



**HAL**  
open science

## Opposition /y/-/u/ chez sept apprenantes japonophones avec et sans aide de l'ultrason lingual: 2. identification perceptive des voyelles isolées par 16 francophones natifs

Takeki Kamiyama, Claire Pillot-Loiseau, Tanja Kocjančič Antolík

### ► To cite this version:

Takeki Kamiyama, Claire Pillot-Loiseau, Tanja Kocjančič Antolík. Opposition /y/-/u/ chez sept apprenantes japonophones avec et sans aide de l'ultrason lingual: 2. identification perceptive des voyelles isolées par 16 francophones natifs. XXXème édition des Journées d'Etude sur la Parole, Laboratoire d'informatique de l'université du Maine, Association francophone de la communication parlée, laboratoire d'informatique de Nantes Atlantique, Jun 2014, Le Mans, France. pp.649-657. hal-01188309

**HAL Id: hal-01188309**

**<https://hal.science/hal-01188309>**

Submitted on 28 Aug 2015

**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.

# Opposition /y/-/u/ chez sept apprenantes japonophones avec et sans aide de l'ultrason lingual: 2. identification perceptive des voyelles isolées par 16 francophones natifs

Takeki Kamiyama<sup>1,2</sup>, Claire Pillot-Loiseau<sup>2</sup>, Tanja Kocjančič Antolík<sup>2</sup>

(1) Linguistique Anglaise Psycholinguistique (LAPS), EA 1569, Université Paris 8 : 2 rue de la Liberté, 93526 Saint-Denis Cedex

(2) Laboratoire de Phonétique et Phonologie (LPP) UMR 7018 CNRS / Université Paris 3 Sorbonne Nouvelle : 19 rue des Bernardins, 75005 PARIS, FRANCE

takeki.kamiyama@univ-paris8.fr, claire.pillot@univ-paris3.fr,  
tkocjancic@googlemail.com

## RESUME

---

Quatre apprenantes japonophones (EXP) ont bénéficié d'un entraînement de prononciation avec rétrocontrôle par ultrason pour produire /u/-/y/ français. Elles ont été enregistrées avant (ENR1), une semaine après (ENR2), et deux mois après l'entraînement (ENR3). Trois autres apprenantes (CTR) ont aussi été enregistrées deux fois. Seize francophones natifs ont écouté des stimuli (4 occurrences de /y/ et /u/ isolés) prononcés par chaque locutrice à ENR1, 2, et 3, et ont identifié la voyelle avec une note. En ENR1, une EXP a produit le /u/ avec une note moyenne supérieure à 2/3 ; en ENR2 et ENR3, le /u/ de 2 EXP (ENR2) et celui de 3 EXP (ENR3) atteint une note supérieure à 2. Celle des CTR n'a pas dépassé 1,3. Le retour ultrason est donc utile pour /u/, dont la position de la langue est invisible.

## ABSTRACT

---

**/y/-/u/ contrast in French by seven Japanese-speaking learners with and without the help of Ultrasound: 2. perceptual identification of isolated vowels by 16 native French listeners**

Four Japanese-speaking learners (EXP) received pronunciation training with ultrasound feedback in producing French /y/-/u/. They were recorded before (REC1), a week after (REC2), and two months after the training (REC3). Three other learners (CTR) were recorded twice at similar intervals. Sixteen French native subjects listened to stimuli (4 tokens of /y/ and /u/ in isolation) pronounced by each learner in REC1, 2, and 3. They identified the vowel, providing a goodness rating. In REC1, one EXP produced /u/ that obtained a mean goodness score above 2/3 for /u/; in REC2 and REC3, the /u/ productions of 2 and 3 EXP's respectively reached a score above 2. Those produced by the CTR's did not obtain a score above 1.3. So ultrasound feedback may be useful especially for /u/, for which the tongue position is invisible.

---

MOTS-CLES : apprenantes japonaises, entraînement phonétique, ultrason, position de la langue, évaluation perceptive, auditeurs francophones, voyelles isolées

KEYWORDS: Japanese-speaking learners, pronunciation training, ultrasound, tongue position, perceptual evaluation, French-speaking listeners, vowels in isolation.

---

# 1 Introduction et but de l'étude

Il a été largement montré que le système vocalique d'une langue maternelle a un impact significatif sur l'acquisition des voyelles de langues étrangères. Les voyelles arrondies fermées antérieures, comme /y/, et postérieures, comme /u/, sont connues pour poser des difficultés aux locuteurs de langues qui ne possèdent pas ces deux types de phonèmes (Rochet 1995, Levy et Law 2010, parmi d'autres ; Flege 1987, Lauret 1998, Kamiyama and Vaissière 2009).

La voyelle antérieure arrondie fermée /y/ et la voyelle postérieure arrondie fermée /u/ sont phonologiquement en opposition dans le français parisien contemporain et parvenir à réaliser cette opposition est d'une grande importance pour les apprenants du français langue étrangère (FLE). Des locuteurs japonophones natifs de Tokyo ont tendance à produire le /u/ français avec un deuxième formant (F2) plus élevé que les locuteurs francophones natifs, ce qui résulte en une voyelle identifiée plutôt comme /ø/ par des auditeurs francophones (Kamiyama et Vaissière 2009). Cela est directement lié à la nature de la voyelle fermée non antérieure /u/ en japonais de Tokyo, communément transcrite [ɯ]. D'un point de vue acoustique, cette voyelle montre un F2 plus élevé (supérieur à 1000Hz) que celui du /u/ français ; d'un point de vue articulatoire, la langue est située moins loin en arrière et les lèvres sont moins arrondies (Bothorel *et al.* 1986; Uemura & Takada 1990).

Il est donc nécessaire, pour les japonophones apprenant le français, de positionner leur langue aussi loin en arrière que le font les francophones natifs quand ils produisent la voyelle /u/. Toutefois, parce que la position linguale en arrière ne peut pas être observée durant la parole continue, la correction de cette position se pratique couramment uniquement sur la base d'une évaluation perceptive. D'un autre côté, l'imagerie ultrason de la langue, qui est la technique la moins invasive, la plus fiable et la plus économique, utilisée pour observer la langue durant la parole, permet le contrôle direct et en temps réel de la position linguale. Cette visualisation a montré son efficacité pour traiter des troubles de la parole (Bernhardt *et al.* 2008), mais il n'existe pas à notre connaissance de recherches prouvant l'efficacité de cet équipement pour améliorer la prononciation du FLE.

Le but de notre recherche est d'évaluer la production de l'opposition entre les voyelles françaises /y/ et /u/ par quatre apprenantes japonaises de niveau intermédiaire ayant suivi des leçons spécifiques à l'aide de l'ultrason linguale pour la production de ces voyelles, et de trois autres apprenantes de même origine n'ayant pas bénéficié de cette aide spécifique. Cet article présente une étude d'évaluation auditive des productions acoustiques des apprenantes par des auditeurs francophones natifs.

## 2 Méthode

Les apprenantes du groupe expérimental (EXP), EXP1 (33 ans, niveau B2), EXP2 (31 ans, B1), EXP3 (28 ans, B1) et EXP4 (29 ans, B2), de niveau de français intermédiaire, ont, en plus de ces trois leçons spécifiques avec l'ultrason, bénéficié de 12 semaines de formation traditionnelle à la phonétique du FLE, comme pour

celles du groupe contrôle (CTR), CTR1 (29 ans, niveau B1), CTR2 (33 ans) et CTR3 (33 ans, niveau B1), qui n'ont pas bénéficié de ces leçons. Toutes sont natives de Tokyo ou des environs, sauf CTR2 de Fukuoka et EXP4 de Kobe. Elles ont commencé à apprendre le français adulte et vivent en France depuis au moins deux ans, sauf CTR1 (5 mois). Une locutrice native française de 42 ans (NAT1) a également été enregistrée.

Pour mettre en évidence les différences dans l'articulation de [y] et [u], les données ultrason et acoustiques ont été enregistrées en même temps avec une synchronisation de ces deux signaux, en utilisant le logiciel *Articulate Assistant Advanced* (AAA, Articulate Instruments Ltd, 2008). La TABLE 1 résume l'ensemble des enregistrements réalisés par sujet et leur chronologie.

Chronologie	1 <sup>er</sup> enregistrement	2 <sup>ème</sup> enregistrement	3 <sup>ème</sup> enregistrement	Groupe
Enregistrements et sujets	ENR 1 (1 semaine avant 1 <sup>ère</sup> leçon et formation 12 semaines)	ENR 2 (1 semaine après 3 <sup>ème</sup> leçon et formation 12 semaines)	ENR 3 (2 mois après le ENR 2)	
NAT1	X			Référence
EXP1	X	X	X	Expérimental
EXP 2	X	X	X	
EXP 3	X	X	X	
EXP 4	X	X	X	
CTR 1	X (sans leçon ultrason)	X (sans leçon ultrason)		Contrôle
CTR 2	X (sans leçon ultrason)	X (sans leçon ultrason)		
CTR 3	X (sans leçon ultrason)	X (sans leçon ultrason)		

TABLE 1 – Enregistrements réalisés pour chaque sujet (19 au total).

Les quatre EXP bénéficié de **trois sessions d'entraînement** de 45 minutes dans lesquelles l'ultrason lingual a été utilisé pour réaliser et contrôler la position de la langue pour les voyelles cibles /y/ et /u/. L'entraînement a débuté avec la répétition avec puis sans image ultrason, puis la lecture de : 1) voyelles /y/ et /u/ isolées puis alternées, 2) syllabes avec puis sans allongement de la voyelle cible, où /y/ et /u/ apparaissaient en position initiale, finale puis médiane. Les voyelles cibles de ces syllabes apparaissaient dans des contextes phonétiques facilitants, neutres puis difficiles tels qu'ils ont été définis pour le français par Callamand (1981). 3) Ces leçons se sont poursuivies avec la répétition puis la lecture de mots respectant ces contextes, puis de phrases et textes contenant ces mots. Le protocole exact de cet apprentissage était ajusté aux capacités et aux préférences de chaque participant (par exemple : position finale des voyelles cibles parfois plus favorisante que la position initiale dans les syllabes, donc abordée avant ; la réalisation du [u] est favorisée par le contexte [uk] chez certaines, [uœ] chez d'autres).

Le corpus était composé de dix répétitions (par deux séries de cinq) de [y] et [u] isolés, ainsi que les voyelles isolées [a], [i] et, pour les sept japonophones, [u]. Des logotomes, mots et phrases avec [y] et [u] ont aussi été enregistrés mais dans l'étude

d'évaluation auditive présentée dans cet article, seules les voyelles isolées ont été utilisées comme stimuli.

Les données acoustiques, enregistrées à 22050Hz, 16bits, ont permis l'analyse des quatre premiers formants mesurés à 25%, 50% et 75% de la durée de chaque voyelle sous Praat (Boersma & Weenink 2013).

Parmi les 10 répétitions de chaque locutrice dans chaque séance d'enregistrement, 4 (les 3<sup>e</sup>, 4<sup>e</sup>, 8<sup>e</sup> et 9<sup>e</sup>) ont été extraites, afin d'éviter que le test soit trop long (ce qui a permis de limiter le temps de passation à 25 minutes environ). Ainsi ont été sélectionnés 228 stimuli composés de 3 voyelles cibles (/u/, /y/ ainsi que /i/) x 4 répétitions x 19 séances d'enregistrement (= 4 EXP x 3 + 3 EXP x 2 + 1 EXP x 1). À cela ont été ajoutées 19 occurrences du /a/ produites par les 8 locutrices, ce qui fait 247 stimuli au total. Ces stimuli ont été placés dans un ordre semi-aléatoire pour que la même voyelle cible ne se succède pas 3 fois ou plus, et que la même locutrice ne se suive pas. Deux listes ont été préparées avec l'ordre inverse des stimuli : l'une a été présentée à la moitié des auditeurs, l'autre aux autres auditeurs.

Un total de 16 auditeurs francophones natifs originaire de la France métropolitaine, âgés de 23 à 63 ans, ont participé à l'expérience. Ils ont été informés qu'ils allaient entendre des voyelles du français prononcées par des étudiantes étrangères, et qu'ils devraient indiquer quelle voyelle ils auraient entendue et si la voyelle aurait été un plus ou moins bon exemple de la catégorie vocalique choisie (mauvais / moyen / bon). Les stimuli ont été présentés sur un ordinateur à travers un casque audio, en utilisant Praat (Boersma & Weenink 2013), et les auditeurs ont répondu en cliquant sur les cases correspondantes sur l'écran. Chaque bouton de réponse contenait le symbole phonémique de la voyelle et un mot exemple (ex. « /i/ (lit) »). Chaque stimulus était précédé par un bip sonore (440 Hz, 50 ms) et un silence d'une seconde environ. Avant de passer le test, les auditeurs ont 1) lu les consignes écrites présentées sur un écran ; 2) passé une phase d'habituation où ils pouvaient entendre les voyelles isolées d'une locutrice native et voir les symboles phonétiques correspondants ainsi que des mots exemples ; et 3) passé un entraînement. Le temps de passation était de 30 minutes environ.

## 3 Résultats

### 3.1 Analyse des réponses

Les réponses enregistrées ont été traitées de manière suivante (cf. Guion et al. 2000). Premièrement, le nombre et les types de réponses (voyelles) obtenus pour chacune des deux voyelles cibles (/u/ ou /y/) prononcées par chaque locutrice dans chaque séance d'enregistrement ont été calculés (64 réponses = 16 auditeurs x 4 productions). Deuxièmement, le score a été calculé pour chaque auditeur en multipliant la note d'évaluation par la fréquence de réponses. Ainsi, la voyelle cible /u/ prononcée par une locutrice dans une séance d'enregistrement a été identifiée 3 fois comme /u/ et 1 fois comme une autre voyelle (/ø/, par exemple) par un auditeur donné, avec des notes de 2, 2, 1, et 1, respectivement, obtient des scores de 1,25 (= (2+2+1)/4) pour la réponse /u/ (correspondance entre la cible et la

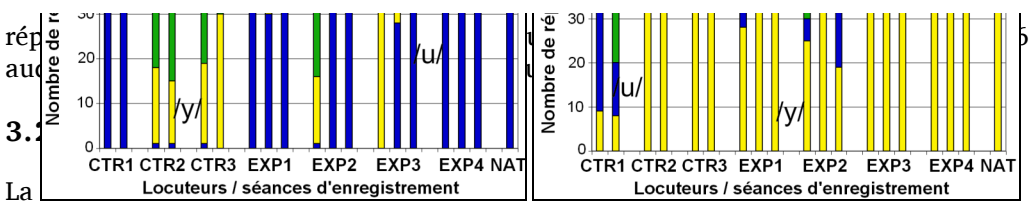


FIGURE 1 montre le nombre et les types de réponses obtenues pour les cibles /u/ et /y/.

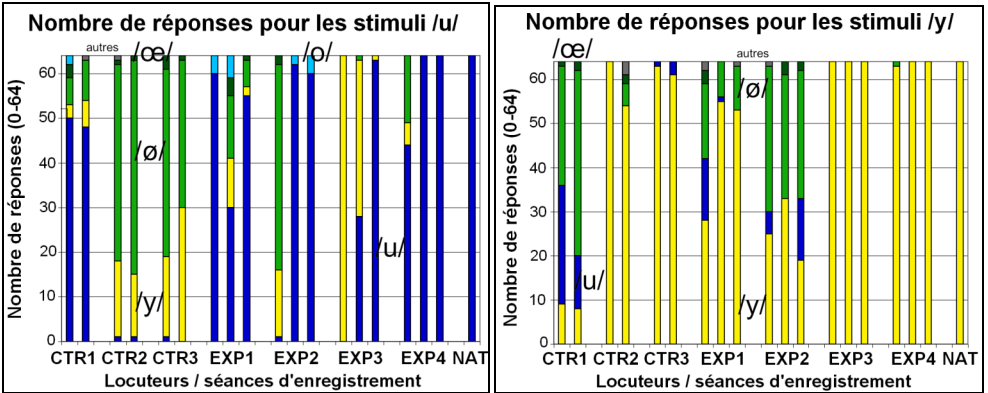


FIGURE 1 – Nombre et types de réponses d'identification aux stimuli des voyelles cibles /u/ (gauche) et /y/ (droite). De gauche à droite : ENR1, ENR2, ENR3 pour chaque locutrice.

### 3.2.1 Voyelle cible /u/

Dans le groupe CTR, seule CTR1 a obtenu la majorité (50 en ENR1, 48 en ENR2) de réponses « attendues » (correspondance entre la voyelle cible et la voyelle identifiée par les auditeurs). CTR2 et CTR3 ont produit des voyelles identifiées comme /ø/ ou /y/ dans les deux séances d'enregistrement. Chez les apprenantes EXP, nous pouvons observer une amélioration, sauf chez EXP1. Le nombre de réponses pour /u/ augmente progressivement (mise à part la diminution de ENR2 à ENR3 chez EXP2 (de 62 à 60 sur 64)), et atteint presque le plafond (entre 60 et 64) en ENR2 (EXP2 et EXP4) et en ENR3 (EXP2, EXP3, EXP4). Notons cependant que 3 apprenantes (CTR1, EXP1 et EXP4) ont obtenu la majorité de réponses « attendues » dès ENR1, mais le score moyen est inférieur à 1,5/3 sauf EXP1 (voir section 3.3).

### 3.2.2 Voyelle cible /y/

Parmi les CTR, CTR2 et CTR3 ont obtenu une grande majorité de réponses attendues malgré une légère baisse de ENR1 à ENR2 (de 64 à 54 chez CTR2, de 63 à 61 chez CTR3), alors que les productions de CTR1 ont été identifiées plutôt comme /ø/ (27 en ENR1 et 42 en ENR2) ou /u/ (27 en ENR1 et 12 en ENR2). Quant aux EXP, EXP3 et EXP4 ont obtenu la quasi-totalité de réponses attendues. Chez EXP1, il y a une amélioration entre ENR1 (28) et ENR2 et 3 (55 et 53, respectivement), alors que EXP2

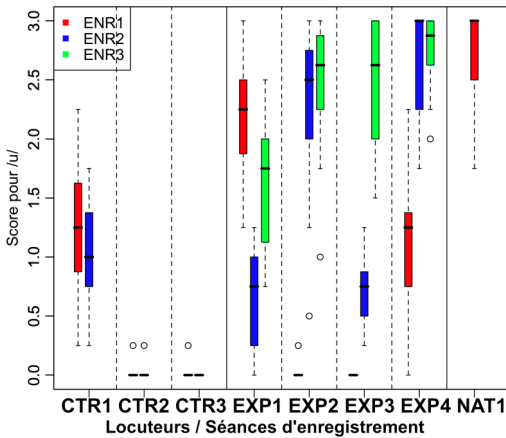
montre des valeurs moins élevées (25, 33 et 19), avec des réponses pour /ø/ aussi nombreuses (33, 28 et 29). Notons cependant que 4 apprenantes (CTR2, CTR3, EXP3 et EXP4) ont obtenu la majorité de réponses « attendues » dès ENR1, avec des scores moyens supérieurs à 2,5/3, sauf CTR3 (voir section 3.3).

### 3.3 Scores de réponses attendues (bonne identification)

Nous nous concentrons maintenant sur les scores obtenus pour les réponses attendues pour les voyelles cibles /u/ et /y/, afin de savoir s'il y a une amélioration de qualité de production (FIGURE 2). Tout d'abord, il n'y a pas d'effet significatif de l'auditeur sur les résultats ( $F_{(15, 592)} = 1,27, p = 0,21$ ). En ENR1, une seule EXP a produit le /u/ obtenant une note moyenne supérieure à 2 (sur 3) ; en ENR2, le /u/ de 2 EXP (EXP 2 et 4), et en ENR3, celui de 3 EXP (EXP 2, 3, 4) atteint une note supérieure à 2. Celui des CTR ne dépasse pas 1,3. Quant à /y/, 2 EXP (EXP 3 et 4) ont obtenu un score supérieur à 2,5 dans toutes les séances d'enregistrement, mais une seule CTR (CTR2) atteint à ce niveau seulement en ENR1.

Entre les séances d'enregistrement, nous n'observons aucune amélioration chez les CTR, même s'il y a des baisses de score en ENR2 par rapport à ENR1. En revanche, les EXP montrent une amélioration progressive pour le /u/ sauf EXP1, et pour le /y/, soit une amélioration progressive (EXP1), soit un effet plafond (EXP3 et 4), sauf EXP2.

Score pour /u/ : 16 auditeurs francophones natifs



Score pour /y/ : 16 auditeurs francophones natifs

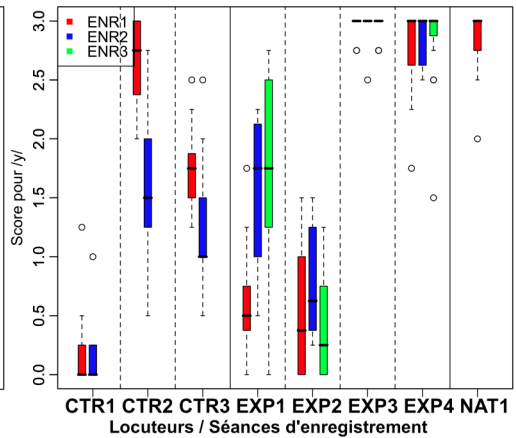


FIGURE 2 – Score moyen (0-3) pour la réponse /u/ obtenue par les stimuli de /u/ (gauche) et pour la réponse /y/ obtenu pour /y/ (droite). De gauche à droite : ENR1, ENR2, ENR3 pour chaque locutrice.

L'ANOVA mesure répétée montre un effet significatif du groupe de locuteur (CTR/EXP/NAT :  $F_{(2, 598)} = 94,61, p < 0,0001$ ), de la voyelle cible ( $F_{(1, 598)} = 40,73, p < 0,0001$ ) et de la séance d'enregistrement ( $F_{(1, 598)} = 33,43, p < 0,0001$ ) sur les scores obtenus.

#### 3.3.1 Groupe EXP : comparaison entre les séances d'enregistrements

### 3.3.1.1 Voyelle cible /u/

Les scores moyens de chaque séance d'enregistrement sur les 4 apprenantes (ENR1 : 0,85 ; ENR2 : 1,60 ; ENR3 : 2,32) étaient significativement différents (ENR1- ENR2 : test t de Student :  $t_{(63)} = 3,91$ ,  $p = 0,0001$  ; ENR2- ENR3 :  $t_{(63)} = 6,93$ ,  $p < 0.0001$ ).

### 3.3.1.2 Voyelle cible /y/

Le score moyen de ENR2 sur les 4 apprenantes (2,04) était significativement différent de celui de ENR1 (1,74 :  $t_{(63)} = 4,58$ ,  $p < 0,0001$ , selon le test t apparié de Student), mais celui de ENR3 (1,97) ne l'était pas de celui de ENR2 ( $t_{(63)} = 1,24$ ,  $p = 0,22$ ). Celui de ENR3 était significativement supérieur à celui de ENR1 ( $t_{(63)} = 2,83$ ,  $p = 0,006$ ).

### 3.3.2 Comparaison entre les EXP et CTR

Les différences de scores entre ENR1 et ENR2 (« amélioration » = ENR2- ENR1) ont été comparées entre les EXP et CTR.

L'amélioration moyenne du /u/ sur toutes les apprenantes de chaque groupe était de -0,10 pour les CTR et de 0,75 pour les EXP, ce qui correspond à une différence significative (test t non-apparié de Student :  $t_{(110)} = 3,75$ ,  $p = 0,0003$ ). Celle du /y/ était de -0,54 pour les CTR et de 0,30 pour les EXP. La différence est également significative ( $t_{(110)} = 7,83$ ,  $p < 0,0001$ ).

### 3.4 Corrélation avec les données acoustiques

La FIGURE 3 montre la distribution des scores de réponses attendues pour les voyelles cibles /u/ et /y/ en fonction des données acoustiques.

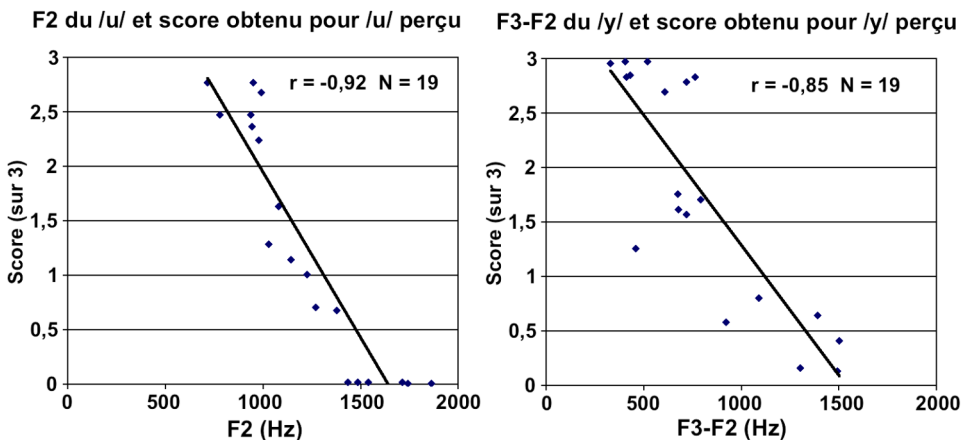


FIGURE 3 – Corrélation entre : gauche : la réponse /u/ obtenue par les stimuli de la voyelle cible /u/ (ordonnée) et F2 moyen des stimuli correspondants (abscisse) (gauche) ; droite : la réponse /y/ que ceux de la voyelle cible /y/ ont obtenu (en ordonnée) et la distance moyenne entre F2 et F3 des stimuli correspondants (en abscisse) (droite).



Les scores pour /u/ que les stimuli du /u/ ont obtenus montrent une corrélation hautement significative avec la moyenne du F2 des stimuli correspondants (coefficient de corrélation de Pearson :  $r = -0,92$ ,  $N = 19$ ,  $p < 0.001$ ). Une corrélation également significative est observée entre les scores pour /y/ que les stimuli du /y/ ont obtenus et la distance entre F2 et F3 des stimuli correspondants ( $r = -0,85$ ,  $N = 19$ ,  $p < 0.001$ ).

Ces résultats corroborent les caractéristiques acoustiques des voyelles isolées du français chez les locuteurs natifs décrites dans la littérature (F1 et F2 proches et bas pour /u/ ; F2 et F3 proches pour /y/ : Liénard 1977, Vaissière 2007, entre autres ; notons qu'une petite distance F1/F2 pourrait également être atteinte par des F1 et F2 plus élevés, dans lequel cas les auditeurs francophones percevraient [o], [ɔ], ou encore [ɑ]).

#### **4 Discussion et conclusion**

L'expérience d'évaluation auditive présentée ci-dessus montre que les apprenantes EXP avaient tendance à obtenir un meilleur score après l'entraînement par ultrason (ENR 2 et 3), alors que la même tendance n'a pas été observée chez les CTR en ENR2. Ces données suggèrent que le retour ultrason est utile pour améliorer la production de /u/-/y/ du français, mais tout particulièrement pour la voyelle postérieure /u/, dont la position de la langue est invisible.

EXP1 présente des données intéressantes : le score du /u/ diminue de ENR1 (2,23 de moyenne) à ENR2 (0,7), alors que celui du /y/ augmente entre ces deux séances (de 0,64 à 1,56) ; en ENR3 elle obtient 1,63 pour /u/ et 1,7 pour /y/. Cela suggère qu'à la fin, cette apprenante a amélioré la production des deux voyelles en les distinguant mieux, ce qui est confirmé par les données articulatoires d'ultrason qui montrent que les positions de langue pour les deux voyelles sont séparées de plus en plus nettement d'une séance d'enregistrement à l'autre (locutrice CS dans Pillot-Loiseau *et al.* 2013).

Il sera donc nécessaire de réexaminer plus en détail les liens entre les données articulatoires ultrason, les données acoustiques et les résultats de l'évaluation auditive.

Le présent article traite uniquement des voyelles isolées, mais des analyses acoustiques des logatomes sont présentées dans la première partie de cette étude, faisant l'objet d'un autre article de ces actes (Pillot-Loiseau *et al.*).

#### **Remerciements**

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'Avenir portant la référence ANR-10-LABX-0083.

## Références

- BERNHARDT, B. M., BACSFALVI, P., ADLER-BOCK, M., SHIMIZU, R., CHENEY, A., GIESBRECHT, N., O'CONNELL, M., SIRIANNI, J., RADANOV, B. (2008). Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clinical Linguistics & Phonetics*, 22(2): 149–162.
- BOERSMA, P., WEENINK, D. (2013). Praat: doing phonetics by computer [logiciel]. Version 5.3.45, retirée en avril 2013 de <http://www.praat.org/>
- BOTHOREL, A., SIMON, P., WIOLAND, F., ZERLING, J.P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Strasbourg, Publications de l'Institut de Phonétique de Strasbourg.
- GUION, S. G., FLEGE, J. E., AKAHANE-YAMADA, R., PRUITT, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107(5): 2711–2724.
- KAMIYAMA, T., VAISSIÈRE, J. (2009). Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Acquisition et interaction en langue étrangère Aile... Lia*, 2: 9-41.
- LIÉNARD, J.-S. (1977). *Les processus de la communication parlée : Introduction à l'analyse et la synthèse de la parole*. Masson.
- PILLOT-LOISEAU, C., KOJANČIČ ANTOLÍK, T. & KAMIYAMA, T. (2013). Contribution of ultrasound visualisation to improving the production of the French /y/-/u/ contrast by two Japanese learners. Poster présenté au Workshop *Phonetics, phonology and languages in contact : Contact varieties, multilingualism, and second language learning*, Paris, France.
- UEMURA, Y., TAKADA, M. (Kokuritsu Kokugo Kenkyūjo [National Institute for the Japanese Language]) (1990). *Nihongo no boin, shiin, onsetsu: chouon undou no jikken-onseigakuteki kenkyuu* [Vowels, consonants and syllables in Japanese: experimental phonetics study in articulatory movements]. Tokyo, Shūei shuppan.
- VAISSIÈRE, J. (2007). Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages. In SOLÉ, M. J., BEDDOR, P. S., OHALA, M. (eds), *Experimental Approaches to Phonology*, p. 54-71, Oxford University Press.