



**HAL**  
open science

# Perception de vocoïdes postérieurs fermés synthétisés : l'effet de la configuration labiale et de la position de la langue sur les auditeurs francophones et japonophones

Takeki Kamiyama

► **To cite this version:**

Takeki Kamiyama. Perception de vocoïdes postérieurs fermés synthétisés : l'effet de la configuration labiale et de la position de la langue sur les auditeurs francophones et japonophones. XXVIIIèmes Journées d'Étude sur la Parole, May 2010, Mons, Belgique. pp.297-300. halshs-00678038

**HAL Id: halshs-00678038**

**<https://shs.hal.science/halshs-00678038>**

Submitted on 11 Mar 2012

**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.

## Perception de vocoïdes postérieurs fermés synthétisés : l'effet de la configuration labiale et de la position de la langue sur les auditeurs francophones et japonophones

Takeki Kamiyama

Laboratoire de Phonétique et Phonologie (UMR7018), CNRS / Sorbonne Nouvelle  
19 rue des Bernardins, 75 005 Paris, France  
takeki.kamiyama@univ-paris3.fr  
[http://lpp.univ-paris3.fr/equipe/takeki\\_kamiyama.htm](http://lpp.univ-paris3.fr/equipe/takeki_kamiyama.htm)

### ABSTRACT

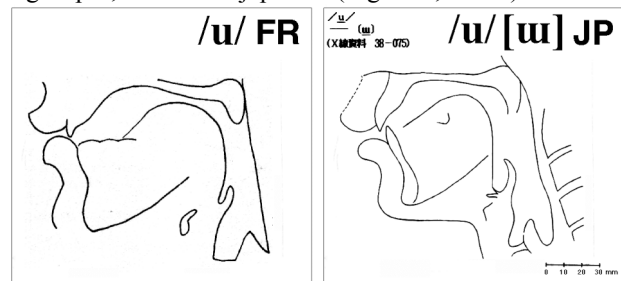
Japanese /u/ is produced with a fronter tongue constriction with less closed lips than French /u/. Such an articulatory configuration results in higher F1 and F2 (F2 > 1000 Hz) than French /u/. Variations in tongue position and lip configuration were simulated using VTCalc (Maeda [10]). A perceptual identification test was conducted using these synthesised vowels with 16 native French-speaking listeners and 16 native Japanese-speaking listeners. The results show that the back tongue position and a high degree of labiality were both essential for French speakers to perceive French /u/, whereas Japanese speakers perceived Japanese /u/ in zones where French speakers identified /u/ and /ø/. These findings explain the fact that Japanese-speaking learners of French have difficulty in distinguishing /u/ and /ø/ in perception and production.

**Keywords:** Japanese /u/, French /u/, tongue position, lips, articulatory synthesis, identification

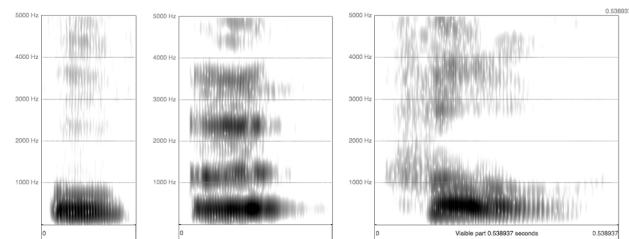
### 1. INTRODUCTION

Il est largement connu que des phonèmes de différentes langues qui sont transcrits phonémiquement avec le même symbole ne correspondent pas nécessairement aux mêmes réalisations phonétiques. C'est le cas du /u/ du français, du japonais et de l'anglais. La réalisation phonétique du /u/ japonais est moins arrondie, plus antérieure, et avec une constriction moins importante que celle du /u/ français (Uemura [14] pour le japonais, Bothorel *et al.* [1] pour le français : Figure 1), et souvent notée [ɯ] en transcription phonétique large. La différence articuloire entre le /u/ français et le /u/ japonais donne lieu à la différence acoustique suivante : le /u/ français en contexte isolé est caractérisé par deux premiers formants proches (il s'agit d'une voyelle focale) et inférieurs à 1000 Hz (Figure 2, gauche ; CALLIOPE [3] pour le contexte [pu] ; Gendrot et Adda [4] pour des données de la parole continue journalistique). Le F2 du /u/ japonais est supérieur à 1000 Hz (Figure 2, milieu ; Sugitô [13], Mokhtari et Tanaka [11] pour 22 mots) et non regroupé avec F1 (il s'agit d'une voyelle « acoustiquement centrale » : formants situés approximativement à équidistance ; Vaissière [15]). Le /u/ (comme dans

« goose ») de l'anglais américain est caractérisé par une légère diphtongaison (Wells [16], entre autres), par un F2 plus élevé que celui du /u/ français (Hillenbrand *et al.* [5], entre autres), et les deux premiers formants ne sont pas regroupés, comme en japonais (Figure 2, droite).



**Figure 1 :** Profil sagittal du /u/ français (gauche : Wioland [17]) et du /u/ japonais (droite : Uemura [14]) en contexte isolé.



**Figure 2 :** Spectrogramme à bande large (Praat [1] ; largeur de fenêtre : 5 ms) du /u/. À gauche, un /u/ français prononcé par un locuteur francophone natif dans la phrase cadre « je dis /u/ comme dans loup » (Kamiyama [7]) ; au milieu, un /u/ japonais prononcé par un locuteur japonophone natif dans la phrase cadre « sore-wa /u/ to iimasu [on appelle cela /u/] » (Kamiyama [7]) ; à droite, un /u/ dans le mot « who » prononcé par un locuteur américain (Jones *et al.* [6]).

Comment les deux paramètres articuloires, la configuration des lèvres et la position de la langue, influencent-ils sur le résultat acoustique et la perception des auditeurs francophones et japonophones ? Nous avons simulé les variations de ces deux paramètres articuloires en utilisant le synthétiseur articuloire VTCalc (Maeda [10]).

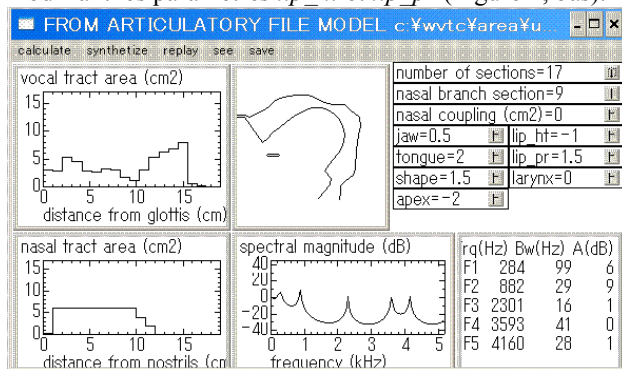
## 2. POSITION DE LANGUE ET LABIALITÉ : SYNTHÈSE ARTICULATOIRE

Le synthétiseur articulatoire VTCalc calcule les caractéristiques acoustiques correspondant à un profil sagittal, depuis la glotte jusqu'aux lèvres, spécifié par les 7 paramètres suivants :

- 1) Position de la mâchoire : *-jaw-*
- 2) Position du dos de la langue : *-tongue-*
- 3) Forme de la langue : *-shape-*
- 4) Position de l'apex de la langue : *-apex-*
- 5) Aperture des lèvres : *-lip\_ht-*
- 6) Protrusion des lèvres : *-lip\_pr-*
- 7) Hauteur du larynx : *-larynx-*

Trois séries de continuum (avec des pas espacés de façon régulière sur les dimensions articulatoires) ont été créées en partant des paramètres articulatoires d'un [u] français (Figure 3) et en en modifiant quelques-uns :

1. modification de l'ouverture des lèvres et de l'avancement de la langue : entre un [u] et un [ø] français et son extension (vers [e]), en modifiant les paramètres *jaw*, *tongue*, *shape*, *apex*, *lip\_ht* et *lip\_pr* (Figure 4, haut) ;
2. avancement de la langue uniquement : [u] français avec la langue mise progressivement vers l'avant, en modifiant les paramètres *jaw*, *tongue*, *shape* et *apex* (Figure 4, milieu) ;
3. ouverture des lèvres uniquement : [u] français avec les lèvres progressivement moins arrondies et protrusées, en modifiant les paramètres *lip\_ht* et *lip\_pr* (Figure 4, bas).



**Figure 3** : Simulation d'un [u] à partir des données articulatoires d'un [u] français sous VTCalc (Maeda [10]). Notons les deux fortes constriction (labiale et vélaire) et le F2 inférieur à 1000 Hz.

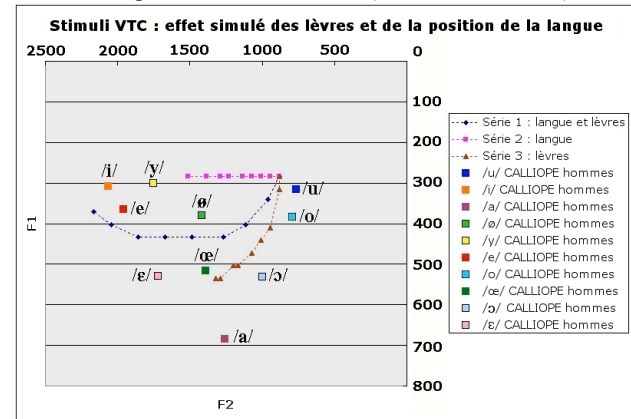


**Figure 4** : Simulation de : 1) avancement de la langue et ouverture des lèvres (en haut) ; 2) avancement de la langue seulement (au milieu) ; 3) ouverture des lèvres

seulement (en bas). Profil sagittal et vue frontale des lèvres.

Nous observons les effets acoustiques suivants sur ces 3 séries de simulation (Figure 5) :

- Série 1 : augmentation de F1, F2 (notamment de F2) ;
- Série 2 : augmentation de F2 ;
- Série 3 : augmentation de F1, F2 (notamment de F1).



**Figure 5** : Simulation d'avancement de la langue et d'ouverture des lèvres (série 1 : losanges), d'avancement de la langue seulement (série 2 : carrés), et d'ouverture des lèvres seulement (série 3 : triangles) : effet acoustique sur le plan F1-F2, en comparaison avec les valeurs formantiques de CALLIOPE [3] (contexte [pV] ; [pVR] pour [i e a œ] ; 10 hommes).

Quels sont les effets de ces deux paramètres articulatoires sur la perception des auditeurs natifs du français et du japonais ?

## 3. PERCEPTION D'AUDITEURS FRANCOPHONES ET JAPONOPHONES

Une expérience d'identification perceptive a été réalisée en utilisant les vocoïdes synthétisés présentés dans la section précédente.

### 3.1. Méthode

Vingt-cinq stimuli ont été ainsi obtenus (= 9 x 3 - 2, celui d'un [u] français étant commun aux trois séries : Figure 5). Ces stimuli ont été disposés dans un ordre semi-aléatoire.

Deux groupes d'auditeurs ont participé à cette expérience de perception. Le premier était composé de 16 auditeurs natifs du japonais n'apprenant pas le français. Ils étaient âgés de 24 à 64 ans. Tous étaient originaires de Tokyo et de ses environs. Aucun d'entre eux n'avait vécu plus d'un an à l'étranger. Le second groupe réunissait 16 auditeurs natifs du français, âgés de 18 à 38 ans, résidant en région parisienne lors de l'expérience.

L'expérience a été effectuée sur un ordinateur. L'auditeur disposait d'une souris et d'un casque. Les tâches de l'auditeur étaient d'écouter les stimuli, de les identifier comme une des voyelles de sa langue maternelle (français

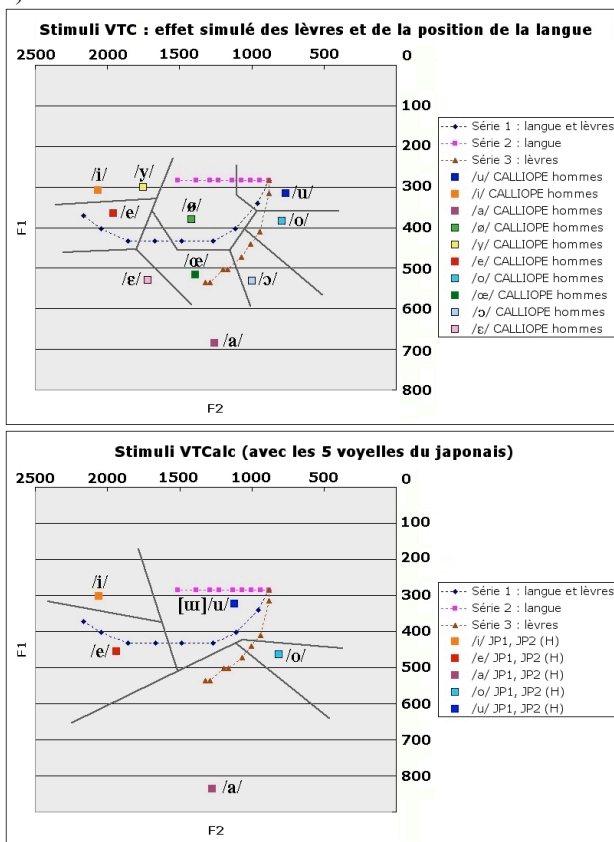
ou japonais), et de les évaluer sur une échelle de 1 (mauvais) à 5 (bon). Il était informé qu'il allait entendre des voyelles prononcées par des étrangers.

Les auditeurs japonais disposaient comme choix de réponse non seulement des cinq voyelles du japonais (/i e a o u/) mais aussi de trois syllabes avec le /j/ (/ja jo ju/). Ces dernières ont été ajoutées afin de permettre la réponse /ju/, séquence utilisée dans des mots d'emprunt pour représenter la voyelle /y/ du français (ex. Hugo /ygo/ > /juRgoR/ [ju:rgo:], ou /jugoR/ [ju:go:] ; voyelle suivie de /R/ représente une voyelle longue). Ces voyelles et syllabes ont été transcrites en syllabaires *katakana* dans l'ordre des syllabaires japonais (アイウエオ, ヤユヨ : /a i u e o/, /ja ju jo/).

L'expérience était précédée par un entraînement pour familiariser l'auditeur avec les tâches. Pendant l'expérience, chaque auditeur a entendu les stimuli dans 4 ordres différents (25 stimuli x 4 répétitions = 100 réponses par auditeur).

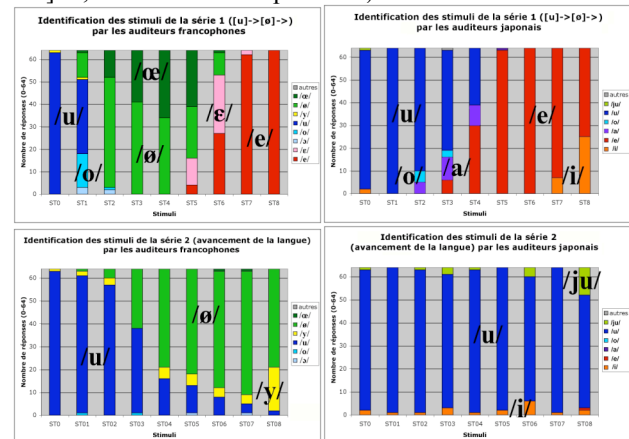
### 3.2. Résultats

Les résultats sont présentés ici sous forme de réponse modale (la réponse la plus fréquemment observée) pour chacun des stimuli et par groupe d'auditeurs (Figure 6) et de nombre de réponse de chaque stimulus (0-64 : Figure 7).



**Figure 6** : Répartition de l'espace acoustique F1-F2 selon la réponse modale des stimuli :1) (en haut) auditeurs francophones (64 réponses par stimulus : 16 auditeurs x 4 réponses), en comparaison avec les valeurs formantiques

de CALLIOPE [3] (contexte [pV] ; [pVR] pour [i e a o œ] ; 10 hommes) ; 2) (en bas) auditeurs japonophones (64 réponses par stimulus : 16 auditeurs x 4 réponses), en comparaison avec les valeurs formantiques de Kamiyama [7] (contexte : « sore-wa /V/ to iimasu [on appelle cela /V/] » ; 2 hommes x 6 répétitions).



**Figure 7** : Nombre de réponses (0-64) d'identification des stimuli des série 1 (langue et lèvres) et 2 (langue) : à gauche, 16 auditeurs francophones ; à droite, 16 auditeurs japonophones. 64 réponses par stimulus (16 auditeurs x 4 réponses).

Les résultats obtenus auprès des auditeurs francophones sont les suivants : i) un fort degré de labialité et la position postérieure de la langue sont nécessaires pour que les auditeurs francophones perçoivent le /u/ français ; ii) des vocoïdes synthétisés en tant que voyelles postérieures fermées non-arrondies ont été perçus majoritairement comme /œ/ ; iii) les stimuli qui ont un F2 supérieur à 1000 Hz ne sont pas perçus comme un bon exemplaire du /u/ français, mais plutôt comme /ø/.

Par contraste, les auditeurs japonais (n'apprenant pas le français) qui ont identifié les mêmes stimuli comme voyelles du japonais ont perçu généralement le /u/ japonais dans les zones où les auditeurs francophones ont entendu /u/ et /ø/, ce qui suggère que les auditeurs japonophones perçoivent le /u/ japonais dans une zone articulaire et acoustique (sur l'axe de F2) plus large.

En ce qui concerne les voyelles perçues en seconde position, nous observons deux cas de figure (Figure 7) : 1) d'une part, on trouve des voyelles qui ont les mêmes caractères d'antériorité et de labialité que la réponse majoritaire, avec une différence d'aperture d'un seul degré (ex. /ø/ et /œ/) ; 2) d'autre part, quand la perception majoritaire passe d'une catégorie à une autre (ex. /u/->/ø/ dans la série 2 : position de la langue), c'est l'un des deux qui constitue la réponse majoritaire, l'autre étant en seconde position.

### 4. DISCUSSION ET CONCLUSION

Les différences de comportements entre les auditeurs francophones et japonophones observées dans cette expérience de perception se résument comme suit. Dans

les zones articulatoires et acoustiques où les auditeurs francophones montrent un passage du /u/ au /ø/, les auditeurs japonophones entendent en général un /u/ japonais, avec une évaluation relativement bonne. Les francophones sont très sensibles à des écarts par rapport au /u/ français aussi bien sur le plan articulatoire (avancement de la langue et délabialisation) que sur le plan acoustique (F2 élevé), tandis que la zone du /u/ japonais est plus large, notamment au niveau de F2.

Ces résultats illustrent des caractères non-linéaires et non-univoques des correspondances entre les domaines articulatoire, acoustique et perceptif. Les voyelles « acoustiquement centrales » (F2 situé vers 1500 Hz, à mi-chemin entre F1 et F3 : Vaissière [15]) ont été observées dans plusieurs séries de continuum synthétisés : 1) autour de [ø] français (série 1) ; 2) avec les lèvres de [u] et la langue en avant (série 2) ; 3) avec la langue de [u] (postérieure) et les lèvres ouvertes (séries 3). Ces vocoïdes ont été tous identifiés majoritairement comme voyelles antérieures arrondies /ø/ (les deux premiers cas) et /æ/ par les auditeurs francophones, quel que soit le lieu d'articulation. Il a été effectivement montré que différentes configurations articulatoires peuvent correspondre à des sons acoustiquement similaires (compensation pour la labialité : Savariaux *et al.* [12]) et que la labialité des voyelles centrales n'est pas toujours identifiée auditivement (Lisker et Rossi [9]).

Les résultats de la présente étude expliquent également le fait que les apprenants japonophones ont des difficultés à distinguer /u/ et /ø/ sur les plans de la perception et de la production et que le /u/ français isolé prononcé par les apprenants japonophones (probablement à un lieu d'articulation proche de celui du /u/ japonais) est identifié majoritairement comme /ø/ par les auditeurs francophones natifs (Kamiyama et Vaissière [8], Kamiyama [7]).

## REMERCIEMENTS

L'auteur remercie Jacqueline Vaissière et Antonia Colazo-Simon ainsi que les deux relecteurs anonymes pour leurs commentaires et suggestions sur les versions antérieures de cet article.

## BIBLIOGRAPHIE

- [1] P. Boersma, D. Weenink. *Praat: doing phonetics by computer* (logiciel). <http://www.praat.org/>
- [2] A. Bothorel, P. Simon, F. Wioland, and J.-P. Zerling. *Cinéradiographie des voyelles et consonnes du français*. Publications de l'Institut de Phonétique de Strasbourg, Strasbourg, 1986.
- [3] CALLIOPE. *La parole et son traitement automatique*. Masson, Paris, Milano, Barcelona, Mexico, 1989.
- [4] C. Gendrot, M. Adda-Decker. Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande. In *Proc. of the Workshop MIDL*, pages 7-12, 2004.
- [5] J. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5): 3099-3111, 1995.
- [6] D. Jones, P. Roach, J. Hartman, and J. Setter. *Cambridge English Pronouncing Dictionary* (17th edition). Cambridge University Press, Cambridge, UK, 2006.
- [7] T. Kamiyama. *Apprentissage phonétique des voyelles du français langue étrangère chez des apprenants japonophones*. Thèse de doctorat de phonétique, Université Paris 3 – Sorbonne Nouvelle, 2009.
- [8] T. Kamiyama and J. Vaissière. Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Revue AILE – LIA*, 2: 9-41, 2009.
- [9] L. Lisker and M. Rossi. Auditory and visual cueing of the [Rounded] feature of vowels. *Language and Speech*, 35(4): 391-417, 1992.
- [10] S. Maeda. A digital simulation method of the vocal tract system. *Speech Communication*, 1(3-4): 199-229, 1982.
- [11] P. Mokhtari and K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue): 57-66, 2000.
- [12] C. Savariaux, P. Perrier, J.-P. Orliaguet, and J.-L. Schwartz. Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis. *The Journal of the Acoustical Society of America* 106(1): 381-393, 1999.
- [13] M. Sugitô. *Ôsaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [Dictionnaire CD-ROM sonore d'accent des japonais d'Osaka et de Tokyo : commentaires]. Maruzen, Tokyo, 1995.
- [14] Y. Uemura (Kokuritsu Kokugo Kenkyûjo [Institut national de la langue japonaise]). *Nihongo no boin, shiin, onsetsu: chouon undou no jikken-onseigakuteki kenkyuu* [Voyelles, consonnes et syllabes en japonais : étude en phonétique expérimentale sur les mouvements articulatoires]. Shûei shuppan, Tokyo, 1990.
- [15] J. Vaissière. *La phonétique*. Presses Universitaires de France, Paris, 2006.
- [16] J. C. Wells. *Accents of English*. Cambridge University Press, Cambridge, UK, 1982.
- [17] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Hachette, Paris, 1991.