Productions vocaliques et consonantiques perturbées dans un cas complexe de chirurgie endobuccale

Florence Fauvet[1&2], Philippe Schultz[1] , Christian Debry[1], Fabrice Hirsch[2], Rudolph Sock[2].

[1]Service O.R.L. - Hôpitaux Universitaires de Strasbourg
1 av. Molière – 67098 Strasbourg – Cédex, France.
[2] E.A. 1339-Linguistique, Langues et Parole (LiLPa) – Composante Parole et Cognition
Institut de Phonétique de Strasbourg – Université Marc Bloch
22 rue Descartes – 67084 Strasbourg– Cédex, France.
Mél : florence.fauvet@chru-strasbourg.fr

ABSTRACT

Total glossectomy with reconstruction produces a serious functional defect of articulators. This study presents articulatory analyses of acoustic speech patterns at different postoperative stages, by comparing the pathological data with normal ones. Some vowels and plosive consonants appearing in VCV sequences (V=a), were recorded. Acoustic signal analysis was based on formant frequencies for vowels, and for consonants on burst spectrum and VOT (in reference to Klatt). Results are interpreted in terms of *adjustments* to the new configuration of the vocal tract, compensation mechanisms in obtaining closer articulatory-acoustic "targets", and evolution of productions through a five month period following surgery and rehabilitation.

1. INTRODUCTION

Partant d'un cas complexe d'opération de la cavité endobuccale avec glossectomie totale et reconstruction, nous nous intéressons aux mécanismes de compensation articulatoire ciblés sur la production des voyelles situées aux extrémités du triangle vocalique, [i], [a] et [u], et à celle des consonnes occlusives [p], [b], [t], [d], [k] et [g]. Nous avons entrepris ce travail avec les objectifs suivants : étudier les caractéristiques fréquentielle et temporelle des sons vocaliques et consonantiques, rendre compte de l'évolution des caractéristiques formantiques et articulatoires au cours et au terme de la réhabilitation orthophonique, et enfin mettre en relation nos observations avec la nature de la chirurgie reconstructrice. Pour cela, nous analyserons le signal acoustique de parole pathologique, en référence à celui d'un locuteur de contrôle.

L'analyse des voyelles est centrée sur les fréquences des deux premiers formants et sur la forme du triangle vocalique. Nous pouvons ainsi relier les valeurs formantiques à l'aperture du conduit vocal pour F1 et à sa variation dans l'axe antéropostérieur pour F2 [1], comparer les résultats à ceux du locuteur de référence et apprécier la capacité du patient à utiliser la configuration modifiée de son tractus vocal.

Pour l'étude des consonnes, et précisément les occlusives, nous cherchons à extraire du signal le burst et la durée du VOT selon Klatt [1] et à repérer les évènements singuliers en relation avec la pathologie, en référence à un sujet contrôle.

Nos hypothèses sont les suivantes : 1) la structure formantique des voyelles devrait être sensiblement différente de celle du locuteur de contrôle, notamment dans la phase post-opératoire immédiate ; 2) le timing du VOT, indice important pour la réalisation des occlusives et pour l'opposition voisée vs. non voisée, serait également perturbé ; 3) les ajustements à la perturbation chirurgicale, couplés à la rééducation orthophonique pourraient contribuer à améliorer les productions du sujet, améliorations qui devraient être visibles dans le signal acoustique.

2. PROTOCOLE

2.1. Corpus

Le corpus comporte la production tenue des voyelles isolées extrêmes [i, a, u]. Les consonnes [p, b, t, d, k,g] sont incluses dans des séquences [VCV], (V=a). L'entourage volontairement symétrique de la consonne minimise la complexité du mouvement articulatoire lingual. Chaque segment est répété dix fois dans un ordre aléatoire en une séance unique pour le locuteur de contrôle ; le nombre de restitutions autorise en conséquence le calcul d'une moyenne. Le patient, enregistré au cours de trois séances, dans un délai d'un mois et demi, à deux mois, puis à cinq mois après opération, n'a pu restituer qu'une ou deux fois chaque séquence. En effet, chez un locuteur pathologique, la maladie et le traitement thérapeutique (intervention, chimio- et radiothérapies) empêchent le recueil d'un grand nombre de répétitions [2].

2.2. Matériel et méthode

Les enregistrements du patient, acoustique et vidéos, ont été réalisés sur un enregistreur Marantz Professional Stereo PMD 60 (fréquence d'échantillonnage = 44,1 kHz), avec un microphone unidirectionnel Sennheiser E 845 S. et sur un camescope numérique Panasonic NV GS 15 (cassette mini D.V.), dans une salle de consultation. Les acquisitions du locuteur de contrôle ont été effectuées dans une chambre anéchoïque, équipée d'un enregistreur Fostex FR2 (fréquence d'échantillonnage = 44,1 kHz) avec un microphone unidirectionnel Sennheiser E 845 S, à l'Institut de Phonétique de Strasbourg.

2.2. Patient

Un patient de 52 ans, présentant une volumineuse tumeur pelvilinguale classée T4N2cM0 a bénéficié d'une pelviglossectomie avec mandibulectomie interruptrice emportant la totalité du plancher de la bouche, la symphyse mandibulaire et la langue. Les sections chirurgicales de la base de langue se situent au niveau des sillons amygdaloglosses et des vallécules qui sont respectées. Une reconstruction mandibulaire est réalisée à l'aide d'une prothèse en titane permettant de suspendre le larynx et de servir de support à la lèvre inférieure, afin de conserver la continence labiale. Un lambeau musculo-cutané de grand pectoral recouvrant la prothèse est suturé entre le vestibule buccal inférieur et les vallécules linguales. Une rééducation orthophonique débute quelques jours après le geste chirurgical et se termine deux mois plus tard.

2.3. *Analyse des données*

Les signaux acoustiques sont analysés spectralement et temporellement selon les techniques classiques de traitement du signal avec le logiciel Praat©.

3. RÉSULTATS DE L'ANALYSE ACOUSTIQUE

3.1. Analyse des voyelles

La figure 1 représente les moyennes des deux premiers formants (F1 et F2) de [i], [a], et [u] prononcées par un locuteur de contrôle (LC) et les mêmes voyelles produites par le patient (P), à un mois et demi (M1.5), à deux mois (M2) et à cinq mois (M5) post-opératoires.

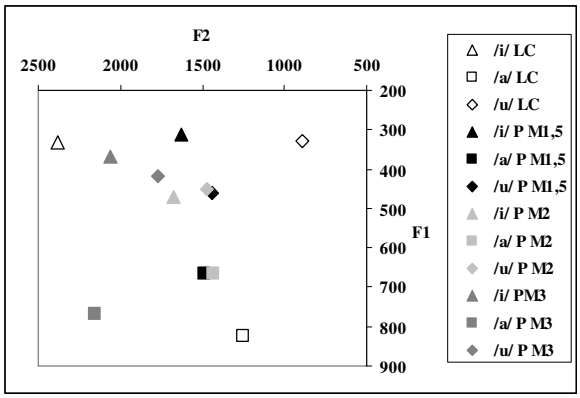

Figure 1 : Extrémités du triangle vocalique du locuteur de contrôle (**LC**) et chez le patient (**P**) à **M1.5**, **M2** et **M3**.

Nous observons, chez le locuteur de contrôle, des valeurs de F1 et de F2 conformes à la littérature [1], reportées ci-dessous (Table 1).

Table 1 : Valeurs des deux premiers formants des voyelles [i, a, u] chez le locuteur de contrôle.

Voyelles	F1	F2
[i]	332 Hz	2386 Hz
[a]	824 Hz	1256 Hz
[u]	329 Hz	892 Hz

Sur la figure 1, nous pouvons remarquer d'emblée les différences et l'évolution de l'aspect des "triangles vocaliques" du patient. L'aire des figures ainsi délimitées augmente progressivement dans l'intervalle d'un mois et demi à cinq mois après chirurgie. Si la surface du triangle vocalique du sujet pathologique reste nettement inférieure cinq mois après opération à celle du locuteur de référence, les voyelles sont mieux différenciées du point de vue acoustique au fur et à mesure des enregistrements.

L'examen détaillé des valeurs de F1 et de F2 extraites du signal pathologique, transposées sur la table 2, fait apparaître :

- une augmentation de F1 et de F2 entre M1.5 et M5 pour chaque voyelle, excepté un seul exemple d'aggravation, celui de F1, pour la voyelle [u], ayant diminué de 50 Hz.

- une variation essentiellement du degré d'aperture de la cavité buccale pour différencier [i, a, u] à M1.5, exprimée par la variation de F1. En revanche, les valeurs de F2 restent très proches pour [a] et pour [u], respectivement de 1486 et 1439 Hz. F2 de [i], voyelle plus aigüe, s'élève à 2062 Hz.

- une tendance à la normalisation de la production du [i], par rapprochement des valeurs de F1 et de F2 à celles du locuteur de contrôle, de M1.5 à M5. Le locuteur pathologique arrive à modifier de façon appropriée la forme de son conduit vocal. Il émet un son de 367 Hz (F1) et 2062 Hz (F2) pour une voyelle « cible » à 332 Hz (F1) et 2386 Hz (F2).
- une tendance au rapprochement de F1 simultanément à un éloignement de F2 des valeurs « cibles » pour [a] et pour [u], comme si l'un induisait l'autre, afin d'obtenir une distance suffisante entre les deux voyelles. Le sujet opéré peut difficilement contrôler les mouvements des articulateurs autres que la « langue » dans le sens antéropostérieur [4], trois mois après la fin de la réhabilitation orthophonique et à un moment où son état de santé s'altère.

3.2. Analyse des consonnes

Nous considérons dans un premier temps les consonnes occlusives de production plus aisée (ne recrutant pas l'intervention active de la « langue », à savoir les bilabiales. Seules les séquences contenant une bilabiale ont été réalisées à chaque fois, avec une fermeture totale et momentanée du conduit vocal [4].

L'analyse de la durée relative du VOT pour l'occlusive [p] dans la séquence [apa], permet de

constater un allongement de la durée du VOT chez le locuteur pathologique lors des deux premiers enregistrements, puis une diminution telle que sa valeur finit par devenir similaire à ce qui a pu être constaté pour le locuteur non-pathologique.

Ainsi, nous observons que le V.O.T. passe de 15 % à M1,5 et M2, avant d'être mesuré à 6 % à M5, résultat comparable au locuteur de contrôle, pour qui le V.O.T. a été évalué à 6% également en moyenne.

Table 2 : Valeurs absolues et relatives du V.O.T. de Klatt pour les séquences [pa] produites par le locuteur de contrôle (LC) et par le patient un mois et demi après l'opération (M1,5), deux mois (M2) puis cinq mois (M5) après l'opération.

pa	Valeurs absolues ms	Valeurs relatives %
M1,5	21	15
M2	27	15
M5	15	6
LC	15	6

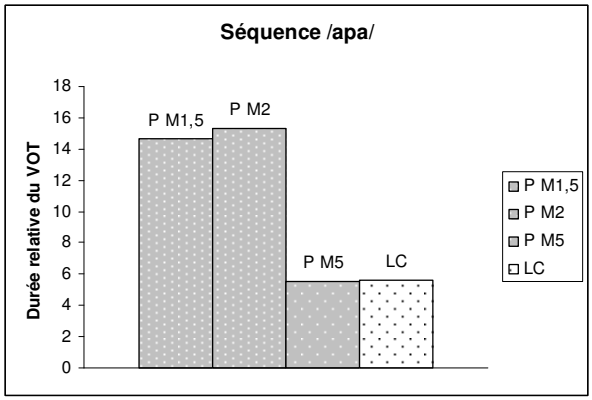

Figure 2 : Étude de la durée relative du VOT chez le patient (P), à un mois et demi (M1.5), à deux mois (M2) et à cinq mois (M5) post-opératoires, en référence à un locuteur de contrôle (LC).

L'articulation d'une occlusive apico-dentale requiert une élévation de l'apex et sa mise en contact avec les dents du maxillaire pour fermer le conduit vocal [5], ce qui est à priori impossible pour un sujet tel que celui de notre étude. Les données montrent une occlusion au premier enregistrement (voir Fig. 3), remplacée par un bruit de friction [6] constant à environ 3900 à 4000 Hz, tout au long de sa production, au dernier enregistrement (voir Figure 4).

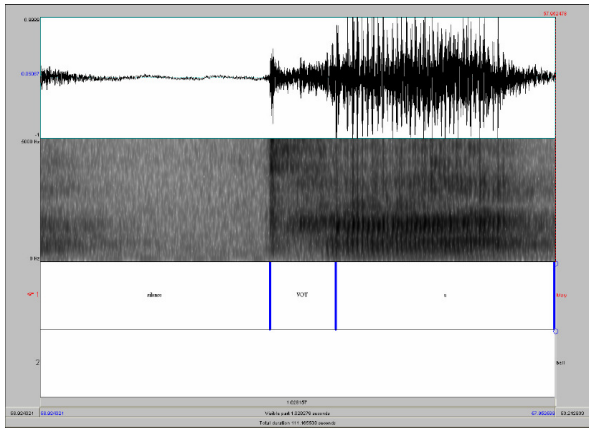

Figure 3 : séquence [ta] à M1.5.

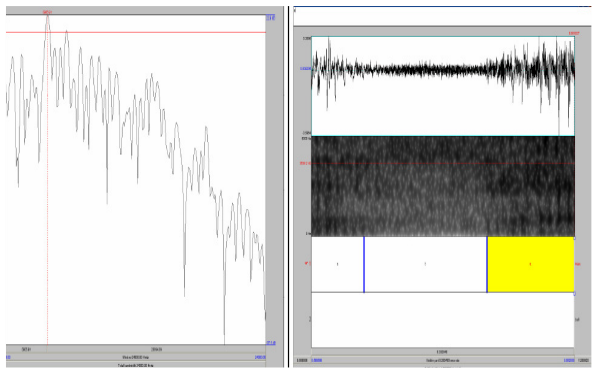

Figure 4 : séquence [ta] à M5. A gauche : une coupe spectrale au milieu du bruit de friction du « [ta] » ; à droite le signal acoustique temporel.

Le remplacement de l'occlusion alvéo-dentale par un bruit de friction se confirme à M5, consonne voisée ou non. Il en va de même pour les occlusives vélaires [k] et [g].

L'observation des signaux de parole lors de la production des consonnes sourdes et sonores ne marque pas de différence notable. Ce trait, non pertinent pour notre patient, ne peut être expliqué, de manière simple, par une conséquence de la chirurgie. Nous pensons que la neutralisation de cette opposition voisée *vs.* non voisée tient, soit à l'influence dialectophone du sujet (l'opposition de sonorité n'étant pertinent en alsacien), soit à une difficulté à retrouver une coordination oro-laryngée propice à l'émergence du contraste phonologique [7].

INTERPRÉTATION DES RÉSULTATS

Plusieurs facteurs peuvent concourir au glissement du triangle vocalique vers l'avant. La nature de la chirurgie a pour conséquence la diminution de la longueur du pharynx : l'augmentation des fréquences de résonances étant directement proportionnelle à la diminution de la longueur du conduit vocal, les valeurs fréquentielles des formants seront plus élevées [1]. La réduction progressive du lambeau myo-cutané

sur une durée de deux à trois mois transforme la géométrie de la cavité buccale, en agrandissant notamment sa partie antérieure. Enfin, les procédés de compensation renforcent les aigus des formants F1et F2.

Logiquement, l'immobilité de la « langue » ne peut rendre possible la configuration requise pour la production des voyelles postérieures comme le [u] avec la masse linguale tirée vers le bas et l'abaissement du larynx. La labialisation se traduit en principe par une baisse de F2 [1], or si nous observons une aggravation, c'est celle de F1 et non celle de F2, dont la valeur augmente (voir table 2).

Le locuteur pathologique, afin de différencier [i] et [u], fait intervenir des variations qui n'apparaissent pas en parole standard, comme celles de F1, soient des variations de l'aperture du conduit vocal.

Les *perturbations* temporelle et articulatoire observées corroborent celles de la littérature avec l'allongement du VOT, et la production d'un bruit de friction en place d'une occlusion [2]. Mais ici, par la réduction progressive de sa durée, le VOT pathologique approche et atteint la valeur d'un VOT habituel.

4. CONCLUSION

Certaines de nos observations confirment nos hypothèses de départ. Les analyses acoustiques montrent une perturbation du signal de parole pathologique, aussi bien spectrale que temporelle, minime ou sévère, selon les phonèmes considérés. Leur examen détaillé fait apparaître une variabilité des caractéristiques dues à la pathologie, lors des enregistrements à plus ou moins longue échéance de la chirurgie. Nous observons, au fur et à mesure des enregistrements, une tendance à la normalité de certaines valeurs formantiques ou du VOT, lorsque c'est possible, par des déformations appropriées du conduit vocal couplées au signal glottique. Ces résultats pertinents justifient le recueil et l'exploitation des données acoustiques, à différents moments postopératoires, pour comprendre les mécanismes de compensation articulatoires dans de tels cas.

Table 2 : Évolution des valeurs de F1 et de F2 chez le patient à M1.5, M2 et M5 de la chirurgie et en comparaison avec le locuteur de contrôle.

Voyelles	**[i]**		**[a]**		**[u]**	
Périodes	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
M1,5	**311 Hz**	**1633 Hz**	**665 Hz**	**1486 Hz**	**461 Hz**	**1439 Hz**
M2	**471 Hz**	**1675 Hz**	**665 Hz**	**1437 Hz**	**451 Hz**	**1472 Hz**
M5	**367 Hz**	**2062 Hz**	**767 Hz**	**2154 Hz**	**418 Hz**	**1768 Hz**
Locuteur de contrôle	**332 Hz**	**2386 Hz**	**824 Hz**	**1256 Hz**	**329 Hz**	**892 Hz**

Remerciements

La participation de A. Gentine et de G. Brock a contribué à la réalisation de ces travaux soutenus par un Programme ACI TTT du Ministère de la Recherche, 2003-2006, attribué à l'Institut de Phonétique de Strasbourg.

BIBLIOGRAPHIE

[1] Calliope. *La parole et son traitement automatique.* Masson, Paris, France, 1989.

[2] J. Lebeau, C. Savariaux, P. Perrier, G. Bettega, B. Raphaël. Évaluation fonctionnelle des reconstructions endobuccales. Un outil intéressant : l'interprétation du signal acoustique. *Revue de Stomatologie et de Chirurgie maxillofaciale,* 101, n°2:60-64, 2000.

[3] W. Wakumoto, K. Ohno, S. Imai, Y. Yamashita, H. Akizuki, K. I. Michi. Analysis of the articulation after glossectomy. *Journal of oral rehabilitation ,* 23:764-770, 1996.

[4] A.J. Greven, M. F. Meijer and F. R. Tiwari. Articulation after total glossectomy: a clinical study of speech in six patients. *European Journal of Disorders of Communication,* 29: 85- 93, 1994.

[5] J. Vaissière. *La phonétique.* Puf, Paris, France, 2006.

[6] C. Savariaux, P. Perrier, J. Lebeau, G. Magan)a, and C. Dorange-Pattoret. Production de parole après traitements de cancers de la cavité endobuccale. In *XXIIIèmes Journées d'Étude sur la Parole,* pages 433-436, 2000.

[7] K. Stevens. *Acoustic Phonetics*. M.I.T. Press, Cambridge, MA, 1999.