

Identification des consonnes du français en syllabe isolée après laryngectomie partielle supracricoïdienne

L Crevier-Buchman^{1&2}, S Hans^{1&2}, J Vaissiere², S Maeda³, D Brasnu^{1&2}

1. Hôpital Européen Georges Pompidou, 20 rue Leblanc, 75015 Paris, France – Tél : + 33 1 56 09 34 63 – Fax : + 33 1 56 09 35 68 – Mél : lise.buchman@wanadoo.fr - 2. Institut de Phonétique, CNRS, UMR 7018, Paris ; 3. ENST, CNRS, Paris

ABSTRACT

This study aimed to determine what patterns of perceptual confusions characterise the voice of patients after supracricoid partial laryngectomy (SCPL) by the identification tests of French consonants. After SCPL, voice is produced by a neoglottis located at approximately 3 cm above the removed vocal folds, thus shortening the vocal-tract length. We first evaluated the voicing distinction, as their vibrator is profoundly modified, and second manner and place of articulation features as their vocal tract is shortened by about 3 cm. Ten male patients were recorded 18 months after SCPL producing 16 French consonants in a syllabic context (CV). Consonant articulation appears to impose certain constraints on voicing ability of SCPL patients, since voiced consonants are predominantly perceived as voiceless consonants.

1. INTRODUCTION

La laryngectomie partielle supracricoïdienne (LPSC) consiste en l'exérèse de la totalité du cartilage thyroïde, de l'espace paraglottique, de l'épiglotte infrahyoïdienne, des deux cordes vocales (CV), des bandes ventriculaires et des ventricules. Sont conservés : l'os hyoïde, le cartilage cricoïde, un ou deux cartilages aryténoïdes mobiles et l'épiglotte suprahyoïdienne. Les seuls muscles intrinsèques laryngés conservés sont ceux de l'unité crico-aryténoïdienne, à savoir le crico-aryténoïdien postérieur (CAP), et latéral (CAL). Le CAP est dilateur de la néoglote [Kon84,Lud91] ; il participe à la réalisation des consonnes sourdes et est antagoniste du muscle inter-aryténoïdien pour la réalisation du trait de voisement dans différentes langues. Ces interventions réalisent un modèle anatomo-physiologique de conduit vocal raccourci d'environ 3 cm, avec suture du cartilage cricoïde à l'os hyoïde et à la base de langue, et capable de produire une voix dite "de substitution", alors que les deux cordes vocales (CV) ont été réséquées. C'est la vibration de la muqueuse des deux aryténoïdes contre l'épiglotte qui constitue la source sonore néoglottique [Cre95,Cre98].

Notre étude perceptive est la première étude prospective, qui décrit et analyse les caractéristiques phonétiques perceptives de la voix et de la parole ainsi que l'étude des phénomènes articulatoires compensatoires dus à la perte du vibreur laryngé et aux modifications du conduit vocal chez dix patients ayant subi une LPSC. Nous avons

réalisé une étude de l'identification des consonnes du français et des tests de perception du trait de voisement de ces consonnes en situation syllabique CV. Le trait de voisement des occlusives et des fricatives est phonologiquement distinctif en français. D'une part, en raison de l'altération de la source vocale, on peut s'attendre à une dégradation du trait de voisement ; d'autre part, la modification du conduit vocal devrait retentir sur le lieu d'articulation ; mais il ne devrait pas y avoir d'altération de mode de constriction ni même de nasalité puisque l'oropharynx et la cavité buccale ne sont pas modifiés par l'intervention.

La perception du voisement en français n'est pas exclusivement le fait de la vibration des cordes vocales ou d'une néoglote. Les corrélats acoustiques de la perception du voisement sont représentés classiquement par l'énergie du burst pour les occlusives [Ste97], la variation de la pente spectrale du burst vers la voyelle qui suit [Ken74,Kew82] et la vitesse de fermeture et d'ouverture du conduit vocal. Pour le lieu d'articulation, le voice onset time (VOT) joue un rôle important [Ste97], et il a tendance à augmenter lorsque le lieu d'articulation recule de la zone labiale, vers la zone vélaire.

2. MATERIEL ET METHODES

2.1 Les patients

Dix patients ont été enregistrés à 18 mois postopératoires. Ce sont tous des hommes âgés de 51 ans à 73 ans (moyenne : 63 ans, médiane : 69 ans).

2.2 Le corpus

Trois répétitions de la syllabe [Ca], où C = une des 16 consonnes du français.

2.3 Le protocole d'écoute

L'écoute se faisait au casque, l'intensité étant pré-réglée (74 dB) et identique pour tous les auditeurs. Les 16 syllabes de chaque patient étaient présentées à la suite avec trois répétitions aléatoires de chaque syllabe soit 3 (répétitions) x 10 (patients) x 16 (consonnes) = 480 stimuli pour chaque auditeur.

2.4 Les auditeurs

Nos trois auditeurs étaient de langue maternelle française, orthophonistes, travaillant dans le service d'O.R.L. de l'hôpital Laënnec depuis plus de deux ans et familiarisés avec ce type de chirurgie et de voix. Il y avait deux femmes et un homme âgés de 25 à 35 ans.

2.5 Les matrices de confusions

Les résultats de chaque test ont été mis sous forme d'une matrice de confusion à 16 entrées (les 16 consonnes) et regroupant les résultats des dix patients évalués par les trois auditeurs. Pour chaque consonne il y avait 30 stimuli (3 répétitions x 10 patients). Il y avait donc 1440 stimuli par matrice (16 syllabes x 3 répétitions x 10 patients x 3 auditeurs). Les résultats sont exprimés en pourcentages.

3. RÉSULTATS

L'analyse des données se fait horizontalement et verticalement. La colonne verticale correspond aux phonèmes cibles, c'est-à-dire aux phonèmes que le patient a voulu prononcer. Les colonnes horizontales correspondent aux catégories phonétiques qui ont été perçues par les auditeurs. A partir de cette matrice, nous avons fait trois autres matrices, correspondant au trait de voisement, au lieu d'articulation : labial (p,b,f,v,m), alvéolaire (t,d,l,s,z,n), palato-vélaire (k,g,r,j,ʒ), et au mode de constriction occlusives (p,t,k,b,d,g) et fricatives (f,s,ʃ,v,z,ʒ). Le nombre de confusions est relevé pour chaque consonne en pourcentage d'erreur. Les résultats sont exprimés en pourcentage qui correspond au nombre de confusions affectant une consonne entendue 90 fois (3 stimuli x 10 locuteurs x 3 auditeurs). Autrement dit, plus le pourcentage d'erreur est élevé, moins cette consonne ou cette classe de consonne est résistante.

3.1 Selon le type de voisement sourd/sonore

Les résultats montrent qu'il existe un réel problème de distinction entre consonnes voisées et non voisées chez les patients. Notre étude montre qu'il existe 24% de confusion dans le sens d'une désonorisation d'une consonne sonore contre 4% dans le sens d'une sonorisation d'une consonne sourde. L'erreur d'identification est la plus sévère pour la consonne [g] avec 31% de confusion avec le [k]. La consonne labiale [b] est confondue avec le [p] dans 31% des cas. Les consonnes alvéolaires [d] ont aussi 29% de confusion en faveur du [t]. Nous observons le même phénomène pour les fricatives avec 14% de confusion du [ʒ] en faveur du [ʃ], et 20% de confusion du [z] en faveur du [s].

Il y a confusion du trait de voisement des consonnes sonores en faveur des consonnes sourdes homorganiques, prédominant pour les occlusives. Les patients favorisent l'articulation consonantique au détriment du voisement.

3.2 Selon le lieu d'articulation : labial, alvéolaire et palato-vélaire

Les occlusives palato-vélares [k] et [g] sont identifiées à des occlusives alvéolaires [t] et [d] dans 15% des cas. La confusion pour les labiales est de 12% en faveur des alvéolaires. Les fricatives ne posent pas de problème d'identification pour le trait de lieu. Pour les nasales, il existe une confusion de lieu avec 7% d'erreur du [m] en faveur du [n]. Il n'y a pas de confusion pour les "glides" [l] et [r].

3.3 Selon le mode occlusif - fricatif

Il n'y a pas de confusion entre occlusives et fricatives hormis 2% de [ʒ] perçu comme [k]. Les classes de consonnes les plus fragiles sont celles des sonores et des palato-vélares. L'analyse des confusions de voisement montre que : les confusions se font quasi - exclusivement dans le sens sonore → sourde ; la plupart des confusions de voisement se font entre consonnes homorganiques et plus particulièrement pour les occlusives que pour les fricatives.

4. DISCUSSION

Après LPSC, le larynx est ascensionné d'au moins 3 cm et le volume du conduit vocal est modifié. Toutes ces modifications peuvent être responsables d'une plus grande tension du vibrateur néolaryngé. Nous n'avons pas retrouvé d'études qui, à ce jour, traitent de l'intelligibilité des consonnes prononcées par des patients ayant eu une LPSC. La comparaison et la discussion de nos résultats sont donc difficiles. Par contre, il existe de nombreuses études concernant les voix alaryngées produites par le sphincter supérieur de l'œsophage [Gan83a,Gan83b, Doy88]. Nous les citerons à titre de réflexion mais se sont deux situations phonatoires très différentes et non comparables. En voix alaryngée, œsophagienne ou trachéo-œsophagienne, la néoglote est un muscle circulaire appartenant au tube digestif. Contrairement à nos résultats, en voix alaryngée, la principale confusion va dans le sens d'une sonorisation des consonnes sourdes, plus marquée en position initiale qu'en intervocalique. De plus, les confusions portent plus sur les fricatives que sur les occlusives.

4.1 Analyse de la méthode

Le choix de la syllabe est motivé par le fait qu'elle représente l'unité phonétique la plus petite. Le choix de la place de la consonne en initiale est aussi un facteur qui peut avoir biaisé nos résultats. Selon Bourciez & Bourciez [Bou95], les consonnes initiales sont plus facilement identifiées que les finales. C'est pourquoi nous avons pris les syllabes avec la consonne en prévoicalique [C-V]. En effet, initier ou arrêter la vibration dans un continuum sonore peut demander un ajustement fin

difficile à réaliser chez nos patients. Cette hypothèse a été confirmée pour les voix œsophagiennes pour lesquelles il existe effectivement une différence, avec une plus grande intelligibilité pour les consonnes en position intervocalique [Hir96].

4.2 Analyse des résultats du trait de voisement

Comme on pouvait le prévoir, la principale confusion dans l'identification des consonnes était pour le trait de voisement. Cela peut s'expliquer car, après LPSC, la modification anatomique porte principalement sur le vibrateur. Cette même constatation d'erreurs de voisement avait été faite pour les voix alaryngées [Doy89,Hir96]. Dans le cas des voix après LPSC, les confusions de voisement affectent principalement les consonnes sonores au profit de leurs homorganiques sourdes : sonore → sourde (même lieu, même mode de constriction). Il s'agit donc d'un problème de dévoisement pur. En effet, pour obtenir un voisement, il faut que les CV vibrent en état de relative détente [Lad83]. Il faut que le larynx soit abaissé ce qui assouplit et détend les CV. Or, après LPSC, le néolarynx est ascensionné d'au moins 3 cm. Par ailleurs, pour sonoriser une consonne, outre l'accolement des CV, il faut élargir les cavités pharyngées et augmenter le volume et la longueur du conduit vocal entre le lieu de constriction et les CV [Vai97]. Cet agrandissement est plus facile pour les consonnes labiales /b/, le volume de la cavité étant plus facilement modulable, que pour les consonnes postérieures /g/, le volume entre la constriction et la source de voisement étant réduit pour le /g/. Après LPSC, le conduit vocal est raccourci les adaptations de taille et de volume sont limitées ; les confusions homorganiques de voisement sont plus fréquentes pour les consonnes palato-vélaires que pour les labiales ou les alvéolaires. Lors du voisement, la muqueuse des cordes vocales est la source de cette vibration. Les caractéristiques de raideur des muqueuses des deux aryténoïdes, qui tiennent lieu chez ces patients de vibrateur laryngien, ont des caractéristiques biomécaniques très différentes. La masse, la viscosité et l'épaisseur sont augmentées. Il est possible que l'inertie de la muqueuse des aryténoïdes soit un facteur supplémentaire d'entrave au voisement. En effet il est fort possible qu'à différence de pression égale entre l'amont et l'aval du vibrateur laryngien, les cordes vocales puissent osciller alors que les muqueuses des aryténoïdes s'écartent sans osciller. Enfin, d'une façon générale, les confusions de voisement seraient aussi dues à un retard de l'adduction par inertie du vibrateur néolaryngé à sonoriser en initiale.

4.3 Les confusions de lieu d'articulation

Contrairement à notre hypothèse, il n'y a pas que le trait de voisement qui soit atteint mais aussi le lieu d'articulation. Les consonnes palato-vélaires et bilabiales sont plus fréquemment perçues comme étant des alvéolaires. On pourrait expliquer l'altération des consonnes postérieures chez nos patients, par une

compensation physiologique ; en effet, pour produire une sonorisation après LPSC, le patient fait un effort de "sphinctérisation" du néolarynx en rétrécissant le conduit pharyngé. Il s'aide de la base de langue qu'il recule ; elle perd alors une partie de son rôle d'articulateur consonantique postérieur et le patient compense par la mise en jeu du dos de la langue. D'où la modification des points d'articulation des consonnes palato-vélaires qui sont perçues plutôt comme alvéolaires. Par ailleurs, la position alvéolaire est aussi considérée par Laver [Lav92] et Klatt & Klatt [Kla90], comme une position neutre ou utilisée par défaut dans des troubles articulatoires. Enfin, une consonne mal perçue par un auditeur pourrait être étiquetée "alvéolaire" par défaut. Cependant, on peut noter que la consonne /r/ est bien identifiée.

4.4 Les fricatives

Elles sont plus facilement perçues que les occlusives et il n'y a pas de confusion quant au lieu d'articulation. Cela peut s'expliquer par le fait que les fréquences mises en valeur lors de la turbulence sont produites par la cavité en avant de la constriction et que les modifications chirurgicales du larynx ne se répercutent pas à ce niveau.

4.5 Le trait de nasalité

En français, le trait de nasalité est très résistant. Il n'y a pas eu de problèmes d'identification du trait de nasalité chez nos patients. Cette constatation avait aussi été faite par Doyle et al. [Doy88] après laryngectomie totale. La consonne /m/ ne requiert qu'une faible pression intra-orale et une durée relativement brève.

4.6 Les "glides"

/l/ et /r/ n'ont pas posé de problème d'identification, constatation faite aussi pour les voix alaryngées [Doy88].

CONCLUSIONS

Dans le cas des voix après LPSC, la modification du vibrateur néolaryngé est responsable de confusions de voisement affectant principalement les consonnes sonores au profit de leurs homorganiques sourdes. La montée du néolarynx et la suture de l'os hyoïde au cartilage cricoïde seraient responsables d'un conflit entre l'articulation et la phonation. La base de langue est indispensable au voisement en participant à l'occlusion néoglottique. Mais la réalisation articulatoire correcte des consonnes avec respect du lieu de constriction demande une certaine liberté de la langue mobile. Les patients semblent favoriser l'articulation consonantique au détriment du voisement.

Une analyse quantitative des indices acoustiques dont les caractéristiques spectrales et temporelles des consonnes

produites par les patients devrait être envisagée pour compléter cette étude.

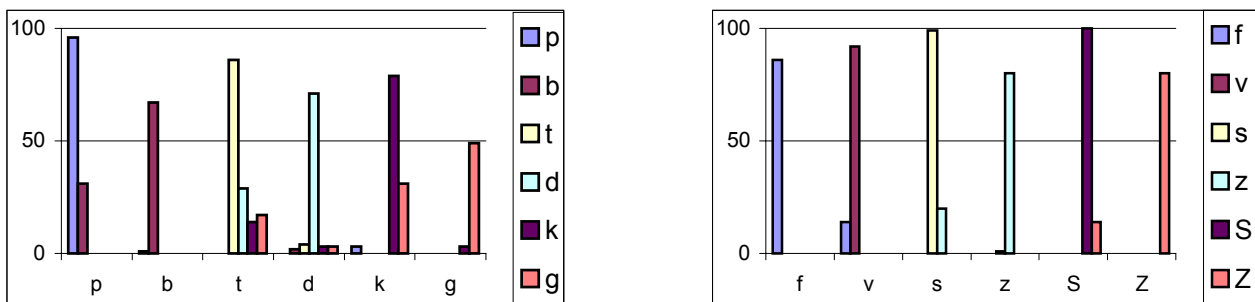


Figure 1 : Histogramme des matrices de confusions correspondant au trait de voisement pour les occlusives [p,b,t,d,k,g] et les fricatives [f,v,s,z,ʃ] en abscisse. En ordonnée, les % de réussite.

BIBLIOGRAPHIE

- [Kon84] Konrad H.R., Rattenbord C.C., Kain M.L., Barton M.D., Logan W.J., Holaday D.A. (1984). "Opening and closing mechanisms of the larynx". *Otolaryngology Head and Neck Surgery*, 92 : 401-405
- [Lud91] Ludlow C.L., Sedory S.E., Fujota M. (1991). Neurophysiological control of vocal fold adduction and abduction for phonation onset and offset during speech. In J. Gauffin & B. Hammarberg (Eds.), *Vocal Fold Physiology. Acoustic, perceptual and physiological aspects of voice mechanisms*. San Diego: Singular Publishing Group.
- [Cre95] Crevier-Buchman L., Laccourreye O., Weinstein G., Garcia D., Jouffre V., Brasnu D. (1995). "Evolution of speech and voice following supracricoid partial laryngectomy". *The Journal of Laryngology and Otology*, 109 : 410-413.
- [Cre98] Crevier-Buchman L., Laccourreye O., Wuyts F.L., Monfrais-Pfauwadel Mc., Pillot C., Brasnu D. (1998). "Comparison and evolution of perceptual and acoustic characteristics of voice after supracricoid partial laryngectomy with cricothyroidopiglottopexy". *Acta Otolaryngologica (Stockh.)*, 118 : 594-599.
- [Ste97] Stevens KN. (1997). Articulatory acoustic auditory relationships. In Hardcastle and Laver, *The handbook of phonetic sciences*, Blakwell, Oxford UK, (pp 464-506).
- Kew82] Kewley-Port D. (1982) "Measurement of formant transitions in naturally produced stop consonant-vowel syllables", *JASA*, 72 (2):379-389.
- [Gan83a] Gandour J., Weinberg B. (1983a). "Perception of intonational contrasts in alaryngeal speech". *Journal of Speech and Hearing Research*, 26 : 142-148.
- [Gan83b] Gandour J., Weinberg B., Garziona B. (1983b). "Perception of lexical stress in alaryngeal speech". *Journal of Speech and Hearing Research*, 26 : 418-424.
- [Doy88] Doyle P.C., Danhauer J.L., Reed C.G. (1988). "Listener's perception of consonants produced by esophageal and tracheoesophageal talkers". *Journal of Speech Hearing Disorder*, 53 : 400-407.
- [Bou95] Bourciez E., Bourciez J. (1995). *Phonétique Française*. Klincksieck Ed, Paris, France.
- [Hir96] Hirose H. (1996). "Voicing distinction in esophageal speech". *Acta Otolaryngologica (Suppl. 524)*, p : 56-63.
- [Doy89] Doyle P.C., Haaf R.G. (1989). "Perception of pre-vocalic and post-vocalic consonants produced by tracheoesophageal speakers". *Journal of Otolaryngology*, 18 : 350-353.
- [Lad83] Ladefoged P. (1983). The linguistic use of different phonation types. In D.M. Bless & J.H. Abbs (Eds.), *Vocal Fold Physiology: contemporary research and clinical issues* (pp. 351-360). San Diego: College Hill.
- [Vai97] Vaissiere J. (1997) "Phonological use of the larynx". *Proceedings Larynx 97*, Marseille, p : 115-125.
- [Lav92] Laver J., Hiller S., Mackenzie B.J. (1992). "Acoustic waveform perturbations and voice disorders". *Journal of Voice*, 6 : 115-126.
- [Kla90] Klatt D.H., Klatt L.C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers". *Journal of the Acoustical Society of America*, 87 : 820-857.