluateurs étaient cohérentes entre elles. Nos données étant de type continu, nous avons appliqué le test de **coefficient de corrélation intraclasse (ICC)**. Tout comme pour la fiabilité intra-juge, nous avons choisi un modèle à deux facteurs à effets aléatoires, avec l'unité «

évaluateur unique » (single) et un type de relation de « cohérence » (consistency). Ainsi, nous avons calculé pour chaque liste le niveau d'accord entre les juges (illustré par l'ICC), et ce pour chaque paramètre du CAPE-Vf. Les guidelines d'interprétation de Koo & Li (2016) s'appliquent comme pour la fiabilité intra-juge.

3.2.2. GRBAS

Comme pour le CAPE-Vf, nous souhaitions déterminer si les évaluations des mêmes voix avec le GRBAS par différents évaluateurs sont cohérentes entre elles. Les données sont ordinales : nous calculons donc **l'alpha de Krippendorff** pour chaque paramètre du GRBAS et pour chaque liste, ce qui permet de rendre compte du niveau d'accord entre les différents juges de chaque liste. Les guidelines d'interprétation de Marzi et al. (2024) s'appliquent comme pour la fiabilité intra-juge.

3.3. Validité de critère

3.3.1. Corrélation CAPE-Vf / GRBAS

Les données issues des évaluations avec le CAPE-Vf sont de type continu ; celles relevées avec le GRBAS sont ordinales. Ainsi, nous avons choisi de mesurer la corrélation entre les scores moyens du CAPE-Vf et du GRBAS pour chaque voix évaluée à l'aide du test de **corrélation de Spearman**. Nous avons choisi d'étudier les corrélations entre les quatre paramètres vocaux équivalents du CAPE-Vf et du GRBAS : les scores de sévérité globale (CAPE-Vf) ont été comparés à ceux du grade G (GRBAS), tout comme les scores de raucité, de souffle et de serrage présents sous le même nom dans les deux outils. Selon Schober et al. (2018), une corrélation de Spearman *r* située entre 0.0 et 0.10 indique qu'il n'y a pas de corrélation. Un *r* entre 0.11 et 0.39 montre une faible corrélation, un r entre 0.40 et 0.69 une corrélation modérée, un *r* entre 0.70 et 0.89 une corrélation forte et un *r* supérieur à 0.90 une corrélation très forte. Ces fourchettes sont cependant à interpréter avec une certaine prudence puisqu'elles ont été établies de manière arbitraire (Schober et al., 2018).

3.3.2. Corrélations CAPE-Vf et GRBAS avec le VHI

Comme précisé ci-dessus, nous souhaitons comparer les données de type continu issues des évaluations avec le CAPE-Vf et les données ordinales du GRBAS avec les données ordinales issues de l'auto-questionnaire VHI. Nous avons donc calculé les corrélations entre les scores de « sévérité globale » du CAPE-Vf et le score total du VHI, puis le score G du GRBAS et le score total du VHI avec le test de **corrélation de Spearman**. Les guidelines d'interprétation de Schober et al. (2018) s'appliquent comme pour la corrélation entre le CAPE-Vf et le GRBAS.

RÉSULTATS ET ANALYSE

1. Statistiques descriptives

1.1. CAPE-Vf

Le tableau 5 décrit les scores de chaque paramètre vocal évalué dans le CAPE-Vf. La **moyenne** illustre la tendance centrale de l'ensemble des données. Le paramètre présentant la moyenne de scores la plus faible est le souffle (18.4) ; la moyenne de scores la plus haute est celle de la sévérité globale (36.2).

La **médiane** correspond à la valeur divisant l'ensemble de données en deux parties égales ; en d'autres termes, c'est la valeur du milieu lorsque les données sont triées par ordre croissant. Elle est moins sensible aux valeurs extrêmes que la moyenne et offre une meilleure représentation de la « valeur centrale » de l'ensemble de données. On retrouve la médiane la plus faible pour la hauteur (7.00) qui est suivie de près par le souffle (8.00). Ainsi, nous pouvons conclure que la moitié des voix ont obtenu un score de hauteur et de souffle inférieur ou égal à 7 et 8 respectivement. On trouve la médiane la plus élevée pour la sévérité globale (29.0).

L'écart-type est une mesure de dispersion qui indique dans quelle mesure les valeurs de l'ensemble de données sont dispersées autour de la moyenne. Plus l'écart-type est grand, plus les valeurs sont dispersées. L'écart-type le plus réduit est celui du souffle (22.7) et le plus grand est celui de la sévérité globale (28.6).

L'écart interquartile (EIQ) est également une mesure de dispersion. Il représente la différence entre le troisième quartile (Q3) et le premier quartile (Q1). Il couvre la moitié centrale des données et est moins sensible aux valeurs extrêmes que l'écart-type. L'EIQ le plus réduit est celui du souffle (27.3) ; le plus large est celui de la sévérité globale (44.5).

Le score **minimum** attribué pour tous les paramètres est 0. Le score **maximum** plus bas est de 92 pour la raucité et le plus élevé est de 100 pour la sévérité globale.

Tableau 5. Description des scores de chaque paramètre vocal évalué dans le CAPE-Vf

	SÉVÉRITÉ GLOBALE	RAUCITÉ	SOUFFLE	SERRAGE	HAUTEUR	INTENSITÉ
N	243	244	243	244	201	191
Moyenne	36.2	24.4	18.4	28.6	21.8	22.2
Médiane	29.0	15.0	8.00	18.0	7.00	12.0
Ecart-type	28.6	27.1	22.7	27.8	27.5	26.3
Ecart interquartile	44.5	31.3	27.3	38.3	36.0	34.0
Minimum	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	100	92.0	93.0	98.0	99.0	99.0

Ces données sont illustrées sous forme de diagramme en boîte dans la figure 9 ci-dessous.

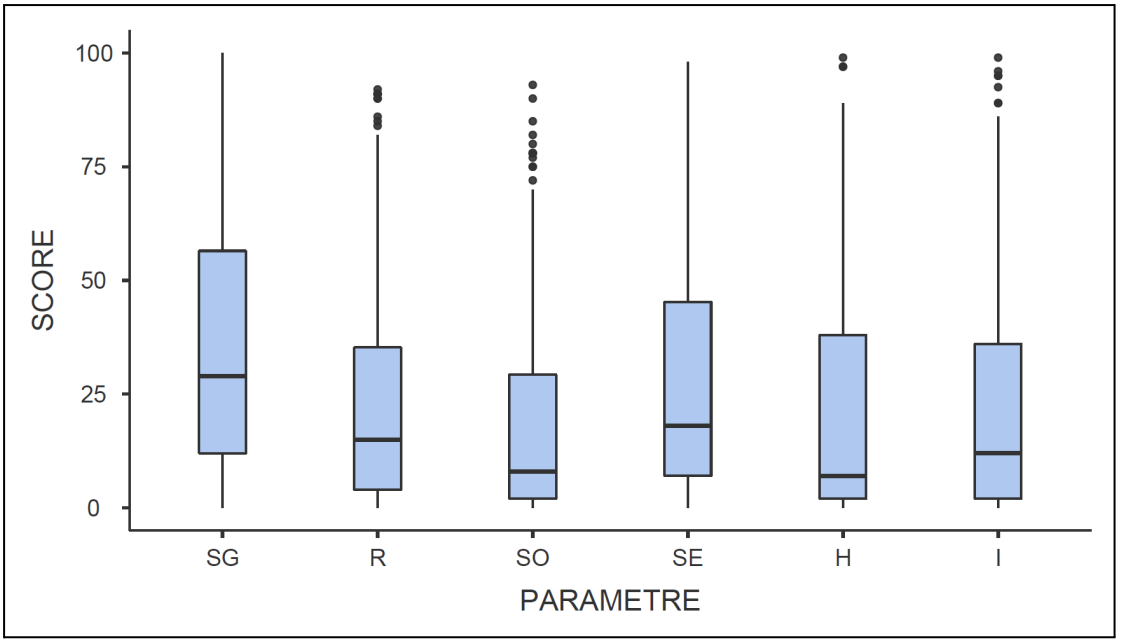


Figure 9. Répartition des scores en fonction des paramètres vocaux évalués dans le CAPE-Vf

Légende : *Les traits horizontaux correspondent aux médianes ; les moustaches représentent l'écart interquartile ; les points représentent les valeurs aberrantes ou* « *outliers* ».

1.2. GRBAS

Afin d'illustrer la répartition des niveaux d'atteinte pour chaque paramètre, nous avons rassemblé nos données dans un tableau de fréquences (Tableau 6). Celui-ci décrit, pour chaque paramètre évalué dans le GRBAS, le nombre (N) de voix pour chaque score possible (0, 1, 2 et 3) ainsi que la proportion que celles-ci représentent. Autrement dit, la proportion correspond au pourcentage de voix ayant obtenu tel score pour tel paramètre parmi toutes les voix évaluées. Nous pouvons donc nous faire une idée du nombre d'échantillons vocaux normaux (0) ou présentant une altération légère (1), modérée (2) et sévère (3) pour chaque paramètre.

La figure 10 permet un aperçu de cette répartition sous forme d'un histogramme à barres.

Tableau 6. Tableau de fréquence des scores du GRBAS

	Score	*N*	*Proportion (%)*
G	0	51	20.6
	1	99	40.1
	2	50	20.2
	3	47	19.0
R	0	122	49.4
	1	73	29.6
	2	27	10.9
	3	25	10.1

B	0	122	49.4
	1	73	29.6
	2	27	10.9
	3	25	10.1
A	0	138	55.9
	1	55	22.3
	2	32	13.0
	3	22	8.9
S	0	86	34.8
	1	98	39.7
	2	33	13.3
	3	30	12.1

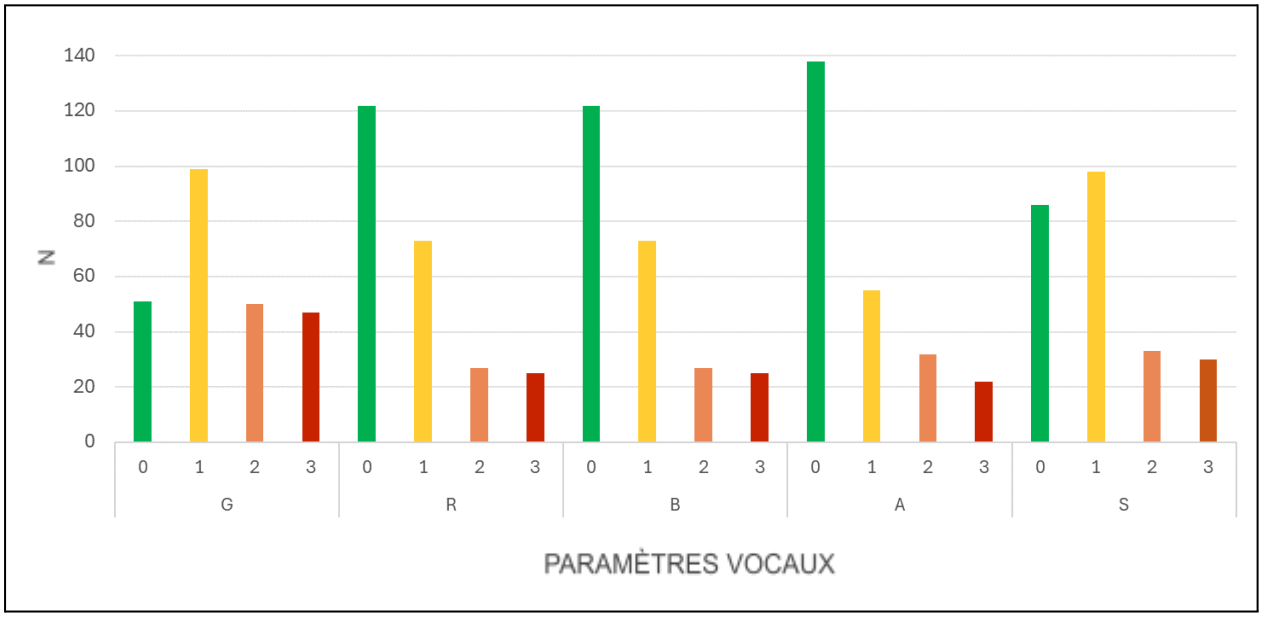


Figure 10. Fréquence d'apparition de chaque score en fonction du paramètre vocal évalué dans le GRBAS

2. Fiabilité intra-juge

2.1. CAPE-Vf

Les résultats pour la fiabilité intra-juge du CAPE-Vf pour chaque paramètre et pour chaque juge sont illustrés dans le tableau 7. Ces résultats indiquent que les paramètres de **sévérité globale**, de **serrage** et de **hauteur** montrent une **bonne fiabilité intra-juge** (ICCm :]0.75 – 0.90]). La **raucité**, le **souffle** et **l'intensité** présentent une **fiabilité modérée** (ICCm :]0.5 – 0.75]).

Concernant le degré de fiabilité des juges, un juge (juge 1) peut être considéré comme très fiable dans ses évaluations ; cinq juges démontrent une bonne fiabilité (juges 2, 3, 7, 9 et 10) et six une fiabilité modérée (4, 5, 6, 8, 11 et 12).

Tableau 7. Fiabilité intra-juge du CAPE-Vf

	SÉVÉRITÉ GLOBALE *ICC (MDA)*	**RAUCITÉ** *ICC (MDA)*	**SOUFFLE** *ICC (MDA)*	**SERRAGE** *ICC (MDA)*	**HAUTEUR** *ICC (MDA)*	**INTENSITÉ** *ICC (MDA)*	*ICCm par juge*
JUGE 1	0.94*** (11)	0.96*** (11)	0.82*** (15)	0.86** (11)	0.99*** (1)	0.83** (2.5)	0.90 (9)
JUGE 2	0.92** (3)	0.69* (6)	*N/A*	0.90** (6)	0.98*** (7)	0.97*** (5)	0.89 (6)
JUGE 3	0.74* (10)	0.75* (4)	0.98*** (4)	0.88** (5)	0.85** (3)	0.82** (6)	0.84 (5)
JUGE 4	0.90** (4)	0.92** (1)	0.87* (2)	0.70* (5)	0.65* (1)	0.35 (4)	0.73 (3)
JUGE 5	0.81** (6)	0.80** (14)	0.37 (19)	0.85** (15)	*N/A*	*N/A*	0.71 (14)
JUGE 6	0.70* (13)	0.84** (7)	0.15 (13)	0.81** (10)	0.86** (1)	0.50 (8)	0.64 (10.5)
JUGE 7	0.92** (6)	0.63* (16)	0.42 (7)	0.83** (8)	0.99*** (1.5)	*N/A*	0.76 (7)
JUGE 8	0.91** (14)	0.81** (6)	0.19 (2)	0.69* (11)	0.99*** (2)	0.84** (8)	0.74 (7)
JUGE 9	0.95*** (8)	0.97*** (4)	0.92** (5)	0.93*** (9)	0.64* (18)	0.87** (8)	0.88 (8)
JUGE 10	0.99*** (3)	0.98*** (4)	0.60 (1)	0.93*** (3)	0.95*** (2)	0.41 (6)	0.81 (2.5)
JUGE 11	0.92** (8)	0.16 (7)	0.77* (6)	0.98*** (6)	*N/A*	*N/A*	0.71 (6.5)
JUGE 12	0.99*** (6)	0.22 (6)	0.15 (6)	0.57 (10)	0.89** (5)	0.77* (2)	0.60 (6)
ICCm par paramètre	**0.89**	**0.73**	**0.57**	**0.83**	**0.88**	**0.71**	

** : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001*

***Légende : ICCm** : ICC moyen ; MDA : médiane des différences absolues*

Les couleurs représentent les niveaux d'interprétation de Koo & Li (2016) :

- *< 0.5 : faible fiabilité*
- *> 0.5 à 0.75 : fiabilité modérée*
- *>0.75 à 0.9 : bonne fiabilité*
- *> 0.90 : excellente fiabilité*

La mention N/A témoigne de l'impossibilité à traiter les données avec le logiciel de statistiques, soit parce que le juge n'en a pas produites ou parce que les données ne présentent pas assez de variabilité (ex : trop de scores 0). Seules les évaluations de douze juges ont pu être incluses, un des participants n'ayant pas évalué les voix répétées.

2.2. GRBAS

Le tableau 8 présente les résultats de la fiabilité intra-juge du GRBAS pour chaque paramètre et pour chaque juge. Le **grade (G)** montre un degré d'**accord satisfaisant** entre les deux évaluations par un même juge (*α moyen ≥ 0.80*). La **raucité** (R), le **souffle** (B) et **l'asthénie** (A) présentent un degré d'**accord modéré** (*α moyen : [0.67 – 0.79]*) alors que le **serrage** (S) montre un **accord faible**.

Le degré de fiabilité des juges est illustré par la valeur de l'alpha de Krippendorff moyen par juge : quatre juges présentent un accord satisfaisant dans leurs diverses évaluations, quatre montrent un accord modéré et cinq un accord faible.

Tableau 8. Fiabilité intra-juge du GRBAS (Alpha de Krippendorff)

	G	R	B	A	S	α *moyen par juge*
JUGE 1	0.63	0.88	0.86	0.75	0.82	0.79
JUGE 2	1	0.39	0.52	*N/A*	0.54	0.61
JUGE 3	0.88	0.67	1	0.83	0.80	0.84
JUGE 4	1	1	0.71	1	0.35	0.81
JUGE 5	0.69	0.39	0.47	0.86	0.76	0.63
JUGE 6	0.77	0.88	0.71	0.85	0.64	0.77
JUGE 7	0.37	0.52	0.45	0.88	0.71	0.59
JUGE 8	0.80	0.77	0.82	1	0.20	0.72
JUGE 9	0.86	0.88	0.25	0.52	0.38	0.58
JUGE 10	0.85	0.77	0.63	0.64	1	0.78
JUGE 11	0.86	0.88	0.99	*N/A*	0.49	0.81
JUGE 12	0.84	0.88	1	0.84	0.77	0.87
JUGE 13	0.85	0.51	0.67	0.33	0.78	0.63
α moyen	**0.80**	**0.72**	**0.70**	**0.77**	**0.63**	

NB : L'alpha de Krippendorff n'a pas de probabilité p associée

***Légende : G** : grade ; **R** : raucité ; **B** : breathiness (souffle) ; **A** : asthenia (asthénie) ; **S** : strain (serrage)*

Les couleurs représentent les niveaux d'interprétation de Marzi et al. (2024) :

- *< 0.67 : accord faible*
- *0.67 - 0.79 : accord modéré*
- *≥ 0.80 : accord satisfaisant*
- *1 : accord parfait*

3. Fiabilité inter-juges

3.1. CAPE-Vf

Le tableau 9 illustre la fiabilité inter-juges rapportée avec le CAPE-Vf pour chaque paramètre et pour chaque liste d'écoute. La **sévérité globale** montre une **bonne fiabilité inter-juges** (ICC_m :]0.75 – 0.90]). La **raucité**, le **souffle**, le **serrage**, la **hauteur** et **l'intensité** démontrent quant à eux une **faible fiabilité** (ICC_m < 0.50), bien que certains scores au sein des listes présentent une fiabilité modérée (raucité listes 1 et 3, souffle liste 4, serrage liste 1, hauteur et intensité liste 3) voire bonne ou excellente (souffle liste 2, serrage liste 4, hauteur liste 1).

Tableau 9. Fiabilité inter-juges du CAPE-Vf

	SÉVÉRITÉ GLOBALE	RAUCITÉ	SOUFFLE	SERRAGE	HAUTEUR	INTENSITÉ
LISTE 1	0.75**	0.67**	0.00	0.59*	0.96***	0.33
LISTE 2	0.64***	0.26	0.77***	-0.15	0.05	0.18
LISTE 3	0.82***	0.68***	0.48**	0.17	0.16	-0.15
LISTE 4	0.85***	0.28**	0.53***	0.79***	0.54***	0.66***
ICC_m	**0.77**	**0.47**	**0.45**	**0.35**	**0.43**	**0.26**

** : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001*

***Légende : ICC_m** : ICC moyen*

Les couleurs représentent les niveaux d'interprétation de Koo & Li (2016) :

- *< 0.5 : faible fiabilité*
- *> 0.5 à 0.75 : fiabilité modérée*
- *>0.75 à 0.9 : bonne fiabilité*
- *> 0.90 : excellente fiabilité*

3.2. GRBAS

Les résultats de l'analyse de fiabilité inter-juges pour les cinq caractéristiques vocales décrites par le GRBAS sont illustrés dans le tableau 10 et montrent tous une **faible fiabilité inter-juges** (α moyen < 0.67).

Tableau 10. Fiabilité inter-juges du GRBAS (Alpha de Krippendorff)

	G	R	B	A	S
LISTE 1	0.61	0.49	0.69	-0.34	0.08
LISTE 2	0.47	0.09	0.59	0.54	-0.11
LISTE 3	0.62	0.37	0.49	0.46	-0.16
LISTE 4	0.75	0.46	0.33	0.47	0.64
α moyen	**0.61**	**0.35**	**0.53**	**0.28**	**0.11**

NB : L'alpha de Krippendorff n'a pas de probabilité p associée

***Légende : G** : grade ; **R** : raucité ; **B** : breathiness (souffle) ; **A** : asthenia (asthénie) ; **S** : strain (serrage)*

Les couleurs représentent les niveaux d'interprétation de Marzi et al. (2024) :

- *< 0.67 : accord faible*
- *0.67 - 0.79 : accord modéré*
- *≥ 0.80 : accord satisfaisant*
- *1 : accord parfait*

4. Validité de critère

4.1. Corrélation CAPE-Vf / GRBAS

Le tableau 11 représente les corrélations de Spearman entre les scores des paramètres comparables du CAPE-Vf et du GRBAS pour chaque juge. En effet, la fiabilité inter-juges ayant montré un faible accord entre les évaluations par les différents juges, nous avons choisi de calculer les corrélations entre les différents paramètres du CAPE-Vf et du GRBAS pour chaque juge et non à partir des moyennes de scores pour chaque voix.

On relève une **corrélation forte** (*r moyen : [0.70 - 0.89]*) entre les scores de **sévérité globale** du CAPE-Vf et le **grade** du GRBAS ainsi qu'une **corrélation modérée** (*r moyen : [0.40 - 0.69]*) entre les scores de **raucité** du CAPE-Vf et du GRBAS, de **souffle** du CAPE-Vf et du GRBAS et de **serrage** du CAPE-Vf et du GRBAS.

Tableau 11. Corrélations de Spearman entre les scores des paramètres comparables du CAPE-Vf et du GRBAS, pour chaque juge

	SG / G	CVR / R	SO / B	SE / S
JUGE 1	0.85***	0.95***	0.93***	0.81**
JUGE 2	0.81**	0.68*	-0.25	0.68*
JUGE 3	0.82**	0.41	0.67*	-0.11
JUGE 4	0.73**	0.70*	0.78**	0.42
JUGE 5	0.91***	0.47	0.78**	0.51
JUGE 6	0.94***	0.47	0.49	0.58*
JUGE 7	0.89***	0.88***	0.63*	0.59*

JUGE 8	0.72**	0.73**	0.75**	-0.28
JUGE 9	0.84***	0.45	0.73**	0.74**
JUGE 10	0.93***	0.85***	0.78**	0.70*
JUGE 11	0.88***	0.58*	0.66*	0.60*
JUGE 12	0.70*	0.62*	0.38	0.61*
JUGE 13	0.94***	0.43	0.51	0.94***
***r*moyen**	**0.84**	**0.63**	**0.60**	**0.52**

** : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001*

***Légende : SG** : sévérité globale du CAPE-Vf ; **G** : grade du GRBAS ; **CVR** : raucité du CAPE-Vf ; **R** : raucité du GRBAS ; **SO** : souffle du CAPE-Vf ; **B** : souffle du GRBAS ; **SE** : serrage du CAPE-Vf ; **S** : serrage du GRBAS*
Les couleurs représentent les niveaux d'interprétation de Schober et al. (2018) :

- *0.0 - 0.10 : pas de corrélation*
- *0.11 - 0.39 : faible corrélation*
- *0.40 - 0.69 : corrélation modérée*
- *0.70 - 0.89 : corrélation forte*
- *0.90 - 1 : corrélation très forte*

4.2. Corrélations CAPE-Vf et GRBAS avec le VHI

Le tableau 12 illustre les corrélations de Spearman entre le score total du VHI, la sévérité globale du CAPE-Vf et le grade du GRBAS. Les scores totaux des **VHI** remplis par les locuteurs sont **modérément corrélés** (*r moyen : [0.40 - 0.69]*) aux scores de **sévérité globale** du CAPE-Vf (**SG**) et au **grade** du GRBAS (**G**).

Tableau 12. Corrélations de Spearman entre le score total du VHI, la sévérité globale du CAPE-Vf et le grade du GRBAS

	SG / score total VHI	G / score total VHI
JUGE 1	0.47	0.49
JUGE 2	0.31	0.34
JUGE 3	0.69*	0.78**
JUGE 4	0.26	0.75**
JUGE 5	0.84**	0.87***
JUGE 6	0.81**	0.71**
JUGE 7	0.71*	0.75**
JUGE 8	0.69*	0.66*
JUGE 9	0.46	0.52
JUGE 10	0.45	0.43
JUGE 11	0.38	0.26
JUGE 12	0.43	0.37
JUGE 13	0.39	0.42
***r*moyen**	**0.53**	**0.57**

** : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001*
***Légende :** **SG** : sévérité globale du CAPE-Vf ; **G** : grade du GRBAS*
Les couleurs représentent les niveaux d'interprétation de Schober et al. (2018) :

- *0.0 - 0.10 : pas de corrélation*
- *0.11 - 0.39 : faible corrélation*
- *0.40 - 0.69 : corrélation modérée*
- *0.70 - 0.89 : corrélation forte*
- *0.90 - 1 : corrélation très forte*

DISCUSSION

1. Etude des objectifs et hypothèses

L'évaluation perceptive menée par les cliniciens de la voix est un élément essentiel du bilan vocal. Face au manque d'outils disponibles en langue française, nous avons pour objectif de valider la version francophone du CAPE-V afin d'élargir le panel d'outils validés à la disposition des cliniciens de la voix.

1.1. Discussion autour de la première hypothèse : fiabilité du CAPE-Vf

Notre première hypothèse s'intéresse à la fiabilité du CAPE-Vf, c'est-à-dire à la reproductibilité des résultats dans différents contextes. Notre hypothèse est la suivante : l'adaptation française du CAPE-Vf est un outil d'évaluation perceptive de la voix montrant une **bonne fiabilité intra-juge** (ICC ⩾ 0.75) ainsi qu'une **bonne fiabilité inter-juges** (ICC ⩾ 0.75) pour les différents paramètres.

La **fiabilité intra-juge** qui témoigne de la cohérence entre plusieurs évaluations d'une même voix par un juge, se révèle plutôt bonne. La sévérité globale est le paramètre montrant la meilleure fiabilité intra-juge (ICC_m = 0.89), une tendance relevée par Mahalingam et al. (2024) dans leur revue de littérature sur les adaptations du CAPE-V effectuées à ce jour. Ce paramètre est suivi de près par la hauteur (ICC_m = 0.88) puis par le serrage (ICC_m = 0.83) qui présentent également une bonne fiabilité intra-juge. La raucité (ICC_m = 0.73) et l'intensité (ICC_m = 0.71) présentent des scores d'ICC_m dans l'extrémité supérieure de la fourchette correspondant à la fiabilité modérée. Le souffle est le paramètre présentant la fiabilité intra-juge la plus faible (ICC_m = 0.57) ; notons cependant qu'il est relevé dans la littérature que ce paramètre peut être influencé par la raucité (Millet & Dejonckere, 1998 ; Park et al., 2022). De plus, il semblerait que les voix que nous avons incluses dans notre étude aient un niveau relativement faible de souffle, la valeur médiane des scores pour ce paramètre étant de 8.00 pour le CAPE-Vf. Nous remarquons la même tendance chez les scores du GRBAS, où environ la moitié des voix (n=122, c'est-à-dire 49.4%) ont obtenu la note de 0 pour le paramètre B.

Plus globalement, la moitié des juges présentent une bonne – voire excellente – fiabilité dans leurs scores.

Cependant, la **fiabilité inter-juges,** qui illustre le degré d'accord entre plusieurs juges évaluant les mêmes voix, se révèle faible pour la plupart des paramètres. La sévérité globale est le seul paramètre présentant une bonne fiabilité (ICC_m = 0.77). Les cinq autres paramètres vocaux évalués dans le CAPE-Vf démontrent une faible fiabilité. En d'autres termes, cela signifie que les juges sont d'accord dans leurs évaluations de la sévérité globale, mais évaluent différemment les voix quant à leur raucité, leur caractère soufflé et serré, leur intensité et leur hauteur.

Nos résultats sont moins bons que dans d'autres études, où des valeurs supérieures à 0.71 et 0.66 ont été trouvées pour le souffle et la raucité, respectivement. (Mahalingam et al., 2024). Seule la version Hindi du CAPE-V (Joshi et al., 2020) montre des résultats plus faibles que les nôtres pour ce dernier paramètre (ICC = -0.36). L'intensité, qui montre une fiabilité relativement bonne dans d'autres études où on retrouve des valeurs supérieures à 0.54 (Mahalingam et al., 2024), est le

paramètre montrant la plus faible fiabilité inter-juges dans la nôtre (ICC_m = 0.26). Notons cependant que plusieurs juges ont peu ou pas évalué l'intensité, ou ont évalué uniquement le type d'atteinte (cf. tableau 5 : n=191 évaluations) ; ainsi, les résultats doivent être interprétés avec une certaine prudence. Ceci pourrait être le reflet d'un manque de pratique dans l'évaluation perceptive de l'intensité, ce paramètre n'étant pas inclus dans le GRBAS, outil avec lequel les juges sont les plus familiers. Nos résultats pour le serrage sont similaires à ceux rapportés par Behlau et al. (2022) dans la version brésilienne (ICC = 0.32) bien qu'ils soient globalement plus faibles que ceux rapportés dans d'autres études (ICC>0.59). La hauteur montre également une faible fiabilité dans l'étude de Behlau et al. (ICC = 0.25) et bien que celle rapportée dans notre étude soit supérieure (ICC_m = 0.43), elle n'atteint pas les degrés de fiabilité inter-juges relevés dans les autres études (ICC>0.57).

La fiabilité inter-juges inférieure à la fiabilité intra-juge, tendance observée dans la majorité des études de validation du CAPE-V dont la nôtre, pourrait ici s'expliquer en partie par l'absence de marqueurs de sévérité sous l'EVA. En effet, cette modification a été effectuée dans l'objectif d'éviter un regroupement de points sous les marqueurs, comme souligné dans l'étude de Pommée et al. (2024). Ce constat est également mis en lumière par Nagle (2022) : l'EVA hybride incluse dans la version originale du CAPE-V (cf. figure 5, p.25), avec ses marqueurs verbaux de sévérité, pourrait pousser les juges à catégoriser les voix selon ces descriptions. L'absence de ces marqueurs sous l'EVA dans notre étude a ainsi entraîné une fiabilité inter-juge réduite, puisque chaque juge s'est référé à ses propres standards quant à relation entre le score numérique sur l'EVA et la sévérité de l'atteinte perçue. De plus, cette modification apportée à l'EVA a permis aux juges d'évaluer l'atteinte de la sévérité sur un réel continuum, rétablissant ainsi les avantages qu'apporte ce type d'échelle, dont sa sensibilité supérieure aux changements. Autrement dit, les marqueurs sous l'EVA entraînent un manque de sensibilité puisqu'un changement significatif pourrait être relevé entre deux évaluations d'une voix sans pour autant que la sévérité en soit modifiée (à titre d'exemple, une voix qui passerait d'un score de souffle de 40/100 à 20/100 montrerait une évolution positive qui ne serait pas forcément relevée si l'on se fie aux marqueurs catégoriels, puisque ces deux scores sont considérés comme des « atteintes modérées »). Ainsi, la possibilité de l'influence des marqueurs catégoriels de sévérité dans les résultats de précédentes études n'est pas exclue.

La variabilité importante des scores illustrée par la faible fiabilité inter-juges peut également s'expliquer par la nature subjective de l'évaluation perceptive. En effet, lors de l'évaluation vocale les juges se réfèrent à leurs propres normes internes quant à la sévérité de l'atteinte de chaque paramètre vocal. Par conséquent, il est intéressant d'envisager notre outil dans le cadre de son utilité clinique : il est plus important que les juges soient cohérents dans leurs propres évaluations, leur permettant ainsi de mesurer de manière fiable l'évolution des performances de leurs patients dans le temps. La faible fiabilité inter-juges rapportée ici pour le CAPE-V*f* est moins gênante puisqu'il est rare que les cliniciens communiquent exclusivement sur la base de l'évaluation perceptive. Rappelons en effet qu'une évaluation vocale de qualité repose sur son aspect multidimensionnel et sur l'inclusion de diverses méthodes comprenant – sans s'y limiter– l'évaluation perceptive.

Ainsi, au vu de la fiabilité intra-juge satisfaisante et de la faible fiabilité inter-juges, **notre première hypothèse est partiellement validée.**

1.1.1. Discussion autour de notre hypothèse secondaire : fiabilités du CAPE-Vf et du GRBAS

L'étude de Zraick et al. (2011) établissant la validité de la version originale du CAPE-V démontre une fiabilité inter-juges et intra-juge plus élevées pour les scores du CAPE-V que pour ceux du GRBAS. La supériorité psychométrique du CAPE-V par rapport au GRBAS est d'ailleurs un des arguments en faveur de son utilisation en clinique.

Notre étude révèle également que les **fiabilités intra-juge et inter-juges du CAPE-Vf sont supérieures à celles du GRBAS, validant ainsi notre hypothèse secondaire**. En effet, le GRBAS présente une faible fiabilité inter-juges pour tous les paramètres vocaux qu'il évalue, alors que le CAPE-Vf présente une bonne fiabilité inter-juges ($ICC_m = 0.77$) pour le paramètre de sévérité globale. Ceci suggère que l'évaluation de la sévérité globale telle que dans le CAPE-Vf serait plus fiable que l'évaluation avec l'échelle de Likert en quatre points du GRBAS.

Les paramètres vocaux évalués dans les deux échelles sont la sévérité globale (CAPE-Vf) / le grade (GRBAS), la raucité, le souffle et le serrage. La sévérité globale / le grade montrent une bonne fiabilité intra-juge pour les deux outils. La raucité et le souffle présentent également des fiabilités intra-juge similaires, de degré modéré. Le serrage présente quant à lui une bien meilleure fiabilité intra-juge lorsqu'il est évalué avec le CAPE-Vf : on trouve une bonne fiabilité avec ce dernier, contre une faible fiabilité lors de l'évaluation avec le GRBAS.

Notons cependant que dans notre étude les juges ont évalué avec les deux outils les mêmes stimuli vocaux suivant le protocole standardisé en trois tâches vocales du CAPE-Vf. Pourtant, le GRBAS n'inclut aucun protocole standardisé dans son utilisation classique et peut être effectué sur la base de n'importe quel stimulus vocal. La « standardisation » de l'évaluation avec le GRBAS menée ici a donc sûrement influencé positivement les résultats de fiabilité de cet outil ; l'utilisation classique du GRBAS aurait probablement montré des résultats inférieurs à ceux rapportés dans notre étude.

Ainsi, du fait de sa fiabilité supérieure à celle du GRBAS, il semblerait que le CAPE-Vf est l'outil à privilégier dans le cadre de l'évaluation perceptive de la voix si l'on souhaite améliorer la robustesse de celle-ci.

1.2. Discussion autour de la seconde hypothèse : corrélation des scores du CAPE-Vf et du GRBAS

La validité de critère du CAPE-Vf repose sur la comparaison des scores obtenus lors des évaluations de voix avec celui-ci, à ceux obtenus avec le GRBAS, l'actuel outil de référence pour l'évaluation perceptive. Lors d'autres études de validation du CAPE-V en d'autres langues, une forte corrélation a été relevée entre les différents paramètres comparables du CAPE-V et du GRBAS. C'est pourquoi nous avons posé comme hypothèse que l'adaptation française du CAPE-V est un outil d'évaluation perceptive de la voix montrant une bonne validité de critère, mesurée par une forte corrélation avec les différents paramètres ($r \geqslant 0.70$) de l'échelle GRBAS.

Tout comme dans les précédentes études de validation (Mahalingam et al., 2024), notre étude révèle que les paramètres de sévérité globale et de grade présentent une forte corrélation (r_{moyen} = 0.84) ; c'est également la plus élevée de toutes. La raucité (r_{moyen} = 0.63), le souffle (r_{moyen} = 0.60) et le serrage (r_{moyen} = 0.52) présentent une corrélation modérée. Ceux-ci suivent le même ordre d'importance que dans les précédentes études, où le souffle et le serrage présentent des corrélations plus faibles que les autres paramètres.

Seule la corrélation entre la sévérité globale et le grade montre une corrélation supérieure à 0.70, notre seconde hypothèse est donc **partiellement validée**.

1.3. Discussion autour de la troisième hypothèse : corrélation des scores du CAPE-Vf et du GRBAS avec ceux du VHI

Le VHI est un autoquestionnaire illustrant le degré de handicap vocal perçu par le locuteur. Le CAPE-V et le GRBAS sont deux outils employés par les cliniciens de la voix tels que les orthophonistes, phoniatres et ORL pour évaluer la voix. Ces deux types d'évaluations reflètent deux dimensions différentes du trouble vocal, essentielles à relever avant toute prise en charge afin d'appréhender la dysphonie dans la diversité de ses manifestations : d'un côté le vécu quotidien du patient vis-à-vis de son trouble, de l'autre le point de vue expert du clinicien. On ne trouve pas de consensus dans la littérature quant à l'accord de ces deux types de mesures. Alors que certaines études trouvent une corrélation modérée entre les mesures d'auto-questionnaires et d'évaluation perceptive (Dejonckere et al., 1996 ; Mateos-Serrano et al., 2021), d'autres émettent davantage de réserves quant à la force du lien qu'elles présentent (Maunsell et al., 2022). Nous posons ainsi l'hypothèse que la corrélation du score de sévérité globale du CAPE-Vf avec le score total du VHI est plus faible qu'avec le G du GRBAS ($r < 0.70$), étant donné que ces deux outils évaluent des construits différents du phénomène vocal multidimensionnel.

Notre hypothèse est validée : bien que les scores de sévérité globale du CAPE-Vf et de grade du GRBAS montrent une corrélation modérée avec les scores totaux du VHI (r_{moyen} = 0.53 et 0.57 respectivement), celles-ci sont plus faibles que les fortes corrélations rapportées entre la sévérité globale du CAPE-Vf et le grade du GRBAS (r_{moyen} = 0.84). Ceci confirme que ces deux outils évaluent deux réalités différentes liées à dysphonie ; d'une part, le fonctionnement vocal d'un point de vue perceptif, résultant du geste vocal lui-même, et d'autre part l'impact de la dysphonie sur le quotidien du patient, qui découle d'une multitude de facteurs personnels, contextuels et sociologiques. Nos résultats offrent également une perspective intéressante quant à la prévention des conséquences de la dysphonie sur la qualité de vie des patients. En effet, la corrélation entre ces deux types d'évaluations suggère que l'évaluation perceptive puisse servir pour prévenir l'évolution de la dysphonie et, par conséquent, contribuer à maintenir la qualité de vie du patient vis-à-vis de son trouble vocal.

1.4. Analyse qualitative des évaluations avec le CAPE-Vf

L'analyse qualitative des formulaires CAPE-Vf et le recueil des commentaires des étudiantes et juges ayant participé à l'étude ont révélé plusieurs éléments intéressants pouvant servir de lignes directrices pour d'éventuelles améliorations du CAPE-Vf.

Certaines fonctionnalités de l'outil ont été peu investies par nos juges, en commençant par la notation multiple. En effet, sur 1366 évaluations, les tâches vocales ont été évaluées séparément seulement à 104 reprises. Cependant, il est intéressant de noter que les juges étaient très précises dans leurs évaluations séparées. Effectivement, la plus petite différence de scores relevée entre les tâches pour un même paramètre, parmi toutes les évaluations multiples, était de l'ordre de 5 mm. Les 2 EVA vierges ont également été peu investies. En effet, elles n'ont été remplies qu'à trois reprises par la même juge, avec l'ajout des paramètres « respiration », « attaque » et « fry ». Cette même utilisation limitée est relevée par Nagle (2022) dans son étude sur l'utilisation clinique du CAPE-V. Le caractère constant/intermittent de chaque paramètre, représenté par les lettres « C/I » à droite de l'EVA, a également été peu évalué par les juges. Ceci pourrait être le reflet d'un manque d'échantillons vocaux représentatifs d'altérations vocales intermittentes.

Enfin, certains juges ont omis de qualifier la nature de l'atteinte de la hauteur et/ou de l'intensité alors même qu'elles les ont quantifiées en plaçant un curseur sur l'EVA. Inversement, certains juges ont laissé un commentaire sur le type d'atteinte sans la quantifier sur l'EVA ; ceci explique le nombre plus faible de scores relevés dans le tableau 5 pour ces deux paramètres (hauteur : n=201 ; intensité : n=191). Au vu de la confusion entraînée par l'évaluation de ces deux paramètres, nous nous questionnons sur la pertinence de l'échelle choisie, questionnement déjà relevé par Mbagira (2023) lors de la conception de l'adaptation francophone du CAPE-V. Ainsi, il paraît intéressant de clarifier les instructions quant à l'évaluation de ces deux paramètres dans notre protocole final, voire de proposer un tout autre type d'échelle pour l'évaluation de la hauteur et de l'intensité, telle qu'une « description verbale qualitative du critère altéré ou une échelle visuelle analogique bipolaire étiquetée par deux mots qui délimitent le continuum (p.ex., pour l'intensité « trop faible – normal – trop fort »). » (Mbagira, 2023).

Cependant, les commentaires sur la résonance et sur d'autres caractéristiques ont été très bien investis, à l'exception d'une juge qui n'a laissé aucun commentaire supplémentaire. En effet, un total de 134 commentaires sur 228 possibles ont été laissés pour les 41 voix incluses dans notre étude. Environ 31% des voix jugées à priori comme « non altérées » ont reçu un commentaire (dont deux commentaires positifs sur le timbre du locuteur). Ce chiffre est de 55% pour les voix légèrement altérées et de 58% pour les voix modérément altérées. Concernant les voix sévèrement altérées, 85% d'entre elles ont reçu un commentaire. Ceci met en évidence un des principaux avantages du CAPE-V par rapport au GRBAS, à savoir la possibilité d'apporter une appréciation qualitative des voix au-delà de la simple apposition d'une note. En effet, cette dernière laisse trop de place à la subjectivité, à la fois du côté de l'évaluateur et de celui du lecteur. Ainsi, l'espace de commentaire libre inclus dans le CAPE-V$_f$ permet de clarifier les intentions de l'évaluateur, compensant en partie les lacunes de l'outil quant à la faible fiabilité inter-juges que présentent ses résultats quantitatifs.

Une dernière remarque qualitative sur le CAPE-V$_f$ concerne le temps de passation plus long par rapport à celui du GRBAS. En effet, l'essai pilote de notre étude a révélé un temps d'évaluation moyen de 2 min 30 pour le CAPE-V$_f$ et de 1 min 30 pour le GRBAS, sans compter la mesure des EVA que nous avons effectuée à posteriori. La différence entre les temps de passation rapportés dans ce

mémoire pourrait également être attribuée à plusieurs facteurs comme le manque de familiarité des juges avec le CAPE-Vf comparé au GRBAS, la modalité numérique nécessitant davantage de manipulations pour remplir le CAPE-Vf que le remplissage du tableau GRBAS, ou encore le nombre supérieur de paramètres vocaux à évaluer avec le CAPE-Vf. La mesure des EVA et la notation sur 100 de chacun des paramètres représentent également du temps en plus dans la passation du CAPE-Vf. Cette différence sera donc à considérer par le clinicien dans son choix de l'outil, selon s'il souhaite gagner en efficacité au niveau du temps ou optimiser la sensibilité et la fiabilité de son évaluation perceptive.

2. Limites et biais de notre étude

Cette étude présente des biais et limites concernant principalement la méthodologie employée, à la fois lors de l'enregistrement des locuteurs, lors de l'évaluation des voix par les juges et lors de notre analyse statistique des résultats.

Premièrement, les quarante-et-un locuteurs sollicités pour l'enregistrement sont francophones natifs. Ainsi, notre échantillon est peu représentatif de la diversité des accents qu'est amené à rencontrer le clinicien de la voix en France. De plus, ces locuteurs ont été enregistrés par trois orthophonistes dans trois cabinets différents ; ainsi les conditions d'enregistrement, bien que spécifiées dans notre protocole, présentent des différences inévitables ayant pu entraîner plus ou moins de variabilité dans la qualité des enregistrements vocaux.

Malgré notre volonté de masquer aux juges les pathologies vocales présentées par les locuteurs, bon nombre d'entre les patients ont mentionné leur suivi orthophonique et/ou la pathologie pour laquelle ils consultent lors de la tâche de parole spontanée. Ainsi, la connaissance des pathologies n'a pas permis une écoute des voix à l'aveugle, laissant place à une potentielle influence de celle-ci sur l'évaluation perceptive.

Une des limites majeures de notre étude provient de l'échantillon de juges que nous avons obtenu. Nous n'avons pu recruter qu'un seul type de cliniciens de la voix, les orthophonistes. Or, les phoniatres et ORL sont également des acteurs importants de l'évaluation vocale et sont régulièrement amenés à effectuer des évaluations perceptives dans le cadre de leur profession. Ainsi, sans leur participation à notre étude, nous risquons de manquer d'éléments importants à la fois d'un point de vue quantitatif (auraient-ils évalué les voix comme les orthophonistes ?) et qualitatif (auraient-ils investi les diverses fonctionnalités du CAPE-Vf différemment ?).

De plus, huit des vingt-et-un juges originaux ont abandonné l'étude après constitution des listes d'écoute, entraînant ainsi un déséquilibre important dans la composition de celles-ci. Ainsi, nous avons été contraints de faire des calculs de fiabilité inter-juges à partir des scores d'évaluations de seulement deux juges pour la liste 1, donnant ainsi des résultats de fiabilité et de corrélation moins robustes comparés à ceux de la liste 4, composée de cinq juges.

La modalité numérique en distanciel que nous avons choisie pour notre étude présente également des biais. Tout d'abord, nous n'avons aucune garantie de la qualité de l'environnement dans lequel les évaluations ont été menées, bien que nous ayons précisé les conditions idéales dans

notre protocole. L'outil informatique présente aussi ses limites quant à son accessibilité et à sa facilité d'utilisation ; tous les juges ne le manient pas aussi aisément, écueil qu'une évaluation sur papier du CAPE-Vf aurait permis d'éviter. Enfin, l'évaluation sur base d'enregistrements élimine la possibilité d'observer les potentielles manifestations physiques de la dysphonie du patient, importantes à considérer dans l'évaluation globale du trouble vocal. En effet, certaines dysphonies comme les DTM sont caractérisées par des tensions visibles et palpables au niveau du larynx (Payten et al., 2022). Ces symptômes physiques sont des éléments essentiels permettant d'agrémenter l'évaluation perceptive, notamment pour les paramètres « forçage » et « asthénie ». Ainsi, il n'est pas exclu que les résultats de fiabilité de notre étude aient pu être affectés par l'absence d'observation directe engendrée par la modalité exclusivement audio. Nous pouvons alors supposer que l'évaluation réelle, multimodale, en clinique permettrait d'augmenter la fiabilité du CAPE-Vf.

Enfin, comme précisé dans la méthodologie, nous avons effectué la moyenne des différents scores pour un même paramètre lorsque les tâches vocales étaient évaluées séparément. Núñez-Batalla et al. (2015) ont choisi, dans leur validation espagnole du CAPE-V, d'évaluer chaque tâche séparément afin d'effectuer isolément les diverses analyses statistiques de fiabilité et de validité. Ainsi, la fiabilité et la validité du CAPE-V espagnol ont été établies pour les trois tâches, soulignant l'intérêt de l'inclusion des divers stimuli vocaux dans le protocole d'évaluation du CAPE-V. Par souci de faisabilité, nous n'avons pas imposé aux juges d'évaluer les trois tâches séparément ; bien qu'il aurait été intéressant d'évaluer les fiabilités et validités respectives des tâches vocales, cela aurait augmenté considérablement le temps de passation de l'évaluation.

3. Perspectives

Au vu de la bonne validité et fiabilité intra-juge constatée pour le CAPE-Vf dans notre étude, ce mémoire établit la **validité préliminaire** de la version francophone du Consensus Auditory Perceptual Evaluation of Voice, auprès d'une population de locuteurs et d'orthophonistes français. Néanmoins, les limites que présente notre étude, notamment concernant le nombre réduit de participants, nous contraignent à interpréter nos résultats avec prudence. En effet, bien que nos résultats offrent une perspective encourageante dans la poursuite de la validation du CAPE-Vf, il serait toutefois souhaitable de réitérer l'expérience sur une plus grande population afin de généraliser nos constats.

Il paraît également pertinent d'élargir la validation du CAPE-Vf à d'autres populations francophones, permettant ainsi une utilisation plus répandue et adaptée aux spécificités culturelles et linguistiques des diverses régions francophones du monde. Une validation auprès d'autres cliniciens de la voix tels que les ORL et phoniatres, ainsi qu'auprès de chercheurs serait également souhaitable afin de confirmer la pertinence de l'utilisation du CAPE-Vf dans divers contextes cliniques et de recherche.

Une validation par rapport à des mesures instrumentales pourrait également être envisagée, permettant ainsi de rendre compte de la mesure dans laquelle une modification physiologique, attestée par les mesures acoustiques ou aérodynamiques, se reflète dans la perception des paramètres vocaux tels qu'évalués par le CAPE-Vf. Toutefois, une telle comparaison pourrait se

révéler compliquée étant donné que la sévérité globale, le serrage, le souffle et la raucité n'ont pas de corrélats acoustiques consensuels, à l'inverse de la hauteur et de l'intensité qui sont facilement mesurables acoustiquement (Nagle, 2016). Il serait également intéressant de déterminer si les phrases incluses dans le protocole du CAPE-Vf suscitent réellement les divers phénomènes et comportements vocaux qu'elles sont censées provoquer. Pour ce faire, nous pourrions imaginer la combinaison d'une évaluation perceptive ciblée (le comportement vocal visé est-il présent/absent ?) avec une analyse visuelle des spectrogrammes (par exemple, le phénomène de voisement de la phrase 2 pourrait être illustré par le pourcentage de segments voisés ou dévoisés) ainsi que des mesures acoustiques (comme la nasalité sur la phrase 5 qui pourrait être objectivée par le low tone-high tone ratio) ou des mesures aérodynamiques (comme la PSGE pour la phrase 4). Cette perspective paraît d'autant plus intéressante à investiguer puisqu'une étude de El-Adawy et al. (2011) auprès de locuteurs arabophones révèle des corrélations significatives entre le CAPE-V et des mesures acoustiques vocales et aérodynamiques.

Pour tenter d'améliorer la fiabilité inter-juge du CAPE-Vf, il serait pertinent de proposer aux juges des voix d'ancrage (« anchor samples » en anglais) et/ou des sessions d'entraînement en amont des évaluations. En effet, il est reconnu dans la littérature que ceux-ci permettent d'améliorer le degré d'accord et réduire la variabilité des scores entre les juges (Eadie & Baylor, 2006 ; Ghio et al., 2011). Il paraît donc pertinent de proposer un protocole d'entraînement tel que celui élaboré par Ghio et al. (2011) afin d'harmoniser les références internes des juges quant au caractère pathologique d'une voix. Cet entraînement pourrait être inclus à une future version digitalisée du CAPE-Vf.

Enfin, au vu des problématiques relevées lors de l'analyse qualitative de nos résultats, il paraît aussi pertinent d'envisager quelques modifications du protocole actuel du CAPE-Vf afin d'améliorer sa spécificité et l'efficacité de son utilisation. Parmi celles-ci nous pouvons citer le changement du type d'échelle pour les paramètres « hauteur » et « intensité » ou encore la modification de l'EVA. Cette dernière suggestion a déjà été mise en œuvre dans la validation du CAPE-V en Malay par Mohd Mossadeq et al. (2022) ainsi qu'en Mandarin par Chen et al. (2018). Ces versions du CAPE-V incluent une échelle à intervalles réguliers (EIR) à la place de l'EVA (Figure 11), permettant ainsi une augmentation de la rapidité et de l'efficacité des cotations. Une autre solution possible serait envisageable dans le cas où une future version digitale du CAPE-Vf viendrait à naître : l'évaluateur pourrait placer un curseur sur l'EVA, permettant ainsi l'extraction automatique de la mesure.

Figure 11. EIR incluse dans la version du CAPE-V en Mandarin de Chen et al. (2018)

CONCLUSION

Ce mémoire vise à répondre à notre problématique concernant la fiabilité et la validité de l'adaptation française du Consensus-Auditory Perceptual Evaluation of Voice. La littérature révélant un manque d'outils francophones standardisés, valides et fiables utilisables pour l'évaluation perceptive de la voix, cette validation s'inscrit dans une dynamique de diversification des supports à disposition des orthophonistes, ORL, phoniatres et chercheurs pratiquant dans le domaine de la voix.

Nous avons donc pensé trois hypothèses et une hypothèse secondaire pour guider notre travail. La première hypothèse concerne la fiabilité de notre outil : à quel point deux évaluations d'une même voix sont-elles cohérentes, qu'elles aient été effectuées par un même juge à deux reprises ou par plusieurs juges différents ? Nous avons également comparé ces résultats à ceux trouvés pour le GRBAS, échelle de référence pour l'évaluation perceptive de la voix. La deuxième hypothèse vise à déterminer la corrélation entre le CAPE-Vf et le GRBAS. Enfin, nous nous sommes questionnés sur le lien qu'entretiennent l'évaluation perceptive et l'autoévaluation de la qualité de vie du patient dans le cadre de la dysphonie.

Afin de répondre à notre problématique, nous avons élaboré une étude en deux parties. La première phase a consisté en l'enregistrement de quarante-et-un locuteurs dysphoniques et normophoniques. Lors de la seconde phase, treize orthophonistes françaises ont évalué ces voix avec le CAPE-Vf et le GRBAS. Nous avons ainsi pu valider l'hypothèse secondaire ainsi que l'hypothèse 3. Cependant, bien que nos résultats soient prometteurs, nous n'avons que partiellement validé les hypothèses 1 et 2, les seuils de validation définis dans nos hypothèses n'ayant pas été atteints pour tous les paramètres du CAPE-Vf.

Au vu des résultats rapportés dans notre étude, nous pouvons considérer que le CAPE-Vf présente une bonne fiabilité intra-juge ainsi qu'une bonne validité. Cependant, l'échantillon limité de participants – à la fois en termes de taille et de professions représentées – nous amène à considérer que ce mémoire établit la validité préliminaire du CAPE-Vf. Le CAPE-Vf reste un outil pertinent quant aux divers avantages qu'il présente, dont notamment sa sensibilité supérieure au changement et l'inclusion d'un éventail standardisé de stimuli vocaux permettant à la fois l'évaluation de la phonation pure grâce à la tâche de voyelle tenue, l'évaluation ciblée de comportements vocaux par le biais de phrases conçues spécifiquement à cet effet et l'évaluation fonctionnelle sur une tâche de parole semi-spontanée.

De belles perspectives s'offrent à nous concernant la poursuite du développement du CAPE-Vf. En effet, des validations auprès d'autres populations francophones ou par rapport à des mesures instrumentales (notamment acoustiques et aérodynamiques) sembleraient à prévoir afin d'élargir la portée de l'outil. L'amélioration de certains défauts que présente l'outil en lui-même, comme le type d'échelles employées ou le temps de passation, serait également envisageable. La création d'une future version digitale du CAPE-Vf semblerait pouvoir pallier une partie des écueils constatés.

BIBLIOGRAPHIE

Baken R.J., Orlikoff R.F. (2000). *Clinical Measurement of Speech and Voice*. San Diego, CA, USA : Singular Publ. Group, Inc., 2e édition, 174.

Baker, J., Ben-Tovim, D. I., Butcher, A., Esterman, A., & McLaughlin, K. (2007). Development of a modified diagnostic classification system for voice disorders with inter-rater reliability study. Logopedics Phoniatrics Vocology, 32(3), 99-112. https://doi.org/10.1080/14015430701431192

Barsties, B., & De Bodt, M. (2015). Assessment of voice quality : Current state-of-the-art. *Auris Nasus Larynx*, *42*(3), 183-188. https://doi.org/10.1016/j.anl.2014.11.001

Barsties, B., Latoszek, B., Maryn, Y., Gerrits, E., & De Bodt, M. (2017). The Acoustic Breathiness Index (ABI) : A Multivariate Acoustic Model for Breathiness. *Journal of Voice: Official Journal of the Voice Foundation*, *31*(4), 511.e11-511.e27. https://doi.org/10.1016/j.jvoice.2016.11.017

Behlau, M., Rocha, B., Englert, M., & Madazio, G. (2022). Validation of the Brazilian Portuguese CAPE-V Instrument—Br CAPE-V for Auditory-Perceptual Analysis. *Journal of Voice*, *36*(4), 586.e15-586.e20. https://doi.org/10.1016/j.jvoice.2020.07.007

Behrman, A. (2005). Common practices of voice therapists in the evaluation of patients. *Journal of Voice: Official Journal of the Voice Foundation,* *19*(3), 454-469. https://doi.org/10.1016/j.jvoice.2004.08.004

Brockmann-Bauser, M., Drinnan, M. J., & Carding, P. (2010, 1er janvier). *Are Jitter and Shimmer comparable to perceptual voice analysis in healthy voices?*

Céruse, P., Ltaief-Boudrigua, A., Buiret, G., Cosmidis, A. (2012). Anatomie descriptive, endoscopique et radiologique du larynx. *Oto-rhino-laryngologie*. http://dx.doi.org/10.1016/S0246-0351(12)56065-4

Chaves, C. R., Campbell, M., & Côrtes Gama, A. C. (2017). The Influence of Native Language on Auditory-Perceptual Evaluation of Vocal Samples Completed by Brazilian and Canadian SLPs. *Journal of Voice*, *31*(2), 258.e1-258.e5. https://doi.org/10.1016/j.jvoice.2016.05.021

Chen, Y., Robb, M. P., & Gilbert, H. R. (2002). Electroglottographic Evaluation of Gender and Vowel Effects During Modal and Vocal Fry Phonation. *Journal of Speech, Language, and Hearing Research*, *45*(5), 821-829. https://doi.org/10.1044/1092-4388(2002/066)

Chen, Z., Rui, F., Zhang, Y., Ge, P., Zhuang, P., Chou, A., & Jiang, J. (2018). The Mandarin Version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) and Its Reliability. *Journal of Speech Language and Hearing Research*, *61*, 1. https://doi.org/10.1044/2018_JSLHR-S-17-0386

Cheng, J., & Woo, P. (2010). Correlation between the Voice Handicap Index and voice laboratory measurements after phonosurgery. *Ear, Nose, & Throat Journal*, *89*(4), 183-188.

Cohen, S. M., Dupont, W. D., & Courey, M. S. (2006). Quality-of-Life Impact of Non-Neoplastic Voice Disorders : A Meta-Analysis. *Annals of Otology, Rhinology & Laryngology*, *115*(2), 128-134. https://doi.org/10.1177/000348940611500209

Cohen, S. M., Jacobson, B. H., Garrett, C. G., Noordzij, J. P., Stewart, M. G., Attia, A., Ossoff, R. H., & Cleveland, T. F. (2007). Creation and Validation of the Singing Voice Handicap Index. *Annals of Otology, Rhinology & Laryngology*, *116*(6), 402-406. https://doi.org/10.1177/000348940711600602

Cohen, S. M., Kim, J., Roy, N., Asche, C., & Courey, M. (2012). Prevalence and causes of dysphonia in a large treatment-seeking population. *The Laryngoscope*, *122*(2), 343-348. https://doi.org/10.1002/lary.22426

Cornut, G. (2009). *La voix parlée: Vol. 8e éd.* Presses Universitaires de France, 45-58. https://www.cairn.info/la-voix--9782130576747-p-43.htm

Cornut, G. (2019). *La voix* (9e édition). Presses universitaires de France.

Dabirmoghaddam, P., Aghadoost, S., Shokuhifar, G., Khoddami, S. M., Aghajanzadeh, M., & Naderifar, E. (2023). The Relationship Between Voice Handicap Index and Voice-Related Quality of Life in Patients with Early Laryngeal Cancer. *Indian Journal of Otolaryngology and Head and Neck Surgery: Official Publication of the Association of Otolaryngologists of India*, *75*(2), 587-593. https://doi.org/10.1007/s12070-022-03356-y

Deary, I. J., Wilson, J. A., Carding, P. N., & MacKenzie, K. (2003). VoiSS : A patient-derived Voice Symptom Scale. *Journal of Psychosomatic Research*, *54*(5), 483-489. https://doi.org/10.1016/s0022-3999(02)00469-5

Deary, I. J., Webb, A., Mackenzie, K., Wilson, J. A., & Carding, P. N. (2004). Short, Self-Report Voice Symptom Scales : Psychometric Characteristics of the Voice Handicap Index-10 and the Vocal Performance Questionnaire. *Otolaryngology–Head and Neck Surgery*, *131*(3), 232-235. https://doi.org/10.1016/j.otohns.2004.02.048

Deary, I., Scott, S., Wilson, I., White, A., Mackenzie, K., & Wilson, J. (2011). Personality and psychological distress in dysphonia. *British Journal of Health Psychology*, *2*, 333-341. https://doi.org/10.1111/j.2044-8287.1997.tb00547.x

Dehqan, A., Scherer, R. C., Dashti, G., Ansari-Moghaddam, A., & Fanaie, S. (2012). The effects of aging on acoustic parameters of voice. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *64*(6), 265-270. https://doi.org/10.1159/000343998

Dejonckere, P. H., Obbens, C., De Moor, G. M., & Wieneke, G. H. (1993). Perceptual Evaluation of Dysphonia : Reliability and Relevance. *Folia Phoniatrica et Logopaedica*, *45*(2), 76-83. https://doi.org/10.1159/000266220

Dejonckere, P. H., Remacle, M., Fresnel-Elbaz, E., Woisard, V., Crevier-Buchman, L., & Millet, B. (1996). Differentiated perceptual evaluation of pathological voice quality : Reliability and correlations with acoustic measurements. *Revue De Laryngologie - Otologie - Rhinologie*, *117*(3), 219-224.

Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., Remacle, M., & Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology, 258(2),* 77-82. https://doi.org/10.1007/s004050000299

Demirhan, E., Unsal, E. M., Yilmaz, C., & Ertan, E. (2016). Acoustic Voice Analysis of Young Turkish Speakers. *Journal of Voice: Official Journal of the Voice Foundation*, *30*(3), 378 https://doi.org/10.1016/j.jvoice.2015.04.018

Eadie, T. L., & Baylor, C. R. (2006). The Effect of Perceptual Training on Inexperienced Listeners' Judgments of Dysphonic Voice. *Journal of Voice*, *20*(4), 527-544. https://doi.org/10.1016/j.jvoice.2005.08.007

Eckel, F. C., & Boone, D. R. (1981). The s/z ratio as an indicator of laryngeal pathology. *Journal of Speech and Hearing Disorders*, *46*(2), 147–149.

El-Adawy, A., Ahmed, M., Hassan, E., & I.R., M. (2011). Modified GRBAS versus Cape-v scale for assessment of voice quality : Correlation with acoustic and aerodynamics measurement for Arabic speaking subjects. *Scientific Medical Journal*.

Engelbert, A. (2014). Cross-linguistic effects on voice quality : A study on Brazilians' production of Portuguese and English. *Proceedings of the International Symposium on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics*, *5*, 157-170.

Ferrand, C. T. (2002). Harmonics-to-noise ratio : An index of vocal aging. *Journal of Voice: Official Journal of the Voice Foundation*, *16*(4), 480-487. https://doi.org/10.1016/s0892-1997(02)00123-6

Francis, D. O., Daniero, J. J., Hovis, K. L., Sathe, N., Jacobson, B., Penson, D. F., Feurer, I. D., & McPheeters, M. L. (2017). Voice-Related Patient-Reported Outcome Measures : A Systematic Review of Instrument Development and Validation. *Journal of Speech, Language, and Hearing Research*, *60*(1), 62-88. https://doi.org/10.1044/2016_JSLHR-S-16-0022

Fritzell, B. (1996). Voice disorders and occupations. *Logopedics Phoniatrics Vocology*, *21*(1), 7-12. https://doi.org/10.3109/14015439609099197

Fujiki, R., & Thibeault, S. (2021). Examining Relationships Between GRBAS Ratings and Acoustic, Aerodynamic and Patient-Reported Voice Measures in Adults With Voice Disorders. *Journal of Voice*, *37*. https://doi.org/10.1016/j.jvoice.2021.02.007

Gelfer, M. P., & Young, S. R. (1997). Comparisons of intensity measures and their stability in male and female sneakers. *Journal of Voice*, *11*(2), 178-186. https://doi.org/10.1016/S0892-1997(97)80076-8

Gelfer, M. P., & Pazera, J. F. (2006). Maximum Duration of Sustained /s/ and /z/ and the s/z Ratio With Controlled Intensity. *Journal of Voice*, *20*(3), 369–379. https://doi.org/10.1016/j.jvoice.2005.03.011

Gerratt, B. R., Kreiman, J., Antonanzas-Barroso, N., & Berke, G. S. (1993). Comparing internal and external standards in voice quality judgments. *Journal of Speech and Hearing Research*, *36*(1), 14-20. https://doi.org/10.1044/jshr.3601.14

Ghio, A., Dufour, S., Rouaze, M., Bokanowski, V., Pouchoulin, G., Revis, J., & Giovanni, A. (2011). Mise au point et évaluation d'un protocole d'apprentissage de jugement perceptif de la sévérité de dysphonies sur de la parole naturelle. *Revue de laryngologie - otologie - rhinologie*.

Ghio, A. (2012) Bilan instrumental de la dysphonie. Dans Garrel, R., Amy de La Bretèque, B., Brun, V., Entretiens de médecine physique et de réadaptation, & Entretiens de rééducation et réadaptation fonctionnelles. (2012). *La voix parlée et la voix chantée*. Sauramps médical.

Ghio, A., Gasquet-Cyrus, M., Roquel, J., Giovanni, A. (2013, août) *Perceptual interference between regional accent and voice/speech disorders.* Interspeech 2013 14th Annual Conference of the International Speech Communication Association, Lyon, France. https://hal.science/hal-01510195

Ghio, A., Cantarella, G., Weisz, F., Robert, D., Woisard, V., Fussi, F., Giovanni, A., & Baracca, G. (2015). Is the perception of dysphonia severity language-dependent? A comparison of French and Italian voice assessments. *Logopedics, Phoniatrics, Vocology*, *40*(1), 36-43. https://doi.org/10.3109/14015439.2013.837503

Ghio, A. (2021). L'évaluation perceptive de la qualité de la voix. Dans Giovanni, A., Bogdanski, E., Bonnaud Tondat, M., & Bouvier, C. *La voix : Anatomie, physiologie et explorations*. (pp. 95-96) De Boeck supérieur.

Ghio, A., Revis, J., Smithson-Barrière, D., Jarzé, S., & Giovanni, A. (2021). Reliability and Correlations Between Overall Severity, Roughness and Breathiness in the Perception of Dysphonic Voices : Investigating Cognitive Aspects. *Journal of Voice: Official Journal of the Voice Foundation*, *38*(1), 136-143. https://doi.org/10.1016/j.jvoice.2021.07.010

Gilman, M. (2021). Revisiting Sustained Phonation Time of /s/, /z/, and /ɑ/. *Journal of Voice*, *35*(6), 935.e13-935.e18. https://doi.org/10.1016/j.jvoice.2020.03.012

Giovanni, A., de Saint-Victor, S. (2013). Bilan clinique de la voix. *Oto-rhino-laryngologie*. http://dx.doi.org/10.1016/S0246-0351(13)51097-X

Giovanni, A. (2021). L'évaluation perceptive de la qualité de la voix. Dans Giovanni, A., Bogdanski, E., Bonnaud Tondat, M., & Bouvier, C. *La voix : Anatomie, physiologie et explorations*. (pp. 95-96) De Boeck supérieur.

Giovanni A, Mattei A. (2021) Physiologie des plis vocaux. EMC - Oto-rhino-laryngologie 2021;36(3):1-15 [Article 20-750-A-10]. https://dx.doi.org/10.1016/S0246-0351(20)90212-X

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative Voice Data for Younger and Older Adults. *Journal of Voice*, *27*(5), 545-555. https://doi.org/10.1016/j.jvoice.2013.03.002

Grant, S., Aitchison, T., Henderson, E., Christie, J., Zare, S., McMurray, J., & Dargie, H. (1999). A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest*, *116*(5), 1208–1217. https://doi.org/10.1378/chest.116.5.1208

Guyatt, G. H., Osoba, D., Wu, A. W., Wyrwich, K. W., Norman, G. R., & Clinical Significance Consensus Meeting Group. (2002). Methods to explain the clinical significance of health status measures. *Mayo Clinic Proceedings*, *77*(4), 371-383. https://doi.org/10.4065/77.4.371

Hammarberg, B. (2000). *Voice research and clinical needs*. Folia Phoniatrica et Logopaedica, 52, 93-102.

Hasson, D., & Arnetz, B. (2005). Validation and Findings Comparing VAS vs. Likert Scales for Psychosocial Measurements. *International Electronic Journal of Health Education*, *8*.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1990). *Phonatory function associated with hyperfunctionally related vocal fold lesions. Journal of Voice, 4(1), 52–63.* doi:10.1016/s0892-1997(05)80082-7

Hirano, M., Koike, Y., & Von Leden, H. (1968). Maximum Phonation Time and Air Usage During Phonation. *Folia Phoniatrica et Logopaedica*, *20*(4), 185-201. https://doi.org/10.1159/000263198

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica*, *26*(2), 89-94. https://doi.org/10.1159/000263771

Hirano M. (1981) Psycho-acoustic evaluation of voice : GRBAS Scale for evaluating the hoarse voice. *Clinical examination of voice.* Wien : Springer Verlag

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice: Official Journal of the Voice Foundation*, *13*(4), 557-569. https://doi.org/10.1016/s0892-1997(99)80010-1

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*(2), 511-529. https://doi.org/10.1121/1.396829

Holmberg, E. B., Doyle, P., Perkell, J. S., Hammarberg, B., & Hillman, R. E. (2003). Aerodynamic and acoustic voice measurements of patients with vocal nodules : Variation in baseline and changes across voice therapy. *Journal of Voice: Official Journal of the Voice Foundation*, *17*(3), 269-282. https://doi.org/10.1067/s0892-1997(03)00076-6

Hsiung, M.-W., Pai, L., & Wang, H.-W. (2002). Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino-Laryngology*, *259*(2), 97-99. https://doi.org/10.1007/s004050100405

Huston, M. N., Puka, I., & Naunheim, M. R. (2023). Prevalence of Voice Disorders in the United States : A National Survey. *The Laryngoscope*. https://doi.org/10.1002/lary.30929

Iwata, S., & Von Leden, H. (1970). Pitch Perturbations in Normal and Pathologic Voices. *Folia Phoniatrica et Logopaedica*, *22*(6), 413-424. https://doi.org/10.1159/000263421

Järvinen, K., Laukkanen, A.-M., & Aaltonen, O. (2013). Speaking a foreign language and its effect on F0. *Logopedics Phoniatrics Vocology*, *38*(2), 47-51. https://doi.org/10.3109/14015439.2012.687764

Johnson, A., Jacobson, B., Grywalski, C., Silbergleit, A., Jacobson, G., & Benninger, M. (1997). The Voice Handicap Index (VHI) : Development and Validation. *American Journal of Speech-Language Pathology*, *6*, 66-70.

Joshi, A., Baheti, I., & Angadi, V. (2020). Cultural and Linguistic Adaptation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Into Hindi. *Journal of Speech, Language, and Hearing Research: JSLHR*, *63*(12), 3974-3981. https://doi.org/10.1044/2020_JSLHR-20-00348

Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice: Official Journal of the Voice Foundation*, *21*(5), 576-590. https://doi.org/10.1016/j.jvoice.2006.05.001

Kempster, G. B., Kistler, D. J., & Hillenbrand, J. (1991). Multidimensional Scaling Analysis of Dysphonia in Two Speaker Groups. *Journal of Speech, Language, and Hearing Research*, *34*(3), 534-543. https://doi.org/10.1044/jshr.3403.534

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice : Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, *18*(2), 124-132. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kim, T. K. (2017). Practical statistics in pain research. *The Korean Journal of Pain*, *30*(4), 243–249. https://doi.org/10.3344/kjp.2017.30.4.243

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual Evaluation of Voice Quality : Review, Tutorial, and a Framework for Future Research. *Journal of Speech, Language, and Hearing Research*, *36*(1), 21-40. https://doi.org/10.1044/jshr.3601.21

Krom, G. de. (1995). Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments. *Journal of Speech, Language, and Hearing Research*, *38*(4), 794-811. https://doi.org/10.1044/jshr.3804.794

Lafon, J.-C., & Cornut, G. (1962). Phonation et fonctionnement impulsionnel du larynx. *Annales des Télécommunications*, *17*(5), 107-116. https://doi.org/10.1007/BF03014951

Lamesch, S. (2010). *Mécanismes laryngés et voyelles en voix chantée. Dynamique vocale, phonétogrammes de paramètres glottiques et spectraux, transitions de mécanismes.* [These de doctorat, Université Pierre et Marie Curie - Paris VI]. https://theses.hal.science/tel-00488701

Le Huche, F., & Allali, A. (2010a). *La voix. Tome 1. Anatomie et physiologie des organes de la voix et de la parole* (4e édition). Elsevier Masson.

Le Huche, F., & Allali, A. (2010b). *La voix. Tome 2. Pathologies vocales d'origine fonctionnelles (3e édition)*. Elsevier Masson.

Leong, K., Hawkshaw, M. J., Dentchev, D., Gupta, R., Lurie, D., & Sataloff, R. T. (2013). Reliability of Objective Voice Measures of Normal Speaking Voices. *Journal of Voice*, *27*(2), 170-176. https://doi.org/10.1016/j.jvoice.2012.07.005

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, 35, 344-353.

Lyberg-Åhlander, V., Rydell, R., Fredlund, P., Magnusson, C., & Wilén, S. (2019). Prevalence of Voice Disorders in the General Population, Based on the Stockholm Public Health Cohort. *Journal of Voice: Official Journal of the Voice Foundation*, *33*(6), 900-905. https://doi.org/10.1016/j.jvoice.2018.07.007

Ma, E. P.-M., & Yiu, E. M.-L. (2006). Multiparametric Evaluation of Dysphonic Severity. *Journal of Voice*, *20*(3), 380-390. https://doi.org/10.1016/j.jvoice.2005.04.007

Ma, E., Robertson, J., Radford, C., Vagne, S., El-Halabi, R., & Yiu, E. (2007). Reliability of speaking and maximum voice range measures in screening for dysphonia. *Journal of Voice: Official Journal of the Voice Foundation*, *21*(4), 397-406. https://doi.org/10.1016/j.jvoice.2006.03.004

Mahalingam, S., Venkatraman, Y., & Boominathan, P. (2024). Cross-Cultural Adaptation and Validation of Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) : A Systematic Review. *Journal of Voice*, *38*(3), 630-640. https://doi.org/10.1016/j.jvoice.2021.10.022

Mansouri, Y., Naderifar, E., Hajiyakhchali, A., & Moradi, N. (2023). The Relationship Between Dysphonia Severity Index and Voice-Related Quality of Life in the Elementary School Teachers with Voice Complaint. *Journal of Voice*, *37*(3), 466.e35-466.e39. https://doi.org/10.1016/j.jvoice.2021.02.017

Marchese, M. R., Longobardi, Y., Di Cesare, T., Mari, G., Terruso, V., Galli, J., & D'Alatri, L. (2022). Gender-related differences in the prevalence of voice disorders and awareness of dysphonia. *Acta otorhinolaryngologica Italica : organo ufficiale della Societa italiana di*

otorinolaringologia e chirurgia cervico-facciale, *42*(5), 458–464. https://doi.org/10.14639/0392-100X-N2018

Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality : A meta-analysis). *The Journal of the Acoustical Society of America*, *126*(5), 2619-2634. https://doi.org/10.1121/1.3224706

Marzi, G., Balzano, M., & Marchiori, D. (2024). K-Alpha Calculator–Krippendorff's Alpha Calculator : A user-friendly tool for computing Krippendorff's Alpha inter-rater reliability coefficient. *MethodsX*, *12*, 102545. https://doi.org/10.1016/j.mex.2023.102545

Mateos-Serrano, B., García-López, I., Toledano, A., & Gavilán, J. (2021). Subjective voice analysis in patients with muscular tension dysphonia : Comparison between clinician and patient evaluation. *The Journal of Laryngology and Otology*, *135*(5). https://doi.org/10.1017/S0022215121001067

Maunsell, R., Lombas de Resende, R., Morioka, A., Pazinatto, D. B., & Constantini, A. C. (2022). Voice outcomes and its impact on quality of life after open airway surgery in children. *International Journal of Pediatric Otorhinolaryngology*, *152*, 110980. https://doi.org/10.1016/j.ijporl.2021.110980

McAdams, S., & Bregman, A. (1979). Hearing musical streams. *Computer Music Journal*, *3*.

Mehta, D., & Hillman, R. E. (2007). Use of Aerodynamic Measures in Clinical Voice Assessment. *Perspectives on Voice and Voice Disorders*, *17*(3), 14-18. https://doi.org/10.1044/vvd17.3.14

Menin-Sicard, A., & Sicard, É. (2016). *Évaluation et réhabilitation de la voix : Approche clinique et objective*. De Boeck Supérieur. 21.

Miller, D. G., & Schutte, H. K. (1993). Physical definition of the "flageolet register". *Journal of Voice*, *7*(3), 206-212. https://doi.org/10.1016/S0892-1997(05)80328-5

Millet, B., & Dejonckere, P. H. (1998). What determines the differences in perceptual rating of dysphonia between experienced raters? *Folia Phoniatrica et Logopaedica: Official Organ of*

the International Association of Logopedics and Phoniatrics (IALP), *50*(6), 305-310. https://doi.org/10.1159/000021472

Mohd Mossadeq, N., Mohd Khairuddin, K. A., & Zakaria, M. N. (2022). Cross-cultural Adaptation of the Consensus Auditory-perceptual Evaluation of Voice (CAPE-V) Into Malay : A Validity Study. *Journal of Voice: Official Journal of the Voice Foundation*, S0892-1997(22)00151-5. https://doi.org/10.1016/j.jvoice.2022.05.018

Morsomme, D. (2001). *Contribution à la détermination de paramètres subjectifs et objectifs pour l'étude de la voix* [These de doctorat, UCL - Université Catholique de Louvain]. ORBi-University of Liège. https://orbi.uliege.be/handle/2268/26003

Mueller, P. B. (1997). The Aging Voice. *Seminars in Speech and Language*, *18*(2), 159-169. https://doi.org/10.1055/s-2008-1064070

Nagle, K. F. (2022). Clinical Use of the CAPE-V Scales : Agreement, Reliability and Notes on Voice Quality. *Journal of Voice: Official Journal of the Voice Foundation*, S0892-1997(22)00366-6. https://doi.org/10.1016/j.jvoice.2022.11.014

Nemr, K., Simões-Zenari, M., Cordeiro, G. F., Tsuji, D., Ogawa, A. I., Ubrig, M. T., & Menezes, M. H. M. (2012). GRBAS and Cape-V Scales : High Reliability and Consensus When Applied at Different Times. *Journal of Voice,* *26*(6), 812.e17-812.e22. https://doi.org/10.1016/j.jvoice.2012.03.005

Núñez-Batalla, F., Morato-Galán, M., García-López, I., & Ávila-Menéndez, A. (2015). Validation of the Spanish adaptation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *Acta Otorrinolaringologica Espanola*, *66*(5), 249-257. https://doi.org/10.1016/j.otorri.2014.07.007

Oates, J. (2009). Auditory-perceptual evaluation of disordered voice quality : Pros, cons and future directions. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP),* *61*(1), 49-56. https://doi.org/10.1159/000200768

Omori, K. (2011). *Diagnosis of Voice Disorders*. 54(4), 248-253.

Park, Y., Anand, S., Kopf, L. M., Shrivastav, R., & Eddins, D. A. (2022). Interactions Between Breathy and Rough Voice Qualities and Their Contributions to Overall Dysphonia Severity. *Journal of Speech, Language, and Hearing Research*, *65*(11), 4071-4084. https://doi.org/10.1044/2022_JSLHR-22-00012

Paya Herrero, Begona (2009): Voice and Identity: A contrastive study of identity perception in voice. Dissertation, LMU München: Faculty for Languages and Literatures

Payten, C. L., Chiapello, G., Weir, K. A., & Madill, C. J. (2022). Frameworks, Terminology and Definitions Used for the Classification of Voice Disorders : A Scoping Review. *Journal of Voice*, S089219972200039X. https://doi.org/10.1016/j.jvoice.2022.02.009

Piron, A. (1998) OSTEOVOX, formation en thérapies manuelles et réhabilitation sensori motrice dédiée à la sphère oro-faciale. https://osteovox.be/

Pommée, T., Maryn, Y., Finck, C., & Morsomme, D. (2020). Validation of the Acoustic Voice Quality Index, Version 03.01, in French. *Journal of Voice: Official Journal of the Voice Foundation*, *34*(4), 646.e11-646.e26. https://doi.org/10.1016/j.jvoice.2018.12.008

Pommée, T., Mbagira, D., & Morsomme, D. (2024). French-Language Adaptation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *Journal of Voice: Official Journal of the Voice Foundation*, S0892-1997(24)00085-7. https://doi.org/10.1016/j.jvoice.2024.03.011

Reghunathan, S., & Bryson, P. C. (2019). Components of Voice Evaluation. *Otolaryngologic Clinics of North America*, *52*(4), 589-595. https://doi.org/10.1016/j.otc.2019.03.002

Révis, J. (2021). L'évaluation perceptive de la qualité de la voix. Dans Giovanni, A., Bogdanski, E., Bonnaud Tondat, M., & Bouvier, C. *La voix : Anatomie, physiologie et explorations*. (pp. 95-96) De Boeck supérieur.

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, *114*(9), 1549-1556. https://doi.org/10.1097/00005537-200409000-00009

Roubeau, B., Henrich, N., & Castellengo, M. (2009). Laryngeal Vibratory Mechanisms : The Notion of Vocal Register Revisited. *Journal of Voice*, *23*(4), 425-438. https://doi.org/10.1016/j.jvoice.2007.10.014

Sabir, B., Touri, B., & Moussetad, M. (2017). Correlation between acoustic measures, voice handicap index and GRBAS scales scores among Moroccan students. *Current Pediatric Research*, *21*, 343-353.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients : Appropriate Use and Interpretation. *Anesthesia & Analgesia*, *126*(5), 1763. https://doi.org/10.1213/ANE.0000000000002864

Scott, S., Robinson, K., Wilson, J. A., & K. MACKENZIE. (1997). Patient-reported problems associated with dysphonia. *Clinical Otolaryngology and Allied Sciences*, *22*(1), 37-40. https://doi.org/10.1046/j.1365-2273.1997.00855.x

Sheyona, V., & Devadas, U. (2022). The Prevalence and Impact of Voice Problems in Nonprofessional Voice Users : Preliminary Findings. *Journal of Voice: Official Journal of the Voice Foundation*, *36*(3), 383-388. https://doi.org/10.1016/j.jvoice.2020.06.010

Sicard, E., Menin-Sicard, A., & Ruffle, L. (2021, septembre). VOCALAB : UN LOGICIEL CONVIVIAL POUR L'ANALYSE ET LA PRISE EN CHARGE ORTHOPHONIQUE DE LA VOIX. *Choice for Voice: Crossing boundaries in voice*. https://hal.science/hal-03670644

Sicard, E., & Menin-Sicard, A. (2023). *Fréquence fondamentale de la parole—Repères pour la prise en soin orthophonique*. https://hal.science/hal-04073008

Skuk, V. G., & Schweinberger, S. R. (2014). Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech, Language, and Hearing Research: JSLHR*, *57*(1), 285-296. https://doi.org/10.1044/1092-4388(2013/12-0314)

Smith E, Verdolini K, Gray S, Nichols S, Lemke J, Barkmeier J, Dove H, Hoffman H. (1994). Effect of voice disorders on quality of life. *NCVS Status Progress Report, 7*, 1–17.

Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology,* 9, 1112-1122. https://doi.org/10.1016/j.protcy.2013.12.124

The jamovi project (2024). jamovi (Version 2.5) [Computer Software]. Retrieved from https://www.jamovi.org

Titze, I. R., Horii, Y., & Scherer, R. C. (1987). Some technical considerations in voice perturbation measurements. *Journal of Speech and Hearing Research*, *30*(2), 252-260. https://doi.org/10.1044/jshr.3002.252

Titze, I. R. (1994). *Principles of Voice Production*. Prentice Hall.

Titze, I. R., Lemke, J., & Montequin, D. (1997). Populations in the U.S. workforce who rely on voice as a primary tool of trade : A preliminary report. *Journal of Voice*, *11*(3), 254-259. https://doi.org/10.1016/S0892-1997(97)80002-1

Titze, I. R. (2001). Acoustic interpretation of resonant voice. *Journal of Voice: Official Journal of the Voice Foundation*, *15*(4), 519-528. https://doi.org/10.1016/S0892-1997(01)00052-2

Traunmüller, H., & Eriksson, A. (1995). *The frequency range of the voice fundamental in the speech of male and female adults*. *2*.

Wagner, A., & Braun, A. (2003, 3 août). *Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages*. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelone. *ICPhS-15*, 651-654.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics, 23*(3), 349-366. https://doi.org/10.1016/S0095-4470(95)80165-0

Wheeler, K. M., Collins, S. P., & Sapienza, C. M. (2006). The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *Journal of Voice: Official Journal of the Voice Foundation*, *20*(2), 308-317. https://doi.org/10.1016/j.jvoice.2005.03.006

Wilson, D. K. (1987). *Voice Problems of Children.* Williams & Wilkins.

Woisard, V., Bodin, S., Yardeni, E., & Puech, M. (2007). The voice handicap index : Correlation between subjective patient response and quantitative assessment of voice. *Journal of Voice: Official Journal of the Voice Foundation*, *21*(5), 623-631. https://doi.org/10.1016/j.jvoice.2006.04.005

Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Raes, J., & Van de Heyning, P. H. (2000). The dysphonia severity index : An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research: JSLHR*, *43*(3), 796-809. https://doi.org/10.1044/jslhr.4303.796

Yamaguchi, H., Shrivastav, R., Andrews, M. L., & Niimi, S. (2003). A Comparison of Voice Quality Ratings Made by Japanese and American Listeners Using the GRBAS Scale. *Folia Phoniatrica et Logopaedica*, *55*(3), 147-157. https://doi.org/10.1159/000070726

Yiu, E. M.-L. (1999). Limitations of perturbation measures in clinical acoustic voice analysis. *Asia Pacific Journal of Speech, Language and Hearing*, *4*(3), 155-166. https://doi.org/10.1179/136132899807557475

Yiu, E. M.-L., Murdoch, B., Hird, K., Lau, P., & Ho, E. M. (2008). Cultural and Language Differences in Voice Quality Perception : A Preliminary Investigation Using Synthesized Signals. *Folia Phoniatrica et Logopaedica*, *60*(3), 107-119. https://doi.org/10.1159/000119746

Yumoto, E., Sasaki, Y., & Okamura, H. (1984). Harmonics-to-Noise Ratio and Psychophysical Measurement of the Degree of Hoarseness. *Journal of Speech Language and Hearing Research*, 27(1), 2. https://doi.org/10.1044/jshr.2701.02

Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., Thrush, C. R., & Glaze, L. E. (2011). Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *American Journal of Speech-Language Pathology*, *20*(1), 14-22. https://doi.org/10.1044/1058-0360(2010/09-0105)

ANNEXES

Annexe I : CAPE-V Procedures and Forms (Kempster et al., 2009)

Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)
ASHA Special Interest Group 3, Voice and Voice Disorders

PURPOSE AND APPLICATIONS

The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) was developed as a tool for clinical auditory-perceptual assessment of voice. Its primary purpose is to describe the severity of auditory-perceptual attributes of a voice problem, in a way that can be communicated among clinicians. Its secondary purpose is to contribute to hypotheses regarding the anatomic and physiological bases of voice problems and to evaluate the need for additional testing.

CAPE-V is *not* intended for use as the only means of determining the nature of the voice disorder. It is not to be used to the exclusion of other tests of vocal function. Finally, it is not expected to demonstrate a 1:1 relation to results from other tests of vocal function.

ORIGIN

The CAPE-V was developed from a consensus meeting sponsored by the American Speech-Language-Hearing Association's (ASHA) Group 3: Voice and Voice Disorders, and the Department of Communication Science and Disorders, University of Pittsburgh, held in Pittsburgh on June 10-11, 2002. Attending this meeting were speech-language pathologists (SLPs) who specialize in voice disorders and invited experts in human perception (see appendix). The participants' charge was to develop standardized guidelines for auditory-perceptual evaluation of voice, based on theory and data in psychoacoustics, psychometric scaling, and voice perception. Clinical practicality and brevity of administration were also considered in developing these guidelines.

A working group was charged to formalize a consensus statement about minimal recommended standards for optimizing auditory-perceptual judgments in the clinical assessment of voice disorders by speech-language pathologists. The CAPE-V is the initial product. The hope is that widespread use of the current CAPE-V and its future development will encourage a more consistent approach and ultimately more research in the perceptual evaluation of voice disorders. The present document is the preliminary result of the consensus meeting. The ultimate goal is standardization of a reliable tool for clinical voice quality measurement.

DESIGN CONSIDERATIONS

The consensus was that the clinical evaluation of auditory-perceptual characteristics of voice should be derived from a tool with the following attributes: (a) perceptual dimensions should reflect a minimal set of clinically meaningful, perceptual voice parameters, identified by a group of expert clinicians; (b) procedures and results should be obtainable expediently; (c) procedures and results should be applicable to a broad range of vocal pathologies and clinical settings; (d) ratings ultimately should be demonstrated to optimize reliability within and across clinicians , and (e) ultimately, exemplars should be available for training.

DESCRIPTION AND INSTRUCTIONS

General Description of the Tool: The CAPE-V indicates salient perceptual vocal attributes, identified by the core consensus group as commonly used and easily understood. The attributes are: (a) Overall Severity; (b) Roughness; (c) Breathiness; (d) Strain; (e) Pitch; and (f) Loudness. The CAPE-V displays each attribute accompanied by a 100-millimeter line forming a visual analog scale (VAS). The clinician indicates the degree of perceived deviance from normal for each parameter on this scale, using a tic mark. For each dimension, scalar extremes are unlabeled. Judgments may be assisted by referring to general regions indicated below each scale on the CAPE-V: "MI" refers to "mildly deviant," "MO" refers to "moderately deviant," and "SE" refers to "severely deviant." A key issue is that the regions indicate *gradations* in severity, rather than discrete points. The clinician may place tick marks at any location along the line." Ratings are based on the clinician's direct observations of the patient's performance during the evaluation, rather than patient report or other sources.

To the right of each scale are two letters, "C" and "I." "C" represents "consistent" and "I" represents "intermittent" presence of a particular voice attribute. The rater circles the letter that best describes the consistency of the judged

parameter. A judgment of "consistent" indicates that the attribute was continuously present throughout the tasks. A judgment of "intermittent" indicates that the attribute occurred inconsistently within *or* across tasks. For example, an individual may consistently exhibit a strained voice quality across all tasks, which include sustained vowels and speech. In this case, the rater would circle "C" to the right of the strain scale. In contrast, another individual might exhibit consistent strain during vowel production, but intermittent strain during one or more connected speech task. In this case, the rater would circle "I" to the right of the strain scale.

Definitions of Vocal Attributes:
OVERALL SEVERITY: Global, integrated impression of voice deviance.
Roughness: Perceived irregularity in the voicing source.
Breathiness: Audible air escape in the voice.
Strain: Perception of excessive vocal effort (hyperfunction).
Pitch: Perceptual correlate of fundamental frequency. This scale rates whether the individual's pitch deviates from normal for that person's gender, age, and referent culture. The direction of deviance (high or low) should be indicated in the blank provided above the scale.
Loudness: Perceptual correlate of sound intensity. This scale indicates whether the individual's loudness deviates from normal for that person's gender, age, and referent culture. The direction of deviance (soft or loud) should be indicated in the blank provided above the scale.

Blank scales and additional features: The six standard vocal attributes included on the CAPE-V are considered the minimal set of parameters for describing the auditory-perceptual characteristics of disordered voices. The form also includes two unlabeled scales. The clinician may use these to rate additional prominent attributes required to describe a given voice. The clinician may indicate the presence of other attributes or "positive signs" not noted elsewhere under "Additional features." If an individual is aphonic, this should be noted under "additional features" and no additional marks should be made on the scales.
Data collection: The individual should be seated comfortably in a quiet environment. The clinician should audio record the individual's performance on three tasks: vowels, sentences, and conversational speech. Standard recording procedures should be used that incorporate a condenser microphone placed 45 degrees off from the front of the mouth and a 4 cm mike-to-mouth distance. Audio recordings are recommended to be made onto a computer with 16 bits of resolution and a signal-sampling rate of no less than 20 Khertz.

Task 1: Sustained vowels: Two vowels were selected for this task. One is considered a lax vowel (/a/) and the other tense (/i/). In addition, the vowel, /i/, is the sustained vowel used during videostroboscopy. Thus, the use of this vowel during this task offers an auditory comparison to that produced during a stroboscopic exam.

The clinician should say to the individual, "The first task is to say the sound, /a/. Hold it as steady as you can, in your typical voice, until I ask you to stop." (The clinician may provide a model of this task, if necessary) The individual performs this task three times for 3-5 sec each. "Next, say the sound, /i/. Hold it as steady as you can, in your typical voice, until I ask you to stop." The individual performs this task three times for 3-5 sec each.

Task 2: Sentences: Six sentences were designed to elicit various laryngeal behaviors and clinical signs. The first sentence provides production of every vowel sound in the English language, the second sentence emphasizes easy onset with the /h/, the third sentence is all voiced, the fourth sentence elicits hard glottal attack, the fifth sentence incorporates nasal sounds, and the final sentence is weighted with voiceless plosive sounds.

The clinician should give the person being evaluated flash cards, which progressively show the target sentences (see below) one at a time. The clinician says, "Please read the following sentences one at a time, as if you were speaking to somebody in a real conversation." (Individual performs task, producing one exemplar of each sentence.) If the individual has difficulty reading, the clinician may ask him or her to repeat sentences after verbal examples. This should be noted on the CAPE-V form. The sentences are: (a) The blue spot is on the key again; (b) How hard did he hit him? (c) We were away a year ago; (d) We eat eggs every Easter; (e) My mama makes lemon jam; (f) Peter will keep at the peak.

Task 3: Running speech: The clinician should elicit at least 20 seconds of natural conversational speech using standard interview questions such as, "Tell me about your voice problem." or "Tell me how your voice is functioning." "

Data scoring: The clinician should have the individual perform all voice tasks—including vowel prolongation, sentence production, and running speech, before completing the CAPE-V form. If performance is uniform across all tasks, the clinician should mark the ratings indicating overall performance for each scale. If the clinician notes a discrepancy in performance across tasks, he or she should rate performance on each task separately, *on a given line.* Only one CAPE-V form is used per individual being evaluated. In the case of discrepancies across tasks, tick marks should be labeled with the task number. Tick marks reflecting vowel prolongation should be labeled #1 (see form). Tick marks reflecting running speech should be labeled #2. Tick marks reflecting story retelling should be labeled #3. In the rare event that the clinician perceives discrepancies within task type (for example, /a/ versus /i/), he or she may further label the ratings accordingly [for example, 1/a/ versus 1/i/ to reflect the different vowels, or 2(a)-(b)-(c)-(d)-(e)- or (f) for the different sentences]. Unlabeled tick marks indicate uniform performance. See examples below. [Note: Using labels to indicate discrepancies/variation across tasks in the severity of an attribute is different than indicating that an attribute is displayed intermittently (I). If an attribute is judged to have equal severity whenever it appears, but it is not present all the time, "I" should be circled to indicate that the attribute is intermittent and no additional labeling needs to be done.]

Scoring: After the clinician has completed all ratings, he or she should measure ratings from each scale. To do so, he or she should physically measure the distance in mm from the left end of the scale. The mm score should be written in the blank space to the far right of the scale, thereby relating the results in a proportion to the total 100 mm length of the line. The results can be reported in two possible ways. First, results can indicate distance in mm to describe the degree of deviancy, for example "73/100" on "strain." Second, results can be reported using descriptive labels that are typically employed clinically to indicate the general amount of deviancy, for example "moderate-to-severe" on "strain." We strongly suggest using both forms of reporting.

It is strongly recommended that for all rating sessions following the initial one, the clinician have a paper or electronic copy of the previous CAPE-V ratings available for comparison purposes. He or she should also rate subsequent examinations based on direct comparisons between earlier and current audio recordings. Such an approach should optimize the internal consistency/reliability of repeated sequential ratings within a patient, particularly for purposes of assessing treatment outcomes. Although difficult, clinicians are encouraged to make every effort to minimize bias in all ratings. We acknowledge that this solution is imperfect.

Other procedures: The clinician can indicate prominent observations about resonance phenomena under "Comments about resonance." Examples include, but are not limited to hyper- or hyponasality, and cul-de-sac resonance.

Cautions: Data available on the reliability of all rating scales for voice assessment indicate that both intra- and inter-judge agreement varies widely. Although we have attempted to limit sources of variability in the present tool, its reliability and validity have not yet been assessed. Special Interest Group 3 is in the process of field-testing the validity of the CAPE-V in a multicenter clinical trial. Details will be reported as data are analyzed and interpreted. Future editions are projected to include referent voice recordings as "anchors" as well as training modules.

© 2009 American Speech-Language-Hearing Association

Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

Name:_____ **Date:**_____

The following parameters of voice quality will be rated upon completion of the following tasks:

1. Sustained vowels, /a/ and /i/ for 3-5 seconds duration each.
2. Sentence production:
   a. The blue spot is on the key again.
   b. How hard did he hit him?
   c. We were away a year ago.
   d. We eat eggs every Easter.
   e. My mama makes lemon muffins.
   f. Peter will keep at the peak.
3. Spontaneous speech in response to: "Tell me about your voice problem." or "Tell me how your voice is functioning."

Legend: C = Consistent I = Intermittent
MI = Mildly Deviant
MO =Moderately Deviant
SE = Severely Deviant

			SCORE
Overall Severity	_____ MI MO SE	C I	/100
Roughness	_____ MI MO SE	C I	/100
Breathiness	_____ MI MO SE	C I	/100
Strain	_____ MI MO SE	C I	/100
Pitch	(Indicate the nature of the abnormality): _____ _____ MI MO SE	C I	/100
Loudness	(Indicate the nature of the abnormality): _____ _____ MI MO SE	C I	/100
_____	_____ MI MO SE	C I	/100
_____	_____ MI MO SE	C I	/100

COMMENTS ABOUT RESONANCE: NORMAL OTHER (Provide description):_____

_____

ADDITIONAL FEATURES (for example, diplophonia, fry, falsetto, asthenia, aphonia, pitch instability, tremor, wet/gurgly, or other relevant terms):

Clinician:_____

© 2009 American Speech-Language-Hearing Association

Annexe II : CAPE-Vf de Pommée et al. (2024)

Consensus Auditory-Perceptual Evaluation of Voice - version Française (CAPE-V F)

Nom : _____ **Date :**_____

Les paramètres vocaux cités ci-dessous seront évalués à l'issue de la passation des tâches suivantes :

1. Tenue de la voyelle /a/ pendant 3 à 5 secondes.
2. Lecture à voix haute des phrases :

   a. Dites-nous les amis, où allez-vous samedi ?
   b. Son chat Fifi saute sur la souris.
   c. Elle va aux Iles de la Madeleine.
   d. Irma, où va Alain aujourd'hui ? À Ottawa.
   e. Lundi matin, maman nous a montré les nénuphars.
   f. Dobi comptait payer les deux pastèques.

3. Discours semi-spontané en réponse à la question : « Présentez-vous brièvement. » (Nom, prénom, région, activité principale).

Consigne de notation : Placer un point de repère sur la droite de 100mm pour déterminer le niveau d'atteinte.
- Extrémité gauche = absence d'altération
- Extrémité droite = altération maximale

Mesurer la distance (en millimètres) entre l'extrémité gauche de l'échelle et le point de repère pour obtenir un score sur 100.

Sévérité globale _____ C I ___/100

Raucité _____ C I ___/100

Souffle _____ C I ___/100

Serrage _____ C I ___/100

Hauteur tonale (Indiquer la nature de l'atteinte) : _____

_____ C I ___/100

Intensité (Indiquer la nature de l'atteinte) : _____

_____ C I ___/100

_____ _____ C I ___/100

_____ _____ C I __/100

COMMENTAIRES SUR LA RESONNANCE : NORMALE AUTRE (Décrivez):_____

AUTRES CARACTÉRISTIQUES (par exemple, bitonalité, fry/voix de friture, voix de fausset/de tête, asthénie, aphonie, instabilité de la hauteur tonale, tremblement, voix mouillée) :

Clinicien : _____

Annexe III : Protocole d'enregistrement

PROTOCOLE D'ENREGISTREMENT DANS LE CADRE DE L'ÉVALUATION PERCEPTIVE DE LA VOIX AVEC L'OUTIL CAPE-V

Veuillez prendre connaissance de ce document dans son intégralité avant de procéder aux enregistrements.

Afin de vous familiariser avec le processus d'enregistrement, ainsi que de nous permettre de vérifier la qualité des enregistrements, nous vous demandons de faire un **enregistrement d'essai avec votre propre voix** que vous nous transférerez via FileSender avant de procéder à l'enregistrement des patients.

Vous recevrez par voie postale :

- un formulaire de **VHI** pour chaque patient
- une **autorisation de réutilisation et de traitement de données à caractère personnel**, en 1 exemplaire par patient → il faudra bien veiller à le scanner avant que le patient reparte avec le document
- un **formulaire de consentement,** en 1 exemplaire par patient → également à scanner et transmettre
- une **notice d'information** pour chaque patient
- des **cartes flash** avec les phrases du protocole CAPE-V

La notice d'information sera à présenter à chaque patient. Il est important qu'il en prenne connaissance avant de donner son accord pour participer à l'étude. Le patient devra ensuite lire et les exemplaires de l'**autorisation de réutilisation et de traitement de données à caractère personnel** et du **formulaire de consentement**. Veuillez garder une copie scannée de chacun de ces deux documents afin de les transmettre au directeur de l'étude via FileSender. Le patient gardera sa copie signée.

Sur le formulaire VHI, **ne renseignez pas le nom** du participant, mais plutôt son **code**.
Pour chaque participant, un code est assigné avec le format suivant :
VosInitiales+SexeParticipant+GroupeParticipant+JJMMAA+COMPTEUR

- VosInitiales : première lettre de votre prénom et de votre nom de famille.
- SexeParticipant : M/F/MtoF[1]/FtoM[2]
- GroupeParticipant : dysphonique = D / normophonique = N
- JJMMAA : date d'enregistrement (jour, mois et deux derniers chiffres de l'année)
- COMPTEUR = premier participant du jour = 1, second = 2, etc.

Par exemple, si Timothy Pommée, en tant qu'orthophoniste, effectue l'enregistrement d'un homme dysphonique le 29 septembre 2023, et que ce patient est le troisième à être enregistré ce jour-là, le code sera : TPMD2909233

Prenez note, dans un **document que vous garderez à part**, du **nom** du participant, du **code** attribué, de son **âge** et de la **cause de sa dysphonie**. Ce document rassemblant les

[1] **MtoF** : patiente transgenre de sexe biologique masculin effectuant une transition "male to female"
[2] **FtoM** : patient transgenre de sexe biologique féminin effectuant une transition "female to male"

noms et codes de tous les participants enregistrés peut être informatisé, ou sur papier. A la suite des enregistrements, il sera transmis au directeur de l'étude via FileSender.

Avant de commencer l'enregistrement, le patient devra remplir le formulaire VHI. Veuillez nous le transmettre via FileSender.

Conditions d'enregistrement :
Pour éviter un effet de fatigue ou d'échauffement vocal, l'enregistrement doit être effectué **en début de séance**. L'individu doit être assis confortablement dans un **environnement calme**.
Le clinicien devra procéder à un enregistrement audio de la performance de l'individu sur trois tâches : la voyelle, les phrases et le discours conversationnel. La procédure standard d'enregistrement doit comporter l'utilisation d'un micro-casque ou micro sur pied placé à 45 degrés et à une distance de **6 à 10 cm de la bouche**. Les enregistrements audio doivent être effectués en mode **mono**, avec une résolution d'au moins 16 bits (idéalement **32 bits**) et un taux d'échantillonnage du signal d'au moins 22 kHz (idéalement **44100Hz**). Veuillez vous référer aux instructions de réglages d'enregistrement en fin de ce document si nécessaire. En cas d'erreur nécessitant un ré-enregistrement, veillez à limiter le nombre de répétitions à **3 fois maximum**.

Tâche 1 : voyelle tenue
Le clinicien dit : "La première tâche consiste à dire le son /a/. Tenez ce /a/ aussi stable que possible, avec votre voix habituelle, jusqu'à ce que je vous demande d'arrêter. Nous allons le faire trois fois." **Le participant effectuera donc la tâche 3 fois**, pendant **3-5 secondes** à chaque fois.
La voyelle laxe /a/ est relativement neutre sur le plan articulatoire et produite avec un tractus vocal ouvert. Elle permet d'obtenir un aperçu de la phonation sans influence de la coarticulation opérant en parole continue.

Tâche 2 : lecture de phrases
Six phrases ont été conçues pour susciter divers comportements laryngés et signes cliniques.
Le clinicien donne au participant évalué des cartes "flash", montrant ainsi une seule phrase cible à la fois. Veillez bien à ce que le participant ait terminé de lire la phrase en entier avant de montrer la prochaine carte, afin que le bruit des cartes qui défilent ne perturbe pas l'enregistrement de la voix.
Consigne : Le clinicien dit : "Veuillez lire les phrases suivantes une à la fois, comme si vous parliez à quelqu'un dans une vraie conversation." (L'individu exécute la tâche, lisant chaque phrase, l'une après l'autre.) Si la personne a des difficultés à lire, le clinicien peut lui demander de répéter les phrases après lui. Cela devrait être noté.

Phrase 1 : "Dites-nous les amis, où allez-vous samedi ?"→ analyser la production des extrêmes du triangle vocalique (i, u, a)
Phrase 2 : "Son chat Fifi saute sur la souris." → analyser les transitions entre les phonèmes voisés et les phonèmes non voisés
Phrase 3 : "Elle va aux Îles de la Madeleine" → analyser la phonation continue, la présence/absence d'arrêts ou de spasmes et la capacité à maintenir le voisement

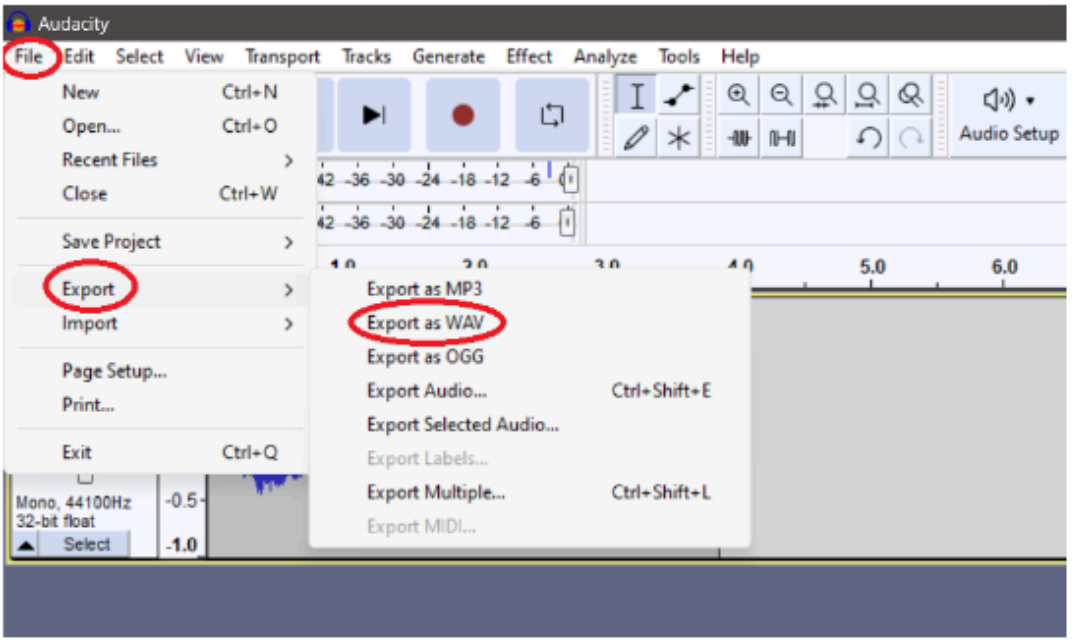
Phrase 4 : “Irma, où va Alain aujourd’hui ? À Ottawa.” → évaluer les attaques glottiques dures
Phrase 5 : “Lundi matin, maman nous a montré les nénuphars.” → évaluer l’hyponasalité
Phrase 6 : “Dobi comptait payer les deux pastèques.” → évaluer les fuites nasales et l’impact des phonèmes occlusifs non voisés sur la phonation

Tâche 3 : discours semi-spontané
Le clinicien doit susciter au moins 20 secondes de discours conversationnel en réponse à la question : “Présentez-vous brièvement en tant que JEAN/JEANNE DUPONT, en indiquant votre région d’origine et votre activité principale.” Le participant est libre d’inventer des éléments s’il préfère ne pas partager d’informations personnelles.

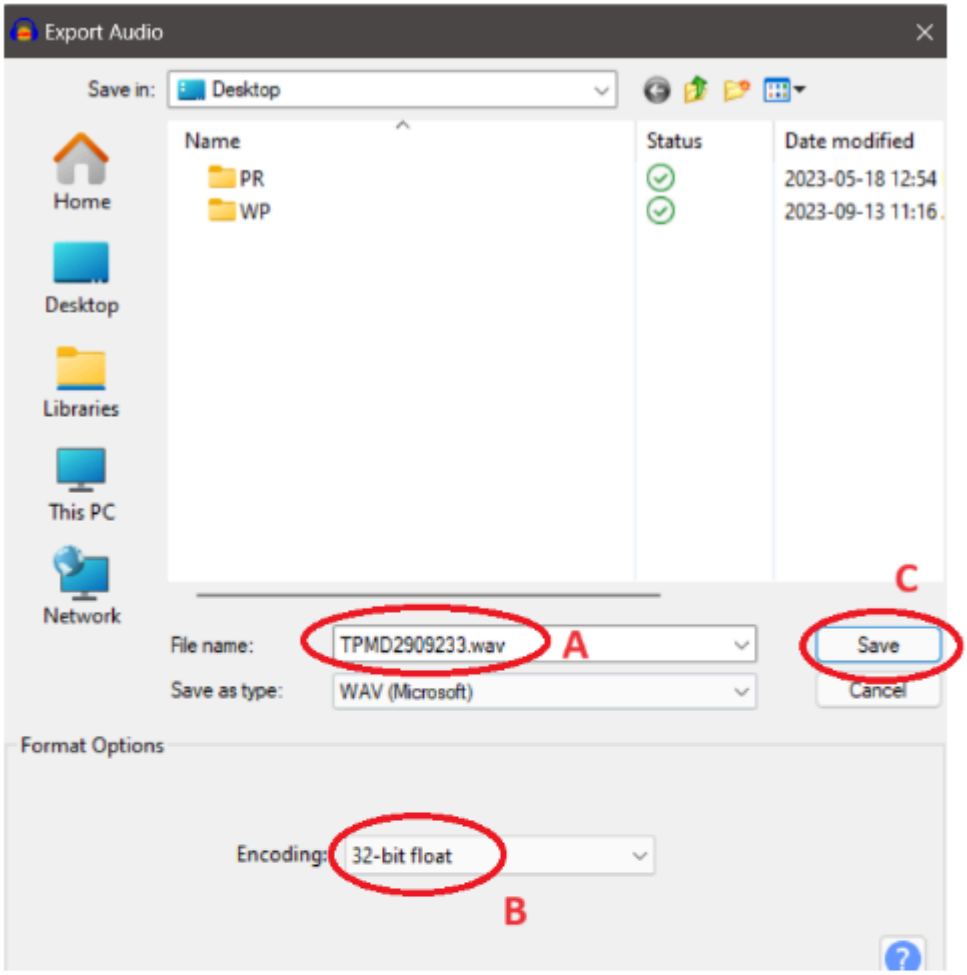
Envoi des données

Merci de sauvegarder ces enregistrements au format WAV dans un dossier de votre ordinateur. Renommez le fichier avec le **code du participant**, respectant le format décrit au début de ce document. **Notez également ce code sur le formulaire VHI du participant.**
Nous vous recontacterons pour collecter les enregistrements ainsi que les copies des formulaires VHI. Pour sauvegarder un enregistrement au format WAV depuis Audacity :

1. Sous “File” > “Export” > “Export as WAV” :



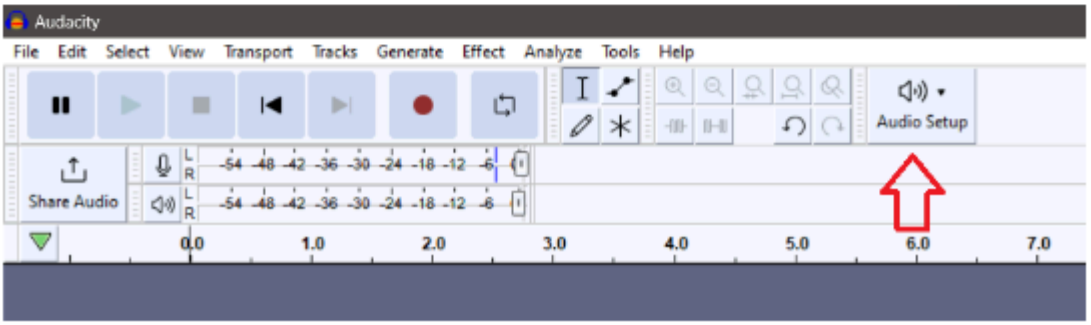
2. Naviguez vers le dossier où vous souhaitez enregistrer le fichier.
   A. Renommez le fichier au format : InitialePrénom+InitialeNom+AnnéeNaissance+Sexe+Jour+Mois+Année
   B. Sélectionnez « 32-bit float » sous « Encoding »
   C. Cliquez sur « Save »



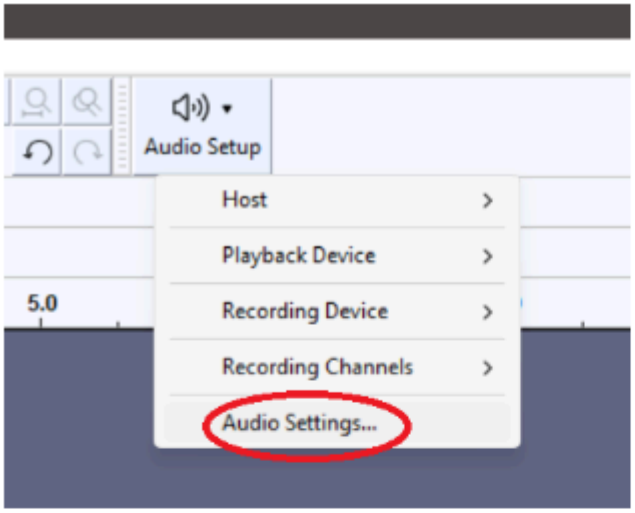
Instructions de réglages d'enregistrement

Les instructions suivantes utilisent le gratuiciel Audacity, téléchargeable ici : https://www.audacityteam.org/download/

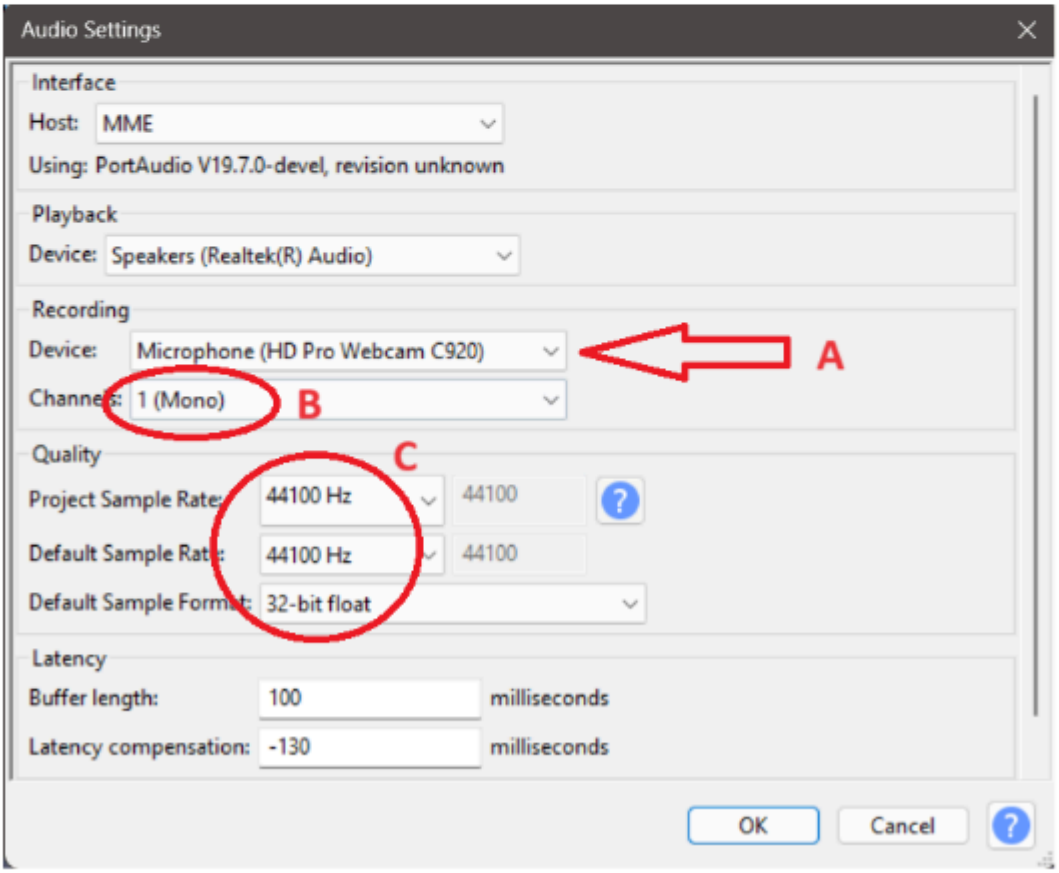
1. Une fois Audacity ouvert, cliquez sur Audio Setup



2. Cliquez sur Audio Settings

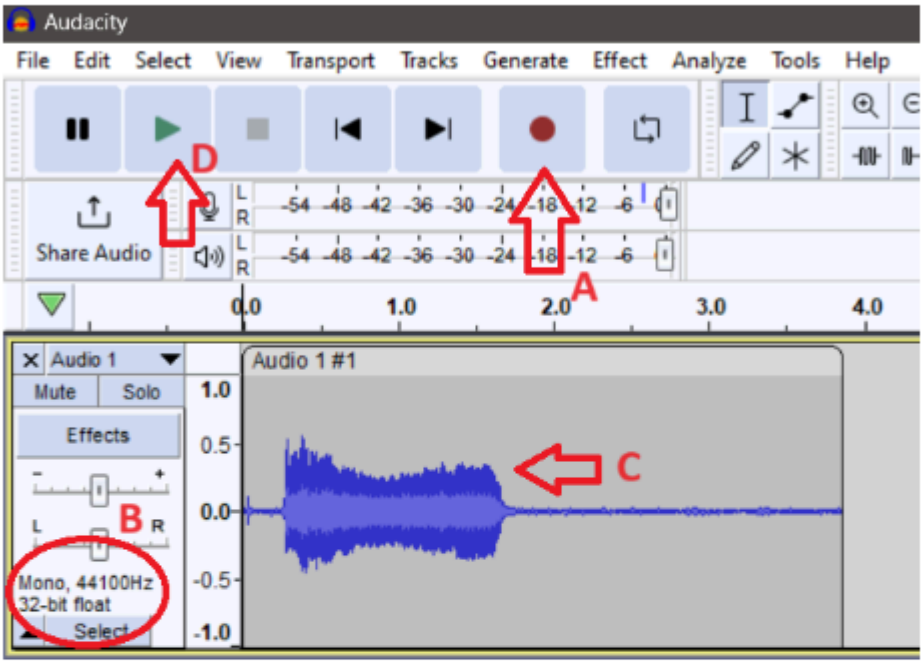


3. A. Vérifiez que la source sélectionnée sous « Recording > Device » correspond au microphone que vous avez branché à votre ordinateur.
   B. Choisissez le mode « 1 (Mono) » sous « Channels »
   C. Sous « Quality », sélectionnez 44100 Hz pour « Project Sample Rate » et « Default Sample Rate », et « 32-bit float » sous « Default Sample Format »
   Si cette rubrique n'apparaît pas, une mise à jour du logiciel est peut être nécessaire.



4. Faites un test avant d'enregistrer les stimuli :

A. Cliquez sur le bouton rouge (Record) et parlez dans le micro.
B. Vérifiez que les réglages correspondent à ceux préconisés.
C. Vérifiez qu'une onde sonore apparaît.
D. Écoutez l'enregistrement pour vous assurer que la qualité vous semble correcte.



RÉCAPITULATIF DES ÉLÉMENTS À TRANSMETTRE :

- votre enregistrement test → à transmettre avant de procéder aux enregistrements
- le formulaire VHI rempli par le patient
- l'enregistrement du patient
- les copies du formulaire de consentement et de l'autorisation de réutilisation et de traitement de données à caractère personnel de chaque patient

Veuillez transférer ces données via le lien d'accès FileSender qui vous a été partagé.

Nous vous remercions vivement de votre participation !

Annexe IV : Version française du VHI de Woisard et al. (2004)

		J	PJ	P	PT	T
F1	On m'entend difficilement à cause de ma voix					
P2	Je suis à court de souffle quand je parle					
F3	On me comprend difficilement dans un milieu bruyant					
P4	Le son de ma voix varie au cours de la journée					
F5	Les membres de la famille ont du mal à m'entendre quand je les appelle dans la maison					
F6	Je téléphone moins souvent que je le voudrais					
E7	Je suis tendu(e) quand je parle avec d'autres à cause de ma voix					
F8	J'ai tendance à éviter les groupes à cause de ma voix					
E9	Les gens semblent irrités par ma voix					
P10	On me demande : "Qu'est-ce qui ne va pas avec ta voix ? »					
F11	Je parle moins souvent avec mes voisins, mes amis, ma famille à cause de ma voix					
F12	On me comprend difficilement quand je parle dans un endroit calme					
P13	Ma voix semble grinçante et sèche					
P14	J'ai l'impression que je dois forcer pour produire la voix					
E15	Je trouve que les autres ne comprennent pas mon problème de voix					
F16	Mes difficultés de voix limitent ma vie personnelle et sociale					
P17	La clarté est imprévisible					
P18	J'essaie de changer ma voix pour qu'elle sonne différemment					
F19	Je me sens écarté(e) des conversations à cause de ma voix					
P20	Je fais beaucoup d'effort pour parler					
P21	Ma voix est plus mauvaise le soir					
F22	Mes problèmes de voix entraînent des pertes de revenus					
E23	Mes problèmes de voix me contrarient					
E24	Je suis moins sociable à cause de mon problème de voix					
E25	Je me sens handicapé(e) à cause de ma voix					
P26	Ma voix m'abandonne en cours de conversation					
E27	Je suis agacé(e) quand les gens me demandent de répéter					
E28	Je suis embarrassé(e) quand les gens me demandent de répéter					
E29	A cause de ma voix je me sens incompétent(e)					
E30	Je suis honteux (se) de mon problème de voix					

J = jamais, PJ = presque jamais, P =parfois
PT = presque toujours, T = toujours
(Cotation : J=0 point, PJ=1 point, P=2 points, PT=3 points, T=4 points, total /120)

Annexe V : Protocole d'évaluation des voix avec le CAPE-Vf

PROCÉDURE D'ÉVALUATION PERCEPTIVE DE LA VOIX AVEC L'OUTIL CAPE-V

Bonjour Madame, Monsieur, Docteur.e,

Nous vous remercions vivement pour votre participation à notre étude !

Pour rappel, notre mémoire a pour objectif la **validation française de l'adaptation francophone du CAPE-V**, effectuée par D. Mbagira (2023).

*Le CAPE-V est un outil **standardisé**, comparable au GRBAS. Les professionnels de la voix ont besoin d'outils sensibles, spécifiques et fiables pour évaluer les dysphonies : à ce jour, les outils disponibles en langue française sont limités. Le CAPE-V est un outil prometteur validé dans de nombreuses autres langues.*
*De précédentes études ont relevé un **degré de corrélation élevé** entre les évaluations d'un même échantillon vocal avec le **CAPE-V et le GRBAS** ainsi qu'une **meilleure fidélité inter-juges pour le CAPE-V** que pour le GRBAS (Nemr et al. 2012). Le CAPE-V serait également un outil **plus sensible aux paramètres fins des troubles vocaux** que le GRBAS (Karnell et al. 2007).*

Dans ce document, vous trouverez les informations nécessaires sur la procédure à mener lors de l'évaluation des voix. Veuillez en prendre connaissance dans son intégralité avant de débuter.

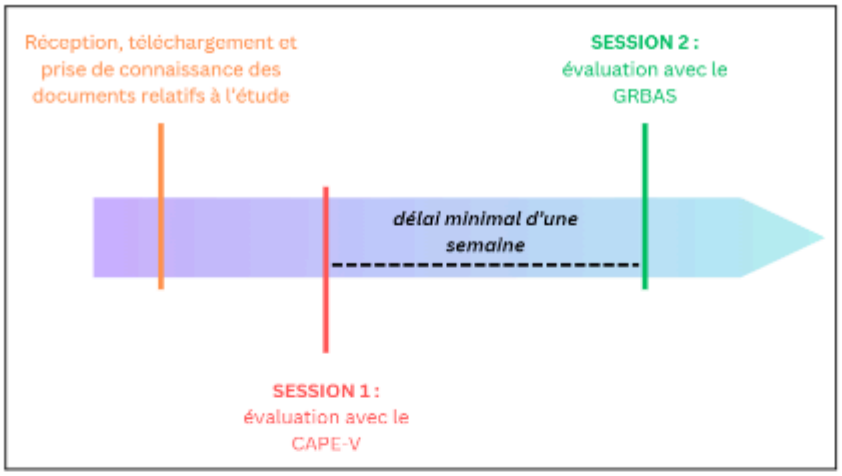
Nous restons à votre entière disposition pour toutes questions, remarques ou précisions complémentaires au **06 56 67 48 76** ou par mail à **margaux.shanks@univ-tlse3.fr**

Bonne écoute !

Vous avez reçu par mail :

- 19 copies du CAPE-V, en format .pdf
- une copie du tableau GRBAS, en format word
- un lien FileSender vous permettant d'accéder aux 19 enregistrements de voix que vous serez amené.e à évaluer

Voici un aperçu de la temporalité des sessions d'évaluation.



Dans la mesure du possible, nous demandons aux participants de séparer les deux sessions d'évaluation d'un laps de temps d'**au moins une semaine**, de manière à ce que les deux outils n'aient pas d'influence l'un sur l'autre.

Veuillez **ne pas écouter les enregistrements** avant le jour de l'évaluation.

SESSION 1 : EVALUATION AVEC LE CAPE-V

Matériel :

- ordinateur
- casque audio
- les 19 copies du CAPE-V
- fichiers audio téléchargés via le lien FileSender

Ouvrez le premier fichier CAPE-V : "**CAPE-V_1".** Ouvrez l'enregistrement correspondant : "**enregistrement_1**"

Il est important que vous écoutiez toutes les tâches vocales avant de remplir le formulaire CAPE-V. Pour des instructions sur comment annoter un document pdf, veuillez vous référer à la fin de ce document.

Si la performance est **uniforme** à travers toutes les tâches, vous pouvez procéder à la cotation en indiquant la performance globale pour chaque caractéristique vocale.

Si vous notez une **divergence** dans les performances entre les tâches, la marque sur l'échelle de gradation sera faite avec le numéro de tâche :

- la tenue des voyelles sera étiquetée #1
- la lecture ou la répétition de phrases sera étiqueté #2
- le discours conversationnel sera étiqueté # 3

Exemple :

Dans les rares cas où le clinicien perçoit des différences significatives entre les items de la tâche “phrases”, il ou elle peut étiqueter les notes en conséquence : 2(a) - (b) - (c) - (d) - (e) - ou (f) pour les différentes phrases.

Lorsqu’aucune étiquette particulière n’est utilisée, cela indique une performance uniforme.

Remarque : Si un paramètre vocal est jugé de même sévérité chaque fois qu'il apparaît, mais il n'est pas présent tout le temps, "I" doit être encerclé pour indiquer son intermittence et aucun étiquetage supplémentaire ne doit être fait.

Autres caractéristiques : Le clinicien peut indiquer des observations importantes sur les phénomènes de résonance "Commentaires sur la résonance." Les exemples incluent mais ne sont pas limités à l'hyper- ou hyponasalité, et une résonance en cul-de-sac.

SESSION 2 : EVALUATION AVEC LE GRBAS

Matériel :

- ordinateur
- casque audio
- la copie électronique du formulaire GRBAS
- fichiers audio téléchargés via le lien FileSender

Ouvrez le fichier “**GRBAS FORMULAIRE**”.
Ouvrez le premier enregistrement : “**enregistrement_1**”
Remplissez les cases correspondantes pour chaque enregistrement.

INFORMATIONS GRBAS :

Extrait du livre "La voix : anatomie, physiologie et exploration" (Giovanni et al., 2021) :

INITIALE	ANGLAIS	FRANÇAIS	DÉFINITION
G	Grade	Grade global	Impression globale du degré d'anormalité de la voix
R	Roughness	Raucité	Impression d'irrégularité des vibrations de la corde vocale qui correspond aux fluctuations irrégulières de la Fo et/ou à l'amplitude du son glottique
B	Breathiness	Souffle	Impression d'une fuite d'air assez importante à travers les cordes vocales, relative à des turbulences
A	Asthenicity	Asthénie	Manque de puissance de la voix relatif à une intensité faible du son et/ou un manque des harmoniques élevés
S	Strain	Serrage	Impression d'un état hyperfonctionnel de phonation relatif à une Fo anormalement haute (bruit dans les hautes fréquences et/ou richesse en harmoniques dans les hautes fréquences)

Tableau 19-1. Définition des paramètres de l'échelle GRBAS

Chaque paramètre est côté selon quatre degrés de sévérité :

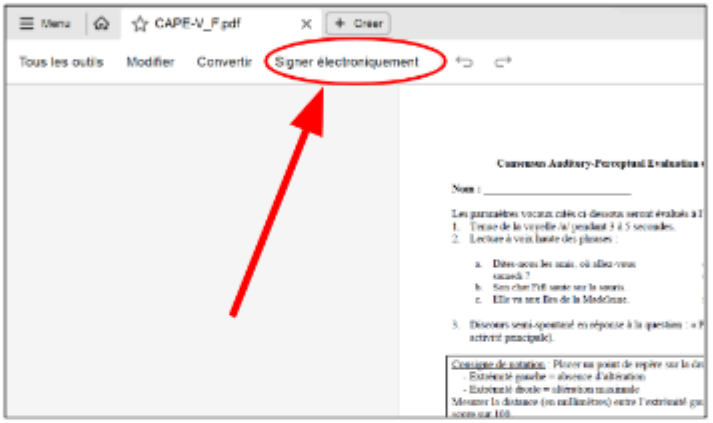
- 0 (normal),
- 1 (légèrement altéré),
- 2 (altéré)
- 3 (sévèrement altéré).

Nous vous invitons à sauvegarder toutes les modifications effectuées sur les documents lors de l'évaluation. **Veuillez nous prévenir par mail lorsque vous aurez terminé vos deux sessions d'évaluation**. Nous vous enverrons ensuite un lien FileSender sur lequel vous pourrez déposer tous les documents.

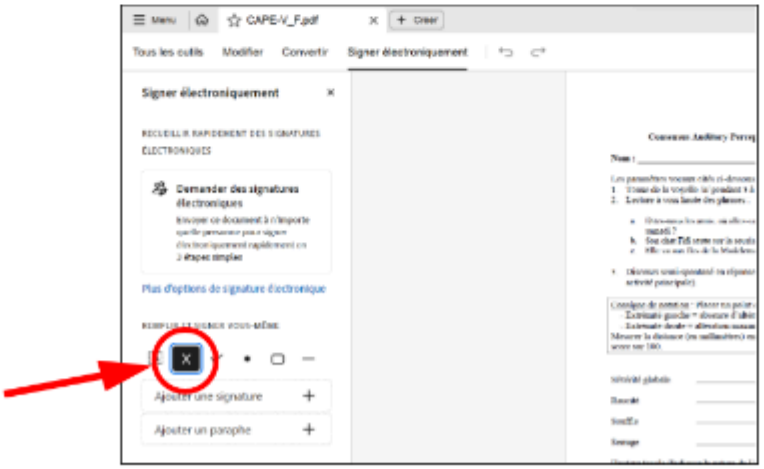
Nous vous remercions de votre participation !

Pour annoter le pdf avec Adobe Acrobat :

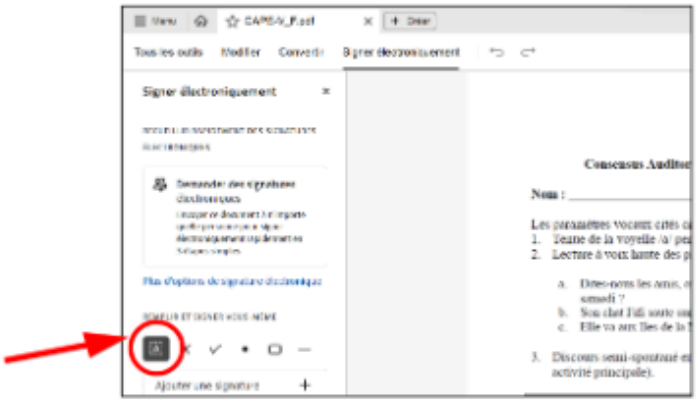
- cliquez sur "Signer électroniquement" en haut à gauche de la page



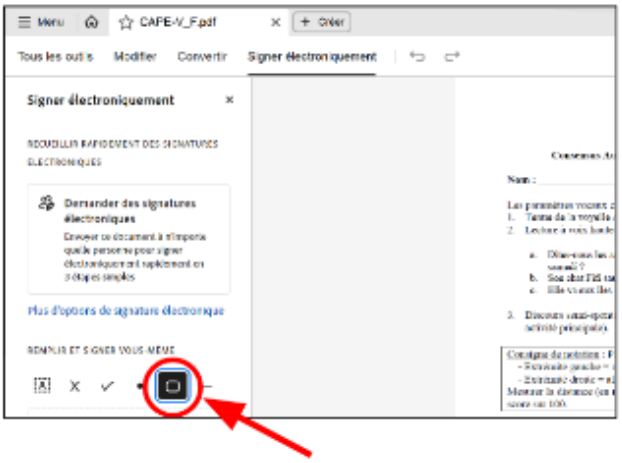
- dans le menu qui s'affiche, cliquez ensuite sur la deuxième icône : "Ajouter une croix" ; puis placez ensuite votre croix sur le CAPE-V aux endroits correspondants



- **pour ajouter du texte**, cliquez sur la première icône : "Remplir les champs de formulaire". Cliquez ensuite à l'endroit où vous souhaitez ajouter du texte

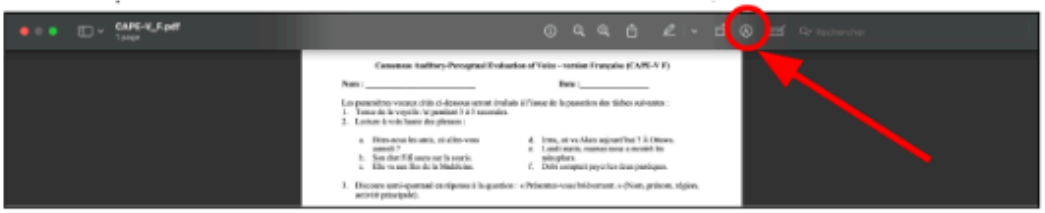


- pour entourer une lettre, cliquez sur la cinquième icône : “Dessiner un rectangle” ; puis cliquez sur la lettre que vous souhaitez entourer

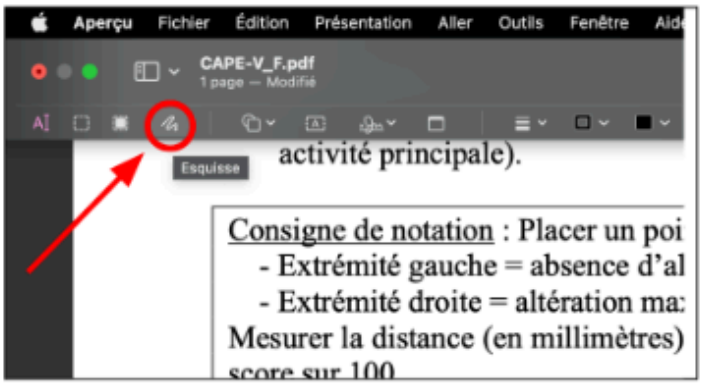


Pour annoter le pdf avec un ordinateur Mac :

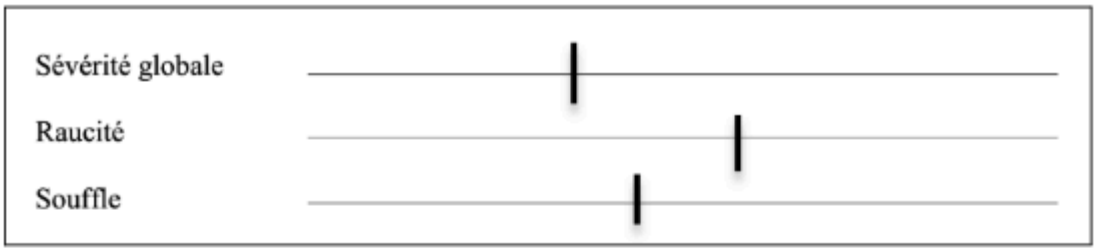
- cliquez sur l'icône “Afficher la barre d’outils d’annotation”, en haut à droite



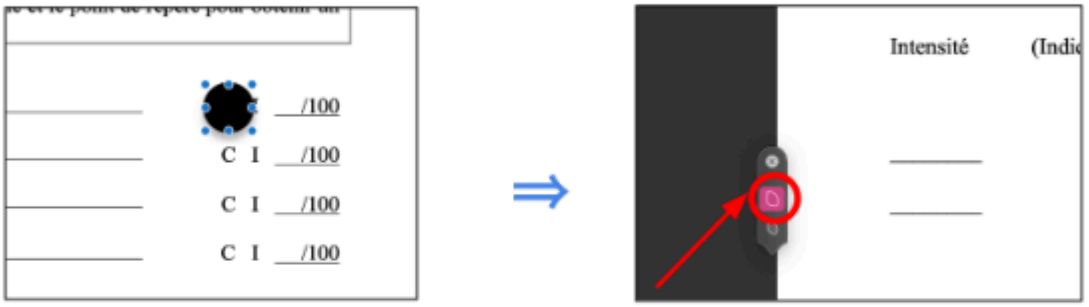
- cliquez ensuite sur “Esquisse”, à gauche sur la barre qui est apparue



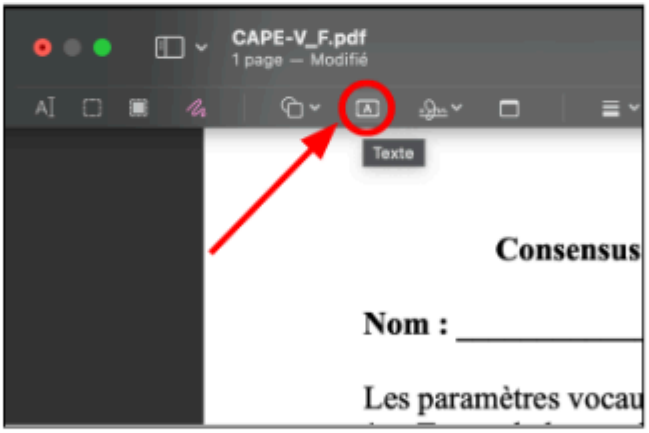
- avec votre souris, tracez vos points de repères



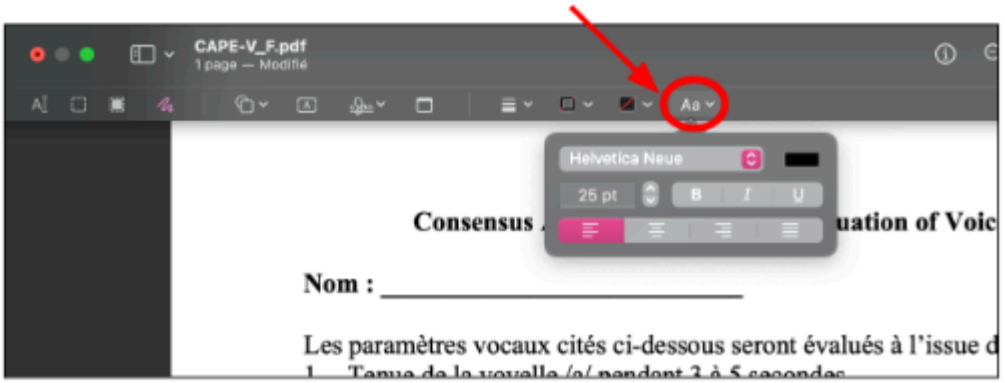
- cette même fonctionnalité permet également d'entourer. Si le cercle se remplit, cliquez sur cette icône en bas à gauche :



- pour ajouter du texte, cliquez sur l'icône "Texte".



- pour changer la police / taille du texte, :



Annexe VI : Tableau GRBAS tel que présenté dans notre étude

ENREGISTREMENT	G	R	B	A	S
enregistrement_1					
enregistrement_2					
enregistrement_3					
enregistrement_4					
enregistrement_5					
enregistrement_6					
enregistrement_7					
enregistrement_8					
enregistrement_9					
enregistrement_10					
enregistrement_11					
enregistrement_12					
enregistrement_13					
enregistrement_14					
enregistrement_15					
enregistrement_16					
enregistrement_17					
enregistrement_18					
enregistrement_19					

RÉSUMÉ

Titre : Validation française du Consensus Auditory Perceptual Evaluation of Voice (CAPE-V)

Résumé :

Contexte : L'évaluation perceptive de la voix est un élément essentiel du bilan vocal multidimensionnel dans le cadre de la dysphonie. Avec les mesures aérodynamiques et acoustiques, l'autoévaluation du patient et l'examen laryngé direct, elle fait partie des examens recommandés par la Société Européenne de Laryngologie. L'objectif de notre étude est d'effectuer la validation psychométrique de l'adaptation francophone du Consensus Auditory Perceptual Evaluation of Voice (CAPE-Vf), dans le but d'élargir le panel restreint d'outils d'évaluation perceptive francophones standardisés, valides et fiables à disposition des cliniciens de la voix français.

Méthodes : Quarante-et-un locuteurs dysphoniques et normophoniques ont rempli un VHI avant d'être enregistrés suivant le protocole du CAPE-Vf. Treize orthophonistes ont ensuite effectué une évaluation perceptive de ces voix avec le CAPE-Vf et l'échelle GRBAS. La fiabilité intra-juge et inter-juges a ensuite été calculée. Pour établir la validité de critère du CAPE-Vf, nous avons comparé les résultats des deux outils. Enfin, nous avons également comparé les résultats de ces évaluations perceptives à ceux du VHI.

Résultats : Le CAPE-Vf présente une bonne fiabilité intra-juge pour la majorité des paramètres vocaux évalués. Cependant, les résultats de la fiabilité inter-juges se révèlent moins significatifs. On retrouve une corrélation modérée entre les quatre paramètres comparables du CAPE-Vf et du GRBAS. L'évaluation perceptive et l'auto-questionnaire VHI révèlent une corrélation plus faible que celle rapportée pour le CAPE-Vf et le GRBAS.

Conclusion : La validité préliminaire du CAPE-Vf est ainsi établie, présentant des résultats encourageants quant à la validation future de l'outil sur une plus grande population.

Mots clés : CAPE-V, voix, dysphonie, évaluation perceptive, orthophoniste, GRBAS, VHI, validation

ABSTRACT

Title : Establishing the validity of the French version of the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V)

Abstract :

Introduction : Auditory-perceptual voice evaluation is a necessary component of the multidimensional voice assessment in the context of dysphonia. Along with aerodynamic and acoustic measures, self-assessment by the patient and videostroboscopy, it is part of the five dimensions to be evaluated as recommended by the European Laryngological Society. The objective of our study was to establish validity of the French version of the Consensus Auditory Perceptual Evaluation of Voice, with the aim of expanding the limited range of standardized, valid and reliable francophone auditory-perceptual evaluation tools available to French-speaking voice clinicians.

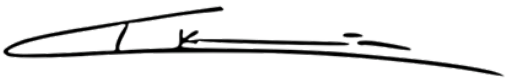
Methods : Forty-one dysphonic and non-dysphonic participants were recorded following the CAPE-Vf protocol, and VHI questionnaires were also collected. Auditory-perceptual evaluation was conducted with both CAPE-Vf and GRBAS at a one-week interval by thirteen experienced speech-language pathologists. Intra-rater and inter-rater reliability was examined for both tools, as well as criterion validity through the comparison of CAPE-Vf and GRBAS scores. Participants' VHI scores were also compared to that of the auditory-perceptual evaluation by professionals.

Results : The CAPE-Vf demonstrated good intra-rater reliability for most voice parameters. However, interrater reliability yielded less significant results. A moderate correlation was found between the four comparable CAPE-Vf and GRBAS voice parameters. The correlation between patients' self-assessed impact of the voice disorder and auditory-perceptual evaluation by the clinician was found to be less significant than the correlation between CAPE-Vf and GRBAS.

Conclusion : This study was able to establish preliminary validity of the French version of the CAPE-V. Further investigations of validity and reliability on a larger population remain essential in establishing the CAPE-Vf 's validity.

Keywords : CAPE-V, voice, dysphonia, auditory-perceptual evaluation, speech-language pathologist (SLP), GRBAS, VHI, validation

SIGNATURE DES ENCADRANTS

Sandrine MICHEL :

Timothy POMMÉE :