



Évaluation perceptive du souffle: influence du type de tâche, du niveau d'expertise et du type d'échantillon.

C.Billiet, J. Schoentgen et M. Remacle, I. Verduyck, .



ULB



UCL
Université
catholique
de Louvain

1. Introduction théorique

1.1 Le jugement perceptif

- Évaluation subjective, dépend de différents facteurs:
 - ✓ expérience du juge
 - ✓ échantillon utilisé
 - ✓ type de tâche
- Pour être une mesure significative:
 - bons niveaux d'accord et de fiabilité inter et intra-juges

1.2 GRBAS et le jugement comparatif par paires

- GRBAS = échelle perceptive subjective la plus utilisée en clinique actuellement
 - ✓ 5 paramètres à évaluer selon 4 degrés de sévérité
 - ✓ avantages: simple, compacte, rapide à faire passer
 - ✓ inconvénients: simplicité de cotation est trop réducteur, manque de fiabilité, fait appel aux représentations standard internes
- Jugement comparatif par paires = présentation de deux stimuli de référence à l'auditeur qui détermine celui qui est le plus soufflé/enroué/féminin...
 - ✓ **avantages:** très bons niveaux d'accords inter et intra-juges, diminution de l'influence de l'expertise, limitation du recours aux représentations prototypiques.
 - ✓ **inconvénients:** temps de passation, limitation du nombre d'échantillons

1.3 La qualité perceptive: souffle

- Dans la littérature, description acoustique ou perceptive plutôt qu'une réelle définition:

✓ "fuite d'air audible qui conduit à une phonation limitée et légère"
Bassich et Ludlow, 1986

✓ "qualité de la voix caractérisée par un bruit de frottement audible qui est le résultat d'une fermeture incomplète des cordes vocales" Jeannon et al., 1998

- Corrélats entre évaluation perceptive et mesures acoustiques:
actuellement, pas de consensus entre les différents auteurs (NHR, ...).

2. Méthodologie

2.1 Participants

- **Juges expérimentés**
 - ✓ Formation dans la pratique vocale (entraînement et expérience)
 - ✓ **4 juges** ont réalisé l'évaluation du **souffle** avec le B de la GRBAS et jugement par paires (2 juges en test-retest)
- **Juges naïfs**
 - ✓ Pas de formation ou d'entraînement dans le domaine vocal
 - ✓ 7 participants ont réalisé l'évaluation du **souffle** avec le B de la GRBAS et le jugement pas paires (test-retest)
 - ✓ 5 participants ont réalisé l'évaluation de l'**enrouement général** avec la GRBAS et le jugement par paires (tps 1&2, pour voyelles et phrases)

2.2 Échantillons

- Voix de la base de données ECLIPSE
 - Critères de sélection dans la base de données
 - ✓ Voix féminines
 - ✓ Suppression des enregistrements avec du bruit ou du grésillement
 - ✓ Écartement des voix avec accents, erreurs de prononciation et de lecture
- ⇒ Groupe de voix de base = 23 voix
⇒ Pour notre étude: 20 voix nécessaires
- Sélection en fonction de la fréquence fondamentale
 - ✓ Écartement des deux voix les plus graves
 - ✓ Écartement de la 2ème voix la plus aiguë
 - 2 types d'échantillons:
 - ✓ Voyelle /a/ soutenue
 - ✓ 2 phrases phonétiquement équilibrées de P. Combescure

2.3 Tâches

- Échelle GRBAS

- ✓ Appréciation du souffle et/ou de l'enrouement général
- ✓ Évaluation de 0 (pas d'altération) à 3 (altération sévère)
- ✓ Ordre de présentation des stimuli est aléatoire

- Logiciel de jugement par paires

- ✓ Présentation de paires de stimuli dont il faut sélectionner « le plus soufflé/enroué »
- ✓ Toutes les comparaisons possibles sont effectuées soit $(20 \times 19) / 2 = 190$
- ✓ Principe de cotation simple:
 - Au départ, toutes les voix ont un score de 0
 - Stimulus choisi dans la paire a son score incrémenté d'un point
 - Possibilité de jugement neutre: deux voix ont leur score incrémenté de 0,5

2.4 Procédure

- Juges naïfs

- ✓ Brève explication du soufflé
- ✓ Réalisation pratique:
 - Casque
 - Pièce calme
 - Pausés
 - Durée d'une session: +/- 2 heures
 - 2 séances à au moins 14 jours d'intervalle

- Juges expérimentés

- ✓ Pas d'explication du soufflé mais juges experts qui ont réalisé le jugement par paires ont suivi la même formation et collaborent
- ✓ Participants reçoivent un CD avec les tâches à réaliser (avec explicatif)
- ✓ Réalisation pratique
 - Jugement réalisé individuellement
 - Casque
 - Parfois fractionnement des jugements en plusieurs séances

3. Résultats

3.1 Type de tâche

- Corrélations

- ✓ Calculées sur base du coefficient de Spearman
- ✓ **Quand coefficient supérieure à 0,6 avec $p.<0,01$, corrélation considérée comme élevée**
- ✓ **Quand coefficient supérieur à 0,8 avec $p.<0,01$, corrélation considérée comme très élevée**

- Corrélations inter-juges

- ✓ GRBAS :
 - **62%** des corrélations (soit 34/55) sont élevées
 - **5%** des corrélations (soit 3/55) sont très élevées
- ✓ Jugement par paires
 - **84%** des corrélations (soit 46/55) sont élevées
 - **35%** des corrélations (soit 19/55) sont très élevées
 - Tous les juges experts: bonnes corrélations entre eux

3. Résultats

3.1 Type de tâche

- Corrélations intra-juges

- ✓ GRBAS

- 75% des corrélations (soit 6/8) sont élevées dont 1/8 très élevée

- ✓ Jugement par paires

- 100% des corrélations (soit 9/9) sont élevées

- 77% des corrélations (soit 7/9) sont très élevées

3.2 Niveau d'expertise

- Comparaison des rangs attribués par les naïfs et les experts

- ✓ Calculée sur base du test U de Mann Whitney avec $p < 0,01$

- ✓ Aucune différence significative de classement entre les naïfs et les experts

- ✓ G B non comparés – corrélations intra et inter-juges trop faibles pour juges naïfs.

3.3 Type d'échantillon: phrases/voyelles

- Comparaison de rangs attribués par les juges pour les voyelles et les phrases
 - ✓ **Test de Wilcoxon** , rangs moyens pour des groupes appariés puisque les échantillons sonores proviennent des mêmes voix
 - ✓ **Significativité <0,01**
- GRBAS
 - ✓ Pas de différence significative de classements entre voyelles et phrases
- Jugement comparatif par paires
 - ✓ Pas de différence significative de classements entre les voyelles et les phrases

3.4 Comparaison entre classement du souffle et de l'enrouement général

- Comparaison des classements faits pour le souffle et pour l'enrouement général.
 - ✓ Test U de Mann Whitney avec $p. < 0,01$. (N=11 pour l'évaluation du souffle, N=5 pour l'évaluation de l'enrouement général)
- Résultats:
 - ✓ 10% des voix (soit 2/20) sont jugées différemment pour le souffle et l'enrouement général
- On sait que le B est fortement corrélé au G – difficile de tirer des conclusions.

4. Discussion

4.1 Influence du type de tâche

- Stabilité inter-juges
 - ✓ **Hypothèse de départ:** le paradigme de comparaison par paires permet une stabilité de jugement plus importante entre les juges par rapport à l'échelle traditionnelle GRBAS
 - ✓ **Résultats:**
 - Stabilité inter-juges plus importante dans le jugement par paires que dans la GRBAS
 - Corrélations élevées et très élevées sont plus nombreuses dans le jugement par paires que dans la GRBAS
- Stabilité intra-juge
 - ✓ **Hypothèse de départ:** jugement par paires stabilise l'évaluation d'un même juge entre les temps 1 et 2 par rapport à la GRBAS
 - ✓ **Résultats:**
 - Très bonne stabilité intra-juge avec le jugement par paires
 - Juges naïfs et experts obtiennent des résultats similaires

4.2 Impact du niveau d'expertise

- **Hypothèse de départ:** influence de l'expérimentation sur les classements puisque les experts ont un entraînement à l'évaluation de voix dysphoniques.

- **Résultats:**

- ✓ Pour 100% des voix: pas de différence de classements entre les naïfs et les experts.
- ✓ Rangs moyens similaires entre les groupes
- ✓ Diminution de l'influence de l'expérience quotidienne des professionnels dans le jugement par paires

4.3 Influence du type d'échantillon

- **Hypothèse de départ:** comparaison par paires permet de juger l'altération du souffle tant sur les voyelles que sur la parole.
- **Résultats:**
 - ✓ Jugement par paires: pas de différences de classement entre les voyelles et les phrases
 - ✓ GRBAS: pas de différence significative mais une tendance à l'évaluation perceptive différente entre les deux types d'échantillons

4.4 Évaluation du souffle

- **Hypothèse de départ:** les classements sont différents lorsque les juges évaluent le souffle ou l'enrouement général
- **Résultats:**
 - ✓ Différence significative entre les classements pour deux voix seulement
 - ✓ Pas de différence significative pour 18/20 voix
 - ✓ MAIS:
 - Le grade général de la voix dépend en partie de la quantité de souffle présent dans la voix
- Pour des recherches futures:
 - ✓ Évaluation par les mêmes juges de différents attributs perceptifs et comparer s'il y a des différences de classements
 - ✓ Recherche similaire avec des échantillons de synthèse dans lesquels le souffle pourrait être manipulé

Conclusion

- Nos résultats confirment que l'utilisation du jugement par paires:

- ✓ augmente la stabilité inter et intra-juges de l'évaluation perceptive du souffle

- ✓ diminue l'influence du niveau d'expertise sur le classement des voix

- ✓ n'est pas sensible au type d'échantillon utilisé

⇒ Mesure fiable

The letters 'SFPPC' are rendered in a bold, 3D, orange font with a slight shadow. They are arranged in a slightly upward-sloping line from left to right. The letter 'S' is partially enclosed by a large, light gray circular shadow that appears to be cast from the top left.

SFPPC

Merci pour votre attention!

The numbers '2010' are rendered in a bold, 3D, orange font with a slight shadow. They are arranged in a slightly downward-sloping line from left to right. The number '0' is partially enclosed by a large, light gray circular shadow that appears to be cast from the top right.

2010