

CORRÉLATIONS ENTRE LES MOUVEMENTS ARTICULATOIRES ET LES VARIATIONS FORMANTIQUES

LAURENT SANTERRE

La corrélation dont je présente ici les grandes lignes a été établie à partir de l'analyse comparée sur les plans articulatoire et acoustique de 280 phrases de quatre syllabes prononcées par deux locuteurs francophones de Montréal.

Les radiofilms, de 48 images à la seconde en moyenne, sur le plan sagittal médian, ont été tracés image par image et mesurés à une dizaine de points à la fois. La prise de son a été faite au moyen d'un microphone protégé contre le bruit des appareils de prise de vue et elle a permis une analyse spectrale satisfaisante; le rapport signal-bruit y est de l'ordre de 40 dB.

La prise de vue et la prise de son se sont trouvées électroniquement synchronisées sur une bande magnétoscopique de circuit fermé de télévision. Cette synchronisation ne dispense pas de celle qu'il faut faire ensuite entre les séquences filmées et l'analyse spectrale. Cette dernière synchronisation consiste à faire coïncider étroitement la course du film avec celle des sonagrammes, de manière à faire correspondre chaque image du film avec la section sonographique qui représente la production acoustique à ce moment précis.

Pour réduire le coefficient d'erreur, j'ai pris, comme points de repère, des occlusives situées aux deux extrémités d'une même séquence sur le film et sur la bande enregistrée. Exemple: soit une séquence dont la première phrase est *Donne un coup de râpe* et la 10e, *Les pêches sont vertes*; entre l'explosion du [d] et celle du [p] de la dernière phrase, je compte 800 images sur le film et, au moyen d'un segmentateur, 317,5 cm sur la bande magnétique. La vitesse exacte de déroulement de bande du magnétophone est facile à préciser au moyen d'un chronographe; il suffit de mesurer la bande enregistrée en 100 ou 200 secondes et de diviser le nombre de millimètres par le nombre de centisecondes; dans l'exemple qui nous occupe, j'ai trouvé 190 mm/sec. Les 3175 mm de bande entre [d] et [p] ont été enregistrés en 16,71 secondes, ce qui donne 47,87 images par seconde sur le film. On peut faire ces calculs sur toutes les séquences du film, si on le juge utile.

Le temps d'une image est ici de 2,089 cs et représente 3,969 mm de bande magnétique. Il faut maintenant connaître la vitesse du sonographe. Il faisait 100 tours en 239 secondes, soit un tour en 2,39 sec. La longueur des sonagrammes est de 318 mm,

ce qui donne une course de 1,33 mm d'enregistrement par centiseconde. Dans ce cas, le temps d'une image de film, soit 2,089 cs, correspond à 2,77 mm sur le sonagramme. Il faut donc faire correspondre chaque image du film avec des segments de 2,77 mm sur le sonagramme.

Dans la synchronisation de chaque sonagramme, on est guidé par les éléments transitoires comme les explosives; une fois fixée une image du film sur sa tranche spectrale, tout le reste de la phrase se trouve synchronisée. D'ordinaire d'un bout à l'autre de la phrase, les contraintes imposées par les distances entre les mouvements articulatoires et les transitoires acoustiques ne laissent pas plus de 1 cs de liberté dans la synchronisation d'une durée de 2 secondes.

Cette méthode permet de suivre point par point les mouvements articulatoires et les variations formantiques correspondantes. L'analyse de l'ensemble du corpus m'a permis de dégager les constantes de la corrélation articulatoire et acoustique, dont je me contenterai de rappeler plus loin les grandes lignes pour vérifier la validité d'un modèle théorique de résonance électroacoustique qui permet de schématiser la corrélation.

1. RÉSONATEUR ACOUSTIQUE ET ÉLECTRIQUE

Chiba et Kajiyama (1941), puis Fant (1960, 1968) et Flanagan (1965) comparent le canal buccal à un résonateur simple de 17,5 cm de longueur. Il est fermé du côté de la glotte et ouvert du côté des lèvres; en termes électroniques, on dit que l'impédance est maximale à la glotte et minimale aux lèvres, ce qui occasionne un maximum de courant ou d'amplitude vibratoire à l'embouchure, et un maximum de pression ou de voltage à la glotte.

On sait que là où le courant est maximum, le voltage est minimum, et que là où le courant est minimum, le voltage est maximum. La résonance doit être de telle sorte que le maximum de courant se trouve du côté ouvert, là où l'air est libre, à la pression atmosphérique; à l'autre extrémité du résonateur, l'air n'est pas libre, mais soumis à un maximum de pression, et le courant vibratoire, ou la vélocité, y est à son point minimum. Ce déphasage de 180° entre la pression et le courant se retrouve à tous les points du résonateur.

Parce que le courant doit être maximum du côté ouvert et minimum du côté fermé, la longueur du résonateur ne peut représenter que le quart de la longueur de l'onde dont il est le siège d'origine; cette condition est vérifiée chaque fois que le résonateur représente 1/4, ou 3/4, ou 5/4 de la longueur d'onde.

$$17,5 \text{ cm} = 1/4 \lambda, \text{ ou } 3/4 \lambda, \text{ ou } 5/4 \lambda; \lambda = 70 \text{ cs, ou } 23 \frac{1}{3} \text{ cm, ou } 14 \text{ cm.}$$

Selon la formule $\lambda = \frac{V}{F}$ (où $V = 350 \text{ cm/sec.}$), $F = 500 \text{ Hz, ou } 1500 \text{ Hz, ou } 2500 \text{ Hz.}$ On sait qu'un résonateur peut avoir simultanément plusieurs fréquences de résonance propres. Ce sont les fréquences de F_1 , F_2 et F_3 de ce résonateur théorique

régulier. On peut considérer qu'il ferait entendre le schwa, comme la bouche qui présente les positions articulatoires sans abaissement du voile du palais, ni labialité, entre [ɛ] et [e] (Figure 1).

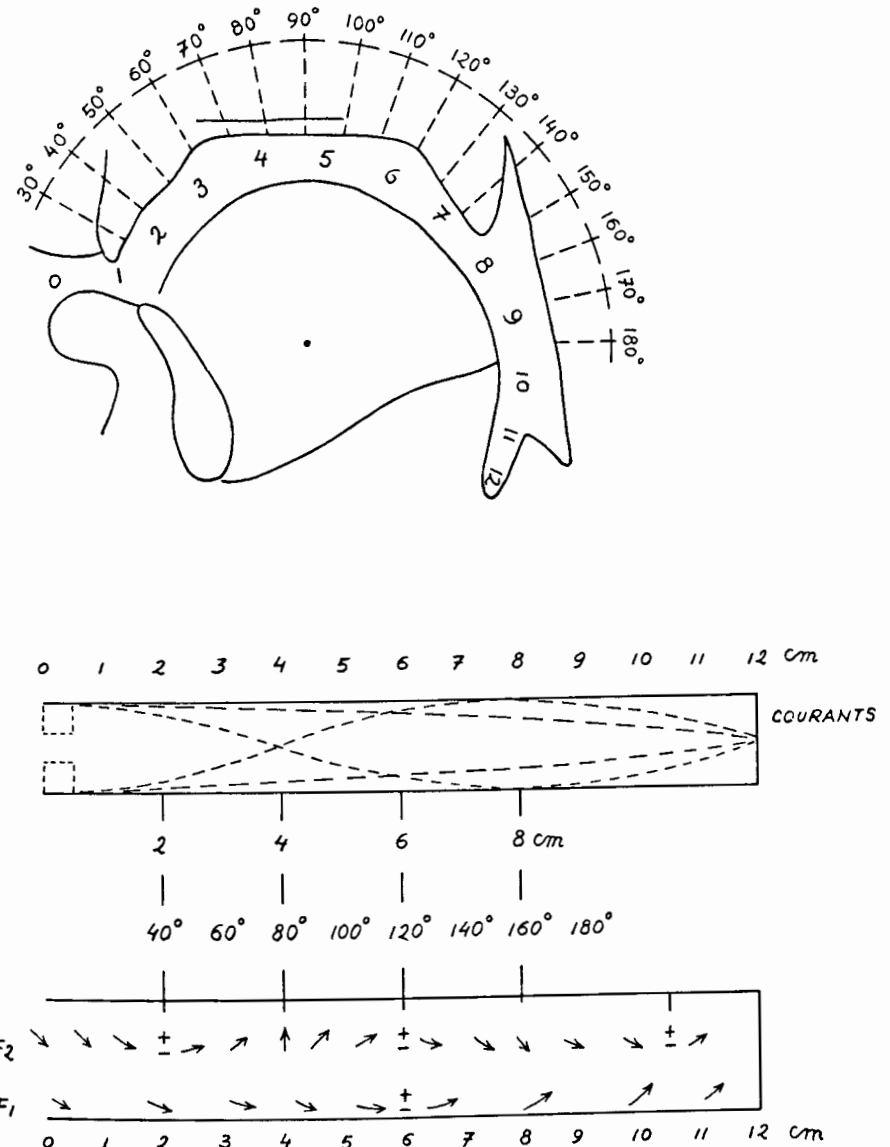


Fig. 1. Coupe radiocinématographique sur le plan sagittal médian des positions articulatoires du Schwa. Les mesures d'angles sont faites à partir d'un point central à égale distance de la pointe des incisives supérieures et du palais dur. Les flèches indiquent pour F_1 et pour F_2 l'abaissement ou le relèvement de la fréquence à partir de 500 Hz et de 1500 Hz, quand le résonateur est rétréci; en cas de dégagement au lieu de rétrécissement aux points marqués, le sens des flèches est à inverser. Les points marqués \pm sont neutres, parce que le courant et la pression y sont égaux.

2. LOIS DE LA RÉSONANCE

Elles sont les mêmes pour les résonateurs mécaniques, acoustiques ou électriques :

$$F = \frac{1}{2\pi\sqrt{\frac{m}{c}}} = \frac{1}{2\pi\sqrt{\frac{i}{v}}} = \frac{1}{2\pi\sqrt{LC}}$$

La fréquence de résonance d'un système oscillant est directement proportionnelle à la racine carrée de la rigidité (C) du système (ressort), ou de la pression ou voltage (V), et inversement proportionnelle à la racine carrée de la masse (m), ou de la quantité d'électrons mise en mouvement (i).

Dans les oscillations électriques, la pression et le courant sont distribués sous forme d'énergie potentielle aux bornes d'un condensateur (C) et d'énergie cinétique dans une bobine de self-induction (L). Dans ces termes, la fréquence est inversement proportionnelle à la racine carrée du produit de L (Henry) et de C (farad).

La relation $\lambda = \frac{V}{F}$ et $F = \frac{V}{\lambda}$ montre que la fréquence est inversement proportionnelle à la longueur du résonateur, puisque cette dernière représente un nombre impair de quarts de longueur d'onde du système. Par exemple, si le résonateur buccal s'allonge par labialité de 17,5 à 18 cm, $\lambda = 18 \times 4 = 72$ cm, et $F = 486$ Hz au lieu de 500 Hz.

Fant (1968:216) dit qu'une constriction en un point du résonateur où prédomine le courant abaisse la fréquence propre du résonateur, tandis qu'une constriction en un point où prédomine le voltage a pour effet de relever la fréquence; on sait

en effet, par la formule $F = \frac{1}{2\pi\sqrt{\frac{i}{V}}}$, que ce qui gêne ou diminue le courant (i)

abaisse la fréquence, et que ce qui ajoute à la pression ou au voltage (V) élève la fréquence. Il s'agit de savoir si un rétrécissement du résonateur en un ou plusieurs points donnés ajoute plus à la pression qu'il n'enlève au courant, ou au contraire, s'il enlève plus au courant qu'il n'ajoute à la pression.

La répartition du courant selon la longueur d'onde est représentée au centre de la Figure 1; on peut en déduire la répartition du voltage, qui est inversement proportionnelle à celle du courant; en certains points, le courant est prédominant; en d'autres points, c'est le voltage qui prédomine; enfin en certains autres points, le courant et le voltage sont égaux. Autrement dit: l'énergie cinétique ($W_{ci} = \frac{1}{2} Li^2$) est supérieure, ou inférieure ou égale à l'énergie potentielle ou statique ($W_{po} = \frac{1}{2} CV^2$).

Ainsi, en régime de résonance de basse fréquence, (17,5 cm = 1/4 λ ; $F = 500$ Hz) soit $F1$, un rétrécissement du résonateur dans la partie qui va de l'ouverture au centre

(points 0 à 6 sur la Figure 1) contrarie un maximum de courant et abaisse la fréquence; cette action est beaucoup plus marquée à l'embouchure et va s'atténuant jusqu'au point 6 où le courant est égal au voltage. Dans la partie postérieure du résonateur, les constriction ont peu d'influence sur le courant qui est déjà plus faible que la pression; c'est alors sur cette pression surtout qu'elles agissent pour faire relever la fréquence.

En régime de $F2$, le rétrécissement du résonateur à l'ouverture abaisse la fréquence; au point 2 (à 40° sur les axes de mesures articulatoires, Figure 1) où le courant est égal au voltage, la constriction est sans effet; il en est de même au point 6 (120°); au point 4 (80°), la constriction élève considérablement la fréquence, le courant étant à peu près nul; au point 8 (160°), le courant est maximal et toute entrave y fait baisser la fréquence.

La Figure 1 met en parallèle le résonateur théorique et le canal buccal. Les lignes pointillées dans le résonateur représentent les valeurs maximales ou minimales du COURANT selon le régime de résonance: $F1$: — — —, $F2$: - - - -. Au bas de la figure, les flèches indiquent les mouvements ascendants ou descendants de $F1$ et de $F2$ selon les points de RÉTRÉCISSEMENT. En cas de dégagement au lieu de rétrécissement, le sens des flèches doit être inversé. La longueur du canal buccal et du résonateur est ici de 12 cm au lieu de 17,5 cm, parce que j'ai réduit aux 2/3, sur les tracés articulatoires, les dimensions réelles de la bouche de mes informateurs.

Comme les mouvements articulatoires n'affectent jamais un seul point, mais plusieurs zones du canal buccal à la fois, il y a lieu de voir dans chaque cas si les variations de fréquences aux différents points s'additionnent ou s'annulent. Prenons comme exemple les zones touchées par les articulations successives des voyelles antérieures de [ε] à [i], (Figure 2). La langue s'élève sur les axes 70° et 80° et les zones adjacentes; un rétrécissement sur ces axes fait monter $F2$ et descendre $F1$; de l'autre côté, l'agrandissement du canal pharyngal aux zones situées autour de 160° et 170° fait aussi monter $F2$ et descendre $F1$: les variations s'additionnent donc pour faire des voyelles antérieures des timbres à formants de plus en plus distants, de [ε] à [i].

Pour les voyelles postérieures [ɔ], [o] et [u], la masse de la langue se déplace vers les zones articulatoires situées autour de l'axe 130°; $F2$ descend et $F1$ monte; le dégagement sur les axes antérieurs fait aussi baisser $F2$ et monter $F1$; mais en s'élevant pour [o] et [u], la langue libère le passage pharyngal inférieur, ce qui fait au contraire baisser $F1$. Cette influence abaissante sur $F1$ est appuyée par la projection et la fermeture labiales qui accompagnent les voyelles postérieures; $F1$ est plus abaissé par l'allongement du canal buccal et la fermeture antérieure qu'il n'est relevé par le rétrécissement vélaire; la tendance de $F1$ à monter par rétrécissement sur 130° est d'ailleurs faible; elle est nulle à 120°, elle serait plus forte à 160°. Du côté des lèvres, les variations sont beaucoup plus marquées que du côté de la glotte où l'impédance ne peut varier dans de grandes proportions.

Les [a] s'expliquent par la grande ouverture; par rapport à la position neutre [ə], les axes antérieurs sont un peu dégagés, ce qui fait monter $F1$ et baisser $F2$. De

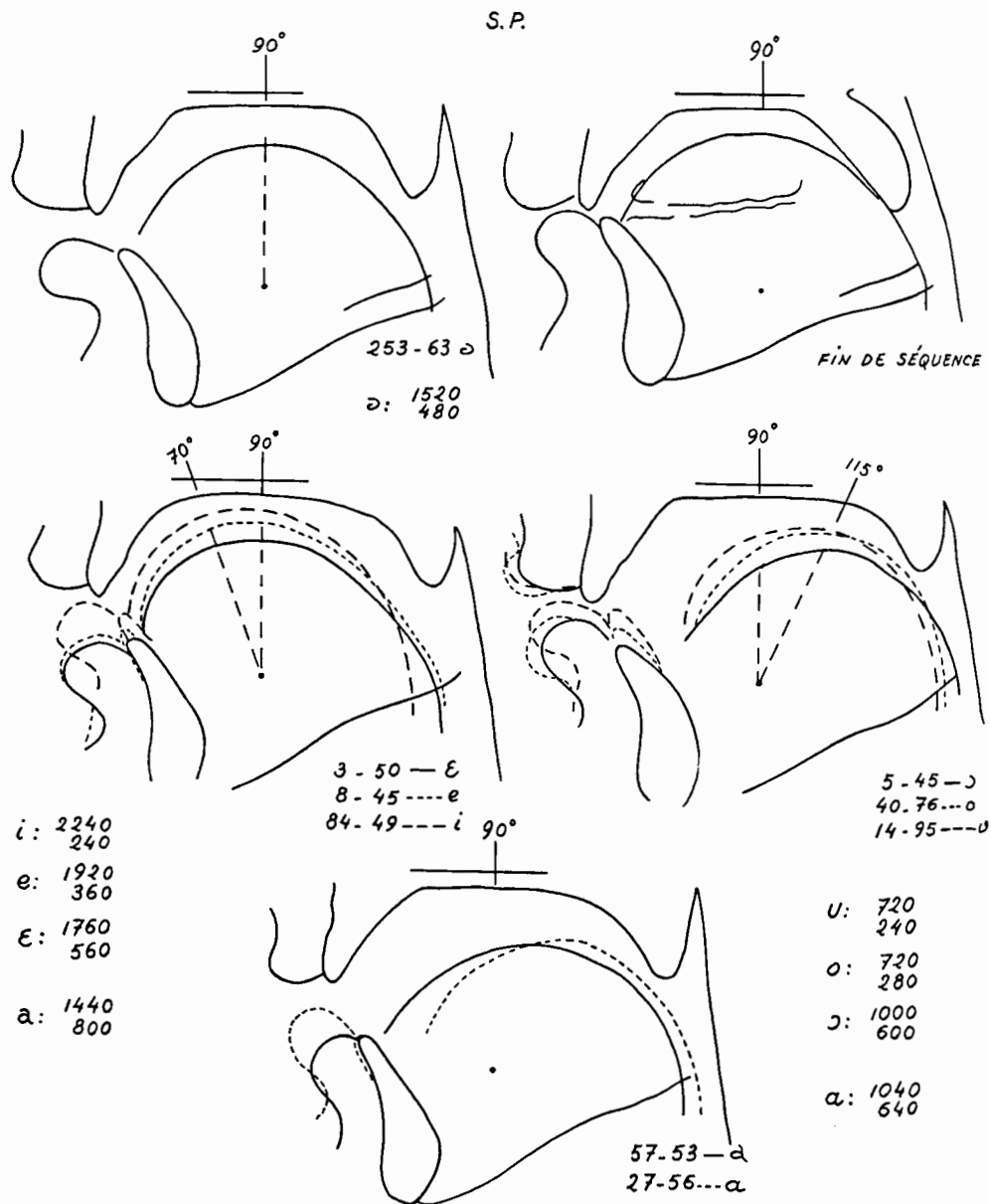


Fig. 2. Les numéros de phrases et d'images renvoient à la documentation filmée dont l'ensemble paraîtra dans *Les voyelles orales dans le français parlé à Montréal*, Paris, Klincksieck, 1972.

[a] à [a], il y a recul de la langue sur les axes pharyngaux, d'où abaissement de $F2$ et relèvement léger de $F1$; mais ce relèvement de $F1$ est contrarié et renversé par la projection labiale.

3. CORRÉLATION ARTICULATOIRE ET FORMANTIQUE

Il serait trop long de présenter ici les variations formantiques que j'ai étudiées parallèlement aux mouvements articulatoires dans l'ensemble des 560 phrases du corpus. Je ne puis que rappeler les grandes lignes de la corrélation pour la rapprocher du fonctionnement du résonateur théorique.

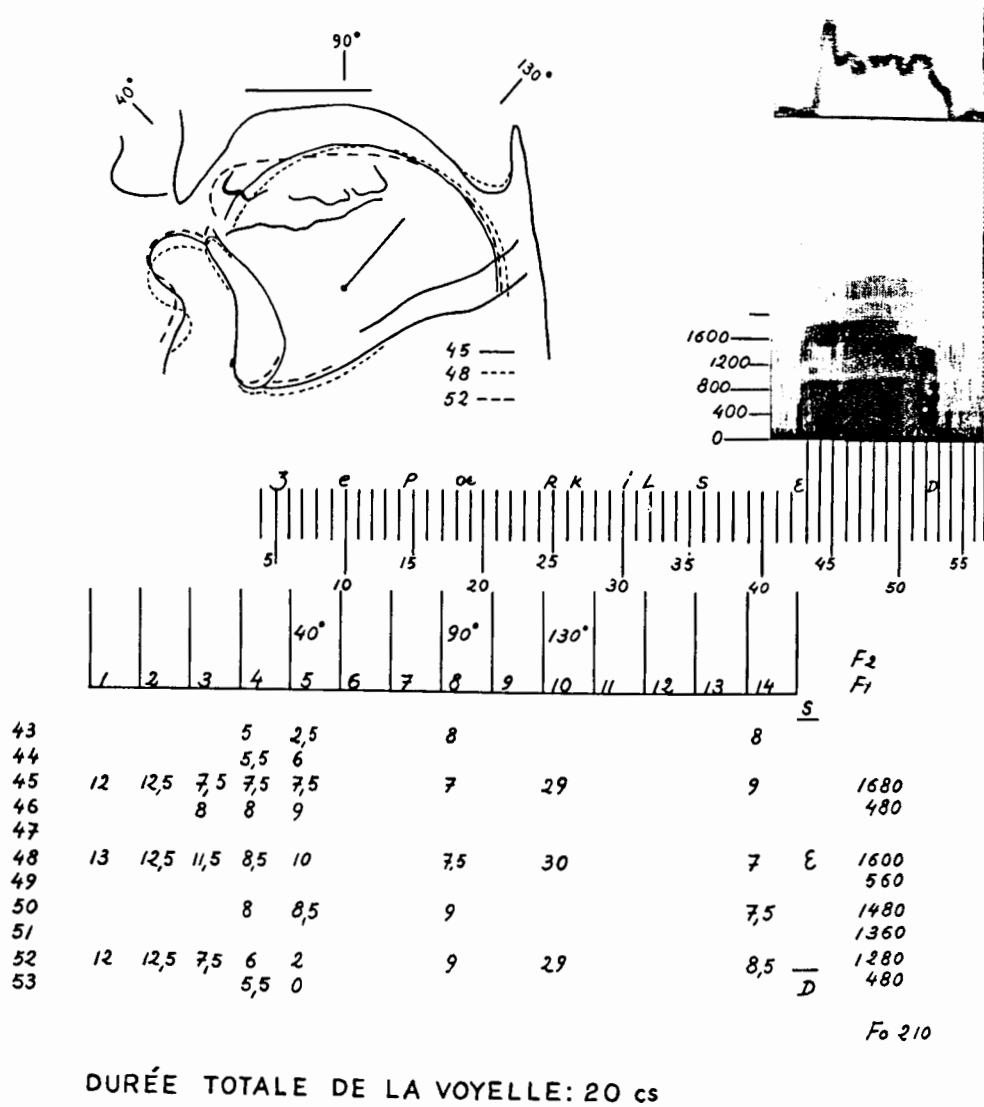
1. *Les lèvres*. — La fermeture labiale abaisse aussi bien $F2$ que $F1$, comme le montre le résonateur. Les consonnes bilabiales et labiodentales n'ont pas de locus acoustique qui serait leurs fréquences propres, mais elles ne font qu'abaisser les résonances imposées par la langue pour les voyelles; on le voit clairement au fait que les fréquences à partir de ces consonnes (CV) et à l'arrivée sur ces mêmes consonnes (VC) se situent entre 800 Hz et 2000 Hz, selon les voyelles d'accompagnement. Je néglige de distinguer entre [P], [B] et [M] ou [F] et [V], parce que je m'en tiens à la corrélation globale.

2. *Axes 30° - 50°, [t, d, n, s, z, l]*. — Les constrictions ou les occlusions du résonateur buccal dans la zone située autour de l'axe de 40° ne sont pas critiques pour les fréquences de $F2$, mais elles ont un effet abaisseur très marqué sur $F1$. Pour les deux régimes vibratoires, c'est la forme et la position de la langue aux autres points, sous le palais dur surtout (axe de 90°) au moment de l'implosion ou de l'explosion qui déterminent les fréquences des transitions. Les locus acoustiques des consonnes, points fixes vers lesquels se dirigeraient les formants de voyelles, ne sont pas des traits consonantiques utiles dans l'analyse de la parole réelle, dans la mesure où c'est toujours la forme et la position de la langue qui déterminent toutes les fréquences, aussi bien celles des transitions que celles des voyelles. Le seul fait de poser l'apex sur les alvéoles abaisse $F1$ aux plus basses fréquences, mais n'influence pas $F2$; quand la transition est remontante après la consonne (CV), c'est que la langue remonte vers le palais dur durant la transition, (Figures 3 et 4).

3. *Axe 50°, [ʃ] et [ʒ]*. — Le rétrécissement du résonateur buccal dans la région située autour de l'axe de 50° influence un peu à lui seul la fréquence de résonance des transitions sur $F2$, et font baisser $F1$. Les transitions vers la consonne (VC) sur $F2$ sont hautes ou basses selon que $F2$ est haut ou bas, et elles sont légèrement remontantes.

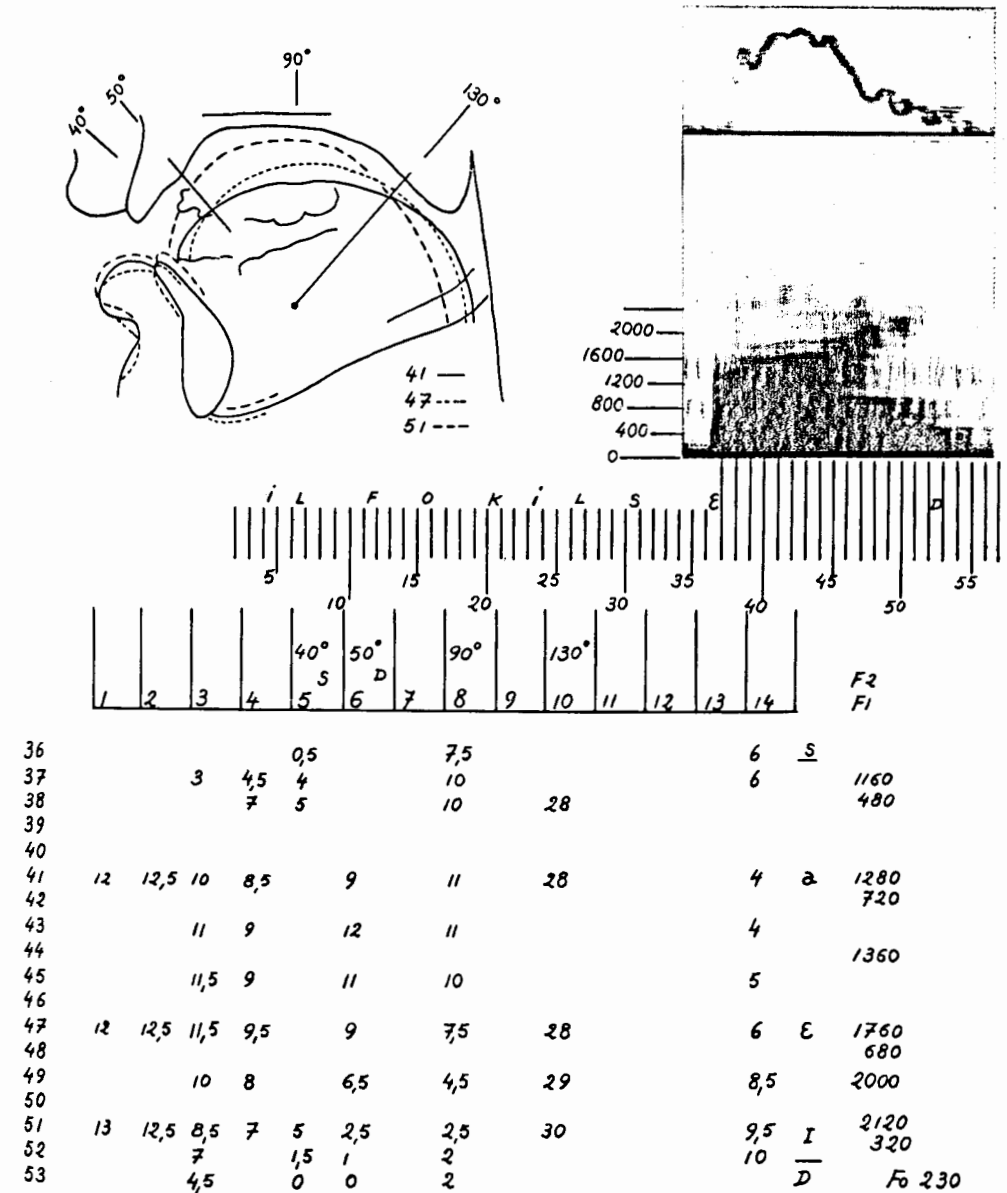
4. *Axes 80° - 110°, [k], [g], [ŋ]*. — Le rétrécissement du canal buccal dans cette région abaisse $F1$.

Pour $F2$, il faut distinguer entre les voyelles antérieures et les voyelles postérieures. Pour les premières, l'occlusion, qui peut se faire autour de l'axe de 85°, occasionne une élévation très marquée de $F2$ et une chute de $F1$; pour les secondes,



DURÉE TOTALE DE LA VOYELLE: 20 cs

Fig. 3. Ordre des mesures articulatoires: (1) projection de la lèvre inférieure; (2) projection de la lèvre supérieure; (3) distance entre les lèvres; (4) distance entre les incisives; (5 à 10) distances de la langue aux incisives, aux alvéoles, au palais dur, au voile du palais, sur des axes déterminés par leur angle; (11) distance de la langue à la luette; (12) ouverture des fosses nasales; (14) distance de la langue à la paroi pharyngale. Les mesures sont exprimées en millimètres et correspondent aux 2/3 des dimensions réelles.



DURÉE TOTALE DE LA VOYELLE: 33cs

Fig. 4. Ordre des mesures articulatoires: (1) projection de la lèvre inférieure; (2) projection de la lèvre supérieure; (3) distance entre les lèvres; (4) distance entre les incisives; (5 à 10) distances de la langue aux incisives, aux alvéoles, au palais dur, au voile du palais, sur des axes déterminés par leur angle; (11) distance de la langue à la luette; (12) ouverture des fosses nasales; (14) distance de la langue à la paroi pharyngale. Les mesures sont exprimées en millimètres et correspondent aux 2/3 des dimensions réelles.

l'occlusion, qui peut se faire autour de l'axe de 105°, fait aussi remonter la fréquence de F_2 , mais moins considérablement; l'action sur F_1 est aussi moins marquée; c'est surtout la projection et la fermeture labiales accompagnant les voyelles postérieures qui abaissent considérablement les fréquences du résonateur.

Cette simplification des points d'articulation des occlusives palatales selon les voyelles d'accompagnement antérieures ou postérieures ne se vérifie pas toujours dans la réalité; j'ai des exemples de [k] articulés sur 110° avec [ɛ], comme dans *Tu viens avec*.

5. *Axes 130° - 140° [R]*. — L'articulation du [R] uvulaire, quand elle est assez ferme, permet d'examiner l'effet de la constriction dans cette zone, indépendamment de la labialité qui accompagne les voyelles postérieures françaises. Les transitions de [R] sur F_1 sont plus hautes que le formant de la voyelle, (VC) à moins qu'il soit déjà très haut; la constriction est abaissante sur F_2 , même quand il est déjà bas.

6. *Axe 180°, passage pharyngal*. — Le recul de la langue de [a] à [ɑ] fait baisser F_2 , mais cela est aussi l'effet de la labialité. F_1 devrait remonter, selon le résonateur théorique, s'il n'était empêché de le faire par cette labialité. Tout ce que je puis affirmer ici, c'est que la parole réelle ne contredit pas les prévisions du résonateur théorique, mais elle ne permet pas non plus de le vérifier, à cause de l'interférence contraignante des lèvres.

4. CONCLUSION

La corrélation dégagée, au moyen de radiofilms de la parole réelle et de sonagrammes synchronisés, entre les mouvements articulatoires et les variations formantiques est parallèle à une autre corrélation établie entre les déformations d'un résonateur théorique simple et les variations de fréquences qui en résultent. Le résonateur buccal peut être symbolisé par un modèle électro-acoustique. Il est bien entendu que la parole réelle n'est pas réductible aux simplifications d'un modèle théorique aussi simple, mais les lois acoustiques qui s'y exercent sont les mêmes partout, quoique dans des conditions d'application plus complexes pour le canal buccal.

Il ressort de la corrélation établie pour la parole réelle qu'à tout moment, les fréquences du résonateur buccal sur F_1 et sur F_2 sont fonctions de sa longueur, de son ouverture et de sa forme, aussi bien au cours des transitions entre consonnes et voyelles que durant les voyelles elles-mêmes. La notion de locus, qui seraient des points fixes propres indépendants des voyelles qui précèdent et qui suivent la consonne, est inutile dans l'analyse de la parole réelle. Les transitions pour une même consonne sont très différentes selon qu'il s'agit d'une position implosive ou explosive et surtout selon la fréquence des formants des voyelles environnantes.

*Département de linguistique et des langues modernes
Université de Montréal*

BIBLIOGRAPHIE

- Chiba, T. et M. Kajiyama
1941 *The Vowel — its Nature and Structure* (Tokyo).
Fant, G.
1960 *Acoustic Theory of Speech Production* (La Haye, Mouton).
1968 "Analysis and Synthesis of Speech Processes", in *Manual of Phonetics*, Bertil Malmberg, ed. (Amsterdam, North-Holland Publishing Co.), pp. 173-277.
Flanagan, J.L.
1965 *Speech Analysis, Synthesis and Perception* (Berlin, Heidelberg, New York, Springer-Verlag).
Voir aussi l'article de B.E.F. Lindblom, "Phonetics and the Description of Language", pp. 63-97 de ce volume.

DISCUSSION

TAGGART (Montréal)

Veillez expliquer l'affirmation que les 'locus' n'existent pas.

SANTERRE

Dans la parole réelle, il n'y a pas, pour la même consonne, un seul point dit 'locus' d'où partiraient et vers lequel se dirigeraient les transitions de F_2 , indépendamment des voyelles. Les points de départ et d'arrivée des TR^2 varient considérablement pour une consonne selon la voyelle en cause, parce que ce sont, à tout moment, les positions articulatoires du canal buccal qui imposent les fréquences de résonance; ces positions sont très différentes pour de nombreuses raisons qui ne tiennent qu'en partie à la consonne. D'ailleurs, pour un même groupe CV ou VC, les transitions varient selon l'entourage vocalique et consonnantique non immédiat, et selon la réduction articulatoire imposée par le débit et les positions en dehors de l'accent, etc. Le 'locus' est une NOTION utile en parole synthétique, mais insatisfaisante pour l'analyse de la parole réelle.

LÉON, P. (Toronto)

Pourriez-vous préciser quelles sont les conséquences pour la programmation de la synthèse?

SANTERRE

Il faut programmer la synthèse avec beaucoup plus de précisions que ne le comporterait l'existence d'un seul 'locus' par consonne; pour le naturel de la parole synthétique, il convient de prévoir chaque transition et chaque formant en tenant compte et de la voyelle et de la consonne, de même que la voyelle qui précède et de celle qui suit, de la position implosive ou explosive de la transition, de la nature de l'accent, de la vitesse du débit et de la réduction vocalique et consonnantique imposée par la coarticulation.

On pourrait mettre en mémoire un nombre considérable de programmes syllabiques

qui tiendraient compte sur commande de ces paramètres. Dans ce domaine, ce n'est pas le phone qui est l'unité de travail, mais la 'syllabe phonétique'.

BEAUCHEMIN (Sherbrooke, P.Q.)

La règle de bouche artificielle est-elle ainsi confirmée?

SANTERRE

Un modèle théorique ne confirme rien; il aide à comprendre ce qui se passe, aussi bien dans la bouche réelle que dans la bouche artificielle. Mon modèle électroacoustique simple m'a permis d'observer, dans ses grandes lignes, la corrélation entre les mouvements articulatoires et les variations formantiques. Les lois semblables de l'acoustique et de l'électroacoustique devraient se retrouver dans la bouche artificielle comme dans la bouche réelle, ou dans n'importe quel système résonant.