



HAL
open science

simulation numérique du conduit vocal

Robert Espesser

► **To cite this version:**

Robert Espesser. simulation numérique du conduit vocal. Travaux interdisciplinaires du Laboratoire Parole et Langage, 1977, 4, pp.86-122. hal-01649595

HAL Id: hal-01649595

<https://hal.science/hal-01649595>

Submitted on 27 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SIMULATION NUMERIQUE DU CONDUIT VOCAL

par

ROBERT ESPESSER

Nous présentons une simulation numérique du conduit vocal basée sur le modèle de Kelly et Lochbaum, dont nous avons affiné le comportement en dynamique (élimination du bruit lors des mouvements du conduit, et prise en compte de ses changements de longueur). Le système est actuellement conçu pour synthétiser les voyelles orales et nasales et les transitions à l'exclusion des consonnes.

Ce type de simulation nécessite un système de conversion numérique-analogique, dont les principes sont exposés dans la première partie de cette étude.

L'ensemble du système est implanté sur un T 1600 télé-mécanique.

I. SYNTHESE DIRECTE

I.1. Rappel du principe

Une onde dont le spectre est limité à F Hz peut être décrite par un échantillonnage à $2F$ Hz (8).

Convertissons de tels échantillons numériques en tension, à la cadence de $2F$ Hz, après lissage par filtrage passe-bas à F Hz et amplification, ce signal alimente un haut-parleur : l'onde a été physiquement reconstituée à partir d'échantillons numériques.

En général, le calcul des échantillons n'est pas effectué en temps réel : il est nécessaire d'utiliser une mémoire de masse pour les stocker. Ces échantillons ne sont pas restitués à un rythme uniforme, et il faut donc régulariser le débit par une mémoire tampon.

I.2. Description du système

Nous avons réalisé un système "en ligne" sur un calculateur T 1600 télémechanique, avec la configuration suivante : 16 kmots de 16 bits, une unité de disque Control Data de 2,5 Mmots, un convertisseur N/A de 10 bits (9 bits + signe) délivrant une tension comprise entre ± 10 v, un canal d'accès direct à la mémoire (ADM) de liaison avec le disque.

En général, de tels systèmes ont deux canaux E/S indépendants : l'un par le disque, l'autre, cadencé par une horloge externe, pour le convertisseur (8,1).

Nous ne disposons pas de ce second canal, le convertisseur ne fonctionnant que sous le contrôle de l'unité centrale (mode programmé simple).

La synchronisation s'effectue par l'intermédiaire d'une entrée binaire du calculateur, connectée à l'horloge externe : selon l'état de cette entrée (à 1 ou à 0) l'unité centrale envoie ou non un échantillon au convertisseur ; entre deux envois successifs, on dispose d'un créneau de travail (pour lancer un échange-disque, etc...).

Après initialisation, on lit l'un des tableaux (mémoires-tampons) tandis que l'autre est rempli avec les échantillons stockés sur disque. Chaque tableau a une taille d'environ 3 Kmots.

Le programme est écrit en assembleur T 1600, et se présente sous la forme d'un sous-programme appellable FORTRAN IV.

I.3. Utilisation et résultats

L'utilisateur entre sur le télétype l'adresse de début et la taille de la séquence à synthétiser ; après audition, il peut soit réécouter cette séquence, soit demander la synthèse d'une autre séquence.

La fréquence maximum d'échantillonnage est de 20 kHz, sur une seule voie ; nous sommes ici limité par le mode de couplage du convertisseur au ordinateur et par la synchronisation "par programme".

Celle-ci n'assure pas la conversion des échantillons à une cadence exacte (présence de fluctuations dues aux temps d'exécution variable des instructions nécessitées par la synchronisation) d'où un bruit "numérique", qui se révèle gênant lors de la synthèse de sinusoides de fréquence supérieure à 5 kHz.

Ce système de synthèse directe, provisoire, s'est avéré de qualité suffisante lors de son utilisation pour la synthèse de voyelles (le bruit n'est alors plus perçu).

II. SIMULATION DU CONDUIT VOCAL

On assimile le conduit vocal à un tuyau acoustique déformable transversalement (variation de section) et longitudinalement (variation de longueur).

Le modèle utilisé est celui de Kelly et Lochbaum (6), repris par plusieurs auteurs (10, 14, 15, 16). Nous reppelons seulement les équations de propagation.

II.1. Equations de propagation

II.1.1. Conduit vocal

Soit Δt le temps mis par l'onde de vitesse volumique pour parcourir chaque tube élémentaire.

Soit $V(+,n,k)$ la vitesse volumique se propageant dans le sens positif, incidente sur la jonction $n+1$ à l'instant $k \Delta t$.

$V(-,n+1,k)$ la vitesse volumique se propageant

dans le sens négatif, incidente sur la jonction $n+1$ à l'instant $k \Delta t$ (fig. 1).

Pour tenir compte des pertes dues à l'absorption d'énergie par les tissus, à la viscosité de l'air, etc..., on considère que la vitesse volumique s'atténue d'un facteur α_n durant sa propagation dans le tube n ; il vient :

$$V(+,n+1,k+1) = \alpha_{n+1} (1+R_n)V(+,n,k) - R_n V(-,n+1,k)$$

$$V(-,n,k+1) = \alpha_n (1-R_n)V(-,n+1,k) + R_n V(+,n,k)$$

avec $R_n = \frac{A_{n+1} - A_n}{A_{n+1} + A_n}$ coefficient de réflexion de la jonction n

On voit sur la fig 1 que pour chaque jonction, il suffit de faire ce calcul tous les $2 \Delta t$ (12,15).

Pour des tubes de 1 cm de long et une vitesse de propagation de 350 m/s (vitesse du son à 37°C) cela correspond à une fréquence d'échantillonnage de 17,5 kHz.

II.1.2. Nasalité

Le conduit nasal est également découpé en 12 tubes élémentaires de 1 cm de long (sa longueur totale est fixe).

Les hypothèses de propagation et les notations sont les mêmes que précédemment.

On a trois ondes incidentes, $V(+,1)$, $V(-,2)$, $V(-,3)$ qui induisent trois nouvelles ondes $V(-,1)$, $V(+,3)$, $V(+,2)$ (Fig.2)

Chacune des ondes incidentes donnent une onde réfléchie et une onde transmise (12,16).

Soient R_i et T_i les coefficients de réflexion et de transmission de l'onde incidente V_i :

$$R_i = \frac{(A_j+A_k)-A_i}{A_i+A_j+A_k} \quad i, j, k \in \{1, 2, 3\} ; i \neq j \neq k$$

$$T_i = 1 + R_i$$

l'onde transmise provenant de chaque tube se partage dans les deux autres proportionnellement à leur aire, d'où :

$$V(-,1) = \frac{\alpha_1 [V(+,1) (A_2+A_3-A_1) + 2A_1 (V(-,2) + V(-,3))]}{A_1+A_2+A_3}$$

$$V(+,2) = \frac{\alpha_2[V(-,2)(A_1+A_3-A_2) + 2A_2(V(+,1)+V(-,3))]}{A_1+A_2+A_3}$$

$$V(+,3) = \frac{\alpha_3[V(-,3)(A_1+A_2+A_3) + 2A_3(V(-,2) + V(+,1))]}{A_1+A_2+A_3}$$

où $\alpha_1, \alpha_2, \alpha_3$ sont les facteurs de perte des tubes 1, 2, 3.

II.2. Terminaison aux lèvres

On considère que le conduit vocal se termine sur un tube d'aire A_r , très supérieure à l'aire aux lèvres, et de longueur infinie, d'où une absence d'onde retour (6).

Le dernier coefficient de réflexion vaut donc :

$$R_t = \frac{A_r - A_n}{A_r + A_n} \quad \text{avec } A_r = 400 \text{ cm}^2$$

A_n : aire aux lèvres

Enfin, l'effet de diffusion autour de la tête (5) donne la pression à une distance hors de la tête par :

$$P(r) = j\omega p \frac{V_1 e^{-jkr}}{4\pi r}$$

où V_1 : vitesse volumique aux lèvres

ω : vitesse angulaire

d'où : $P(k) = V_1(k) - V_1(k-2)$

$P(k)$ échantillon stocké sur disque, k indice temporel.

II.3. Le larynx

On considère que la glotte a une impédance infinie ; on ajoute à l'onde réfléchie un échantillon d'onde glottale.

L'onde glottale est obtenue à la fréquence et l'amplitude désirées au moyen d'un oscillateur, au sens de MUSIC V(8).

Cet oscillateur peut utiliser les résultats d'un système de détection de la mélodie (16) (nous verrons, au III.1, la communication entre ce système et l'oscillateur).

La forme d'onde utilisée est une de celles données par ROSENBERG (13) (forme d'onde dite "polynomiale").

II.4. Traitement des mouvements du conduit vocal

Nous nous sommes limité à la simulation dynamique du conduit vocal au sens strict, à l'exclusion donc des voies nasales.

Dans la suite, nous appelons configuration un ensemble de n sections d'aire, où n est la longueur du conduit vocal en cm, chaque aire étant celle d'un tube élémentaire, on le notera :

$$\{C\} = \{C_i, i \in \{1, n\}\}$$

On parlera de la configuration $\{C\}$, constituée des n aires C_i ; i va croissant de la glotte vers les lèvres.

II.4.1. Simulation des mouvements transversaux

On considère qu'entre deux configurations "cibles" (phonémiques par exemple), chacune des aires évolue linéairement en fonction du temps.

Le changement de forme du conduit peut conduire à des effets acoustiques gênants (présence de "clics" durant la transition) d'autant plus prononcés qu'il y a plus d'énergie dans le conduit et que les changements d'aires sont plus importants. Pour les éviter, on peut attendre la fin de l'impulsion glottale (12) (d'où une contrainte de chronométrage) ou reformuler les équations (15) (calculs plus longs, utiles surtout pour les consonnes).

Nous proposons une troisième solution, consistant à calculer un nombre suffisant d'interpolations entre deux configurations cibles. Nous l'avons visé à 24, quelles que soient les configurations cibles et la durée de la transition. Nous obtenons ainsi des transitions sans bruits perceptibles, même en présence d'une énergie importante dans le conduit vocal.

II.4.2. Simulation des mouvements longitudinaux

Le conduit vocal est affecté de mouvements longitudinaux, localisés aux lèvres et au larynx ; les modèles articulatoires actuels prennent en compte ces mouvements qui peuvent atteindre au total 2 à 3 cm d'amplitude (2, 7, 9), mais le changement de longueur n'était jusqu'à présent pas traité dans les simulations où vitesse et/ou pression sont calculées explicitement (cf. réalisations fondées sur le modèle de Kelly et Lochbaum (6, 9, 12, 14, 15, 16) et la simulation de Flanagan, Ishizaka & Shipley (1975) (4))

En première approximation, on considère que les variations de longueur sont reportées en totalité aux lèvres.

La discrétisation spatio-temporelle choisie impose une variation brusque de la longueur. La figure 3 schématise l'évolution choisie dans le cas d'un allongement de l c : la longueur vaut n durant la première moitié de la transition, et $n+1$ durant la seconde.

Chacune des aires évolue linéairement par rapport au temps. La valeur de l'aire $n+1$ à i/ϵ est prise égale à celle de l'aire n au même instant (fig. 4.1. pour un allongement, fig. 4.2. pour un raccourcissement).

Sur les fig 5.1. et 5.2. on voit que le changement de longueur introduit un retard Δt dans le calcul de l'échantillon de pression, ce qui est négligeable.

Si l'on veut changer de longueur quelle que soit l'énergie présente dans le conduit, il faut prendre de nouvelles précautions pour éviter des "clics" importants.

Cas d'un allongement

Donner la valeur 0 à $V(-,n+1)$ (fig. 5.1.) introduit une forte perturbation qui va se traduire acoustiquement quelques dizaines d'échantillons plus tard par un "clic" violent. Cette valeur nulle serait justifiée si nous simulions en détail la variation continue de longueur ; dans notre cas, celle-ci passe brutalement de n à $n+1$ cm ; à $i/2$, nous devons plutôt considérer la situation comme stabilisée, et donner à $V(-,n+1)$ la valeur qu'elle aurait eue si l'allongement s'était produit

antérieurement. Calculons $V(-,n+1)$ dans le cas où l'allongement avait eu lieu Δt avant. Soient A_n et A_{n+1} les aires des deux derniers types ; on a :

$$R_{n+1} = \frac{400 - A_{n+1}}{400 + A_{n+1}}$$

$$R_n = \frac{A_{n+1} - A_n}{A_{n+1} + A_n} = 0 \quad \text{car } A_n = A_{n+1}$$

$$V(-,n+1) = \alpha_{n+1} R_{n+1} \alpha_{n+1} (1+R_n) V(+,n)$$

Cas d'un raccourcissement

Le raisonnement est le même que précédemment, on se place dans la situation où le raccourcissement a déjà eu lieu, ce qui implique l'inexistence de $V(-,n+1)$ (fig. 5.2).

Nous obtenons ainsi des transitions entre configurations de longueur différentes, sans bruits perceptibles.

Pour des variations de plus de 1 cm, il est nécessaire de calculer des configurations intermédiaires.

Variation de 2 cm

Soient $\{C_{1i}, i \in \{1, \dots, n\}\}$ et $\{C_{2i}, i \in \{1, \dots, n+2\}\}$ les deux configurations initiales ; la configuration intermédiaire :

$\{D_i, i \in \{1, \dots, n+1\}\}$ est définie par :

$$1 \leq i \leq n \quad D_i = \frac{C_{1i} + C_{2i}}{2}$$

$$D_{n+1} = \frac{C_{1n} + C_{2n+1}}{2}$$

$\{D\}$ est ainsi utilisable dans le sens $\{C_1\} \rightarrow \{C_2\}$ et $\{C_2\} \rightarrow \{C_1\}$ et sera introduite au milieu de la transition.

Le procédé est le même pour des variations de 3 cm : on utilise deux configurations intermédiaires (l'une au premier tiers de la transition, l'autre au second tiers).

III. PRESENTATION DES PROGRAMMES

L'ensemble des programmes a été rédigé en FORTRAN IV.

III.1. Programmes relatifs à la source vocale

a) Calcul de la forme d'onde

Ce programme génère une période d'onde glottale et range les échantillons sur un fichier disque.

b) Mélodie, amplitude

Le système de détection de la mélodie inscrit la séquence des valeurs du fondamental sur un fichier. L'utilisateur peut aussi générer une ligne mélodique, à partir d'une série de couples (valeur en Hz, temps en ms), entre lesquels sont calculées des valeurs intermédiaires par interpolation linéaire.

Ce dernier programme procède de même pour le calcul de l'évolution de l'amplitude, à partir des couples (amplitude en dB, temps en ms) donnés par l'utilisateur ; les amplitudes calculées sont inscrites sur fichier.

Ces trois fichiers sont récupérés par le programme de calcul des échantillons.

III.2. Programme d'interpolation

L'utilisateur dispose des configurations sous la forme de fichiers symboliques, comprenant la longueur du conduit et la valeur des aires. Les configurations peuvent être spécifiées à volonté. Ces fichiers sont accessibles par l'éditeur de texte du système, afin de faciliter les modifications.

L'utilisateur communique au programme le nom des fichiers décrivant les configurations ; le programme interpole linéairement les aires, en tenant compte des changements de longueur et construit un fichier en écrivant successivement les configurations calculées.

III.3. Programme de calcul des échantillons

On distingue deux phases dans les mouvements du conduit : la transition entre deux configurations-cibles, et le maintien de celles-ci, appelé tenue.

L'utilisateur entre sur le télétype les durées des tenues et des transitions en alternance, ces dernières sont exprimées en millisecondes.

On peut affecter à une tenue une durée comprise entre 0 et 1870 ms ; la transition peut durer de 1 à 1870 ms.

On communique ensuite au programme les noms des fichiers "amplitude", et "mélodie".

L'essentiel du temps de calcul est consacré à la résolution des équations aux jonctions ; le sous-programme correspondant utilise l'arithmétique entière.

Le temps de calcul est d'environ 500 fois le temps réel ; la réduction de ce délai nécessite la présence de la multiplication-division entière, câblée, d'une unité de flottant câblé et la rédaction du principal sous-programme de calcul en assembleur.

III.4. Audition et visualisation

L'audition est assurée par le système de conversion N/A vu en I. Il est possible de visualiser sur imprimante* une portion quelconque de la séquence d'échantillons ; on peut également y tracer le spectre (obtenu par FFT*) d'une fenêtre de 256 ou 512 points.

IV. RESULTATS

Les documents présentés ont été réalisés avec les conditions suivantes :

- les configurations vocaliques sont celles proposées par Mrayti et Guerin (11).

- lors de la conversion numérique-analogique, le filtrage était fixé par :

* Sous-programmes communiqués par Descout R., C.N.E.T. (Lannion).

- . la fréquence de coupure à -3 dB : 5 kHz
- . la chute de 48 dB/octave

- le coefficient de perte α_i est constant, et vaut 0,991 dans le conduit vocal (valeur choisie par comparaison des formes d'ondes synthétiques et naturelles).

Nous n'avons pu actuellement obtenir de résultats satisfaisants pour la synthèse des voyelles nasales ; nous attribuons cette difficulté à l'imprécision des sections d'aires utilisées.

IV.1. Conduit statique

La qualité acoustique des 11 phonèmes vocaliques que nous avons synthétisés est satisfaisante. Nous avons comparé nos valeurs de fermants, obtenus par FFT de la réponse impulsionnelle du conduit (ensemble B) avec celles (ensemble A) données par Mrayati & Guerin (11) (assimilation du conduit à une ligne de transmission et détermination de la fonction de transfert) et celles (ensemble C) que nous avons obtenues par la méthode de Coker (1968, cité dans 5), reprise par Zerling (1974) (17). Ces trois méthodes, utilisées avec les mêmes fonctions d'aire, donnent des résultats très proches (voir tableau 1).

IV.2. Enchaînement de voyelles

Comme il est prévisible, la qualité acoustique est meilleure, les transitions entre configurations améliorant le naturel de la synthèse. Cependant, celle-ci reste tributaire de la qualité des voyelles statiques, bien qu'à la faveur du contexte, certaines d'entre elles, par effet de contraste, soient identifiées sans ambiguïté.

Exemple 1 [aøioy] (fig.6)

La comparaison des sonagrammes, en bande étroite et large de [aøioy], naturels et synthétiques, donne une idée de la qualité de la synthèse. Les paramètres utilisés sont les suivants :

mélodie :

fondamental statique à 120 Hz

chronométrage :

Les transitions durent 112 ms, les tenues 110 ms
(échelle de temps : 1 cm = 80 ms).

amplitude :

gain de 25 dB sur 4 périodes
tenue sur 110 périodes
chute de 6 dB sur 6 périodes

Chacun des autres exemples présentés a été synthétisé avec une ligne mélodique "naturelle", acquise à l'aide du système de détection de la mélodie (16). En amplitude, l'attaque et la chute ont été ajustées à partir des données fournies par oscillogramme.

Bien que limitée au début et à la fin de la séquence synthétique, la commande d'amplitude s'est révélée suffisante pour les exemples choisis : en effet, mises à part l'attaque et la chute, les variations d'amplitude glottale sont faibles, et les écarts de pression buccale observés sur les séquences naturelles reflètent essentiellement l'intensité spécifique des voyelles, qui, étant donné ce type de synthèse est respectée dans le calcul.

Le chronométrage est dérivé des sonagrammes des séquences naturelles (échelle de temps : 1 cm = 80 ms)

Exemple 2 [ai:jajajə] (fig. 7)

Les configurations utilisées correspondent aux voyelles suivantes : [a], [i], [ø], avec une configuration intermédiaire pour la transition [i]→ [ø] les [j] sont réalisés avec la configuration correspondant à [i].

Pour la voyelle d'appui [ə], on a utilisé la configuration [ø], avec une tenue très brève (20 ms). Les caractéristiques de l'amplitude sont les suivantes :

- gain de 25 dB sur 3 périodes
- tenue sur 128 périodes
- chute de 23 dB sur 36 périodes

Cette chute importante et progressive permet la perception de la finale [jə], malgré un [i] dont la tenue est assez longue (120 ms).

Exemple 3 [u j u j u: j ə] (fig.8)

On ne peut utiliser ici un [i] pour réaliser les [j] ; en effet, passer de [u] à [i] nécessite un raccourcissement important du conduit vocal, surtout aux lèvres ; or, on observe que lors de la prononciation de [ujujujə], les lèvres gardent la position correspondant à [u]. La configuration [i] n'est pas atteinte, à la différence de [aja], où un simple resserrement de la mâchoire à partir de [a] conduit à une configuration proche de [i].

Nous avons donc utilisé la cible suivante, dérivée de [i] et [u] :

- longueur : 18 cm (la configuration de [u] a 19 cm de long, celle de [i], 17 cm).

- les 16 premières aires ont été calculées par :

$$A_k^{[j]} = (95 A_k^{[i]} + 5 A_k^{[u]}) / 100 \text{ avec } k \in 1, \dots, 16$$

où $A_k^{[j]}$, $A_k^{[i]}$ et $A_k^{[u]}$ sont les k èmes premières aires des configurations de [j], [i] et [u] respectivement :

$$A_{17}^{[j]} = 1 \text{ cm}^2$$

$$\text{et } A_{18}^{[j]} = 0,75 \text{ cm}^2$$

Nous obtenons ainsi des suites [uju] satisfaisantes, alors qu'en utilisant des [i], "l'hyperarticulation" simulée conduisait à percevoir plutôt [wi] que [uj] même en utilisant un [i] de durée bien inférieure à celle du [j] réel (essais avec un [i] d'environ 10 ms, alors que le [j] dure de 40 à 50 ms). Comme dans l'exemple 2, [ø] est utilisé comme voyelle d'appui pour le [jə] final.

Les valeurs de l'amplitude sont les suivantes :

- gain de 25 dB sur 8 périodes
- tenue sur 68 périodes

- chute de 25 dB sur 14 périodes.

Comme précédemment, cette chute importante est nécessaire pour la réalisation de [jə].

Exemple 4 [awi] [ewi] (fig. 9)

Nous avons utilisé les configurations correspondant à [a],[u],[e] et [i], et les configurations intermédiaires dues aux changements de longueur (de [a] à [u], + 2 cm, de [u] à [i], - 2 cm, de [e] à [u], + 3 cm).

[w] est réalisé avec un [u] brièvement tenu (35 à 50 ms). Les valeurs d'amplitude sont les suivantes :

- pour [awi]
 - . gain de 15 dB sur 6 périodes
 - . tenue sur 56 périodes
 - . chute de 10 dB sur 6 périodes

La chute de l'intensité est rapide (les périodes sont brèves) : cette perte permet d'obtenir une bonne qualité de l'interrogation.

- pour [ewi]
 - . gain de 25 dB sur 16 périodes
 - . tenue sur 6 périodes
 - . chute de 9 dB sur 6 périodes

La présence de la mélodie améliore nettement la qualité de la synthèse et les réalisations 1, 2, 3, 4 ont été jugées plus naturelles que les autres.

- 0 -

En conclusion, le système mis au point peut d'ores et déjà, malgré ses limites, être utilisé sur plusieurs plans.

- perception des éléments prosodiques des voyelles statiques, l'amplitude et le fondamental pouvant être précisément contrôlés.

- influence de la forme de l'onde glottale sur la qualité de la voix.

- approche articulatoire des transitions voyelle-voyelle (consonnes vocaliques incluses).

Le premier développement à entreprendre nous paraît être la production des consonnes constrictives, puis ultérieurement celles des occlusives. Il sera alors nécessaire de travailler avec le couple pression-vitesse, au moins au voisinage de la constriction, si on envisage l'introduction automatique du bruit dans le conduit (3).

glotte

lèvres

jonction 0 jonction 1 jonction 2

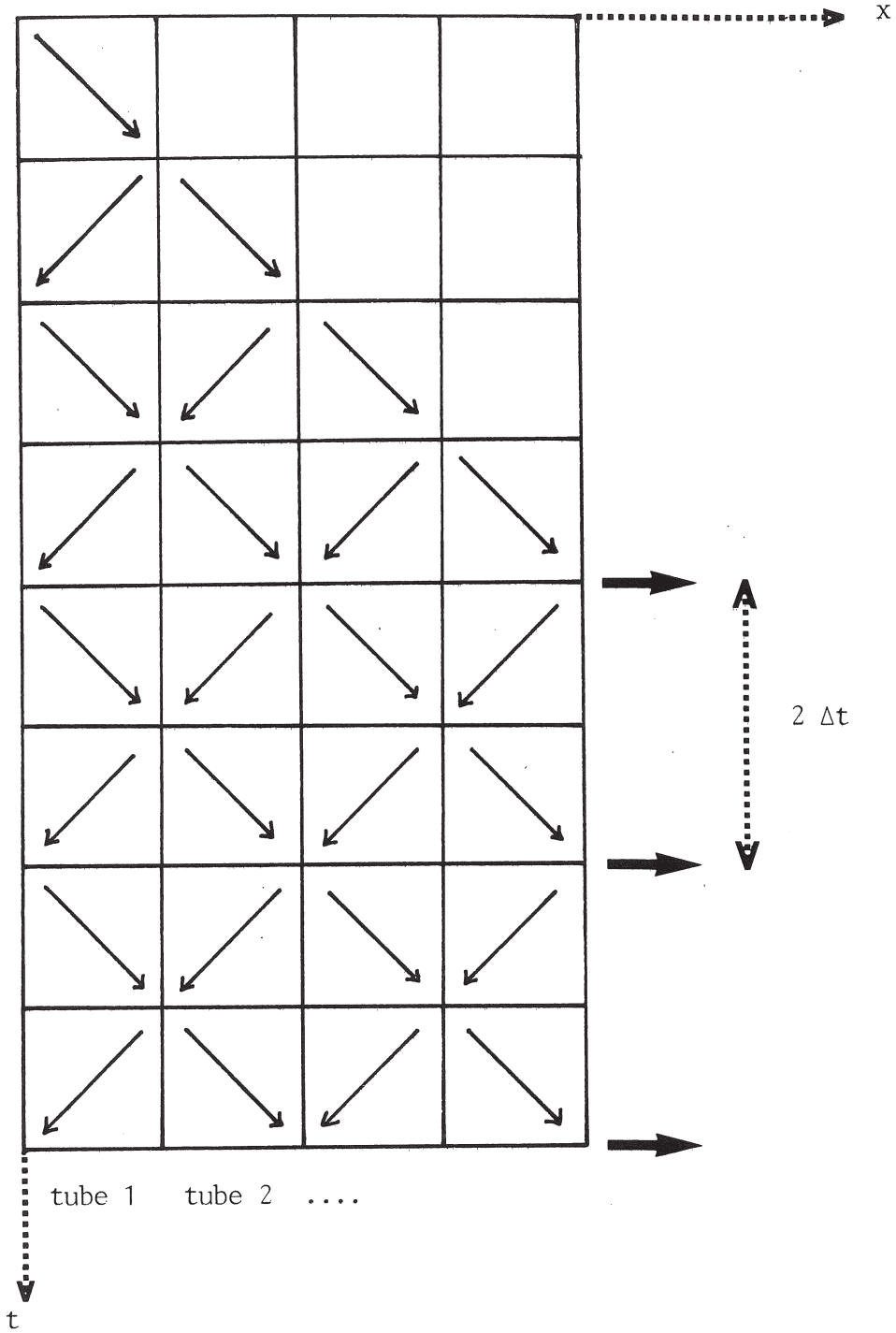
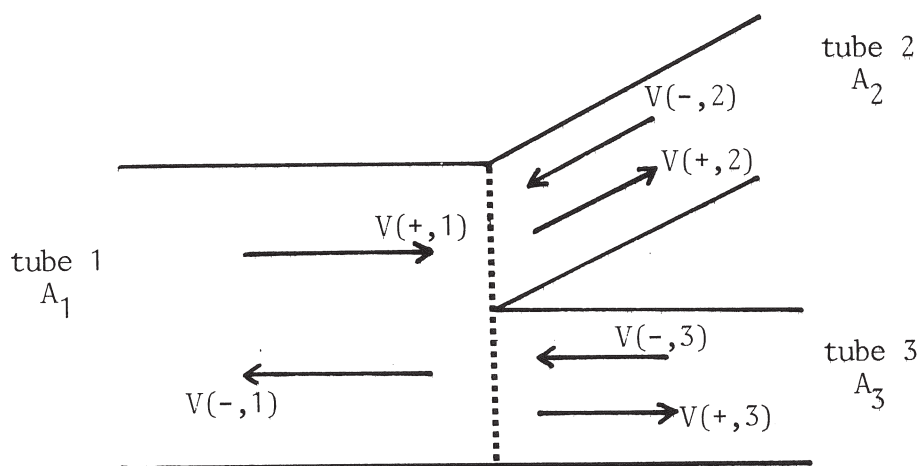


fig.1

Schéma général de propagation (d'après RICE)



tube 1:pharynx
tube 2:début cavité nasale
tube 3:début cavité buccale

fig.2

situation lors du couplage
des cavités nasales

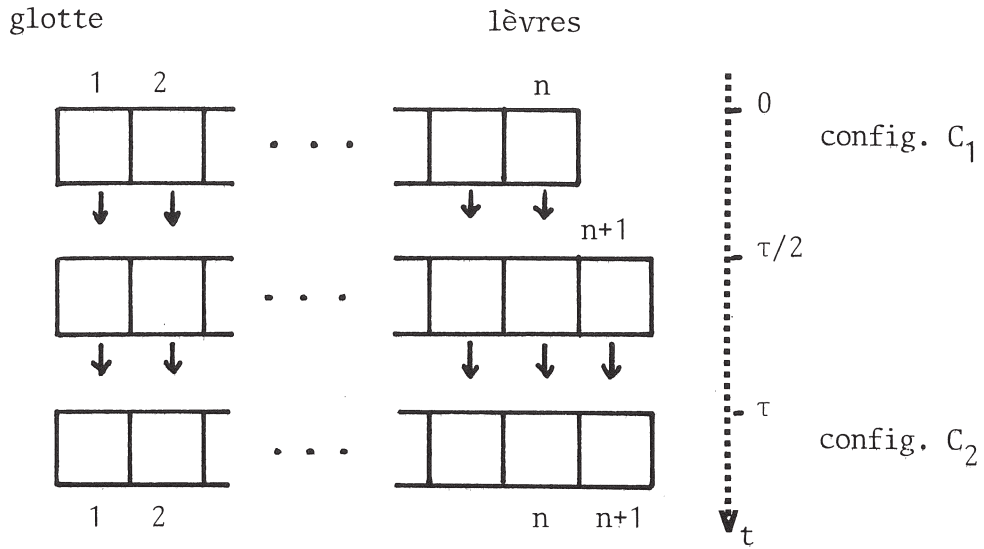


fig. 3

traitement du changement de longueur (allongement)

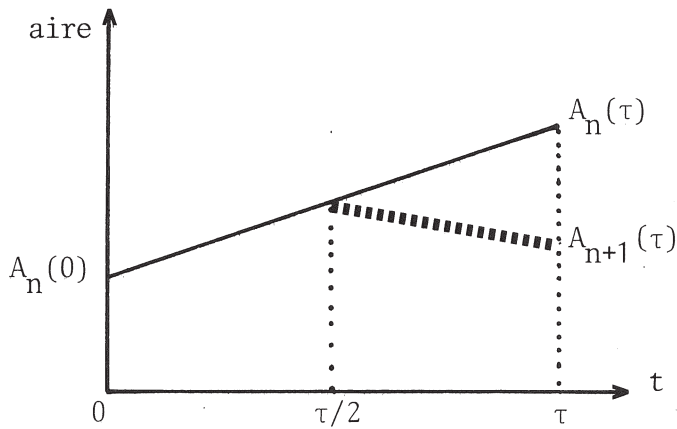


fig. 4.1.

----- évolution de A_{n+1}

— évolution de A_n

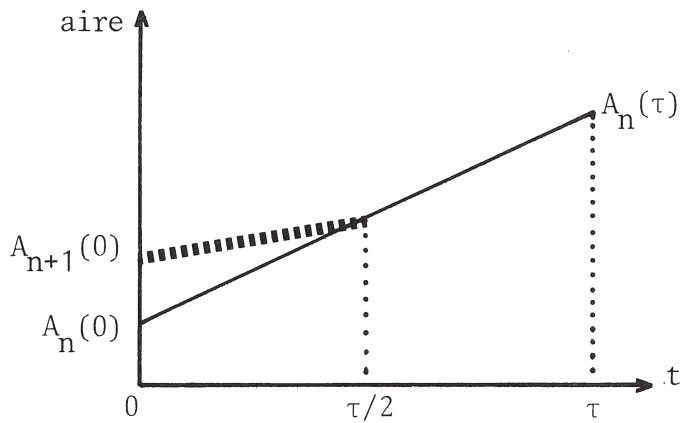


fig. 4.2.

évolution des aires aux lèvres lors d'un changement de longueur

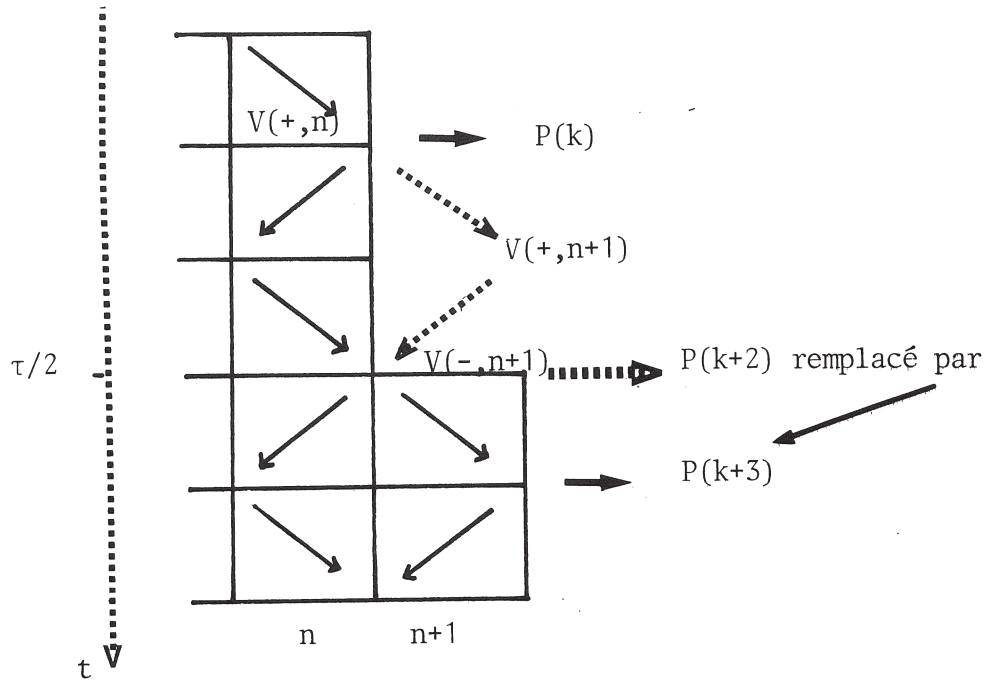


fig. 5.1.

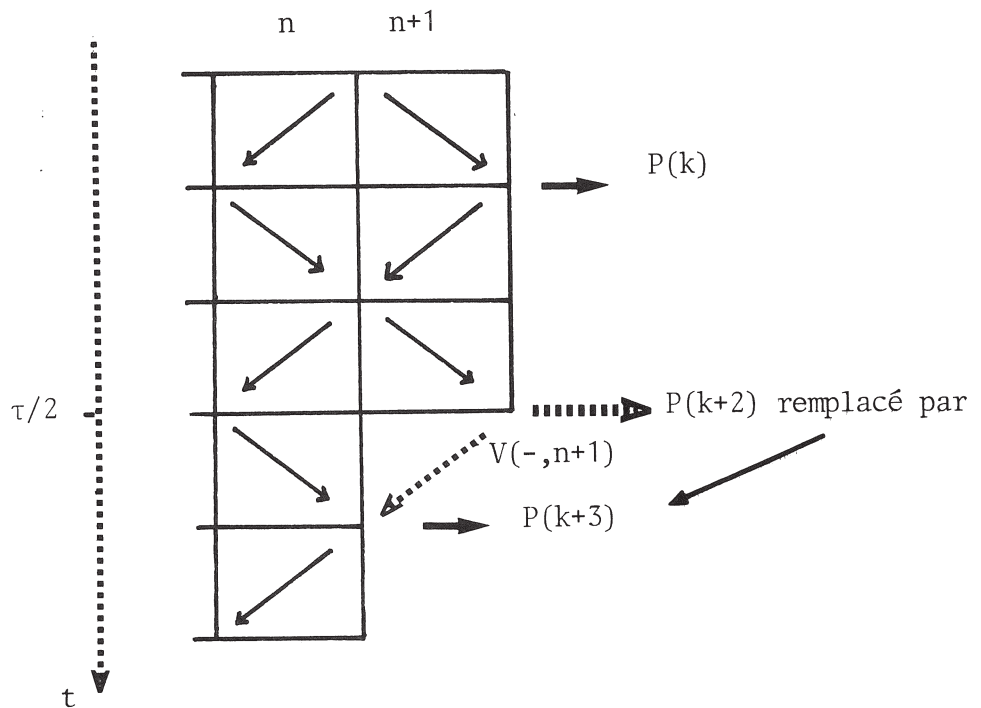


fig. 5.2.

calcul des échantillons de vitesse et de pression lors d'un changement de longueur

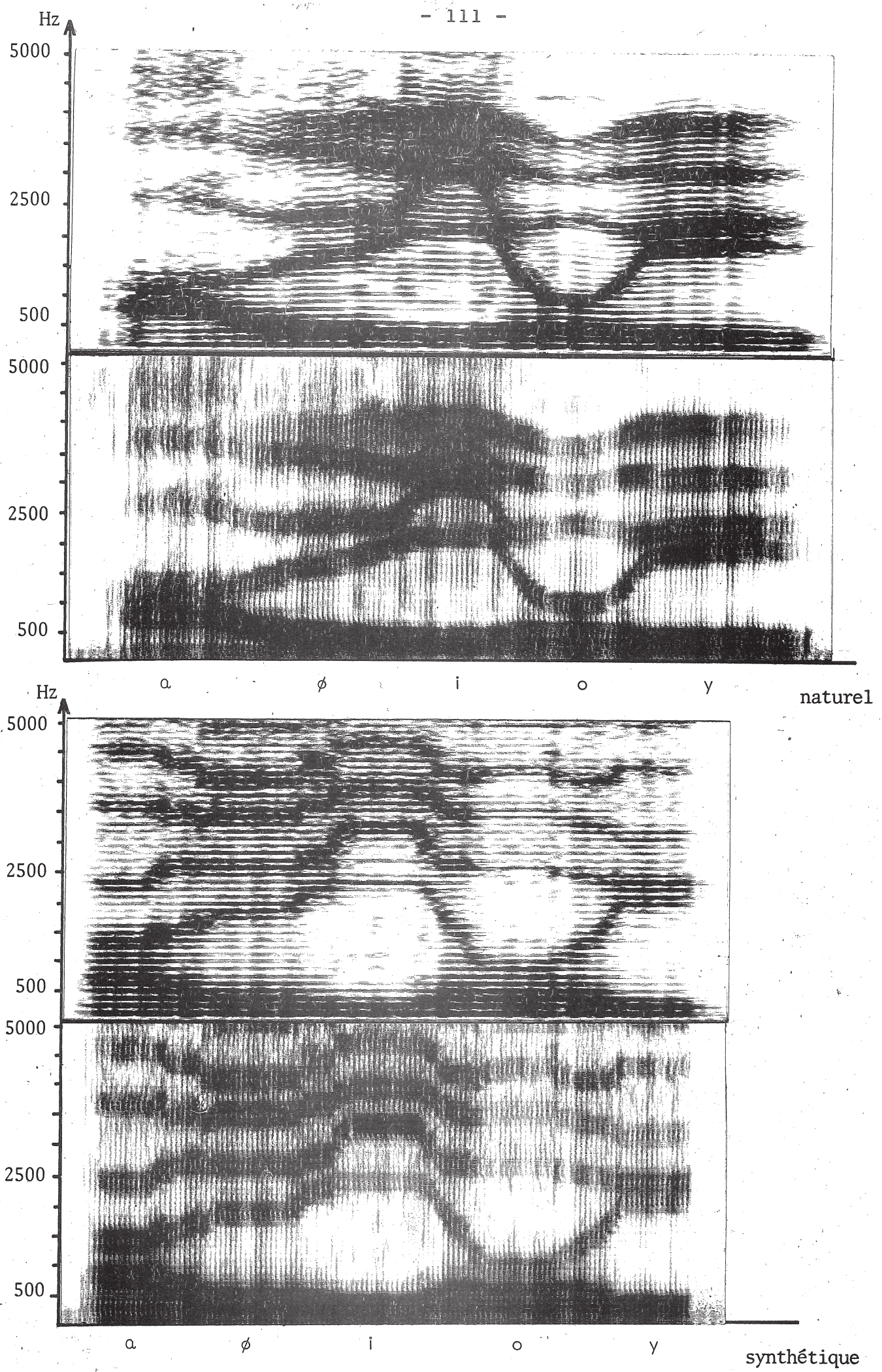
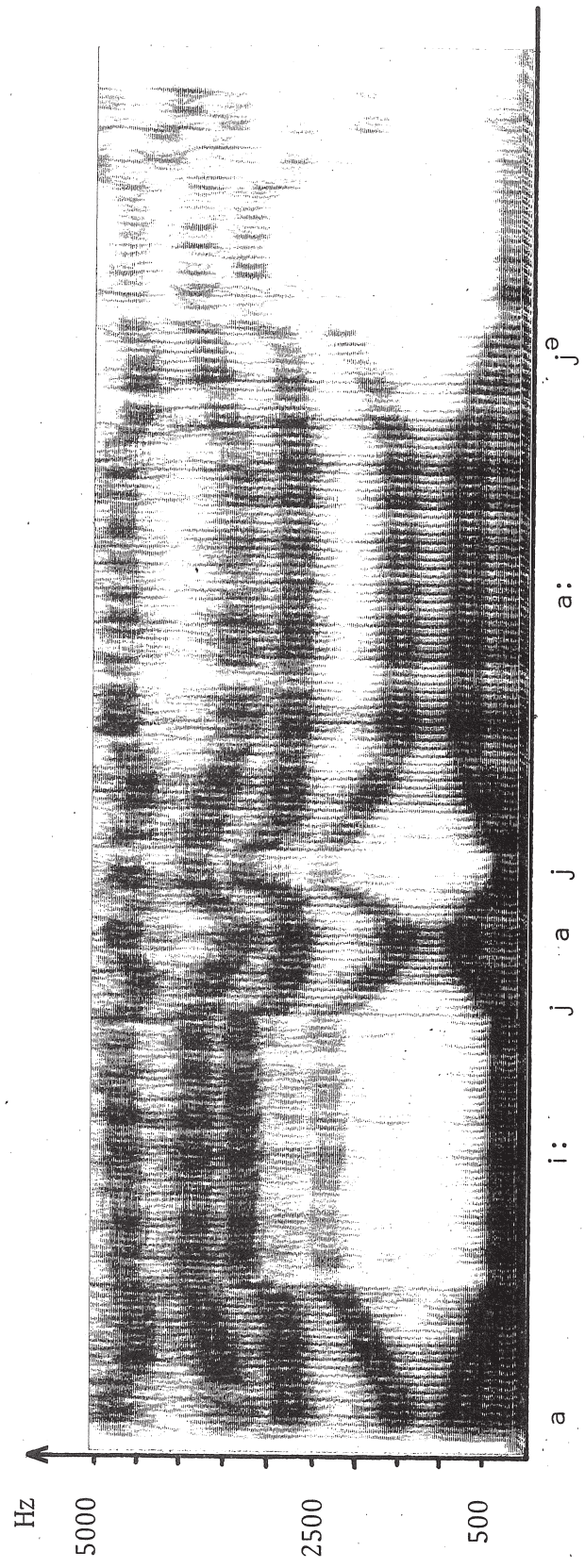
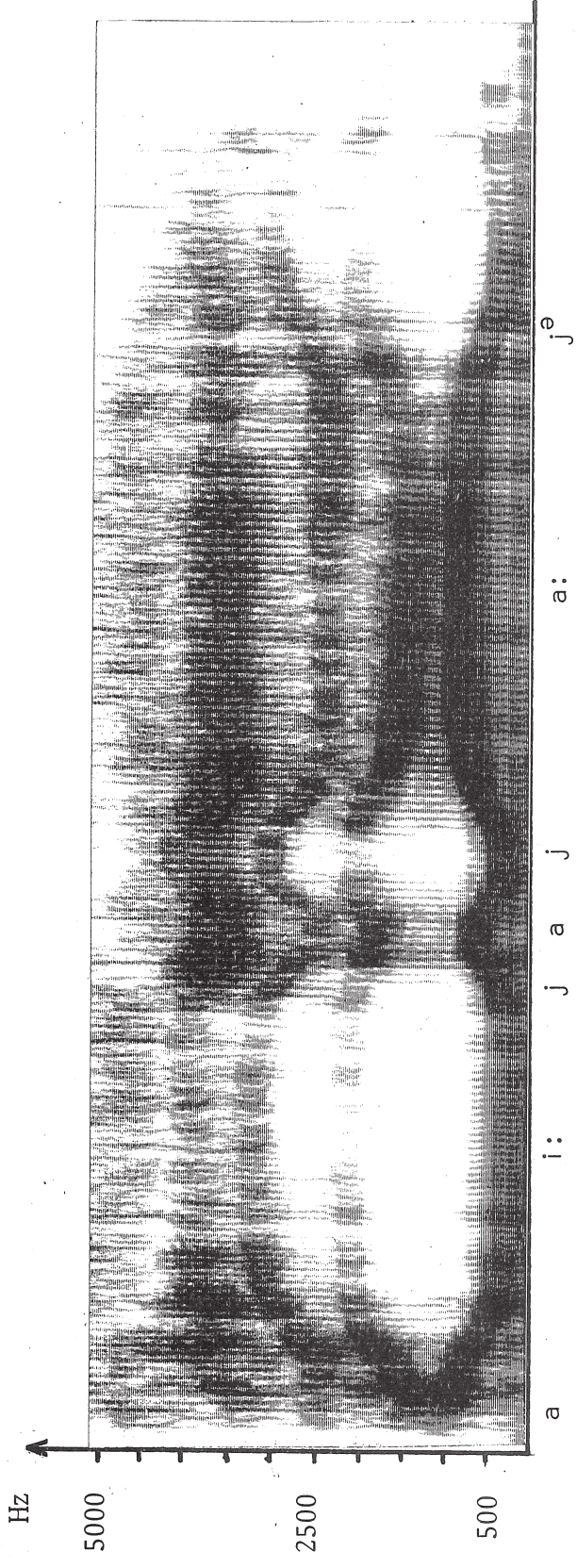


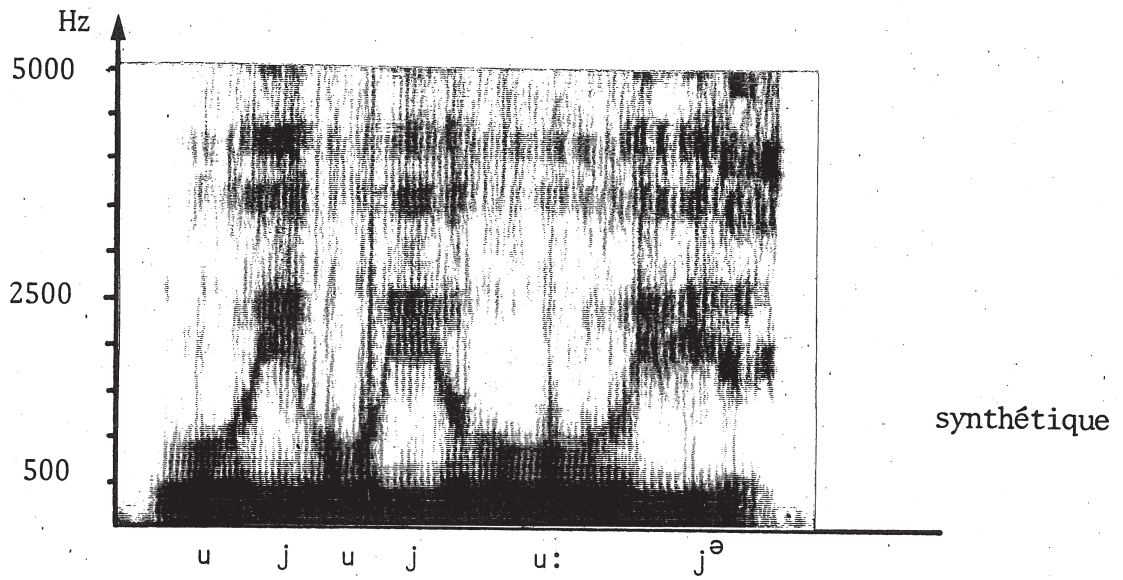
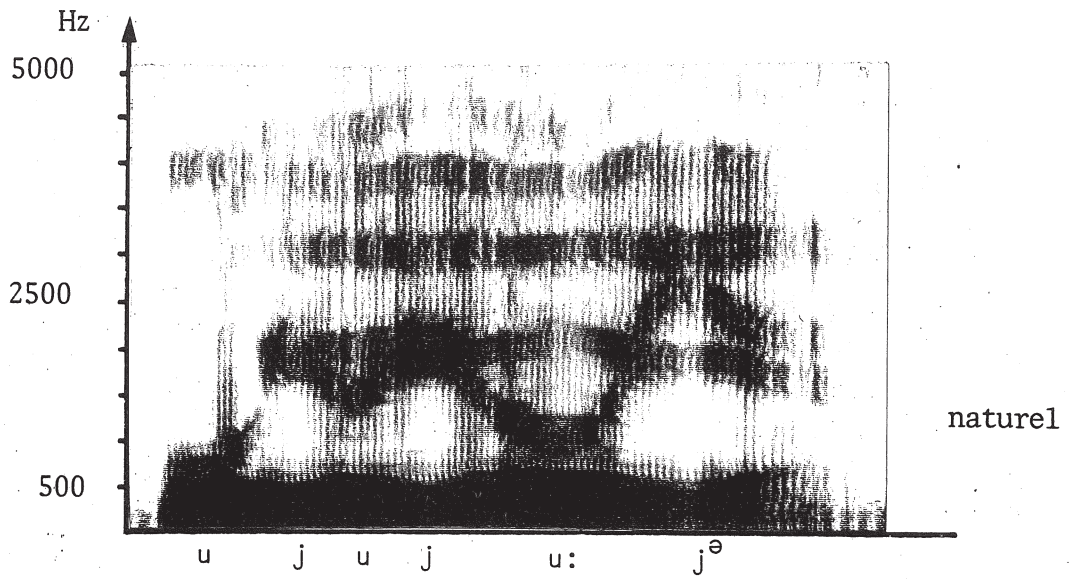
fig.6

synthétique

naturel

synthétique





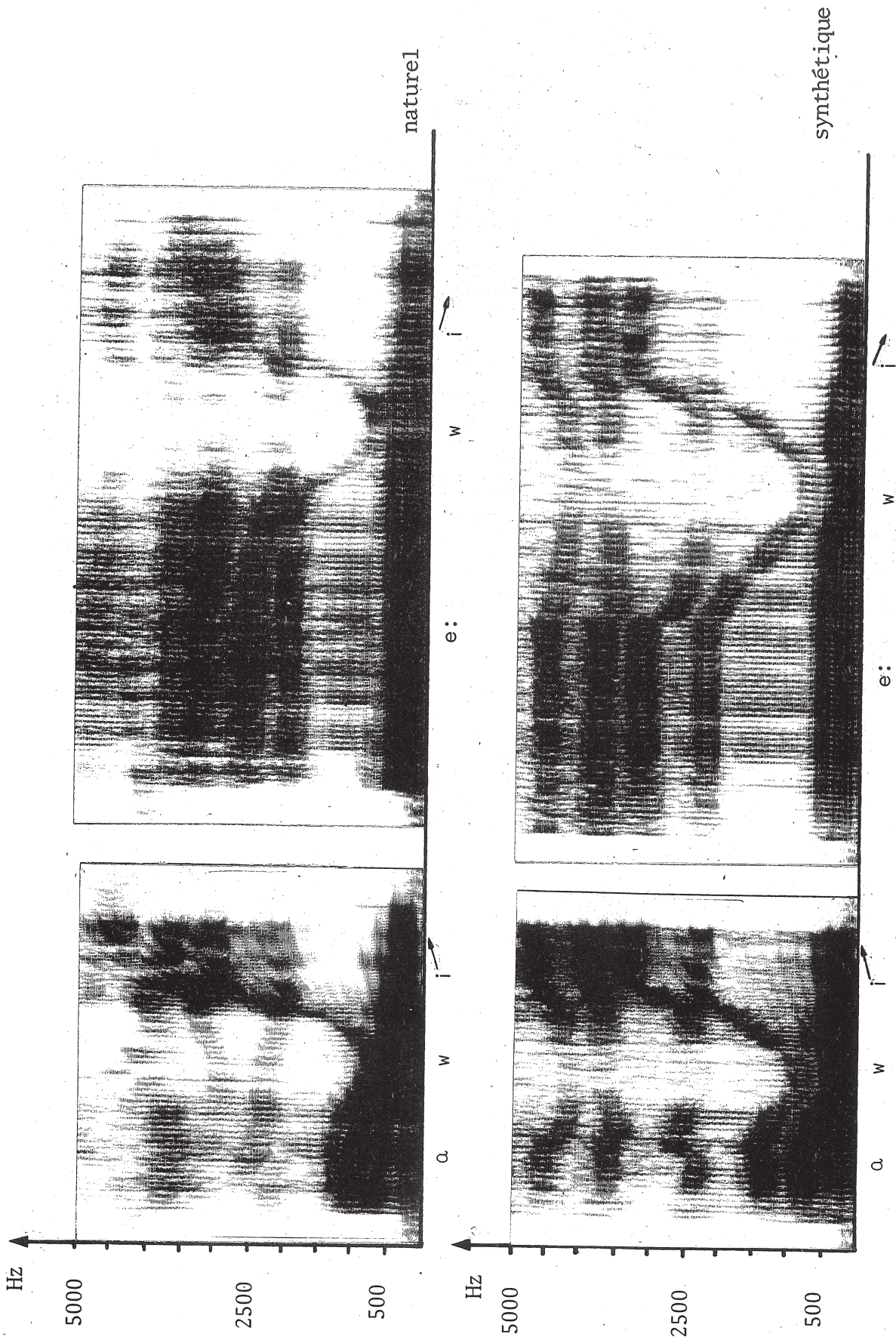


fig. 9

TABEAU 1

Phonèmes	[o.]	[a]	[e]	[ε]	[o]	[ɔ]	[φ]	[œ]	[i]	[y]	[u]
Calculs de MRAYATI & GUERIN (A)											
F ₁	676	678	418	458	402	540	384	535	244	280	296
F ₂	1 192	1 300	2 111	1 881	797	1 007	1 592	1 411	2 250	1 846	783
F ₃	2 208	2 583	2 665	2 584	2 417	2 480	2 461	2 269	3 082	2 230	2 286
Nos résultats (B)											
F ₁	685	685	380	450	410	545	380	550	220	270	275
F ₂	1 290	1 430	2 220	1 970	820	1 160	1 800	1 500	2 255	1 980	770
F ₃	2 270	2 666	3 110	2 830	2 560	2 600	2 530	2 390	3 260	2 290	2 290
Calculs d'après ZERLING (C)											
F ₁	699	705	399	448	378	525	346	529	195	231	248
F ₂	1 304	1 475	2 226	1 995	875	1 175	1 786	1 542	2 257	1 986	770
F ₃	2 313	2 759	3 201	2 856	2 530	2 672	2 572	2 432	3 271	2 348	2 293

Moyenne des écarts

- B comparé à A

sur F₁ = 3,9 %

sur F₂ = 6,2 %

sur F₃ = 5 %

- B comparé à C

sur F₁ = 7 %

sur F₂ = 1,5 %

sur F₃ = 1,7 %

BIBLIOGRAPHIE

- 1 - CHARBONNEAU, G., KARATCHENTZEFF, P. et RISSET, J.C.
Un système de synthèse directe des sons à l'aide d'ordinateurs.
Revue d'Acoustique 21, 1972 : 289-296.
2. FANT, G.
Acoustic theory of speech production, Mouton, 'S Gravenhage,
1960.
- 3 - FLANAGAN, J.L., CHERRY, L.
Excitation of vocal-tract synthesizers, *IASA*, 1969. 45 = 764-769.
- 4 - FLANAGAN, J.L., ISHIZAKA, K. & SHIPLEY, K.L.
Synthesis of speech from a dynamic model of vocal cords and
vocal tract, *The Bell Syst. Tech. J.*, 1975, 54, n° 3 : 485-506.
- 5 - FLANAGAN, J.L.
Speech analysis synthesis and perception, Springer-Verlag,
Berlin, 1972.
- 6 - KELLY, J.L., LOCHBAUM, J.R. & C.C.
Speech synthesis, *Proc. 4th ICA Congress*, Copenhagen, 1962.
- 7 - LINDBLUM, B.E.F., SUNDBERG, J.E.F.
Acoustical consequences of lip, tongue, jaw and larynx move-
ments, *JASA*, 1971, 50 = 1166-1179.
- 8 - MATHEWS, M.V.
The technology of computer music, MIT Press Cambridge, 1969.
- 9 - MERMELSTEIN, P.
Articulatory model for the study of speech production, *JASA*,
1973, 53 : 1070-1082.
- 10 - MERMELSTEIN, P.
Computer simulation of articulatory activity in speech produc-
tion. *Joint conf. on Artificial Intelligence*, Washington, 1969.
- 11 - MRAYATI, M., GUERIN, B.
Etude des caractéristiques acoustiques des voyelles orales
françaises par simulation du conduit vocal avec pertes, *Revue
d'Acoustique*, 1976, 36 : 18-32.

- 12 - RICE, L.
A new line analog speech synthesizer for the PDP 12, *Working papers in phonetics*, 1971, 17 : 58-75.
- 13 - ROSENBERG, A.E.
Effect of glottal pulse shape on the quality of natural vowels
JASA, 1971, 49 : 583-588.
- 14 - RUIZ, P., MERMELSTEIN, P.
Speech generator with a computer simulated vocal tract analog
JASA, 1969 : 110 (A).
- 15 - RUIZ, P.
A digital simulation of the time-varying vocal tract, *JASA*,
1971, 49 : 123 (A)
- 16 - TESTON, B. & ROSSI, M.
Un système de détection automatique des éléments prosodiques,
8ème Journée d'Etude sur la Parole, Aix-en-Provence, mai 1977.
- 17 - TITZE, I.R.
The human vocal cords : a mathematical model part I, *Phonetica*,
1973, 28, n° 3-4 : 129-170.
- 18 - ZERLING, J.P.
Etude d'un programme d'ordinateur permettant de calculer les
trois premiers formants à partir de la fonction d'aire,
Travaux de l'Institut de Phonétique de Nancy, 1974, vol. 1 :
257-291.