M. Lennig

Université de Montréal
Département de linguistique

Analyse Perception **Articulation** **Reconnaissance** **Acoustique** Synthèse **Analyse** **Perception** Articulation Reconnaissance Acoustique **Synthèse** **Analyse** **Perception** Articulation Reconnaissance **Acoustique** **Synthèse** Analyse Perception **Articulation** **Reconnaissance** Acoustique Synthèse **Analyse** **Perception**

XIIe Journées d'études sur la parole

Analyse Perception **Articulation** **Reconnaissance** **Acoustique** Synthèse **Analyse** **Perception** Articulation Reconnaissance Acoustique **Synthèse** **Analyse** **Perception** Articulation Reconnaissance **Acoustique** **Synthèse** Analyse Perception **Articulation** **Reconnaissance** Acoustique Synthèse **Analyse** **Perception**

organisées à l'Université de Montréal
avec la collaboration du GALF
les 25, 26 et 27 mai 1981

Laboratoire de phonétique
Département de linguistique
Université de Montréal

XIIèmes Journées d'études sur la parole

Organisées à l'Université de Montréal
avec la collaboration du GALF
les 25, 26 et 27 mai 1981

Comité d'organisation

Danièle Archambault
Raymond Descout
Alain Marchal
André Marti
Philippe Martin
Laurent Santerre

12èmes JOURNEES D'ETUDE SUR LA PAROLE

MONTREAL 25.26.27 mai 1981

SIMULATION D'UN LECTEUR AUTOMATIQUE DU FRANCAIS

O'SHAUGHNESSY, Douglas
LENNIG, Matthew
MERMELSTEIN, Paul

INRS-Télécommunications
3, Place du Commerce
Ile des Soeurs, Québec

DIVAY, Michel

Institut Universitaire de
Technologie de Lannion-France

RESUME

Nous exposons les résultats concernant la simulation sur ordinateur d'un système qui permet de convertir un texte français dactylographié en discours sonore. La version pratique d'un tel lecteur devrait pouvoir accepter une page d'un livre ou d'un journal et lire le texte de cette page à haute voix, à une vitesse pouvant atteindre 200 mots par minute et avec un naturel raisonnable. Ces lecteurs existent actuellement pour l'anglais, mais pas pour le français. Pour effectuer ce project il a fallu développer un logiciel pour la reconnaissance optique des caractères (ROC) afin de reconnaître la sortie d'un dispositif d'exploration, écrire un nouveau logiciel et modifier des algorithmes existants pour le traitement linguistique et la génération de la parole. La recherche a été axée principalement sur quatre domaines:

a) Le développement d'algorithmes pour convertir la sortie d'un dispositif ROC en texte français.
b) Le développement d'un ensemble adéquat de règles pour la conversion de graphèmes en phonèmes pour le français.
c) L'élaboration d'un modèle d'intonation pour le français.
d) La modification d'un synthétiseur actuel de parole de l'anglais pour l'appliquer au français.

Le système permet actuellement la génération en temps non réel, mais d'un manière complètement automatique, d'un discours continu hautement intelligible.

. ABSTRACT

SIMULATION D'UN LECTEUR AUTOMATIQUE DU FRANCAIS

O'SHAUGHNESSY, Douglas
LENNIG, Matthew
MERMELSTEIN, Paul

INRS-Télécommunications
3, Place du Commerce
Ile des Soeurs, Québec

DIVAY, Michel

Institut Universitaire de
Technologie de Lannion-France

SUMMARY:

A reading machine for the blind generates speech output from typed text input. It consists of: an optical character recognizer, a linguistic processor, and a sound generation algorithm. In the first component, a hardware scanner produces a binary image from the printed page, and software interprets the image as a string of letters and punctuation, using template matching techniques. Assuming a single-font as input, the program gives virtually error-free performance in non-real time. The second component translates letters into phonemes, and provides intonation appropriate to the syntax of the message. In French, this requires deciding when to pronounce usually-silent word-final letters, and interpreting letters in general by rules different from English, such as with nasalized vowels. Similarly, the pitch and duration of individual sounds is different for French. Syllables have more regular durations in French, and pitch has a greater tendency to rise at the end of a French word than an English word. Without such rules, the French speech would have an English accent. The final component interprets the phonemes as variations in vocal tract resonances and amplitudes, and drives a speech synthesizer to yield sound output. It is based on a formant synthesis algorithm developed by Klatt of M.I.T., but modified for French by the addition of nasalized and front rounded vowels, as well as other changes related to different pronunciations in French and English. Without making such acoustic rule modifications, the speech would have a heavy English accent. The final stage of actual synthesis is currently in software, but will soon be replaced by an LSI chip.

SIMULATION D'UN LECTEUR AUTOMATIQUE DU FRANCAIS

O'SHAUGHNESSY, Douglas
LENNIG, Matthew
MERMELSTEIN, Paul

INRS-Télécommunications
3, Place du Commerce
Ile des Soeurs, Québec

DIVAY, Michel

Institut Universitaire de
Technologie de Lannion-France

1) INTRODUCTION

Un lecteur automatique acceptant à l'entrée un texte imprimé (sous forme de feuilles détachées, de journaux, de livres, ou même d'images de télévision) et produisant à la sortie un discours sonore (la version parlée du texte), est une entreprise socialement désirable et représente un défi technique. Ce lecteur serait d'un grande utilité pour les personnes dont la vue est handicapée, puisque de nos jours une grande partie de l'information est imprimée. Un petit nombre d'aveugles seulement peuvent lire le Braille, mais virtuellement tous comprennent le langage parlé et un lecteur automatique permettrait l'accès à l'information qui doit pour le moment être enregistrée sur bandes par des personnes dont la vue n'est pas handicapée et qui lisent le texte à haute voix.

Il y a environ trois ans, un lecteur a été développé pour la langue anglaise (KURZWEIL, 1976) et un deuxième dispositif similaire est sur le point d'apparaître (CALDWELL, 1978). Cependant, il n'existe actuellement aucun lecteur automatique parlant français, et à notre connnaissance personne n'en a commencé le développement. Nous nous sommes efforcés de simuler un lecteur automatique afin de démontrer la qualité de la parole synthétique quand l'entrée consiste en un texte français.

2) LES COMPOSANTS D'UN LECTEUR AUTOMATIQUE

Un lecteur automatique comprend trois composants: reconnaissance optique des caractères, traitement linguistique et synthèse de la parole.

2.1 Reconnaissance optique des caractères (ROC)

Ce composant effectue la conversion image-graphème, c'est à dire qu'il convertit l'image d'un texte imprimé en une suite de caractères (lettres, chiffres et symboles) contenus dans l'image. Dans notre système le premier étage du ROC utilise un explorateur automatique standard de page, qui produit une représentation numérique de la page, en interprétant l'image comme un ensemble constitué par des millions de points noirs et blancs. Sur une page blanche, évidemment, des ensembles avoisinants de points noirs sont interprétées par l'oeil humain comme des lettres. Notre dispositif d'exploration peut résoudre une image avec une précision de 82 points par centimètre; si l'on considère un exemple standard de 2.5 millimètres par caractère, chaque symbole serait échantillonné 20 fois dans sa largeur. Chaque échantillon est quantifié au moyen de 8 bits sur une échelle noir-blanc. Un seuil permet de décider quels sont les points gris que l'on devrait considérer comme étant noirs (faisant partie du symbole) ou blancs (faisant partie de l'arrière fond). Deux cents points par pouce sont suffisants pour résoudre les variations fines de presque tous les caractères, bien qu'un texte écrit fin comme par exemple un annuaire téléphonique et les avis publicitaires d'un journal pourrait exiger une resolution plus fine.

On a développé un logiciel faisant la conversion de l'image numérique en une suite de graphème. Nos programmes acceptent une fonte standard IBM, et produisent un taux d'erreur d'environ 0.2%, avec une méthode de reconnaissance de forme utilisant un gabarit de comparaison standard.

Les programmes segmentent l'image en régions correspondant approximativement aux lettres, en respectant les constraintes provenant de la dimension des lettres et en tenant compte du fait que des pages dactylographiées consistent généralement en lignes horizontales espacées d'une manière uniforme le long des dimensions verticales et horizontales. Dans un texte français, les accents se trouvent souvent dans une région séparée du reste de la lettre, mais un autre programme permet de les joindre ensemble, en se servant du contexte. Chaque région est comparée à une série de gabarits mis en mémoire, mais certains symboles spéciaux (comme "$"), dont la représentation numérique peut être très différente à cause de la résolution du dispositif d'exploration, possèdent deux gabarits mémorisés. Il est possible d'utiliser un ensemble simple de gabarits uniquement parce qu'on utilise à l'entrée du système une seule fonte.

Les gabarits sont ordonnés selon leur probabilité d'apparition, déduite d'une étude statistique grossière de la langue française. Ainsi, les voyelles qui surviennent plus fréquemment comme "e" et "a" sont les premières à être considerées, et les consonnes rares comme "x" et "k" sont les dernières. Puisque la comparaison se termine quand une correspondance adéquate a été trouvée, en considérant d'abord les candidats les plus probables, il est possible de réduire le temps moyen de calcul.

Au cours de l'exploration, certaines versions du même symbole peuvent être plus grandes que d'autres. Aussi la comparaison commence par centrer toutes les régions selon leurs limites maximum et ensuite elle les décale horizontalement et verticalement pour obtenir une

meilleure correspondance si l'entrée et les gabarits mémorisés sont suffisamment identiques.

2.2 Traitement linguistique

Ce composant du système effectue la conversion graphème-phonème, c'est à dire qu'il convertit la suite des lettres et des signes de ponctuation en une suite de phonèmes du français. Il produit aussi un ensemble de marqueurs d'intonation suffisamment détaillés pour pouvoir spécifier la hauteur et la durée de chaque phonème. La première tâche est effectuée au moyen de règles de transformation lettre-phonème, qui permettent la conversion de graphèmes en phonèmes en tenant compte des propriétés inhérentes du langage. On peut décrire un langage par un ensemble de lois qui lui sont propres. Certains langages peuvent être décrits d'un manière plus compacte que d'autres. Par exemple, l'Espagnol possède un ensemble de règles lettre-phonème relativement petit, alors qu'un ensemble de règles complet pour l'Anglais serait considérablement plus complexe. Le Français se trouve en quelque sorte entre les deux.

2.2.1 Conversion lettre à phonème

Une méthode lettre-phonème avec un petit lexique constitue probablement le meilleur moyen pour la conversion. Un ensemble compact de règles lettre-phonème peut tenir compte de la majorité des mots d'un langue, tandis que le lexique consiste principalement en mots que les règles lettre-phonème n'arrivent pas à décrire facilement (ce qui inclus souvent des mots très répandus, ainsi que des mots étrangers). Pour augmenter l'efficacité du lexique, on pourrait mémoriser les préfixes et les suffixes répandus, et les mots pourraient être d'abord décomposés en affixes et en racines avant de consulter le dictionnaire. Il n'est donc pas nécessaire d'enregistrer toutes les dérivations et conjugaisons d'un mot du lexique, mais de conserver uniquement ses composants de base. Cette méthode s'applique particulièrement bien au français dont les verbes possèdent de nombreuses terminaisons.

Le système de conversion lettre-phonème développé à l'Université de Rennes en collaboration avec le CNET-Lannion (Centre National d'Études des Télécommunications) (DIVAY et GUYOMARD, 1977) consiste en un ensemble de règles qui s'appliquent dans une situation orthographie donnée. Les règles sont ordonnées pour qu'on puisse examiner d'abord les liaisons et les enchaînements, puis le /s/ final du mot (qui peut être silencieux, par exemple "chats", "chiens", ou que l'on doit prononcer, par exemple "hélas"), ensuite les mots répandus, les mots réguliers et finalement l'élision de certaines lettres à la fin d'un mot. Cette manière de procéder spéciale pour les mots répandus permet au programme de fonctionner plus rapidement, puisqu'un grand pourcentage des mots n'auront pas à passer à travers l'ensemble des règles lettre-phonème. Les règles sont fonctions du contexte et se présentent sous la forme:

$$a \longrightarrow b \;/\; c + d$$

(la suite de symboles "a" est transformée dans la suite "b" si on trouve "a" dans le contexte "c-d", c'est à dire que "c" est à gauche et "d" est à droite). Par exemple, une voyelle suivie par "m" ou "n" serait considérée comme une voyelle nasalisée si la lettre suivante consiste en une consonne non-nasalisée, mais ne serait pas considérée

comme une voyelle nasalisée si elle était suivie par une voyelle (par exemple "lent" par apposition à "ennemi").

2.2.2 Intonation

Les paramètres fondamentaux de l'intonation, le hauteur et la durée, varient avec plusieurs paramètres du texte entrant (BEAUCHEMIN, 1972; BENGUEREL, 1971). La variation phonétique (par exemple les voyelles basses présentent une hauteur faible et de plus longues durées) peut être traitée directement à partir de la liste d'entrée de phonèmes. Mais puisque la hauteur et les durées des composants d'un mot dépendent aussi de la structure syntaxique et de la signification sémantique (CHOPPY et LIENARD, 1975), il est nécessaire d'effectuer une autre analyse linguistique du texte. L'absence de modèle adéquat de l'intonation pour de nombreux synthétiseurs est une cause principale de l'accent "étranger" de la parole synthétisée au moyen de règles.

On a développé un algorithme pour accepter à l'entrée la ponctuation et une analyse simple des parties du discours, ainsi que les phonèmes d'une phrase, et cet algorithme produit à sa sortie la fonction d'intonation. Le lexique comprend une liste de mots syntaxiques (les articles, les pronoms, et les prépositions) qui permet de distinguer les mots sémantiquement moins importants des mots de contenu (substantifs, verbes, adjectifs); ceux-ci ont des durées plus longues et des variations de hauteur plus importantes.

Nous avons modelé la fréquence fondamentale et les durées des phonèmes d'après le modèle décrit dans une analyse de la parole naturelle (O'SHAUGHNESSY, 1981; SMITH, 1977; WAJSKOP, 1979; METTAS, 1969). On a trouvé que la durée d'une voyelle varie énormément en fonction de la nasalité et des caractéristiques des consonnes suivantes (les voyelles nasalisées et les voyelles qui se trouvent devant les consonnes voisées sont allongées). Aussi, les consonnes pouvaient former des agrégats plus ou moins longs, dépendant de la difficulté d'articulation et de la proximité des points d'articulation (par exemple les durées de /s/ et de /t/ dans "reste" sont plus courtes que dans "messe" ou "mette", à cause de la similarité de l'articulation). La fréquence fondamentale tend à augmenter pour les mots importants, et demeure constante pour les mots moins importants, mais la rapidité des variations semble plus petite qu'en anglais.

L'algorithme de durée attribue une durée à chaque phonème d'une phrase donnée, et la modifie ensuite selon son contexte phonétique. Ainsi, par exemple, on attribue aux voyelles en général et aux voyelles nasalisées en particulier de longues durées (voir Tableau I), tandis que des durées plus courtes sont attribuées aux consonnes non voisées. Les durées de base sont alors ajustées par le contexte immédiat:

a) si une consonne apparaît dans un agrégat, sa durée est écourtée si les consonnes ont des caractéristiques phonétiques similaires, ou est allongée si l'agrégat de consonnes est difficile à prononcer

b) les voyelles sont écourtées si elles sont suivies de consonnes non voisées et allongées si elles sont suivies par des consonnes voisées (surtout les sons fricatifs). Pour le contexte de plus haut niveau, celui du mot, les durées, sont écourtées pour les mots

de plusieurs syllabes et pour les mots que l'on présume être moins importants et par conséquent moins appuyés.

La ponctuation affecte la durée, car les durées sont allongées pour le mot qui vient juste devant un signe de ponctuation (on suppose que ce signe nécessite une courte pause dans le discours). Puisque de longues phrases ne comportent pas nécessairement des signes de ponctuation, une analyse syntaxique simple choisit les candidats les plus probables pour les frontières phonologiques, et allonge les mots qui les précèdent.

L'algorithme de la fréquence fondamentale (F_0) attribue une direction générale, appelée ligne de déclinaison, aux membres principaux de la phrase. Dans ce cas F_0 est décalé dans le sens positif et diminue graduellement. Les modifications principales apportées à cette méthode générale sont dues à la ponctuation: un point indiquant la fin d'une phrase entraîne une diminution importante de F_0 tandis que pour un point d'interrogation F_0 demeure élevé tout au long, avec une augmentation majeur vers la fin de la phrase. (Si la question contient un mot d'interrogation, par exemple "qui", "pourquoi", "comment", le F_0 diminue). D'autres signes de ponctuation, comme les virgules, entraînent une plus petite augmentation de F_0 juste avant la pause qui marque l'emplacement de la ponctuation. A l'intérieur d'une phrase (unités phonologiques entre les pauses présumées), F_0 varie selon l'importance présumée du mot et le nombre de syllabes qui composent le mot. L'importance d'un mot dépend du niveau d'emphase de la phrase: soit appuyé, soit non appuyé. Les mots non appuyés possèdent un F_0 qui diminue (mais qui n'atteint pas une valeur assez basse pour être au-dessous de la ligne de base de déclinaison, que la fin d'une phrase pourrait indiquer). Les mots appuyés nécessitent une augmentation de F_0 dans leur syllabe finale et parfois dans la première syllabe aussi (pour les cas contenant plusieurs syllabes, les syllabes du milieu nécessitent une légère diminution de F_0).

2.3 Synthèse de la parole

La synthèse de la parole (QUINIO et TEIL, 1970; VASSIERE, 1971; FLANAGAN et RABINER, 1973; FLANAGAN et al, 1970) convertit une suite de phonèmes et de marqueurs d'intonation en un discours, et consiste en deux parties: un convertisseur phonème-paramètre et un modèle du conduit vocal. Le modèle du conduit vocal est actuellement simulé au moyen d'un logiciel FORTRAN, mais sera remplacé bientôt par une pastille LSI pouvant effectuer la synthèse en temps réel. Le convertisseur phonème-paramètre constitue un choix de méthode fondamental. Il serait possible d'enregistrer les élements de la voix humaine et de les traiter au moyen d'une analyse prédictive linéaire (PL) pour réduire la mémoire utilisée. Alors, lorsque différents sons devraient apparaître en séquence dans une phrase donnée, les composants mémorisés pourraient être transformés et enchaînés pour produire le

discours de sortie. Alternativement, on peut effectuer la synthèse phonétiquement en utilisant une simulation directe des paramètres établis de la parole (c'est à dire, manipulation des résonnances des formants, des bandes passantes et des amplitudes).

La méthode d'enchaînement par PL possède un accent machine moins prononcé mais elle est limitée à une simulation de la voix du locuteur original. Elle est sujette aux dégradations du codage PL et elle nécessite une mémoire plus grande. La méthode phonétique peut potentiellement produire un discours parfait mais pratiquement, elle possède en général un accent machine. Nous avons choisi la deuxième méthode et avons modifié un logiciel existant de synthétiseur développé par KLATT (1980) au M.I.T. (et utilisé actuellement pour le lecteur automatique de Telesensory Systems) pour s'appliquer aux phonèmes de français. Par exemple, nous avons ajouté les voyelles nasalisées et les voyelles arrondies antérieures que l'on trouve dans la langue française mais pas dans la langue anglaise, et nous avons éliminé le phénomène anglais de diphtongaison. Un autre changement majeur a été apporté aux consonnes "liquides" /l/ et /r/. Le /l/ français est articulé en avançant plus la langue vers l'avant et avec un contact moindre avec le palais. Ceci tend à augmenter le deuxième formant comparé avec le /l/ anglais. Le /r/ anglais est un son rétrofléchi avec un troisième formant très bas. Le /r/ français, d'autre part, est prononcé sans rétroflexion du bout de la langue et généralement avec un point de constriction vélaire. L'allophone non voisé du /r/ est souvent accompagné d'un son fricatif. Cet allophone apparait quand le /r/ est adjacent à une consonne non voisée de la même syllabe (par exemple "parc", "trac" - non voisés: "large", "vrai" - voisés). Le Tableau II indique les valeurs couramment utilisées pour les trois premiers formants de toutes les sonorités françaises, et le Tableau III indique les valeurs pour les consonnes ordonnées selon les points d'articulation (puisque les consonnes ayant la même articulation, comme /t/, /d/ et /n/, possèdent les mêmes cibles).

3) CONCLUSIONS

Au cours de la première année consacrée au développement du lecteur automatique parlant français, la recherche et le développement ont été orientés vers la possibilité d'un démonstration en temps non réel. Nous pouvons maintenant démontrer une simulation de lecteur automatique pour lequel:

a) Le ROC fournit une suite de symboles à partir d'une page dactylographiée.

b) Cette suite de symboles est ensuite traitée par des logiciels successifs qui effectuent une reconnaissance des formes, un traitement linguistique et l'élaboration des paramètres de la parole.

c) La suite de paramètres sert à commander un synthétiseur logiciel qui calcule le signal numérique de parole. La parole synthétique, produite en temps non réel, a été jugée intelligible et de qualité acceptable au cours de plusiers tests d'audition. D'autres essais perceptuels objectifs seront effectués très prochainement.

Au cours de l'année qui vient, nous construirons les dispositifs nécessaires pour obtenir une lecture en temps réel et pour remplacer certaines fonctions du logiciel. L'utilisation d'un petit microprocesseur devrait aboutir à une économie dans les temps de calcul. Lorsque l'interface de tous les composants sera réalisé, il sera possible d'obtenir une démonstration en temps réel, produisant un discours alors que la page du texte est en train d'être explorée. Cette deuxième démonstration devrait mettre en valeur le fonctionnement du système dans les conditions d'utilisation désirées, c'est à dire, avec une personne manoeuvrant le ROC et écoutant le discours de sortie simultanément.

REFERENCES

Beauchemin, N. (1972) "Corrélation des durées sous l'accent en français", 7th Intern. Cong. on Phonetic Sciences (1971), A. Rigault et R. Charbonneau (Mouton: the Hague), 860-865.

Benguerel, A., (1971) "Duration of French vowels in unemphatic stress", Language & Speech 14, #4, 383-391.

Caldwell, J., (1978) "Flexible, high performance speech synthesizer", J. Acoust. Soc. Am. 64, S1, S72.

Choppy, C. et Lienard, J.S., (1975) "Un algorithme de prosodie automatique sans analyse syntaxique", J.E. Toulouse.

Divay, M. et Guyomard, R., (1977) "Conception et réalisation sur ordinateur d'un programme de transcription graphémo-phonétique", thèse de 3ème cycle, Université de Rennes.

Flanagan, J., Coker, C., Rabiner, L., Shafer, R., et Umeda, N., (1970) "Synthetic voices for computers", IEEE Spectrum 7, 22-45.

Flanagan, J. et Rabiner, L., (1973) "Speech Synthesis", (Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania).

Klatt, D., (1980) "Software for a cascade/parallel formant synthesizer", J. Acoust. Soc. Am. 67, 3, 971-995.

Kurzweil, R., (1976) "The Kurzweil reading machine: a technical overview", Science, Technology, and the Handicapped, AAAS Meeting, 3-7.

Mettas, O., (1969) "Etude sur la durée des consonnes dans l'un des parlers parisiens", Studia Linguistica 22, #2, 91-103.

O'Shaughnessy, D., (1981), "A Study of French vowel and consonant durations", J. of Phonetics, à apparaître.

Quinio, J. et Teil, D., (1970) "La synthèse de la parole à partir des diagrammes phonétiques", Revue Acoustique 3, no. 9.

Smith, A., (1977) "The timing of French, with reflections on syllable timing", Works in Progress, U. of Edinburgh 10, 97-108.

Vassière J., (1971) "Contribution à synthèse par règles du français", thèse de 3ème cycle, Université de Grenoble.

Wajskop, M., (1979) "Segmental durations of French intervocalic plosives", Frontiers of Speech Communication Research, B. Lindblom & S. Ohman (Academic Press: N.Y.), 109-123.

TABLEAU I - Durées de base des phonèmes (en millisecondes)

PHONEMES	DUREE
i,y,u	90
ø,ɛ,ɔ,a,oe	90
e,o,ɑ	140
ə	70
ɛ̃,ã,ɔ̃,œ̃	200
p,t,k	116
b,d,g	109
f,s,ʃ	176
v	115
z	128
ʒ	150
m	108
n	123
ɲ	142
l	101
r	113
w,j,ɥ	143

TABLEAU II - Valeurs cibles des 3 premiers formants pour les sonorités (en Hz)

PHONEME	EXEMPLE	F1	F2	F3
i	lit	250	2500	3000
e	bébé	360	2010	2600
ɛ	lait	480	1820	2470
ɑ	pâte	660	1160	2410
a	patte	720	1320	2400
ɔ	bonne	630	920	2430
o	beau	420	690	2400
u	bout	290	810	2130
ə	debout	470	1320	2070
y	sur	300	1800	2200
ø	feu	420	1510	2120
oɛ	boeuf	560	1300	2120
ɛ̃	fin	500	1600	2500
œ̃	brun	500	1500	2500
ɑ̃	sans	600	1000	2250
ɔ̃	bon	400	750	2200
w	bois	285	610	2150
j	biais	240	2070	3020
ɥ	huile	330	1400	2200
l	lit	330	1400	2400
r	rouge	450	900	2350

TABLEAU III - Valeurs cibles des 3 premiers formants pour les consonnes (en Hz)

PHONEMES	F1	F2	F3
p,b,f,v,m	200	900	2100
t,d,s,z,n	200	1400	2700
ʃ, ʒ	200	1650	2400
k,g, ɲ	250	1600	1900