HAL
archives-ouvertes.fr

# Apprentissage phonétique des voyelles du français langue étrangère chez les apprenants japonophones

Takeki Kamiyama

▶ **To cite this version:**

Takeki Kamiyama. Apprentissage phonétique des voyelles du français langue étrangère chez les apprenants japonophones. Linguistique. Université de la Sorbonne nouvelle - Paris III, 2009. Français. NNT : 2009PA030148 . tel-00473029v2

**HAL Id: tel-00473029**
**https://tel.archives-ouvertes.fr/tel-00473029v2**

Submitted on 6 Feb 2014

**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.

**UNIVERSITÉ SORBONNE NOUVELLE – PARIS 3**

**École Doctorale 268 « Langage et Langues : description, théorisation, transmission »**

# Apprentissage phonétique des voyelles du français langue étrangère chez des apprenants japonophones

## THÈSE

pour obtenir le grade de

Docteur de l'Université Sorbonne Nouvelle – Paris 3

Discipline : Phonétique

présentée et soutenue publiquement le 28 novembre 2009

par

**Takeki KAMIYAMA**

Thèse dirigée par Madame Jacqueline VAISSIÈRE, professeur

Membres du jury :

Monsieur Rudolph SOCK, pré-rapporteur

Monsieur François WIOLAND, pré-rapporteur

Monsieur Jean-Yves DOMMERGUES

Monsieur Pierre HALLÉ

Monsieur Shinji MAEDA

Madame Jacqueline VAISSIÈRE, directeur

**PRESENTATION DE L'ETUDE** ..... **7**

**LISTE DES SIGLES ET ABREVIATIONS** ..... **10**

**PREMIERE PARTIE : INTRODUCTION** ..... **11**

RÉSUMÉ ..... 11

1.1. LE FRANÇAIS LANGUE ÉTRANGÈRE ENSEIGNÉ AU JAPON ..... 12

1.2. L'ENSEIGNEMENT ET L'APPRENTISSAGE DE LA PRONONCIATION DES LANGUES ÉTRANGÈRES ..... 14

1.3. ÉTUDES ANTÉRIEURES SUR L'ACQUISITION DE LA PHONÉTIQUE ET DE LA PHONOLOGIE DES LANGUES ÉTRANGÈRES / SECONDES ..... 30

1.4. PHONÉTIQUE ET PHONOLOGIE DU JAPONAIS DE TOKYO -- LANGUE SOURCE (L1) DE NOTRE ÉTUDE ..... 37

*1.4.1. Les voyelles* ..... *37*

*1.4.2. Les consonnes* ..... *52*

*1.4.3. Les structures syllabiques* ..... *59*

*1.4.4. La prosodie : l'accent lexical et l'intonation* ..... *65*

*1.4.5. Convention de transcription* ..... *74*

1.5. PHONÉTIQUE ET PHONOLOGIE DU FRANÇAIS PARISIEN ..... 76

*1.5.1. Les voyelles : généralité* ..... *77*

*1.5.2. Les voyelles : des données acoustiques* ..... *80*

*1.5.3. Les consonnes* ..... *82*

*1.5.4. Les structures syllabiques* ..... *83*

*1.5.5. La prosodie* ..... *85*

1.6. PHONÉTIQUE ET PHONOLOGIE DE L'ANGLAIS AMÉRICAIN (*GENERAL AMERICAN*) ..... 90

*1.6.1. Le système vocalique de l'anglais américain (GA :* General American*)* ..... *90*

*1.6.2. D'autres caractéristiques à noter* ..... *102*

1.7. COMPARAISON DES TROIS LANGUES CONCERNÉES : LE JAPONAIS DE TOKYO (L1), L'ANGLAIS AMÉRICAIN (L2), ET LE FRANÇAIS PARISIEN (L3) ..... 103

1.8. L'APPRENTISSAGE DES VOYELLES L2 CONCERNANT LE FRANÇAIS, L'ANGLAIS, ET LE JAPONAIS ..... 105

*1.8.1. Types d'expérience de perception et de production* ..... *105*

*1.8.2. Études inter-langues antérieures concernant les voyelles du français, de l'anglais, et du japonais* ..... *108*

1.9. HYPOTHÈSE DE NOTRE ÉTUDE ..... 116

**DEUXIEME PARTIE : DES ETUDES PRELIMINAIRES SUR L'APPRENTISSAGE DES DIX VOYELLES ORALES ISOLEES DU FRANÇAIS PARISIEN PAR DES APPRENANTS JAPONOPHONES.117**

RÉSUMÉ ..... 117

2.1. LES VOYELLES DU JAPONAIS DE TOKYO ET DU FRANÇAIS PARISIEN ..... 118

*2.1.1. Les 5 voyelles isolées du japonais prononcées par des locuteurs natifs du japonais de Tokyo (Expérience 1).* ..... *118*

*2.1.2. Les voyelles isolées du français prononcées par des locuteurs natifs (Expérience 2)* ..... *120*

2.2. ANALYSE PERCEPTIVE : PERCEPTION DES VOYELLES DU FRANÇAIS PAR DES APPRENANTS JAPONOPHONES : IDENTIFICATION DES VOYELLES ISOLÉES (EXPÉRIENCE 3) ..... 125

*2.2.1. Les stimuli* ..... *125*

*2.2.2. Les procédures* ..... *126*

*2.2.3. La perception d'auditeurs natifs* ..... *126*

*2.2.4. Les apprenants japonophones* ..... *129*
2.3. ANALYSE ACOUSTIQUE : LA PRODUCTION DES VOYELLES ORALES DU FRANÇAIS PAR DES APPRENANTS JAPONOPHONES : LECTURE ET RÉPÉTITION (EXPÉRIENCE 4) ..... 132
*2.3.1. Lecture* ..... *133*
*2.3.2. Répétition* ..... *142*
2.4. IDENTIFICATION : LES VOYELLES ORALES ISOLÉES PRODUITES PAR DES APPRENANTS JAPONOPHONES ET PERÇUES PAR DES AUDITEURS FRANCOPHONES NATIFS : TEST D'IDENTIFICATION ET D'ÉVALUATION (EXPÉRIENCE 5) 151
*2.4.1. Les stimuli* ..... *151*
*2.4.2. Les auditeurs* ..... *152*
*2.4.3. Les procédures* ..... *152*
*2.4.4. Les résultats* ..... *153*
2.5. CONCLUSION PARTIELLE ..... 174

**TROISIEME PARTIE : LA PERCEPTION ET LA PRODUCTION DES VOYELLES /U/, /Y/ ET /Ø/ EN CONTEXTE ISOLE ..... 176**

RÉSUMÉ ..... 176
3.1. ANALYSE PERCEPTIVE : PERCEPTION DES VOYELLES /u y ø/ PAR DES APPRENANTS JAPONOPHONES : DISCRIMINATION AXB SUR DES VOYELLES ISOLÉES (EXPÉRIENCE 6) ..... 178
*3.1.1. Les stimuli* ..... *178*
*3.1.2. Les procédures* ..... *179*
*3.1.3. La perception par des apprenants japonophones* ..... *180*
*3.1.4. La perception par des non-apprenants : locuteurs natifs du français et natifs du japonais n'apprenant pas le français* ..... *185*
3.2. ANALYSE ACOUSTIQUE : PRODUCTION DES VOYELLES /u y ø/ PAR DES APPRENANTS JAPONOPHONES DE DIVERS NIVEAUX D'APPRENTISSAGE (EXPÉRIENCE 7) ..... 188
*3.2.1. Le corpus* ..... *188*
*3.2.2. Les locuteurs* ..... *189*
*3.2.3. Les procédures* ..... *190*
*3.2.4. Les analyses* ..... *191*
*3.2.5. Les résultats* ..... *191*
3.3. IDENTIFICATION : PERCEPTION DES VOYELLES /u y ø/ PRODUITES PAR DES APPRENANTS JAPONOPHONES AUPRÈS D'AUDITEURS FRANCOPHONES NATIFS (EXPÉRIENCE 8) ..... 205
*3.3.1. Les stimuli* ..... *205*
*3.3.2. Les auditeurs* ..... *209*
*3.3.3. Les procédures* ..... *209*
*3.3.4. Les analyses des réponses* ..... *210*
*3.3.5. Les résultats* ..... *210*
3.4. PERCEPTION DE VOYELLES SYNTHÉTISÉES (SYNTHÈSE ARTICULATOIRE ET SYNTHÈSE À FORMANTS) AUPRÈS D'AUDITEURS NATIFS FRANCOPHONES ET JAPONOPHONES (EXPÉRIENCE 9) ..... 218
*3.4.1. Les stimuli* ..... *218*
*3.4.2. Les auditeurs* ..... *232*
*3.4.3. Les procédures* ..... *232*
*3.4.4. Les analyses des données* ..... *233*
*3.4.5. Les résultats* ..... *234*
3.5. DISCUSSION ..... 246

3.6. CONCLUSION PARTIELLE ..... 264

**QUATRIEME PARTIE : L'ENSEIGNEMENT DE LA PRONONCIATION DES VOYELLES /U/, /Y/ ET /Ø/ AUX APPRENANTS JAPONOPHONES ..... 265**

RÉSUMÉ ..... 265

4.1. ANALYSE DE MANUELS GÉNÉRALISTES ..... 266

*4.1.1. Les manuels généralistes analysés ..... 266*

*4.1.2. Les points étudiés ..... 268*

*4.1.3. Les résultats ..... 269*

*4.1.4. Synthèse et discussion ..... 275*

4.2. PRÉSENTATION D'UN MANUEL DE PRONONCIATION PUBLIÉ AU JAPON ..... 277

*4.2.1. Les phénomènes traités ..... 277*

*4.2.2. L'organisation des dossiers ..... 278*

4.3. OUTILS FACILITANT LA PRISE DE CONSCIENCE DES APPRENANTS ..... 279

*4.3.1. La méthode verbo-tonale ..... 280*

*4.3.2. Continuum de vocoïdes postérieurs (Lauret) ..... 283*

*4.3.3. Image acoustique : la gravité de /u/ ..... 285*

*4.3.4. Synthèse articulatoire (VTCalc et TractSyn) ..... 293*

*4.3.5. Spectrogramme en temps réel (WaveSurfer) ..... 297*

4.4. CONCLUSION PARTIELLE ..... 298

**CONCLUSION GENERALE ..... 300**

**BIBLIOGRAPHIE ..... 303**

# Présentation de l'étude

## Apprentissage phonétique des voyelles orales du français langue étrangère chez des apprenants japonophones

L'objet d'étude de cette thèse ressort du domaine de l'application de la phonétique expérimentale à la didactique de la prononciation des langues étrangères.

La production et la perception des langues secondes sont influencées par l'expérience linguistique que les apprenants ont vécue dans leur langue native (L1), dans d'autres langues acquises et/ou apprises, et par le système phonétique et phonémique des langues cibles (L2, L3 ou plus).

Cette étude traite essentiellement des voyelles. Il est particulièrement intéressant d'étudier l'acquisition, l'apprentissage, et l'enseignement des *voyelles* des langues étrangères et secondes, car l'articulation des voyelles est plus continue que celle de la plupart des *consonnes* (notamment les occlusives), qui sont plus faciles à discriminer. Il est donc particulièrement difficile de définir, d'expliquer, et d'apprendre à reproduire les cibles articulatoires qui correspondent aux voyelles. Sur le plan perceptif, il a été montré que les voyelles sont perçues de manière moins catégorielle que les consonnes (Schouten et van Hessen 1992[1] pour une comparaison de voyelles et d'occlusives, entre autres). De plus, on observe une grande variété dans la réalisation des voyelles dans les divers parlers d'une langue, comme d'un français régional à l'autre, et les auditeurs sont particulièrement sensibles à des déviations subtiles, révélatrices d'un accent régional, ou socio-professionnel.

Les timbres des voyelles ne transmettent pas uniquement des informations linguistiques. Des timbres inattendus de voyelles, dus à une déformation du conduit vocal, sont perçus comme portant des informations attitudinales ou émotives (phrayngalisation associée au dégoût, antériorisation et palatalisation à la petitesse, *etc*.). Fónagy (1983 : 27-30) a montré que la distance minimale entre le dos de la langue et le palais pendant l'articulation de /e/ hongrois dans *Én* « moi » prononcé avec indignation était plus petite que celle de /i/ hongrois dans *Igen* « oui » prononcé avec indifférence, et donc que certaines réalisations de /e/ pourraient être plus fermées que /i/ selon les émotions[2]. Cela suggère que les écarts de timbre vocalique par rapport à l'attente des auditeurs transmettent des attitudes ou des émotions, voulues ou non par les locuteurs.

---

[1] M. E. H. Schouten, A. J. van Hessen. Modeling phoneme perception. I: Categorical perception. *Journal of the Acoustical Society of America* 92(4) : 1841-1855, 1992.

[2] I. Fónagy. *La vive voix : essais de psycho-phonétique*. Paris, Payot, 1983.

Notre objectif est d'examiner la production et la perception des voyelles arrondies fermées et mi-fermée du français /u y ø/ en contexte isolé, et de relier la production à la perception, lors de l'apprentissage phonétique et phonémique du français langue étrangère par des apprenants de langue maternelle japonaise. Le choix de ces trois voyelles sera expliqué dans la première partie, notamment dans les chapitres 1.8. et 1.9.

Le niveau des apprenants japonais étudiés dans cette étude varie entre débutant (trois mois d'apprentissage) et intermédiaire (trois-quatre ans d'apprentissage) et sera détaillé dans les deuxième et troisième parties.

Dans nos études acoustiques, les quatre premiers formants sont mesurés à plusieurs moments durant la voyelle. Pourquoi à plusieurs moments ? Les voyelles du français sont caractérisées par leur nature non diphtonguée (voir les spectrogrammes de Liénard 1977[3], reproduits à la Figure 23, p. 78), et il faut vérifier si les francophones natifs prononcent véritablement des monophtongues. Il est également nécessaire d'observer des tendances possibles de diphtongaison auprès des apprenants : il a été montré que les apprenants de certaines langues maternelles produisent des vocoïdes diphtongués pour les voyelles du français (voir Lauret 1998[4] pour les locuteurs anglophones américains).

Pourquoi les quatre premiers formants (et non seulement les deux premiers) ? De nombreuses études acoustiques qui portent sur les voyelles, notamment sur celles de l'anglais, ne présentent que les deux premiers formants sur un triangle vocalique à deux dimensions. Certes, les deux premiers formants permettent de représenter les réalisations des phonèmes vocaliques des langues telles que l'anglais et le japonais. Cependant, le troisième formant (F3) est indispensable pour représenter des systèmes avec une opposition entre les séries antérieure non arrondie et antérieure arrondie, ce qui est le cas du français. Certains locuteurs francophones montrent des valeurs similaires des deux premiers formants pour /i/ et /y/, et c'est F3 seul qui permet de distinguer ces deux voyelles. Chez d'autres encore, /i/ et /e/ sont réalisés avec des formants F1 et F2 similaires (voir la Figure 41 à la page 123). Il a été également montré que le quatrième formant (F4) permet de décrire des différences des phonèmes similaires dans différentes langues. Le /i/ français est caractérisé par un rapprochement de F3 et F4, alors que le /i/ anglais est marqué par un F2 élevé (plus élevé que celui du /i/ français, mais F3 et F4 ne sont généralement pas regroupés : voir Vaissière 2007[5], Gendrot *et al*. 2008 pour une confirmation statistique sur une

[3] J.-S. Liénard. *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole*. Paris, Masson, 1977.

[4] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français*. Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[5] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, and M. Ohala, *Experimental Approaches to Phonology*. Oxford University Press, Oxford, 54-71, 2007.

grande base de données[6]). Par ailleurs, le meilleur modèle de régression sur la forme du conduit vocal pendant la production de voyelles est obtenu quand les quatre premiers formants sont pris en compte, même pour les langues comme le japonais de Tokyo qui n'a que 5 voyelles (Mokhtari *et al*. 2007[7]). Dans la présente étude, les quatre premiers formant seront analysés et présentés afin de mieux caractériser et visualiser les voyelles produites par les locuteurs francophones natifs et les apprenants (ex. F3 et F4 sont situés loin des deux premiers formants pour le /u/ français prononcé par les locuteurs natifs, alors que les formants sont à peu près équidistants pour la même voyelle prononcée par les apprenants, comme nous le verrons dans les parties 2 et 3).

Dans la première partie, nous présenterons : 1) la situation de l'enseignement du français au Japon ; 2) diverses questions sur l'enseignement et l'apprentissage de la prononciation des langues étrangères ; 3) des études antérieures et des modèles de perception et de production des sons des langues étrangères et secondes ; 4) les caractéristiques phonémiques et phonétiques du japonais de Tokyo, du français parisien, et de l'anglais américain ; 5) des études antérieures sur la perception non native et l'acquisition des voyelles orales concernant le français, l'anglais, et le japonais. Cette partie sera conclue par les hypothèses à la base de notre recherche.

La deuxième partie traite une série d'expériences préliminaires effectuées sur les 10 voyelles orales du français auprès d'un nombre limité d'apprenants japonophones.

La troisième partie présente une série d'expériences de perception et de production des voyelles /u y ø/ auprès d'un plus grand nombre d'apprenants japonophones et de locuteurs natifs du japonais n'apprenant pas le français. Nous montrerons en particulier que la qualité de la production, la discrimination et l'identidication d'une voyelle donnée ne sont pas entièrement corrélées.

La quatrième et dernière partie traite de l'aspect didactique des phénomènes étudiés dans les parties précédentes. Nous présenterons une analyse de manuels de français langue étrangère publiés au Japon concernant le traitement des voyelles /u y ø/, ainsi que des méthodes susceptibles d'améliorer la perception et la production des voyelles.

---

[6] C. Gendrot, M. Adda-Decker, J. Vaissière. Les voyelles /i/ et /y/ du français : focalisation et variations formantiques. *XXVIIèmes Journées d'Étude sur la Parole*, juin 2008, Avignon, 205-208, 2008.

[7] P. Mokhtari, T. Kitamura, H. Takemoto, K. Honda. Principal components of vocal-tract area functions and inversion of vowels by linear regression of cepstrum coefficients. *Journal of Phonetics* 35(1) : 20-39, 2007.

# Liste des Sigles et abréviations

F0 : fréquence du fondamental

F1, F2, F3, F4 : premier, deuxième, troisième, et quatrième formants

Hz : Hertz

ms : milliseconde(s)

C : consonne

V : voyelle

H : ton haut

B : ton bas

API : Alphabet Phonétique International / Association Phonétique Internationale

L1 : langue première, langue native, langue source

L2 : deuxième langue

L3 : troisième langue

jp : japonais

fr : français

GA : *General American*

RP : *Received Pronunciation*

VOT : *Voice Onset Time* « temps d'établissement de voisement »

FLE : français Langue Étrangère

ALE : anglais Langue Étrangère

# Première partie : introduction

## RÉSUMÉ


Dans le chapitre 1.1., nous allons situer brièvement l'enseignement du français langue étrangère (FLE) au Japon. Le FLE est enseigné comme troisième langue L3, après l'anglais L2.

Puis, dans 1.2., nous discutons de diverses questions concernant l'enseignement et l'apprentissage de la prononciation des langues étrangères et secondes, en particulier :

- pourquoi enseigner la prononciation ?

- comment l'enseigner ?

- doit-on limiter le but final de l'apprentissage à l'intelligibilité, ou peut-on viser à la perfection ?

- que peut apporter la phonétique expérimentale à la didactique de la prononciation ?

Le chapitre 1.3. présente quelques-uns des modèles de production et de perception des langues étrangères.

Dans les chapitres 1.4.-1.6., nous présentons les caractéristiques phonétiques et phonémiques du japonais de Tokyo, du français parisien, et de l'anglais américain. En résumé :

**- Les références** : le nombre de phonèmes vocaliques variant d'une région à l'autre en France (Martinet 1969, Vaissière 2006, entre autres), nous prenons comme référence le français parisien. Cette variété compte, en plus des trois voyelles nasales, dix voyelles orales (y compris les trois voyelles antérieures arrondies /y ø œ/ ; quatre degrés d'aperture). Le japonais de Tokyo a cinq voyelles /i e a o u/ (Sugitô 1995, entre autres). L'anglais américain (GA : *General American*) compte 15 voyelles, mais ne contient pas une série de voyelles antérieures arrondies, ni de nasales (Wells 1982, entre autres).

**- Les différences de réalisation de la voyelle /u/ (parfois noté /uː/ pour l'anglais) dans les trois langues** : la réalisation phonétique du /u/ japonais est moins arrondie et plus antérieure (avec une constriction moins importante) que celle du /u/ français (Uemura 1990 pour le japonais, Bothorel *et al.* 1986 pour le français), et souvent décrite comme [ɯ]. Cette différence articulatoire 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 et inférieurs à 1000 Hz (CALLIOPE 1989 pour le contexte [pV] et [pVR] ; Gendrot et Adda 2004 pour les données de la parole continue), alors que le F2 du /u/ japonais est supérieur à 1000 Hz (Sugitô 1995, Mokhtari et Tanaka 2000 pour 22 mots différents pour chaque voyelle). Le /uː/ de l'anglais américain est caractérisé par une légère diphtongaison (Wells 1982, entre autres) et par un F2 plus élevé que celui du /u/ français (Hillenbrand *et al*. 1995, entre autres).

**- La voyelle /y/** : parmi les 3 voyelles antérieures arrondies du français parisien, le /y/ se distingue des deux autres par le regroupement de F2 et F3, tandis que l'énergie se répartit uniformément pour /ø œ/ (voyelles « acoustiquement centrales » : Vaissière 2006 : 73), avec F2 vers 1500 Hz. Le japonais de Tokyo et l'anglais américain n'ont pas de phonème vocalique correspondant au /y/ (à la différence des langues comme le suédois, l'allemand, le néerlandais, le finnois, le cantonais, *etc*.).

Les caractéristiques acoustiques des voyelles /u/, /y/ et /ø/ sont ainsi bien distinctes en français. Le /u/ [ɯ] japonais, qui ne présente aucun regroupement de formants, serait donc

acoustiquement proche du /ø/ français. Les trois voyelles /u/, /y/ et /ø/ sont donc nouvelles soit phonémiquement soit phonétiquement pour les apprenants japonophones.

L'apprentissage de ces trois voyelles n'est pas facilité par l'expérience de l'apprentissage de l'anglais L2, car la série antérieure arrondie est absente, et le /u:/ anglais n'est pas une voyelle fermée aussi postérieure et aussi arrondie comme le /u/ français (Ostiguy *et al*. 1996, entre autres).

Flege (1987) a examiné la production de /u/ et /y/ du français par des locutrices natives de l'anglais américain, et a montré que même des locutrices expérimentées (adultes vivant à Paris depuis 11,7 ans en moyenne, et utilisant le français comme langue principale) ne parvenaient pas à produire le /u/ comme les locuteurs natifs du français (c'est-à-dire avec un F2 bas), alors qu'elles n'avaient pas de difficulté à produire le /y/ comme les natifs (avec F2 et F3 proches). L'auteur a ainsi suggéré qu'il serait plus difficile de produire de manière authentique des phones « similaires » en L2 (*i.e.* /u/) que des phones « nouveaux » (*i.e.* /y/).

Ces constatations nous mènent aux interrogations suivantes, qui constituent le corps de la thèse : les apprenants japonophones présentent-ils une tendance similaire à celle des apprenants anglophones, étant donné que la réalisation phonétique du /u/ japonais, équivalent phonémique du /u/ français, est bien éloignée de celle du /u/ français, et que le /y/ français n'a pas de phonème correspondant en japonais ? Qu'en est-il du /ø/, qui serait considéré phonémiquement nouveau, mais acoustiquement similaire au /u/ japonais ? Une série d'expériences seront présentées dans les deuxième et troisième parties afin de répondre à ces questions.

### 1.1. LE FRANÇAIS LANGUE ÉTRANGÈRE ENSEIGNÉ AU JAPON

Ce n'est que depuis 2002 que l'apprentissage d'une langue étrangère est officiellement obligatoire au collège (durée : trois ans) et au lycée (durée : trois ans) au Japon. La version révisée du texte officiel 『学習指導要領』 « *gakushuu shidou youryou* » [Le guide d'instruction d'apprentissage] pour le collège, publiée le 14 décembre 1998, et mise en vigueur le 1er avril 2002[8], compte parmi les matières obligatoires l'enseignement d'une langue étrangère, et c'est l'anglais qui devrait être choisi en principe comme langue étrangère enseignée en priorité[9]. En réalité, l'enseignement d'une langue étrangère était pratiqué dans presque tous les établissements secondaires même avant cette révision du « guide », et la langue choisie était presque unanimement l'anglais, mises à part un petit nombre d'écoles privées où les élèves pouvaient choisir le français ou l'allemand comme première langue étrangère. Même dans ces derniers cas, l'anglais était enseigné en tant que deuxième langue étrangère. Le Tableau 1 ci-dessous représente le nombre d'établissements de l'enseignement secondaire qui assurent des cours de langues autres que l'anglais et le nombre d'élèves qui suivent ces cours de langues au Japon[10].

---

[8] Une version électronique du texte original en japonais est disponible sur le site du ministère de l'Éducation, de la Culture, des Sports, des Sciences et de la Technologie (文部科学省) : http://www.mext.go.jp/b_menu/shuppan/sonota/990301/03122602.htm

Une traduction anglaise est disponible également sur le site du ministère sous le titre « *the Course of Study for Foreign Languages* » : http://www.mext.go.jp/english/shotou/030301.htm

[9] « *For compulsory foreign language instruction, English should be selected in principle* » (la traduction anglaise du « *gakushuu shidou youryou* » pour le collège).

[10] 文部科学省初等中等教育局国際教育課 Mombu-kagakushou shotou chuutou kyouikukyoku kokusai kyouikuka [Division de l'éducation internationale, Département de l'éducation primaire et secondaire, ministère de l'Éducation, de la Culture, des Sports, des Sciences et de la Technologie]. 『平成18年度高等学校等における国際交流の状況について』 *Heisei 18 nendo koutougakkou-tou ni okeru kokusai kouryuu no joukyou ni tsuite* [Sur la situation des échanges internationaux aux lycées pendant l'année scolaire 2006-

**Tableau 1** : Langues autres que l'anglais enseignées dans les établissements de l'enseignement secondaire au Japon. Le nombre total des établissements (entre parenthèses) représente le nombre réel, sans compter plusieurs fois les établissements qui assurent des cours de plusieurs langues (source : *Heisei 18 nendo koutougakkou-tou ni okeru kokusai kouryuu no joukyou ni tsuite* [Sur la situation des échanges internationaux aux lycées pendant l'année scolaire 2006-2007], p. 29).

| LYCÉES | Nombre d'établissements | | Nombre d'élèves |
|---|---|---|---|
| | mai-05 | mai-07 | mai-07 |
| Chinois | 553 | 819 | 21264 |
| Français | 248 | 393 | 10059 |
| Coréen | 286 | 426 | 8865 |
| Allemand | 105 | 157 | 3898 |
| Espagnol | 105 | 135 | 2632 |
| Russe | 25 | 39 | 544 |
| Italien | 10 | 21 | 387 |
| Portugais | 11 | 17 | 157 |
| Indonésien | 1 | 9 | 29 |
| Thaï | 2 | 10 | 22 |
| Esperanto | 2 | 2 | 21 |
| Latin | 1 | 1 | 13 |
| Philippin (tagalog, pilipino) | 3 | 3 | 3 |
| Vietnamien | 1 | 9 | 2 |
| Arabe | 1 | 1 | 2 |
| Autres | 1 | 0 | 0 |
| Total | | (788) | 47898 |

| COLLÈGES | Nombre d'établissements | | Nombre d'élèves |
|---|---|---|---|
| | mai-05 | mai-07 | mai-07 |
| Français | 12 | 17 | 2099 |
| Chinois | 14 | 24 | 921 |
| Coréen | 13 | 13 | 354 |
| Espagnol | 4 | 4 | 168 |
| Allemand | 3 | 6 | 150 |
| Autres | 1 | 0 | 0 |
| Total | | (41) | 3692 |

Il existe un petit nombre d'établissements secondaires (collèges et lycées) où le français est enseigné comme deuxième langue étrangère : selon le sondage du ministère de l'Éducation, 10059 élèves (0,3% des lycéens japonais) apprenaient le français dans 393 lycées (7,4% de tous les établissements), et 2099 élèves (0,1% des collégiens japonais) dans 17 collèges (0,2% des établissements) au 1er mai 2007. En revanche, le français se trouve parmi les langues les plus apprises, après le chinois.

L'apprentissage et l'enseignement du français et d'autres langues, à l'exception de l'anglais, sont effectués principalement après le collège et le lycée, c'est-à-dire à l'université, après au moins six ans d'apprentissage de l'anglais dans l'enseignement secondaire, quel que soit le niveau de compétences en anglais alors atteint par chaque élève[11].

---

2007], p. 29. Le document est disponible sur le site du ministère : http://www.mext.go.jp/b_menu/houdou/19/11/07103102/001.pdf

[11] Les leçons d'initiation à l'anglais sont effectuées désormais à l'école primaire. Pendant l'année scolaire 2006-2007, 95,8% des 22031 établissements publics ont assuré en moyenne 14,8 heures par an d'enseignement de l'anglais (soit 1 à 1,5 heures par mois). Source : 『小学校英語活動実施状況調査（平成18年度）』 *Shougakkou eigo katsudou jisshi joukyou chousa (Heisei 18 nendo)* [Enquête sur la situation de pratique d'activités d'anglais dans les écoles primaires (année scolaire 2006-2007)], effectuée par la Division de l'éducation internationale, Département de l'éducation primaire et secondaire, Ministère de l'éducation, de la culture, des sports, des sciences et de la technologie (文部科学省初等中等教育局国際教育課). http://www.mext.go.jp/b_menu/houdou/19/03/07030811/001.htm

C'est ainsi que la quasi-totalité des japonophones natifs apprennent le français en tant que troisième langue (L3), ou deuxième langue étrangère (第二外国語 *daini gaikokugo*), après l'anglais appris comme deuxième langue (L2) ou première langue étrangère.

Dans cette étude, l'apprentissage et l'enseignement de la prononciation du français langue étrangère concerne uniquement des apprenants japonophones dont le français est L3. Étant donné le nombre important des études qui portent sur l'acquisition de l'anglais langue seconde en Amérique du Nord, le cas des langues secondes sera également mentionné, mais essentiellement par rapport à des études antérieures citées dans l'introduction. Nos données expérimentales portent exclusivement sur des apprenants qui apprennent le français en tant que langue étrangère au Japon.

### 1.2. L'ENSEIGNEMENT ET L'APPRENTISSAGE DE LA PRONONCIATION DES LANGUES ÉTRANGÈRES

Avant de poursuivre sur le cas particulier de l'enseignement de la prononciation du français par des apprenants japonais, nous pouvons nous poser un certain nombre de questions générales sur l'apprentissage et l'enseignement de la prononciation. Qu'apprend-on, et qu'enseigne-t-on concernant la prononciation ? Pourquoi apprend-on et enseigne-t-on la prononciation des langues étrangères ? Est-il possible d'acquérir et d'enseigner une « bonne » prononciation des langues étrangères ? Est-il utile de le faire ? Les questions et les réponses proposées par Lauret (2007 :13-39)[12] nous fournissent des arguments très étayés. Nous en reprenons ici quelques-uns et les développons.

**Pourquoi apprend-on et enseigne-t-on la prononciation des langues étrangères ? Est-ce utile ?**

L'insistance sur l'importance de la prononciation varie selon le courant méthodologique. Dans la méthode grammaire-traduction, centrée sur l'apprentissage de la grammaire de l'écrit, la prononciation n'est pas considérée comme une composante importante. La méthode directe est caractérisée par « l'input » oral des enseignants natifs, même si l'apprentissage de la prononciation même n'a pas une place particulièrement importante. L'enseignement de la prononciation est un élément essentiel dans l'approche SGAV (Structuro-Globale Audio-Visuelle), qui a élaboré la méthode verbo-tonale (cette méthode sera détaillée dans la quatrième partie). Par contraste, la prononciation a été négligée dans les premières versions de l'approche communicative, au profit de l'importance attachée à la communication du sens plutôt qu'à la production des formes correctes, qu'il s'agisse de la prononciation ou de la grammaire. Galazzi-Matasci et Pedoya (1983)[13] témoignent de la disparition quasi totale de la composante phonétique dans l'approche communicative. Or, il semblerait que l'apprentissage de la prononciation par les apprenants soit important si leur objectif principal est de communiquer avec succès.

Les motivations varient selon les apprenants. Ils ne souhaitent pas toujours atteindre l'idéal de la prononciation d'un locuteur natif de la langue. Si la motivation principale de l'apprenant est une motivation instrumentale (c'est-à-dire qu'il apprend la langue purement comme un outil pour

---

[12] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[13] E. Galazzi-Matasci, E. Pedoya. Et la pédagogie de la prononciation ? *Le français dans le monde*, 180, 39-44, 1983.

réaliser d'autres fins), qu'il n'a aucune motivation d'intégration à la communauté de la langue cible, et que l'acquisition de la bonne prononciation n'est pas avantageuse pour lui, l'apprenant ne verra pas d'intérêt à apprendre à parler la langue cible comme les natifs, ou presque. Imaginons le cas d'un apprenant qui apprend une des deux langues officielles de son pays qu'il ne maîtrise pas, car il est natif de l'autre langue officielle ou natif d'une autre langue ; pour travailler en tant que fonctionnaire, les compétences des deux langues officielles sont demandées, mais l'apprenant lui-même ne s'intéresse pas à la communauté de la langue cible. Dans ce cas-là, il ne sera probablement pas prêt à faire l'effort nécessaire pour apprendre à parler comme les locuteurs natifs. Cependant, la situation sera différente si l'apprenant est soumis à une position vulnérable sur le plan sociolinguistique. Derwing (2003)[14] a effectué des interviews auprès de 100 apprenants de l'anglais langue seconde (et non étrangère) du niveau de compétence intermédiaire. La majorité des apprenants ont donné une réponse affirmative à la question suivante : « *seriez-vous mieux respectés si vous prononciez mieux l'anglais ?* » Ces résultats suggèrent que l'acquisition d'une bonne prononciation d'une langue seconde dans un contexte endolingue peut du moins égelement intéresser les apprenants qui n'ont pas nécessairement choisi la langue étudiée avec une motivation d'intégration[15].

Dans le cas des japonophones apprenant le français langue étrangère, la motivation instrumentale serait assez rarement présente, à la différence des situations décrites ci-dessus, sauf pour valider des cours obligatoires. Ce serait assez souvent le cas pour l'apprentissage de l'anglais ou du chinois, langues considérées comme « utiles » du point de vue économique, alors que les étudiants japonais ont tendance à choisir le français plutôt par curiosité, ou par appréciation ou admiration de la culture française. Nous pouvons y voir une motivation identitaire ou d'intégration (une envie d'être comme les locuteurs natifs). Cela expliquerait d'ailleurs le « syndrome de Paris[16] », observé chez des Japonais visitant Paris, notamment chez des femmes, qui avaient développé une image idéalisée de cette ville, et fortement déçus par la réalité.

Une autre question qui pourrait se poser est celle de l'identité. Selon Lauret (2007 : 21), les nouvelles habitudes articulatoires de la langue cible font peur à certains apprenants, surtout parmi les adolescents et les adultes, comme si elles risquaient de « fragiliser la personnalité » (Lauret 2007 *op. cit.* : 21). Il est donc important, comme cet auteur le souligne, de comprendre que « l'adoption d'une nouvelle musique et de nouveaux sons ne présente **aucun risque identitaire**[17] » et que « savoir être différent est au contraire un jeu, un plaisir et un enrichissement. » Lauret cite par la suite un passage de Laroy (1995)[18], qui compare le risque identitaire et les inconvénients que l'accent étranger pourrait générer :

---

[14] T. M. Derwing. What do ESL students say about their accents? *Canadian Modern Language Review / La Revue canadienne des langues vivantes* 59(4) : 547-567, 2003.

[15] Pour les études qui portent sur la motivation des apprenants des langues étrangères et secondes, voir le compte-rendu d'Ellis 1994 : 508-517.

R. Ellis. *The study of second language acquisition*. Oxford, Oxford University Press, 1994.

[16] A. Viala, H. Ôta, M.-N. Vacheron, P. Martin, F. Caroli. Les Japonais en voyage pathologique à Paris : un modèle original de prise en charge transculturelle. *Nervure* 12(5) : 31-34, 2004.

[17] Mis en gras par l'auteur.

[18] C. Laroy. *Pronunciation*. Oxford, Oxford University Press, 1995.

> *« La manière dont les élèves parlent est une expression de leur identité. On peut respecter leur choix d'un accent non standard mais on doit attirer leur attention sur le fait que l'accent peut amener des inconvénients (comme le risque d'être mal compris), aussi bien que de réels avantages personnels.*
>
> *Ensuite, la question est de savoir : que va-t-il arriver à l'identité personnelle de l'élève quand il a à produire un son non familier de la langue étrangère ? Cela nous amène à un point important. Non seulement il sera presque impossible que les professeurs amènent les élèves à prononcer mieux qu'ils ne le souhaitent eux-mêmes, mais il sera tout aussi difficile de les amener à prononcer mieux que ce qu'ils pensent pouvoir faire, même s'ils le doivent pour un diplôme ou pour leur travail. »*[19]

Plus concrètement, quels sont les inconvénients de garder un accent étranger ? Lauret (2007 *op. cit.* p. 19) illustre des témoignages typiques d'apprenants concernant l'importance de la prononciation dans le bon déroulement de la communication :

> *« Je veux changer ma prononciation parce qu'on n'écoute pas ce que je dis, mais comment je le dis. »*
>
> *« Après quelques minutes, les personnes ne veulent plus m'écouter. »*

et de la part d'auditeurs natifs :

> *« Écouter quelqu'un qui a un fort accent, pour quelques phrases, ça va, mais après une minute, ça peut être insupportable. »*

L'anecdote suivante et les commentaires de Wioland (2005 : 14)[20] soulignent également le rôle essentiel de la prononciation :

> « Anecdote 10
>
> Un témoignage révélateur d'un professeur de français détaché en Hongrie qui prolongeait volontairement son séjour après un an :
>
> « J'aime beaucoup le hongrois, je n'ai qu'un seul petit problème, on ne me comprend pas bien quand je parle. Le nouveau stagiaire français qui vient d'arriver se fait mieux comprendre que moi qui connais la langue hongroise ! »
>
> *De quelle langue parle ce professeur ? De la grammaire de l'écrit, du vocabulaire, de l'orthographe, de l'histoire de la langue ? Est-ce suffisant pour communiquer ? Apparemment non. »*

Cette anecdote illustre bien le fait que les connaissances explicites ou implicites de la grammaire, du lexique, *etc.*, ou même les connaissances explicites sur le système phonologique et la prononciation, de la langue ne suffisent pas pour se faire comprendre, si l'on ne prononce pas d'une manière satisfaisante.

Il existe également des études expérimentales qui soutiennent ces témoignages. Les trois articles présentés ci-dessous montrent que les locuteurs natifs, que ce soit des auditeurs entraînés ou non, sont bien conscients de l'accent étranger, tant au niveau segmental que prosodique.

Flege (1984)[21] a effectué une série d'expériences perceptives auprès de plusieurs groupes d'auditeurs natifs de l'anglais, en utilisant des extraits de l'anglais parlé par huit locutrices natives

---

[19] B. Lauret *op. cit.* (p. 21), traduit par l'auteur.

[20] F. Wioland. *La vie sociale des sons du français*. Paris, l'Harmattan, 2005.

de l'anglais américain venant du centre des Etats-Unis (quatre de Chicago, deux de Denver, une de Rochester, une de Milwaukee), et huit locutrices natives du français européen (deux venaient de la Belgique, six de la France métropolitaine : quatre de Paris, une d'Annecy, une de St. Etienne) habitant aux Etats-Unis depuis six mois à 31 ans. Les auditeurs, quelle que soit leur expérience en phonétique et la familiarité avec l'accent français, ont reconnu l'accent étranger des locuteurs francophones aussi bien dans des phrases lues que dans la parole spontanée. Ils ont bien identifié les locuteurs francophones (entre 63% - 95% des cas) quelle que soit la longueur de l'extrait présenté : syntagmes entiers (ex. « *two little dogs* »), syllabes (ex. /tuː/ ou /tiː/), portions de syllabes correspondant aux segments (ex. /t/, /iː/, /uː/), ou même les 30 premières millisecondes de « *two* » (correspondant grosso modo au relâchement, à la friction du /t/, et à l'aspiration). Il a ainsi montré que les locuteurs natifs détectent bien même de petites déformations de prononciation par rapport à la forme attendue.

Le facteur segmental n'est pas seul en jeu. Munro (1995)[22] a utilisé un filtrage passe-bas (la bande passante était de 225 Hz pour les voix d'homme, 300 Hz pour les voix de femme) afin de masquer les moyennes et hautes fréquences, qui contiennent des informations segmentales, tout en conservant la fréquence du fondamental (F0)[23]. Vingt auditeurs (dix femmes et dix hommes) anglophones américains (de l'Alabama) non-entraînés linguistiquement ont écouté des phrases simples lues et des énoncés narratifs ainsi filtrés, et leur ont donné une note de degré d'accent étranger. Les énoncés produits par dix locuteurs natifs (cinq femmes et cinq hommes) de l'anglais canadien, énoncés ensuite filtrés, ont constamment obtenu une meilleure note que ceux produits par dix locuteurs natifs (cinq femmes et cinq hommes) du chinois mandarin apprenant l'anglais, énoncés également filtrés. Les analyses acoustiques des stimuli suggèrent que ce seraient les différences de débit de la parole, de patterns d'intonation, et d'organisation temporelle (ex. réduction ou non des mots grammaticaux comme *and*) qui ont joué un rôle dans l'évaluation par les auditeurs natifs.

Boula de Mareüil et Vieru-Dimulescu (2006)[24] ont utilisé la synthèse par diphone et la modification de durée et de la fréquence du fondamental – ou F0 – (algorithme PSOLA sur Praat[25]) de productions naturelles pour générer des stimuli de 14 paires de phrases semblables en italien et en espagnol (ex. *Ha visto la casa del presidente americano*. « Vous (*sg*.) avez / il a / elle a vu la maison du président américain. »). Ils ont effectué deux tests d'évaluation. Pour chacun des deux tests, 40 auditeurs (20 italiens et 20 espagnols : aucun auditeur n'a participé aux deux tests) ont écouté les

---

[21] J. E. Flege. The detection of French accent by American listeners. *The Journal of the Acoustical Society of America* 76(3) : 692-707, 1984.

[22] M. J. Munro. Nonsegmental Factors in Foreign Accent. *Studies in Second Language Acquisition* 17 : 17-34, 1995.

[23] Selon CALLIOPE (1989 : 133), la F0 varie aux alentours de 100 à 150 Hz pour l'homme adulte, et de 140 à 240 Hz pour la femme adulte.

« *La gamme de variation moyenne de la fréquence fondamentale dépend, essentiellement, de l'âge et du sexe du locuteur. Elle se situe aux alentours de 100 à 150 Hz pour l'homme adulte, et de 140 à 240 Hz pour la femme adulte, mais peut présenter des variations considérables chez un même locuteur selon le type de phrase prononcée, son attitude et état émotif lors de l'acte de parole, etc.* » (CALLIOPE 1989 : 133)

CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

[24] P. Boula de Mareüil, B. Vieru-Dimulescu. The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica* 63(4) : 247-267, 2006.

[25] P. Boersma, D. Weenink. Praat: doing phonetics by computer (logiciel). http://www.praat.org/

stimuli et jugé s'il s'agissait de l'italien, de l'italien prononcé avec un accent espagnol, de l'espagnol prononcé avec un accent italien, ou de l'espagnol. Les résultats suggèrent que, malgré les différences segmentales (les consonnes géminées de l'italien, la spirantisation des occlusives sonores de l'espagnol, la spirantisation des occlusives sourdes des locuteurs toscans, *etc*. sont des informations contenues dans les stimuli naturels modifiés prosodiquement), la prosodie jouerait un rôle important dans l'identification de l'italien prononcé avec un accent espagnol, et de l'espagnol prononcé avec un accent italien.

Kamiyama (2004)[26] a utilisé également un synthétiseur par diphone (Mbrola[27]) pour générer huit syntagmes courts en français (ex. « l'économie du Canada » ; « une photo de chameaux ») en utilisant un dictionnaire de diphones issus de locuteurs français et japonais. La durée segmentale et la F0 de locuteurs natifs français et de japonophones apprenant le français ont été conservées. Quatre types de stimuli ont été ainsi créés : la voix française avec la durée segmentale et la F0 françaises, la voix française avec la durée segmentale et la F0 des apprenants, la voix japonaise avec la durée segmentale et la F0 françaises, la voix japonaise avec la durée segmentale et la F0 des apprenants. Le résultat du test de perception dans lequel 17 auditeurs natifs francophones de la France métropolitaine ont jugé les stimuli synthétisés (la question posée : « Est-ce qu'il/elle a bien dit ? ») sur une échelle de 1 à 7 montre que non seulement le segmental (différence de voix diphones) mais aussi la durée segmentale et la F0 ont joué un rôle dans le jugement des stimuli par les auditeurs francophones. Kamiyama (2005a)[28] compare les résultats de la même expérience effectuée auprès d'auditeurs natifs francophones et d'apprenants japonophones. Malgré le fait que les apprenants testés habitaient tous dans la région parisienne, la différence de notes d'évaluation des stimuli qu'ils ont trouvée entre la prosodie française et la prosodie des apprenants était moins importante que celle des auditeurs francophones natifs, ce qui n'était pas le cas pour le segmental (différence de voix diphones). Cette dernière étude suggère que les apprenants sont moins sensibles à la « bonne » prosodie de la langue cible, et ce même en ce qui concerne les apprenants qui habitent en France et qui entendent fréquemment la prosodie des locuteurs natifs.

Les études expérimentales présentées ci-dessus confirment que les locuteurs natifs seraient sensibles à l'accent étranger, c'est-à-dire aux écarts de prononciation par rapport aux formes attendues.

Cependant, les apprenants non natifs ne sont pas les seuls à produire des formes sonores éloignées de celles qui sont attendues par les auditeurs natifs d'une variété donnée. Il faudra considérer également des accents régionaux et socio-professionnels. Les locuteurs natifs d'une langue donnée sont confrontés aux locuteurs de la même langue qui parlent avec un accent

---

[26] T. Kamiyama. Perception of foreign accentedness in L2 prosody and segments: L1 Japanese speakers learning L2 French. *Proceedings of the International Congress: Speech Prosody 2004*, 721-724, 2004.

[27] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken. The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of the International Conference on Spoken Language Processing (ICSLP) '96*, 1393-1396, 1996.

[28] T. Kamiyama. Perception of foreign accentedness by native speakers and L2 learners: a preliminary study. *Actes des 7èmes Rencontres Jeunes Chercheurs de l'École Doctorale 268 'Langage et langues', Université Paris III*, 37-41, 2005.

régional ou social. Ne pourrait-on pas considérer l'accent étranger comme les accents régionaux ou les accents socialement marqués ? Il n'y a aucune raison d'imposer un accent particulier, qu'il s'agisse de l'accent « standard » ou non, aux locuteurs natifs, sauf pour ceux qui occupent certaines professions liées à la parole (acteurs, présentateurs, orthophonistes, et enseignants de langues notamment secondes ou étrangères, *etc*.). Comment pourrait-on justifier l'enseignement de la prononciation des langues étrangères comme étant une activité qui cherche à modifier la prononciation des apprenants ?

Nous trouvons certes des études qui montrent l'adaptation perceptive des locuteurs natifs aux accents étrangers (Clarke et Garrett 2004[29], Sidaras *et al*. 2008[30]), mais Evans et Adank (2008)[31] ont montré que, dans un environnement bruité, des auditeurs anglophones s'adaptent plus rapidement à un accent régional qui ne leur est pas familier (l'accent de Glasgow) qu'à un accent étranger (l'accent espagnol), et qu'ils ont mieux identifié les stimuli de l'accent régional dans l'ensemble. Ces résultats semblent justifier la pratique de l'enseignement de la prononciation des langues étrangères.

**Est-il possible d'acquérir et d'enseigner une bonne prononciation des langues étrangères ? Si possible, comment ?**

L'apprentissage de la « bonne » prononciation proche de celle des locuteurs natifs est souvent considéré comme une mission impossible, particulièrement par les partisans de la théorie de la période critique (« *critical period hypothesis* » : Patkowski 1990[32], DeKeyser 2000[33], par exemple ; pour une revue des recherches antérieures, Scovel 2000)[34]. Selon cette théorie, les apprenants adolescents et adultes ne peuvent rarement ou jamais apprendre à prononcer les langues non-maternelles comme les locuteurs natifs, car la capacité d'un apprentissage réussi des langues décline après un certain âge. Par exemple, DeKeyser (2000 : 518-519) suggère que :

> *« Quelque part entre les âges de 6-7 et 16-17 ans, tous perdent l'équipement mental requis pour les patterns abstraits sous-jacents à une langue humaine, et la période critique mérite véritablement son nom ... Il se peut que le fort déclin en capacité d'induire implicitement des patterns abstraits soit une conséquence inévitable des aspects assez généraux de la maturation neurologique et que simplement il se manifeste le plus clairement dans l'acquisition du langage.*[35] »

---

[29] C. M. Clarke, M. F. Garrett. Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America* 116(6) : 3647-3658, 2004.

[30] S. K. Sidaras, J. E. Alexander, L. C. Nygaard. Talker-specific and accent-general perceptual adaptation to Korean-accented English. *The Journal of the Acoustical Society of America* 123(5) : 3887, 2008.

[31] B. G. Evans, P. Adank. Differences in the time-course of accent adaptation: a comparison of adaptation to foreign-accented and unfamiliar regionally-accented speech. *The Journal of the Acoustical Society of America* 123(5) : 3073, 2008.

[32] M. Patkowski. Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics* 11 : 73-89, 1990.

[33] R. DeKeyser. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22 : 499-534, 2000.

[34] T. Scovel. A Critical Review of The Critical Period Research. *Annual Review of Applied Linguistics* 20 : 213-223, 2000.

[35] « *Somewhere between the ages of 6-7 and 16-17, everybody loses the mental equipment required for the abstract patterns underlying a human language, and the critical period really deserves its name ... It may be that the severe decline of the ability to induce abstract patterns implicitly is an inevitable consequence of fairly general aspects of neurological maturation and that it simply shows up most clearly in language acquisition.* »

Cependant, l'existence de la période critique d'apprentissage a été mise en question par certains auteurs. Celce-Murcia *et al.* (1996)[36] résument les facteurs négligés dans l'hypothèse de la période critique : « *les différences entre les paramètres d'acquisition d'une nouvelle langue entre enfants et adultes tels que l'exposition à la langue-cible, les attentes linguistiques des interlocuteurs natifs, la perméabilité de l'ego, l'attitude vis-à-vis de la nouvelle langue et le type de motivation* » (traduit et cité par Lauret 2007 *op. cit.* : 27). Nous pouvons également rajouter la différence de stratégie d'apprentissage et de communication : les enfants prêteraient plus attention à la forme (grammaticale et sonore) qu'à la transmission du message, tandis que les adultes prêteraient moins attention à la forme. Au début de l'acquisition de la langue première, les jeunes enfants sont très sensibles à la façon avec laquelle on s'adresse à eux, alors qu'ils ne sont pas encore capables de comprendre ce qui est dit. Selon Vaissière (communication personnelle), la prosodie typique du conte de fées produite dans une interaction avec des très jeunes enfants (d'un à deux ans) peut leur plaire, quel que soit le contenu du message linguistique (ex. une leçon de phonétique pour apprendre à lire des spectrogrammes, produite avec la prosodie d'un conte de fées, attire leur attention, voire les captive).

Il semble y avoir une condition d'apprentissage qui permettrait aux apprenants adultes et adolescents d'acquérir une prononciation native ou quasi-native. Neufeld (1979)[37] présente une expérience dans le cadre de laquelle la moitié des adultes testés sont parvenus à imiter des énoncés des langues cibles d'une manière qui a été jugée comme celle des natifs, après 18 heures d'entraînement qui portaient uniquement sur la prononciation (segmentale et prosodique) :

> *« Vingt sujets ont suivi trois cours de langue, en japonais, chinois et en esquimau. Chaque programme était enregistré sur vidéo et durait 18 heures. Seules la prononciation et l'intonation étaient enseignées. Pendant la première moitié de chaque programme, les sujets devaient écouter des énoncés en langue étrangère, pour en tracer ensuite les courbes intonative et rythmique. Les leçons comportaient aussi des exercices conventionnels de discrimination auditive où il s'agissait de détecter des différences d'une subtilité variable. A cette étape de l'expérience, il était formellement interdit aux sujets de reproduire à haute voix les sons et les énoncés qu'ils écoutaient.*
>
> *Dans la seconde moitié de chaque programme, les sujets devaient répéter, d'abord à mi-voix puis à haute voix, des énoncés comprenant entre une et seize syllabes, dans la langue-cible. La dernière partie de la leçon finale consistait en un exercice d'imitation, les sujets ayant à reproduire de leur mieux dix énoncés. Ces tentatives étaient enregistrées sur bande magnétique pour être analysées par la suite.*
>
> *Dans chaque langue-cible, la performance des sujets était évaluée selon une échelle de cinq points par trois juges autochtones qui ignoraient tout des raisons d'être et des objectifs de la recherche. Des difficultés d'ordre technique ont malheureusement empêché l'analyse des résultats du programme d'esquimau. Dans les deux autres programmes, les juges ont identifié onze des vingt sujets comme de langue maternelle japonaise, et neuf comme de langue maternelle chinoise, d'après leur habileté à imiter les dix énoncés en langue-cible. »* (Neufeld 1979 : 12)

Comme l'auteur l'admet dans le paragraphe suivant, « *l'expérience ne faisait intervenir que la faculté d'imitation; les sujets n'avaient pas à utiliser la langue de façon créatrice* ». Cela ne correspond évidemment pas à la condition normale de l'enseignement et de l'apprentissage des langues

---

[36] M. Celce-Murcia, D. M. Brinton, J. M. Goodwin. *Teaching Pronunciation: a Reference for Teachers of English to Speakers of Other Languages*. Cambridge, Cambridge University Press, 1996.

[37] G. G. Neufeld. Vers une théorie de la capacité d'apprentissage linguistique. *Encrages*, numéro spécial de linguistique appliquée, Université de Paris VIII : 9-15, 1979.

étrangères, car les sujets de cette expérience ont travaillé uniquement sur les formes sonores de surface, sans aucune considération pour les autres composantes des langues cibles (lexique, morphologie, syntaxe, sémantique, pragmatique, …). Ces données « *semblent indiquer que l'adulte ne perd ni la capacité de percevoir les différences subtiles de son, de rythme et de hauteur, ni celle de reproduire ces sons et ces contours*. » (Neufeld *op. cit.* p. 12)

Malgré la particularité de la condition d'apprentissage expérimentale, les résultats de cette étude suggèrent qu'il est possible d'acquérir une bonne prononciation grâce à des entraînements phonétiques. Il existe d'ailleurs des méthodes d'apprentissage des langues étrangères qui proposent un entraînement intensif de la prononciation au tout début de l'apprentissage, avant de commencer à travailler sur les autres aspects de la langue. Lauret (2007) présente la méthode russe et la réputation des résultats obtenus :

> *« Les écoles de langue de Russie, et en particulier à Moscou, sont réputées pour leur réussite dans l'enseignement des langues et de la prononciation. Des échanges avec des enseignants russes montrent que les programmes eux-mêmes insistent parfois exclusivement sur la prononciation en début de l'apprentissage. Une période initiale de phonétique avec supports technologiques (représentations graphiques de certains paramètres de la prononciation) est parfois uniquement consacrée à l'entraînement à l'imitation de syllabes isolées, puis de syllabes composant des groupes rythmiques non-porteurs de sens »* (Lauret *op. cit.* : 29).

En revanche, le manque ou l'insuffisance de la composante phonétique en début de programme d'apprentissage serait un obstacle à l'apprentissage ultérieur de la prononciation, comme cela risque d'être le cas pour des approches communicatives, selon Lauret (2007) :

> *« Les approches communicatives qui visent à mettre l'apprenant immédiatement au contact de la langue sous ses diverses formes, écrites en particulier, et qui incitent dès le début de l'apprentissage à la production à visée communicative, en minimisant de fait le rôle de la prononciation, tendent probablement à empêcher un développement optimal des possibilités phonétiques »* (Lauret *op. cit.* : 29).

Ce qui est également important pour la réussite d'apprentissage de la prononciation serait d'éviter le plus possible l'influence de l'écrit orthographique.

> *« […] il est possible d'acquérir une bonne prononciation quand on n'est* ***pas exposé à l'écrit orthographique***[38]*. L'exposition à l'écrit active longtemps irrépressiblement des réflexes de lecture dans la langue maternelle ou dans une autre langue de même alphabet. […] L'œil contrarie l'oreille et la bouche »* (Lauret *op. cit.* : 29).

L'apprentissage de l'opposition /l/ et /ʀ/ du français par les apprenants japonophones est un bon exemple qui semble confirmer cette prise de position. Yamasaki et Hallé (1999)[39], et Ooigawa (2008)[40] ont montré que l'opposition /l/-/ʀ/ du français, considérée communément comme une des difficultés majeures auxquelles les apprenants japonophones sont confrontés, a en effet été bien distinguée perceptivement (à la différence de l'anglais), même par des japonophones qui n'apprenaient pas le français. Ces résultats suggèrent que la confusion entre /l/ et /ʀ/ est

---

[38] Mis en gras par l'auteur.

[39] H. Yamasaki, P. A. Hallé. How do native speakers of Japanese discriminate and categorize French /r/ and /l/? *Proceedings of the XIVth International Congress of Phonetic Sciences*, 909-912, 1999.

[40] T. Ooigawa. Nihon-jin, Amerika-jin, Furansu-jin choushusha ni yoru eigo, futsugo, seigo ni okeru boinkan no ryuuon no benbetsu : yobi kenkyuu [Discrimination de liquides intervocaliques en anglais, en français, et en espagnol chez des auditeurs japonais, américains et français : étude préliminaire]. *Sophia University Working Papers in Phonetics* 2008, Sophia University Phonetics Laboratory, Tokyo : 2-20, 2008.

plutôt due à la difficulté de production, ou encore, au niveau du lexique mental, car les deux phonèmes qui se trouvent dans des mots appris par l'écrit risquent d'entrer dans la même catégorie phonémique de l'interlangue, qui est la liquide unique (c'est aussi le cas pour des mots qui sont déjà stockés dans le lexique mental comme mots similaires du japonais ou de l'anglais : ex. *table* /teɪbl/ (en anglais) > /teRburu/ (emprunt japonais) dans le lexique mental avec le /r/ comme le liquide unique, d'où les prononciations susceptibles [tabʁ] ou [tabl] en français). À moins que les apprenants aient appris à distinguer parfaitement (ou presque) bien /l/ et /ʀ/ sur le plan perceptif et productif avant d'affronter l'écrit, la confusion de ces deux phonèmes due à l'orthographe nous semble difficile à éviter, étant donnée la difficulté de distinction perceptive de /l/ et /r/ de l'anglais.

L'importance de l'influence de l'écriture orthographique sur la prononciation est illustrée également par notre expérience personnelle. Quand nous avons assuré un cours d'anglais langue étrangère pour des adolescents japonophones qui venaient tout juste de commencer à apprendre l'anglais (en première année de collège), nous avons tâché, avec un collègue canadien anglophone, à ce que les élèves s'habituent à la prononciation des natifs dès le début. Pendant le premier cours, ils répétaient bien, en prononçant relativement bien à l'américaine. Pour la semaine suivante, ils ont eu des exercices écrits de grammaire à faire à la maison, conformément au programme officiel de l'école. Pendant le deuxième cours, les élèves devaient prononcer les phrases des exercices, mais la plupart d'entre eux les ont prononcées avec une prononciation japonaise, même si nous avions fait répéter le vocabulaire et les phrases de base à haute voix à la fin du cours précédent. Les élèves, qui ne disposaient pas d'un support sonore correspondant aux exercices, dépendaient uniquement de l'orthographe quand ils ont préparé leurs devoirs ; ils étaient de ce fait exposés à l'influence de la prononciation de leur langue maternelle, plus que si l'introduction de l'écrit orthographique avait été retardée.

Dans tous les cas, la pratique d'une écoute fréquente nous semble être un élément essentiel pour un apprentissage réussi de la prononciation. Lauret (2007)[41] compte « une écoute fréquente (la plus fréquente possible), précise ou non, de la musique et des sons de la langue » parmi les quatre éléments qui permettent l'acquisition de la prononciation d'une langue étrangère (p. 169). Nous avons discuté ci-dessus de l'importance du travail de prononciation sans support écrit, notamment en début d'apprentissage, mais une écoute avec support écrit faciliterait l'apprentissage de la graphie-phonie, une fois que la prononciation est maîtrisée.

L'exercice d'écoute et de répétition peut être également combiné avec une représentation prosodique. Dans le cas de l'apprentissage d'une langue avec une écriture nouvelle, la pratique de l'écoute régulière peut aider l'apprentissage de l'écriture. Nous avons développé une série de matériels pédagogiques de révision en ligne pour les apprenants débutants du japonais langue étrangère (Kamiyama, à paraître)[42]. Les utilisateurs peuvent écouter des fichiers sons de phrases ou de syntagmes segmentés en petits éléments accompagnés de leur représentation visuelle (voir ci-

[41] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[42] T. Kamiyama. Hi-senmon katei no sentaku jugyou ni okeru onsei imputto-busoku o oginau tameno intânetto fukushuu kyouzai [Matériels pédagogiques de révision pour compenser le manque d'input sonore dans un cours optionnel de japonais]. *Furansu Nihongo Kyouiku* [Enseignement du Japonais en France] 4, Association des enseignants de japonais en France, 9 pages, à paraître.

dessous), afin de permettre une confrontation entre la perception auditive et visuelle. Deux versions sont disponibles pour cette dernière : 1) Un schéma rythmique et prosodique représentant le nombre de mores (*haku*), avec des mores spéciales (*tokushu-haku*) et celles qui contiennent une voyelle dévoisée en couleurs différentes, ainsi qu'une indication de l'accent lexical et de l'intonation interrogative (Figure 1, à gauche) ; 2) Le schéma décrit en 1) accompagné de l'orthographe en *kana* (Figure 1, à droite). La combinaison de la pratique d'écoute régulière avec la visualisation de l'écrit permettrait de développer non seulement des compétences de prononciation, de compréhension, et d'expression orales, mais aussi l'apprentissage de l'écriture, comme le témoignent les étudiants qui ont utilisé régulièrement ces matériels : « *la manière de prononcer les mots ou les phrases permettait de s'entraîner à lire, à parler et à écrire* », « *très utiles, surtout pour la compréhension orale et écrite des dialogues* », « *une écoute régulière permettait de s'améliorer et également de retenir des exemples qui pouvaient et peuvent encore s'avérer très utiles* ».




**Figure 1 :** Exemple des matériels pédagogiques de révision en ligne qui permettent aux apprenants d'écouter des syntagmes et des phrases en visualisant la représentation prosodique (à gauche) et la représentation prosodique avec l'orthographe en syllabaires *kana* (Kamiyama, à paraître).

Par ailleurs, des études expérimentales ont montré l'efficacité de l'entraînement de perception et de production même auprès d'apprenants non-débutants. Il est établi que l'opposition /l/-/r/ de l'anglais est difficile à percevoir (Yamasaki et Hallé 1999[43], Ooigawa 2008[44], entre autres) et à produire pour les apprenants japonophones, mais il a été montré que l'on peut améliorer la performance à l'aide d'entraînements. Iverson, Hazan et Bannister 2005[45] ont testé un entraînement de perception avec la méthode « *High Variability Phonetic Training* » (les sujets écoutent des enregistrements de mots prononcés par de nombreux locuteurs), et trois différentes méthodes de renforcement acoustique d'indices phonétiques, et montré que toutes les techniques d'entraînement ont permis à des auditeurs japonais d'améliorer l'identification des consonnes /l/

[43] H. Yamasaki, P. A. Hallé. How do native speakers of Japanese discriminate and categorize French /r/ and /l/ ? *Proceedings of the XIVth International Congress of Phonetic Sciences*, 909-912, 1999.

[44] T. Ooigawa. Nihon-jin, Amerika-jin, Furansu-jin choushusha ni yoru eigo, futsugo, seigo ni okeru boinkan no ryuuon no benbetsu : yobi kenkyuu [Discrimination de liquides intervocaliques en anglais, en français, et en espagnol chez des auditeurs japonais, américains et français : étude préliminaire]. *Sophia University Working Papers in Phonetics* 2008, Sophia University Phonetics Laboratory, Tokyo : 2-20, 2008.

[45] P. Iverson, V. Hazan, K. Bannister. Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, 118(5) : 3267-3278, 2005.

et /r/. L'efficacité de l'entraînement perceptif serait encore plus importante si des informations visuelles étaient combinées avec des informations auditives. Hazan *et al.* (2005)[46] ont montré que des japonophones apprenant l'anglais ont davantage amélioré la perception et la production de paires de consonnes en anglais (/v/-/b/-/p/ pour la perception, /l/-/r/ pour la production) après un entraînement audio-visuel qu'après un entraînement auditif seul. En ce qui concerne la production, l'étude de Hattori et Iverson 2008[47], dans laquelle des apprenants japonophones adultes ont participé à dix séances d'entraînement sur /l/ et /r/ de l'anglais, indique que la production de ces consonnes a été améliorée.

Un entraînement intensif de la prononciation en début d'apprentissage et dans une étape ultérieure permettrait donc la maîtrise d'une bonne prononciation, mais le fruit de cet entraînement intensif de prononciation est-il applicable à une production plus spontanée ?

D'après Krashen (1982)[48], les apprenants possèdent deux systèmes totalement séparés : le « système acquis », qui correspond aux connaissances implicites et le « système appris », qui correspond aux connaissances explicites. Le premier est développé par l'acquisition (naturelle), procédé subconscient qui survient quand les apprenants utilisent la langue pour communication. Le dernier est le résultat de l'apprentissage (formel), consistant à prêter attention consciemment à la langue dans un effort de comprendre et mémoriser les règles. Est-il possible d'appliquer les connaissances explicites apprises dans l'enseignement aux connaissances implicites au niveau pratique ?

Selon la théorie du moniteur (« *monitor theory* ») de Krashen, les connaissances « apprises » sont complètement séparées et ne sont pas convertibles en connaissances « acquises ». Les exercices formels ne permettraient donc pas aux apprenants de convertir leurs connaissances explicites en connaissances implicites. Le système « appris » entrerait uniquement en jeu quand les apprenants surveillent (*monitor*) leurs propres énoncés générés par le système « acquis ».

La prise de position « interface » de Bialystock (1978)[49] est différente de la position « non interface » (entre les deux systèmes) de Krashen. Son modèle prend en compte une interaction entre les deux types de connaissances. Les exercices formels (*formal practising*) permettraient aux connaissances explicites de devenir implicites, tandis que l'inférence permet la dérivation des connaissances explicites à partir des implicites (Figure 2).

---

[46] V. Hazan, A. Sennema, M. Iba, A. Faulkner. Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication* 47(3) : 360-378, 2005.

[47] K. Hattori, P. Iverson. English /r/-/l/ pronunciation training for Japanese speakers. *The Journal of the Acoustical Society of America* 123(5) : 3327, 2008.

[48] S. D. Krashen. *Principles and Practice in Second Language Acquisition*. Oxford, Pergamon, 1982.

[49] E. Bialystok. A theoretical model of second language learning. *Language Learning* 28 : 69-84, 1978.

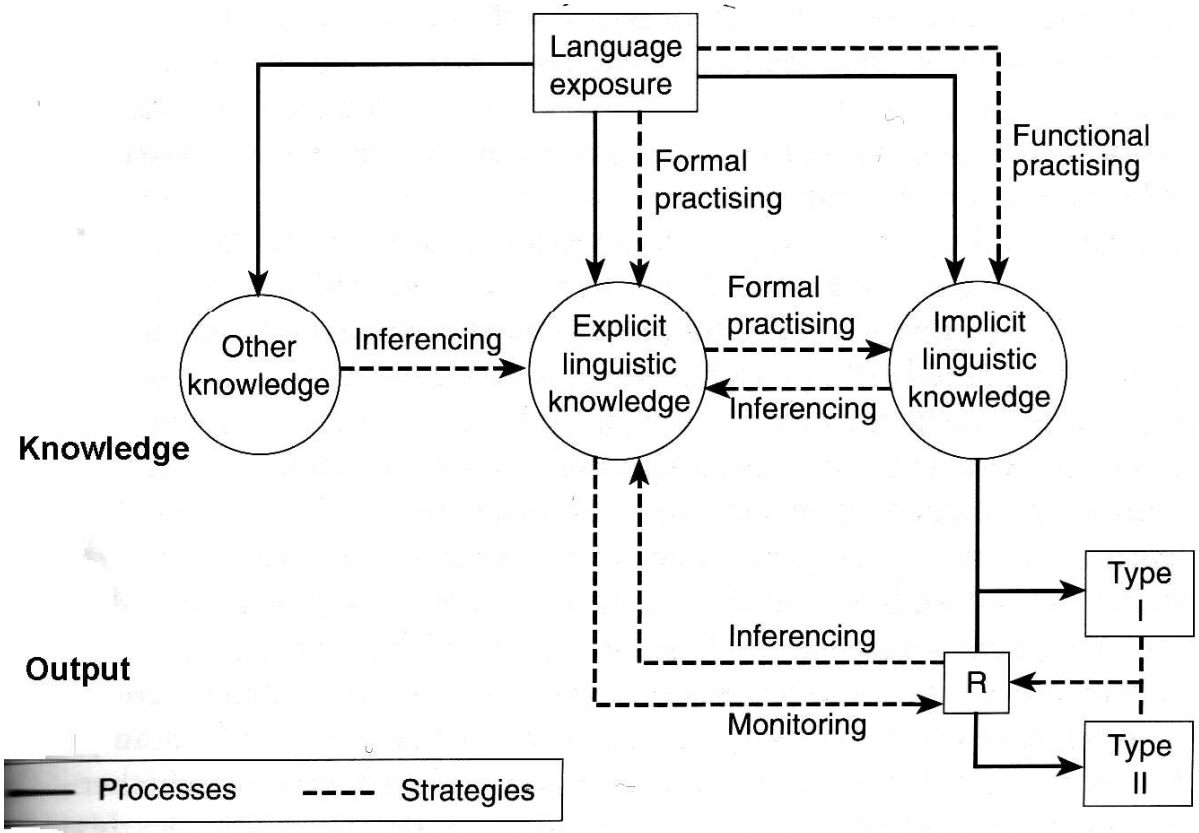



*Figure 9.2: Model of Second Language Learning (from Bialystok 1978)*

**Figure 2 :** Modèle d'apprentissage des langues secondes selon Bialystock (Ellis 1994 : 357)[50].

Dans le cas de l'acquisition et de l'apprentissage de la prononciation, nous pourrons considérer comme connaissances explicites les compétences de prononciation travaillées dans des exercices dans un contexte contraint (ex. répétition immédiate des syllabes), compétences qui seraient appliquées aux connaissances implicites, utilisables dans un contexte réel de communication. Afin d'assurer le passage des premières aux dernières, nous devrions proposer aux apprenants la progression suivante dans des exercices et des activités : de l'auditif au linguistique, du plus contraint au moins contraint (Lauret 2007 : 139)[51]. Ainsi, les apprenants pourraient passer progressivement d'une production contrôlée, avec la conscience dirigée vers la forme sonore (la prononciation), à une production plus spontanée, avec la conscience dirigée plus vers le sens, après avoir automatisé les nouvelles habitudes articulatoires apprises.

**Qu'enseigne-t-on ? Quelle(s) variété(s) ?**

L'enseignement et l'apprentissage de la prononciation des langues étrangères incluent ceux du segmental (les voyelles et les consonnes), leur interaction (telle que la coarticulation), et de la prosodie (rythme, intonation), de ses fonctions linguistique et paralinguistique. Cependant, quelle(s) variété(s) de la langue cible présente-t-on comme modèle d'enseignement et d'apprentissage à suivre ?

---

[50] R. Ellis. *The Study of Second Language Acquisition*. Oxford, Oxford University Press. 1994.

[51] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

En pratique, il s'agit en général du français parisien pour l'enseignement du français (même pour les Canadiens anglophones, dans certains cas, plutôt que la variété de leurs compatriotes francophones), et le « *Received Pronunciation* » (RP) ou le « *General American* » (GA) pour l'anglais. Qu'est-ce qui peut justifier ces choix ? Pourquoi n'enseigne-t-on pas d'autres variétés, puisqu'il existe des locuteurs natifs qui parlent des variétés diverses, et qui se font comprendre par d'autres locuteurs natifs ?

Nous pouvons penser à plusieurs raisons (*i.e.* accent considéré comme neutre, sans aucune caractéristique régionale ou sociale ; accent prestigieux, donc convenant aux apprenants de divers milieux sociaux, dont les milieux intellectuels, *etc.*), mais la plus importante semble être la possibilité d'être compris par le plus grand nombre d'interlocuteurs, natifs ou non natifs.

La RP est le modèle presque unique d'enseignement de l'anglais britannique malgré le fait qu'elle ne soit parlée que par une petite minorité de la population britannique (3 à 5 %, selon Hughes *et al.* 2005 : 3)[52]. L'une des raisons évoquées par Hughes *et al.* (2005)[53], qui expliquent cet écart, est le fait que la RP est probablement devenue l'accent le plus largement compris, grâce à son utilisation à la radio et à la télévision, et donc que les apprenants qui parviennent à la parler ont la meilleure probabilité d'être compris partout dans les Îles Britanniques.

Dans l'étude expérimentale de Clopper et Bradlow (2006)[54], des auditeurs anglophones de divers dialectes américains ont transcrit des phrases prononcées par des locuteurs de quatre accents régionaux différents (*General American*, *Southern*, *Northern*, et *Mid-Atlantic*). Avec un niveau de bruit relativement important (rapport signal/bruit de -6 dB), l'intelligibilité était significativement meilleure pour les locuteurs de *General American* que pour ceux des trois autres dialectes. En revanche, aucun effet n'a été observé concernant la différence entre les groupes d'auditeurs. Ces résultats suggèrent que le General American, qui est un accent standard aux Etats-Unis, serait compris le plus facilement, quel que soit l'accent de l'interlocuteur, malgré une déformation du signal. Les locuteurs non-natifs qui ont acquis cet accent standard, de manière plus ou moins précise, auraient donc plus de chance d'être compris que ceux qui ont acquis d'autres accents, toutes conditions égales par ailleurs.

Les arguments ci-dessus nous semblent justifier le choix de la/les variété(s) enseignée(s) effectivement dans la pratique didactique.

**Objectif de l'apprentissage : peut-on limiter le but final à l'intelligibilité, sans penser à la perfection ?**

Pour les tenants de l'approche communicative dans ses premières versions, l'objectif de l'apprentissage est la maîtrise des compétences de communication de messages plutôt que celle des formes (sonores et grammaticales) correctes. Peut-on limiter l'objectif final de l'apprentissage

---

[52] A. Hughes, P. Trudgill, D. Watt. *English Accents and Dialects*. London, Hodder Arnold, 2005.

[53] *op. cit.* « *Because of its use on radio and television, within Britain RP has become probably the most widely understood of all accents. This in turn means that the learner who succeeds in speaking it, other things being equal, has the best chance of being understood wherever he or she goes in the British Isles.* »

[54] C. G. Clopper, A. R. Bradlow. Effects of dialect variation on speech intelligibility in noise. *The Journal of the Acoustical Society of America* 119(5) : 3424, 2006.

de la prononciation à des compétences suffisantes qui permettent aux apprenants de transmettre leur message correctement, et non à celles qui sont proches du niveau des locuteurs natifs ?

Lauret (1998 : 26-27)[55] discute les notions de fautes phonémiques et phonétiques, en citant Di Cristo (1975)[56], qui reprend la distinction traditionnelle entre ces deux termes : « … on entend par *faute phonémique* une erreur de prononciation qui fait obstacle à la compréhension du message, et par *faute phonétique* une erreur qui révèle un accent étranger, sans poser toutefois à l'auditeur un problème de décodage ». Lauret (1988 et 2007) met en question le critère souvent retenu de « faire obstacle à la compréhension du message ». Il nous fait remarquer que le rapport entre l'intelligibilité et la distinction entre les deux types de fautes est très complexe : certaines fautes phonétiques peuvent bien faire obstacle à la compréhension du message, comme la substitution du /ʀ/ français par le /r/ de l'anglais américain GA (*General American*) dans un énoncé tel que « c'est rare » ou « en RER » (1998 : 26-27 ; 2007 : 22-23). Joue également l'effet de fatigue auprès des auditeurs : les auditeurs natifs ne restent pas nécessairement patients quand ils écoutent des apprenants qui produisent des énoncés intelligibles mais fortement marqués par un accent étranger, comme nous l'avons déjà vu dans les témoignages du même auteur (Lauret 2007 : 19). Cet auteur propose ainsi que l'on considère les deux catégories de fautes en se fondant simplement sur les systèmes des langues cibles et maternelles : les fautes phonémiques comme des cas de confusion entre des unités distinctives de la langue cible, et les fautes phonétiques comme ceux de transfert d'unités de la langue maternelle non-distinctives en langue cible (1988 : 26-27).

Par ailleurs, étant donné que la distinction entre les phénomènes des niveaux phonétique et phonémique reste problématique en ce qui concerne la prosodie, les notions de fautes phonémiques et phonétiques sont également difficiles à définir (comme Lauret 1998 le signale, Di Cristo 1975 traite la prosodie à part).

Si l'on considère d'autres aspects des langues, la grammaire par exemple, l'objectif d'apprentissage ne se limite communément pas à la production d'énoncés intelligibles. Si c'était le cas, les apprenants n'auraient pas à apprendre, par exemple, la conjugaison des verbes du français, car dans une grande majorité des cas, les informations que les terminaisons verbales portent sur la personne sont redondantes en présence du sujet exprimé, et même celles du temps sont redondantes quand le temps est exprimé par un syntagme adverbial (ex. *hier*, *l'année dernière* : la phrase **Je vais au supermarché hier soir* prononcée à la française serait aussi intelligible que la phrase *Je suis allé au supermarché hier soir* prononcée avec un [eɪ] pour /e/, [ɹ] ou [ɻ] pour /ʀ/, ou encore, sans distinguer /u/ et /y/). Mais en réalité, malgré l'intelligibilité assurée sans conjugaison, la maîtrise des terminaisons verbales est considérée comme souhaitable, voire obligatoire. En outre, la conjugaison du français est difficile à acquérir, tout comme la déclinaison des déterminants de l'allemand, et pourtant, cette difficulté n'est généralement pas acceptée comme une excuse pour ne

[55] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français.* Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[56] A. Di Cristo. Présentation d'un test de niveau destiné à évaluer la prononciation des anglophones. *Revue de Phonétique Appliquée* 33-34 : 9-35, 1975.

pas apprendre ou enseigner cet aspect de la langue. La maîtrise de la prononciation est aussi considérée comme une tâche difficile, mais pourquoi faire une exception à la prononciation ?

Nous pourrions dire que la conjugaison et la déclinaison peuvent être maîtrisées par mémorisation, tandis que la maîtrise de la prononciation demande une maîtrise de contrôle moteur. Mais la mémorisation de la conjugaison et de la déclinaison ne signifie pas nécessairement la maîtrise de la mise en pratique de ces connaissances explicites, comme le montre le témoignage typique d'un(e) étudiant(e) français(e) qui a appris l'allemand pendant cinq ou six ans : « j'ai bien appris la déclinaison des déterminants, mais je suis capable de produire juste quatre ou cinq phrases ».

Qu'il s'agisse de la grammaire ou de la prononciation, il est indispensable de s'entraîner à automatiser les nouvelles habitudes (par le processus de passage des connaissances explicites aux connaissances implicites, que Bialystock appelle « *formal practising* », s'il s'agit de l'apprentissage formel plutôt que de l'acquisition naturelle), et de retenir que l'apprentissage de la prononciation demande encore plus de performance physique que d'autres composantes de langue. En revanche, il n'y aurait pas de raison de faire une exception à l'apprentissage de la prononciation. Tout comme pour l'apprentissage de la grammaire, nous pouvons viser une « performance maximale » comme objectif, qui permet de « savoir mobiliser toutes les connaissances acquises pour former un message clair, le plus proche possible de celui d'un natif » (Lauret 2007 : 23), d'autant plus que les apprenants apprennent rarement à prononcer mieux que les enseignants le souhaitent, comme nous l'avons déjà vu dans la citation de Laroy (1995[57]), que nous reprenons ici :

> *« Non seulement il sera presque impossible que les professeurs amènent les élèves à prononcer mieux qu'ils ne le souhaitent eux-mêmes, mais il sera tout aussi difficile de les amener à prononcer mieux que ce qu'ils pensent pouvoir faire, même s'ils le doivent pour un diplôme ou pour leur travail. »*

Il est donc souhaitable que les enseignants et les apprenants visent une prononciation native ou quasi native comme objectif final de l'apprentissage.

**Que peut apporter la phonétique expérimentale à l'apprentissage et à l'enseignement de la prononciation ?**

Les savoirs des spécialistes ne sont pas tous transmissibles directement aux enseignants de langues et aux apprenants. Selon Cuq (dir. : 2003 : 240)[58], le concept de transposition didactique a été proposé par le sociologue Michel Verret (1975)[59] et repris dans le cadre de la didactique des mathématiques par Yves Chevallard (1985)[60]. Ce dernier explique qu'une notion issue de l'extérieur de la didactique est transformée d'abord en un objet enseignable (sélection, programmation), puis en un objet enseigné (présentation, explication, évaluation) et enfin éventuellement, par extension, en un objet d'apprentissage (Cuq 2003, *op. cit.*).

---

[57] C. Laroy. *Pronunciation*. Oxford, Oxford University Press, 1995.

[58] J.-P. Cuq (dir.). *Dictionnaire de didactique du français langue étrangère et seconde*. Paris, Clé International/Asdifle, 2003.

[59] M. Verret. *Le temps des études*. Paris, Champion Honore, 1975.

[60] Y. Chevallard. *La transposition didactique*. Grenoble, La Pensée sauvage, 1985.

Dans le cas de l'enseignement et de l'apprentissage de la prononciation des langues étrangères et secondes, quel sera l'apport des savoirs savants des phonéticiens à l'enseignement de la prononciation ? Il s'agit de découvrir et faire découvrir l'importance relative des écarts entre la production des apprenants et l'attente des auditeurs natifs ou non natifs de la langue cible. Selon Wioland (2007)[61] :

> *« Il est capital de faire découvrir à l'apprenant le fonctionnement de la prononciation : non seulement les structures rythmiques, syllabiques, phonématiques et autres, mais également la hiérarchie d'importance – tout ne revêtant pas la même importance – en fonction de ce qu'attend sans en être forcément consciente « l'oreille francophone ». L'objectif premier est de sensibiliser les enseignants comme les apprenants à ce qui, dans la communication orale en français, est primordial pour « l'oreille francophone ».* (p. 6)

En général, les enseignants natifs ne sont pas en mesure d'expliciter les écarts de prononciation perçus à moins d'avoir reçu un entraînement phonétique. En outre, les enseignants non natifs et les apprenants ne perçoivent pas toujours les mêmes écarts que les auditeurs natifs. Les savoirs savants des phonéticiens et les résultats des études expérimentales permettront aux enseignants natifs et non natifs de minimiser ce décalage de perception entre eux, et de s'en servir pour les activités en classe, et aux apprenants de prendre conscience des difficultés. Si nous prenons un des cas traités dans la présente étude, les enseignants natifs de français ne trouveraient pas facilement comment corriger le /u/ prononcé par des apprenants japonophones et comment expliquer et expliciter l'écart entre la production des apprenants et le son attendu (sauf peut-être le manque de labialisation). D'autre part, les enseignants non natifs (japonophones) et les apprenants ne sont souvent pas conscients du décalage entre le /u/ français et le /u/ japonais, et la difficulté de la maîtrise du /u/ français qui en résulte. C'est ici que les résultats des études expérimentales sur le /u/ français appris par des apprenants japonophones peuvent apporter une contribution de nature à faciliter et à améliorer l'enseignement et l'apprentissage de cette voyelle. Certes, les données obtenues dans une condition expérimentale ne sont pas directement applicables aux activités de production (semi-)spontanée. Mais comme nous l'avons vu ci-dessus, les activités traitant de l'auditif dans des circonstances idéalisées de classe peuvent être suivies de celles qui relèvent plus du niveau linguistique (par rapport à l'auditif) dans des circonstances plus proches de la communication réelle (« de l'auditif au linguistique », « du plus contraint au moins contraint » : Lauret 2007 : 139[62]).

En outre, il faudra également considérer les connaissances dont les enseignants doivent disposer, mais qui ne conviennent pas à la transmission directe aux apprenants. Les connaissances sur les différences acoustiques entre le /u/ et le /ø/ français, par exemple, ne seraient probablement pas utiles à transmettre directement aux apprenants japonophones qui ont des difficultés à distinguer ces deux voyelles, mais il est souhaitable que ceux qui enseignent le français aux japonophones les connaissent. Par ailleurs, nous pourrons transformer ces connaissances en les simplifiant (ex. avec une représentation visuelle ou une « image acoustique » qui correspondrait à notre intuition), afin de faciliter l'apprentissage. Nous verrons des exemples de la voyelle /u/ dans la quatrième partie de notre étude.

---

[61] F. Wioland. "Préface", in B. Lauret, *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[62] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

## 1.3. ÉTUDES ANTÉRIEURES SUR L'ACQUISITION DE LA PHONÉTIQUE ET DE LA PHONOLOGIE DES LANGUES ÉTRANGÈRES / SECONDES

Les phonéticiens et les psycholinguistes se sont intéressés depuis longtemps à la production et à la perception des langues non-natives du locuteur.

De nombreux phonéticiens célèbres de la fin du XIXème au début du XXème siècle étant également des enseignants de langues à l'origine, ils se sont intéressés à l'enseignement et à l'apprentissage de la prononciation des langues étrangères, et ont écrit des ouvrages sur l'enseignement des langues (Paul Passy[63], Henry Sweet[64], Otto Jespersen[65], et Harold E. Palmer[66], qui fut nommé conseiller linguistique du ministère japonais de l'Éducation en 1922, entre autres : rappelons également que le Journal de l'Association Phonétique Internationale était intitulé *Maître phonétique*). Plus près de nous, J. C. Catford, qui a enseigné l'anglais langue étrangère à Athènes au début de sa carrière académique, dit qu'à l'époque (1938-1939), les phonéticiens étaient les plus concernés dans l'enseignement des langues étrangères[67].

**L'explication de Polivanov (1931)**

Polivanov (1931)[68] fournit des exemples de difficultés de perception et de production des langues étrangères, ainsi que ceux de mots d'emprunt qui en résultent, et signale l'influence de la langue maternelle des locuteurs :

> *« Les phonèmes et les autres représentations phonologiques élémentaires de notre langue maternelle [...] se trouvent si étroitement liés avec notre activité perceptive que même en percevant des mots (ou phrases) d'une langue avec un système phonologique tout différent, nous sommes enclins à décomposer ces mots en des représentations phonologiques propres à notre langue maternelle. En entendant un mot inconnu étranger (ou d'une manière générale un fragment de langue étrangère susceptible par son volume d'être saisi par la perception auditive), nous tâchons d'y retrouver un complexe de nos lois de groupement des phonèmes. »* (Polivanov 1931 : 79-80)

---

[63] P. Passy. *Le français parlé. Morceaux choisis à l'usage des étrangers avec la prononciation figurée*. Heilbronn, Henninger, 1886.

P. Passy. *Les éléments d'anglais parlé*. Paris, Firmin-Didot. 1886.

[64] H. Sweet. *Practical study of languages*. New York, Henry Holt and Company, 1900.

[65] O. Jespersen. *How to teach a foreign language*. London, George Allen and Unwin, 1904.

[66] H. E. Palmer. *A grammar of spoken English on a strictly phonetic basis*. Cambridge, England, W. Heffer and sons, ltd., 1939.

[67] Dans le « cours 3 » (*lecture 3*) de « *Catford Tapes* », l'enregistrement d'une série de conférences assurée par J. C. Catford entre le 7 février et le 18 avril 1985 en commémoration de son départ à la retraite. La version électronique est disponible sur le site :

http://www.pagliere.net/alan/blog/2008/02/02/the-catford-tapes/

Selon Catford, David Abercrombie était parmi ses collègues enseignants à Athènes.

Notons en passant que Catford a appris la phonétique expérimentale de Pierre Fouché et de Marguerite Durand pendant son séjour à Paris en 1937-1938.

[68] E. Polivanov. La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague* 4 : 79-96, 1931.

**Le crible phonologique de Troubetzkoy**

Troubetzkoy a développé l'idée de « crible phonologique » (1938/2005 : 54-56[69]). Selon cet auteur, le système phonémique de la langue maternelle fonctionnerait comme un crible à travers lequel les sons des langues étrangères devraient passer :

> *« Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phonèmes. Tout le reste tombe dans un autre crible où restent les marques phoniques ayant une valeur d'appel ; plus bas se trouve encore un crible où sont triés les traits phoniques caractérisant l'expression du sujet parlant. Chaque homme s'habitue dès l'enfance à analyser ainsi ce qui est dit et cette analyse se fait d'une façon tout à fait automatique et inconsciente. Mais en outre le système des cribles, qui rend cette analyse possible, est construit différemment dans chaque langue. L'homme s'approprie le système de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend le « crible phonologique » de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue étrangère entendue, il se produit de nombreuses erreurs et incompréhensions. Les sons de la langue étrangère reçoivent une interprétation phonologiquement inexacte, puisqu'on les fait passer par le « crible phonologique » de sa propre langue. »* (Troubetzkoy 1938/2005 : 54)

L'auteur mentionne également l'effet de l'apprentissage, et signale que les « erreurs grossières » disparaissent peu à peu, mais qu'il reste un accent étranger, qui résulterait du crible phonologique, plutôt que de la difficulté d'articulation :

> *« [les exemples cités] prouvent que ce qu'on appelle l'« accent étranger » ne dépend pas du fait que l'étranger en question ne peut pas prononcer un certain son, mais plutôt du fait qu'il n'apprécie pas correctement ce son. Et cette fausse appréciation des sons d'une langue étrangère est conditionnée par la différence existant entre la structure phonologique de la langue étrangère et celle de la langue maternelle du sujet parlant. »* (Troubetzkoy 1938/2005 : 56)

**L'analyse contrastive**

Selon les tenants de l'analyse contrastive développée dans les années 1950 (Weinreich 1953/1968[70], Lado 1964[71], entre autres), les différences qui se trouvent entre la langue source (L1) et la langue cible (L2) permettront de prédire des difficultés que les apprenants rencontrent lors de l'apprentissage.

L'analyse contrastive a été critiquée par la suite, particulièrement concernant l'acquisition d'autres composantes linguistiques comme la morphosyntaxe, car les différences entre la langue source et la langue cible ne peuvent pas toujours expliquer les phénomènes observés dans le développement de l'interlangue des apprenants. Il a été montré, par exemple, que les apprenants des langues comme l'anglais et l'allemand présentent des tendances similaires dans leur acquisition de l'ordre des mots de la construction négative, quelle que soit leur langue maternelle (les études sont résumées dans Ellis 1994[72]).

---

[69] N. S. Troubetzkoy. *Principes de phonologie* (traduction de Jean Cantineau, revue et corrigée par Luis Jorge Prieto). Paris, Klincksiek, 1938/2005.

[70] U. Weinreich. *Languages in Contact*. The Hague, Mouton, 1953/1968.

[71] R. Lado. *Language Teaching: A Scientific Approach*. New York, San Francisco, Toronto, London, McGraw-Hill. Inc., 1964.

[72] R. Ellis. *The study of second language acquisition*. Oxford, Oxford University Press, 1994.

Cependant, concernant l'acquisition de la phonétique et de la phonologie, les modèles influents récents considèrent l'expérience de la L1 comme un élément important.

**Le modèle PAM de Catherine Best *et al.***

Selon le *Perceptual Assimilation Model* (PAM : Modèle d'assimilation perceptive) de Catherine Best *et al.*, les auditeurs assimilent perceptivement les segments des langues non natives en fonction de la correspondance avec le système phonémique de leur langue maternelle. Il existe trois cas de figure (Best 1995 : 194-195[73]) :

> *1. Le phone en question est* assimilé à une catégorie de la langue native *: assimilé clairement à une catégorie de segment spécifique de la langue native, ou peut-être à un groupe ou une suite de celles-là. Dans ce cas-là, il peut être perçu comme :*
>
> *a. un bon exemplaire de cette catégorie, ou*
>
> *b. un exemplaire acceptable mais non idéal de la catégorie, ou*
>
> *c. un exemplaire notablement déviant de la catégorie*[74]*.*
>
> *2. Le phone en question est* assimilé comme un son de la parole non catégorisable *: assimilé dans l'espace phonologique de la langue native comme une constellation de gestes qui correspondent à quelque chose qui ressemble à la parole, mais non comme un exemplaire clair d'aucune catégorie de la langue native (c'est-à-dire qu'il tombe dans l'espace phonologique de la langue native, mais entre des catégories spécifiques)*[75]*.*
>
> *3. Le phone en question* n'est pas assimilé à la parole (son qui ne relève pas de la parole) *: non assimilé du tout à l'espace phonologique de la langue native, et perçu en revanche comme une sorte de son qui n'appartient pas à la parole*[76]*.*

Ces trois cas de figure peuvent être schématisés comme suite (Figure 3).

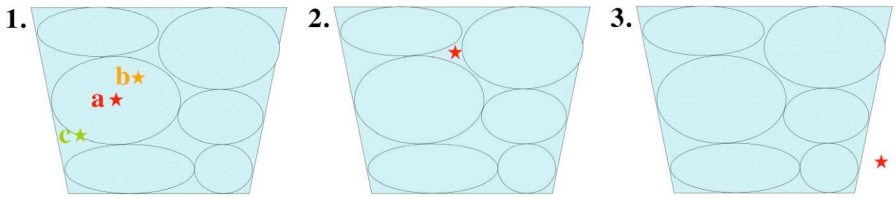



**Figure 3** : Trois patterns d'assimilation perceptive des segments non natifs selon Best (1995). 1. Assimilé à une catégorie native : a. un bon exemplaire ; b. un exemplaire acceptable mais non idéal ; c. un exemplaire notablement déviant. 2. Assimilé comme un son de la parole non catégorisable. 3. Non assimilé à la parole. Figure fondée sur les critères de Best (1995).

---

[73] C. T. Best. "A Direct Realist View of Cross-Language Speech Perception", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 171-204, 1995.

[74] « *1.* Assimilated to a native category: *clearly assimilated to a particular native segmental category, or perhaps to a cluster or string, in which case it may be heard either as:*

*a. a good exemplar of that category*

*b. an acceptable but not ideal exemplar of the category*

*c. a notably deviant exemplar of the category.* »

[75] « *2.* Assimilated as uncategorizable speech sound*: assimilated within native phonological space as a speechlike gestural constellation, but not as a clear exemplar of any particular native category* (i.e., *it falls within native phonological space but in between specific native categories).* »

[76] « *3.* Not assimilated to speech (nonspeech sound)*: not assimilated into native phonological space at all; heard, instead, as some sort of nonspeech sound*. »

En se fondant sur le pattern d'assimilation de chacun des deux membres du contraste, le degré de facilité de différenciation ou de discrimination perceptive d'un contraste dans une langue non native est prédictible. L'auteure compte 6 différents types d'assimilation perceptive concernant les contrastes non-natifs (Best 1995 : 195) :

> - Assimilation à deux catégories (type TC : *two-category*) *Chacun des deux segments non-natifs est assimilé à une catégorie native différente : il est attendu que la discrimination soit excellente*[77].
>
> - Différence de degré de bon exemplaire (type CG : *category-goodness*) *Les deux sons non-natifs sont assimilés à la même catégorie native, mais ils diffèrent dans l'écart par rapport à l'"idéal" natif (ex. l'un est acceptable, l'autre est déviant). Il est attendu que la discrimination soit modérée à très bonne, selon la magnitude de la différence de degré de bon exemplaire de chacun des sons non-natifs*[78].
>
> - Assimilation à une seule catégorie (type SC : *single-category*) *Les deux sons non-natifs sont assimilés à la même catégorie native, mais ils sont également écartés de "l'idéal" natif, c'est-à-dire que les deux sont également acceptables ou les deux sont également déviants. Il est attendu que la discrimination soit mauvaise (bien qu'elle puisse être quelque peu au-dessus du niveau du hasard)*[79].
>
> - Les deux sont non catégorisables (type UU : *uncategorizable-uncategorizable*) *Les deux sons non-natifs tombent dans l'espace phonétique, mais en dehors d'aucune des catégories natives spécifiques, et ils peuvent varier dans leur discriminabilité en tant que sons de la parole non catégorisables. Il est attendu que la discrimination varie de mauvaise à très bonne, selon la proximité entre eux, et la proximité aux catégories natives dans l'espace phonologique natif*[80].
>
> - Non catégorisé versus catégorisé (type UC : *uncategorized-categorized*) *L'un des deux sons non-natifs est assimilé à une catégorie native, et l'autre tombe dans l'espace phonétique, en dehors des catégories natives. Il est attendu que la discrimination soit très bonne*[81].
>
> - Non assimilable (type NA) *Les deux catégories non-natives tombent en dehors du domaine de la parole, étant perçues comme des sons qui ne relèvent pas de la parole, et la paire peut varier dans leur discriminabilité en tant que sons qui n'appartiennent pas à la parole. Il est attendu que la discrimination soit bonne à très bonne*[82].

Ces six types d'assimilation des contrastes non natifs peuvent être schématisés comme suit (Figure 4).

---

[77] « Two-Category Assimilation (TC Type) *Each non-native segment is assimilated to a different native category, and discrimination is expected to be excellent.* »

[78] « Category-Goodness Difference (CG Type) *Both non-native sounds are assimilated to the same native category, but they differ in discrepancy from native "ideal" (*e.g., *one is acceptable, the other deviant). Discrimination is expected to be moderate to very good, depending on the magnitude of difference in category goodness for each of the non-native sounds.* »

[79] « Single-Category Assimilation (SC Type) *Both non-native sounds are assimilated to the same native category, but are equally discrepant from the native "ideal", that is, both are equally acceptable or both equally deviant. Discrimination is expected to be poor (although it may be somewhat above chance level).* »

[80] « Both Uncategorizable (UU Type) *Both non-native sounds fall within phonetic space but outside of any particular native category, and can vary in their discriminability as uncategorizable speech sounds. Discrimination is expected to range from poor to very good, depending upon their proximity to each other and to native categories within native phonological space.* »

[81] « Uncategorized versus Categorized (UC Type) *One non-native sound assimilated to a native category, the other falls in phonetic space, outside native categories. Discrimination is expected to be very good.* »

[82] « Nonassimilable (NA Type) *Both non-native categories fall outside of speech domain being heard as nonspeech sounds, and the pair can vary in their discriminability as nonspeech sounds; discrimination is expected to be good to very good.* »

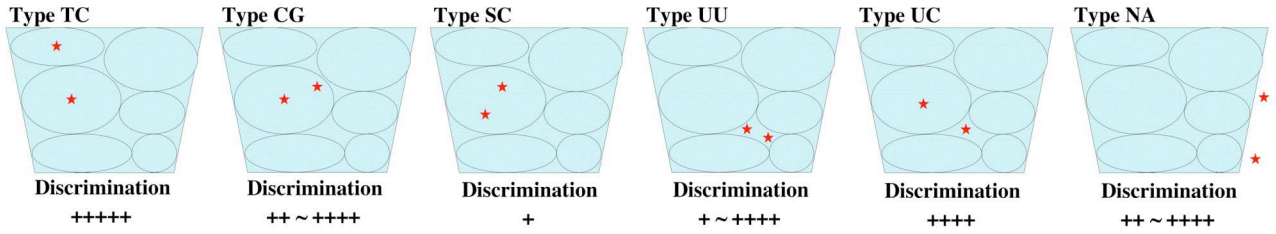



**Figure 4** : Six types d'assimilation perceptive des contrastes non natifs selon Best (1995) : type TC (assimilation à deux catégories), type CG (différence de degré de bon exemplaire), type SC (assimilation à une seule catégorie), type UU (les deux sont non catégorisables), type UC (non catégorisé versus catégorisé), type NA (non assimilable). Qualité de discrimination attendue : + (mauvaise). ++ (modérée), +++ (bonne), ++++ (très bonne), +++++ (excellente). Figure fondée sur les critères de Best (1995).

Ce modèle a été conçu à l'origine pour le cas d'auditeurs qui ne sont pas des apprenants de la langue non native en question, mais selon l'auteure, les difficultés que les apprenants rencontrent seront prédictibles à partir des patterns d'assimilation présentés ci-dessus.

**Le modèle PME de Patricia Kuhl *et al.***

L'effet magnétique de la perception (*Perceptual Magnet Effect* : PME) de Patricia Kuhl *et al.* (Kuhl *et al.* 1992[83], Kuhl et Iverson 1995[84], Kuhl 2000[85], entre autres) est illustré à la Figure 5 ci-dessous (Kuhl et Iverson 1995 : 124). Les stimuli autour d'un prototype phonétique de la langue native sont perceptivement attirés vers le prototype, ce qui réduit la distance perceptive entre le prototype et les autres stimuli.

[83] P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens, B. Lindblom. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255 : 606 – 608, 1992.

[84] P. K. Kuhl, P. Iverson. "Linguistic experience and the 'Perceptual Magnet Effect'", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 121-154, 1995.

[85] P. K. Kuhl. A new view of language acquisition. *Proceedings of the National Academy of Science USA* 97 : 11850-7, 2000.

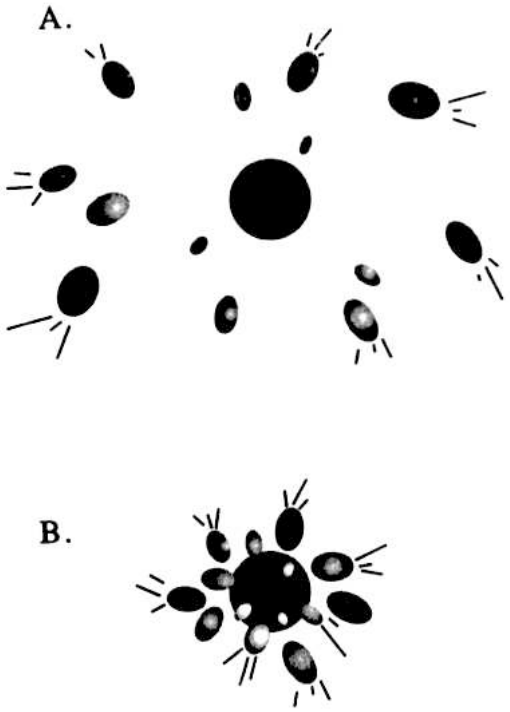



Figure 1. The perceptual magnet effect. Stimuli surrounding a phonetic prototype (A) are perceptually drawn toward the prototype (B), thereby shrinking the perceived distance between the prototype and other members of the category (from Kuhl 1993a).

**Figure 5 :** Effet magnétique de la perception selon Kuhl et ses collègues. Les stimuli autour d'un prototype (A., en haut) sont perceptivement tirés vers le prototype (B., en bas), réduisant ainsi la distance perçue entre le prototype et d'autres membres de la catégorie (Kuhl et Iverson 1995 : 124).

L'espace perceptif ainsi déformé par l'expérience linguistique de la langue native donne lieu à une attraction perceptive des sons autour des prototypes de la langue native. Par conséquent, le modèle prédit que la difficulté rencontrée par les apprenants concernant un segment donné d'une langue non-native dépendra de sa proximité des prototypes magnétiques de leur langue native, qui masquent en quelque sorte la distance entre le segment cible et le prototype le plus proche.

**Le modèle SLM de James Flege *et al.***

À la différence des deux modèles présentés ci-dessus, le modèle de l'apprentissage de la parole (*Speech Learning Model* : SLM) de James Flege *et al*. concerne directement les locuteurs adultes d'une langue seconde. Les études sur lesquelles le modèle est fondé portent majoritairement sur des locuteurs natifs de langues diverses (italien, entre autres, dans Flege *et al*. 1995[86], Flege *et al*. 1997[87], Flege *et al*. 1999[88]) vivant en Amérique du Nord et parlant anglais comme

[86] J. E. Flege, M. J. Munro, I. R. A. Mackay. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16 : 1-26, 1995.

[87] J. E. Flege, E. M. Frieda, T. Nozawa. Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25 : 169–186, 1997.

[88] J. E. Flege, I. R. A. MacKay, D. Meador. Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106(5) : 2973-2987, 1999.

langue seconde (non étrangère). Une autre particularité de ce modèle se trouve dans le fait qu'il traite principalement de la production, tandis que les deux autres sont focalisés sur la perception.

Ce modèle est composé des postulats et des hypothèses suivants (Flege 1995 : 239[89]) :

> *P1 (postulat 1) : Les mécanismes et les procédés utilisés lors de l'apprentissage du système des sons de la L1, y compris la formation des catégories, restent intacts toute la vie, et peuvent être appliqués à l'apprentissage de la L2*[90].
>
> *P2 : Les aspects des sons de la parole qui sont spécifiques à une langue sont spécifiés dans les représentations de la mémoire à long-terme, appelées catégories phonétiques*[91].
>
> *P3 : Les catégories phonétiques établies dans l'enfance pour les sons de la L1 évoluent pendant toute la vie afin de refléter les propriétés de tous les phones de la L1 ou de la L2 qui sont identifiés en tant que réalisation de chaque catégorie*[92].
>
> *P4 : Les bilingues s'efforcent de maintenir le contraste entre les catégories phonétiques de la L1 et de la L2, qui existent dans un espace phonologique commun.*[93]

> *H1 (hypothèse 1) : Les sons de la L1 et de la L2 sont liés perceptivement entre eux à un niveau des allophones (dont la position est prise en compte), plutôt qu'à un niveau phonémique plus abstrait*[94].
>
> *H2 : Une nouvelle catégorie phonétique peut être établie pour un son L2 qui diffère phonétiquement du son de la L1 le plus proche si les bilingues discernent au moins quelques-unes des différences phonétiques entre les sons de la L1 et de la L2*[95].
>
> *H3 : Plus la dissimilarité phonétique perçue entre un son L2 et le son le plus proche de la L1 est grande, plus il est probable que les différences phonétiques entre les sons seront discernées*[96].
>
> *H4 : Il devient de moins en moins probable que les différences phonétiques entre les sons de la L1 et de la L2, et entre les sons de la L2 qui ne sont pas contrastifs en L1, soient discernées, au fur et à mesure que l'âge du début d'apprentissage (AOL :* Age of Learning*) augmente*[97].
>
> *H5 : La formation de catégorie d'un son L2 peut être bloquée par le mécanisme de classification par équivalence. Quand cela arrivera, une seule catégorie phonétique sera utilisée pour procéder des sons L1 et L2 perceptivement liés (diaphones). En définitive, les diaphones se ressembleront en production*[98].

---

[89] J. E. Flege. "Second language speech learning: Theory, findings, and problems". in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 233-277, 1995.

[90] « *P1 The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.* »

[91] « *P2 Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.* »

[92] « *P3 Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.* »

[93] « *P4 Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.* »

[94] « *H1 Sounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level.* »

[95] « *H2 A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.* »

[96] « *H3 The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.* »

[97] « *H4 The likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are noncontrastive in the L1, being discerned decreases as AOL increases.* »

[98] « *H5 Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production.* »

> *H6 : La catégorie phonétique établie par un bilingue pour les sons de la L2 peut être différente de celle d'un monolingue si ; 1) la catégorie chez le bilingue est « déviée » par rapport à une catégorie L1 afin de maintenir un contraste phonétique entre des catégories dans un espace phonologique commun L1-L2 ; ou 2) la représentation chez le bilingue est fondée sur des traits différents, ou sur des poids différents de traits, par rapport à celle de chez le monolingue*[99].
>
> *H7 : La production d'un son correspond en définitive aux propriétés représentées dans la représentation de la catégorie phonétique*[100].

Ce modèle porte ainsi non seulement sur la perception mais aussi sur la production, et cette dernière serait conditionnée par la première. L'auteur affirme qu'une corrélation significative est observée entre les mesures de la production et de la perception des voyelles de la L2 et que cette corrélation est conforme à une autre hypothèse du modèle SLM : « l'exactitude avec laquelle les voyelles de la L2 sont produites est limitée par l'exactitude avec laquelle elles sont perçues » (Flege *et al*. 1999)[101].

Nous pouvons ainsi constater que les trois modèles présentés dans le présent chapitre, le PAM de Best *et al*., le PME de Kuhl *et al*., et le SLM de Flege *et al*., accordent une importance considérable au système phonémique non seulement de la langue cible mais aussi de la langue source. C'est pour cette raison que nous regarderons les systèmes phonémiques des trois langues concernées dans notre étude.

## 1.4. PHONÉTIQUE ET PHONOLOGIE DU JAPONAIS DE TOKYO -- LANGUE SOURCE (L1) DE NOTRE ÉTUDE

Nous allons présenter dans les trois chapitres suivants la phonétique et la phonologie des trois langues concernées : le japonais de Tokyo (langue maternelle des apprenants : L1), le français (langue cible : L3), et l'anglais (langue étrangère déjà apprise : L2). Nous commençons par le japonais de Tokyo (pour une description détaillée, voir Shinohara 1997[102], Labrune 2006[103], Vance 2008[104], entre autres).

### *1.4.1. Les voyelles*

Le système vocalique du japonais est composé de 5 voyelles /i/ /e/ /a/ /o/ /u/. Notons au passage que le système à 5 voyelles est le plus fréquemment observé dans les langues du monde (environ un cinquième des 451 langues enregistrées dans la base de données UPSID :

---

[99] « *H6 The phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if; 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's*. »

[100] « *H7 The production of a sound eventually corresponds to the properties represented in its phonetic category representation*. »

[101] « *the accuracy with which L2 vowels are produced is limited by how accurately they are perceived* ».

J. E. Flege, I. R. A. MacKay, D. Meador. Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106(5) : 2973-2987, 1999.

[102] S. Shinohara. *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Thèse de doctorat, Université de la Sorbonne Nouvelle, 1997.

[103] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

[104] T. J. Vance. *The sounds of Japanese*. Cambridge, Cambridge University Press, 2008.

Maddieson et Precoda 1990 : 109-111[105]). Le système phonémique du japonais ne possède donc ni deux degrés d'aperture moyenne (mi-fermé / mi-ouvert), ni la série antérieure arrondie, ni de voyelles nasales, que possède le français.

Les formants des 5 voyelles japonaises sont indiqués à la Figure 6 (Sugitô 1995[106]), et les profils sagittaux de ces 5 voyelles prononcées isolément sont montrés à la Figure 7 (Uemura 1990[107]).

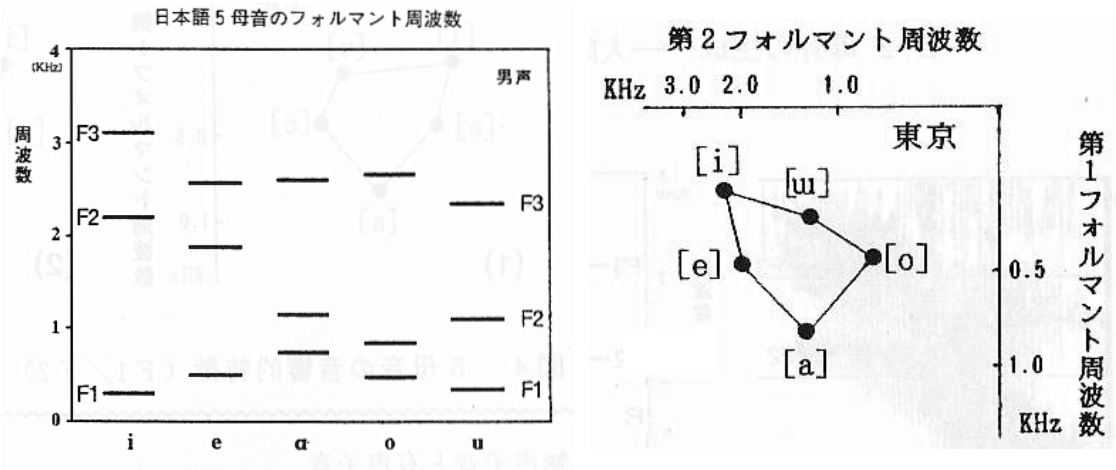



**Figure 6** : Les trois premiers formants (à gauche) et les deux premiers formants représentés en 2 dimensions (à droite ; F1 sur l'axe vertical et F2 sur l'axe horizontal) des 5 voyelles (voix masculine) du japonais de Tokyo (Sugitô 1995).

[105] I. Maddieson, K. Precoda. Updating UPSID. *UCLA Working Papers in Phonetics* 74 : 104–111, 1990.

[106] M. Sugitô. *Ôsaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [Dictionnaire CD-ROM sonore d'accent des japonais d'Osaka et de Tokyo : commentaires]. Tokyo, Maruzen, 1995.

[107] 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]. Tokyo, Shûei shuppan, 1990.

Le film cinéradiographique « *Cineradiograph of Japanese Pronunciation* [Cinéradiographe de la prononciation du japonais] » est disponible sur le site de l'institut national de la langue japonaise (en japonais) : http://www6.kokken.go.jp/x-sen/

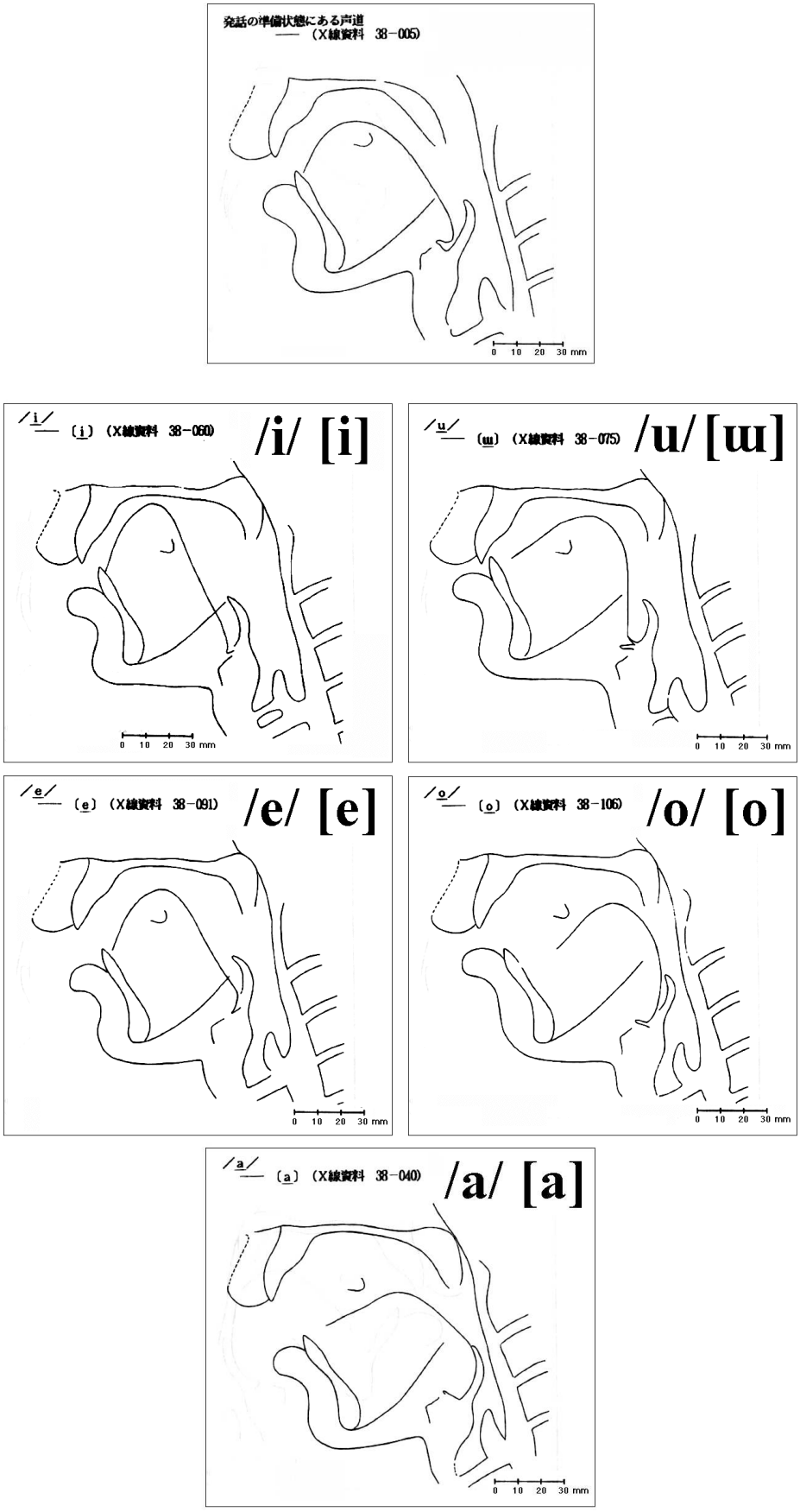
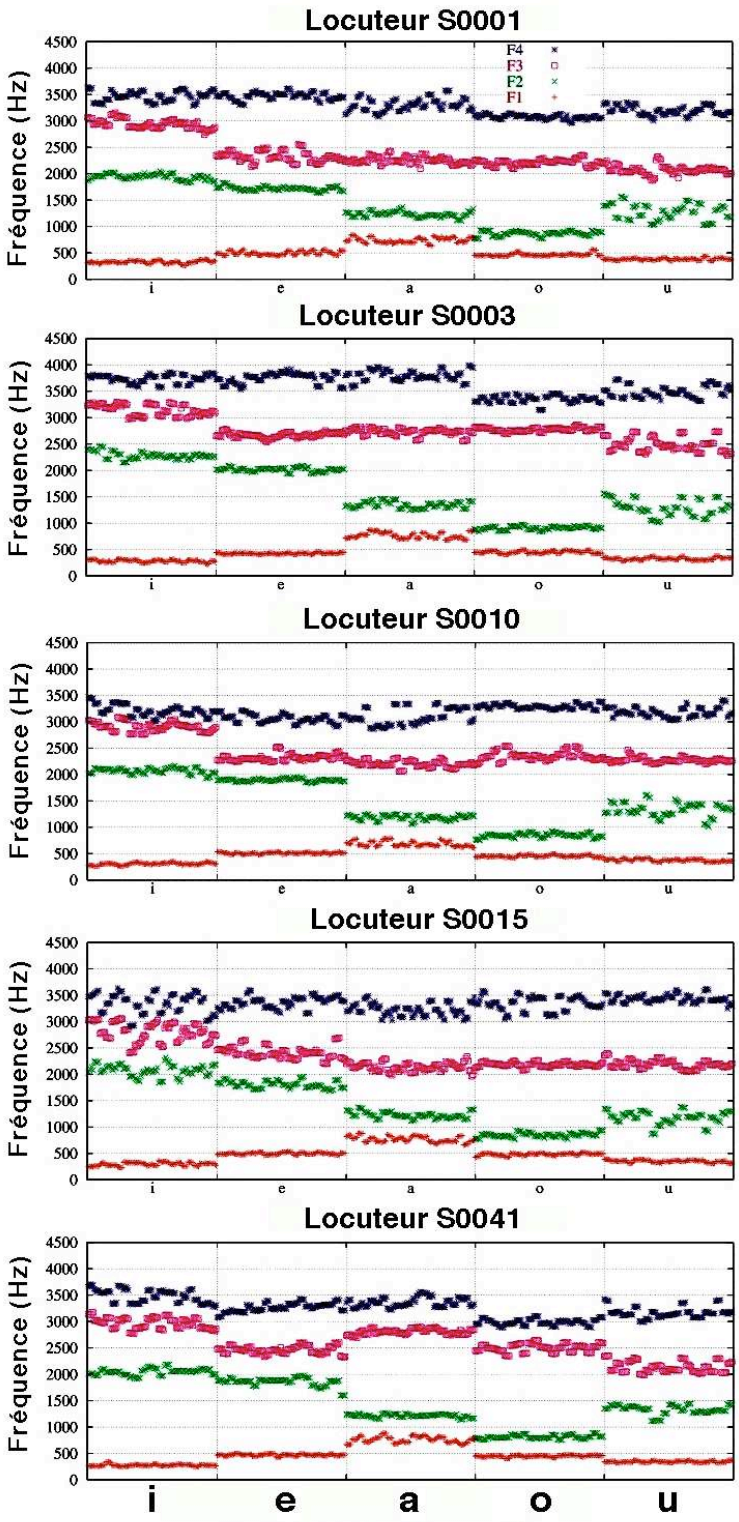



**Figure 7** : Profils sagittaux des 5 voyelles du japonais de Tokyo (locuteur masculin) prononcées isolément, et celui de l'état de préparation d'articulation (position neutre) en haut (d'après Uemura 1990 : 58-60).

Les voyelles moyennes /e/ et /o/ seraient réalisées entre les deux degrés d'aperture, mi-fermé [e]/[o] et mi-ouvert [ɛ]/[ɔ] (Labrune 2006 : 38, Vance 2008 : 53-54). La voyelle fermée antérieure /i/ serait réalisée avec un rapprochement de F3 et F4 comme le /i/ français (Chapitre 1.5., Vaissière 2007[108]), au moins chez certains locuteurs, comme nous pouvons le constater dans les données de Mokhtari et Tanaka (2000)[109] présentées ci-dessous (Figure 8).

[108] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

[109] P. Mokhtari, K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66, 2000.




**Appendix B - Vowel Sequence Charts**
Vowel sequence charts for each of the 5 speakers, showing the distribution of the first four formant frequencies {F1, F2, F3, F4}. Within each vowel category, the horizontal sequence of 110 formant patterns comprises 5 adjacent steady-state frames × 22 vocalic nuclei.

**Figure 8 :** Les 4 premiers formants des 5 voyelles /i e a o u/ du japonais dans 22 mots (différents pour chaque voyelle) lus par 5 hommes japonophones. Pour chacune des 5 voyelles, 110 valeurs (5 mesures x 22 occurrences de la voyelle) de chacun des 4 formants sont représentés (d'après Mokhtari et Tanaka 2000 : 65).

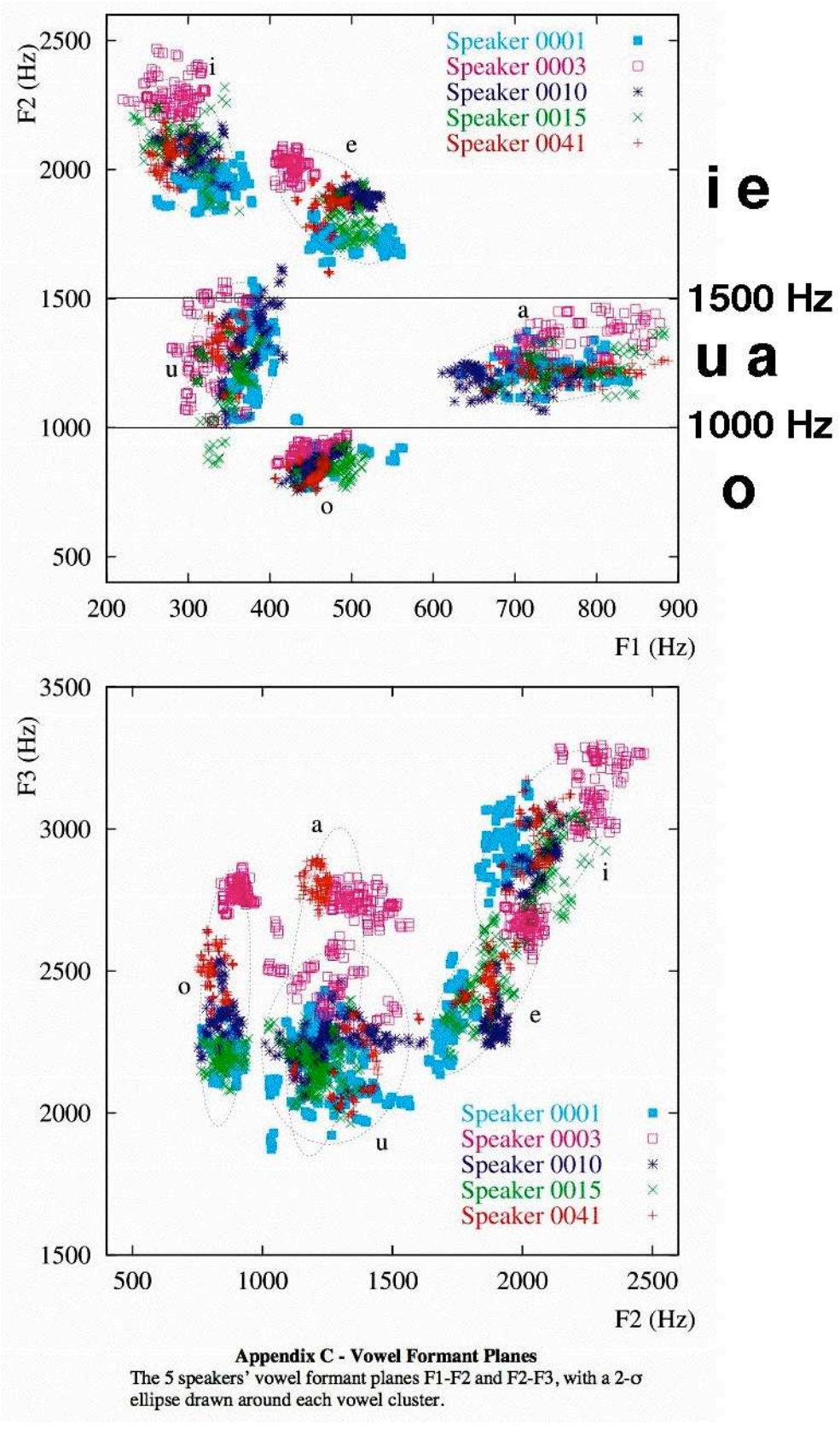



**Figure 9 :** Les 3 premiers formants des 5 voyelles /i e a o u/ du japonais dans 22 mots pour chacune des voyelles lus par 5 hommes japonophones, représentés sur les plans F1-F2 (en haut) et F2-F3 (en bas). Les ellipses correspondent à 2 écarts-types. D'après Mokhtari et Tanaka (2000 : 66).

La réalisation phonétique de la voyelle postérieure fermée /u/ est communément décrite [ɯ] en transcription phonétique large, avec l'allophone central [ɨ], qui apparaît dans un contexte antérieur (avec des obstruentes coronales : Shinohara 1997 : 27[110]). Il est intéressant de noter que le /u/ japonais dans les noms propres japonais est transcrit de manière variable en coréen, tantôt /u/ : postérieur arrondi ; et tantôt /ɯ/[111] : postérieur/central non arrondi. Il est transcrit /ɯ/ (一) dans les syllabes japonaises ス・ツ・ズ /su tu zu/ (>스 쓰 즈 /sɯ s'ɯ cɯ/ en coréen : ex. 鈴木／スズキ Suzuki /suzuki/ > 스즈키 /sɯcɯkʰi/). En revanche, il est transcrit /u/ (ㅜ) dans les autres contextes (ex. 福岡 Fukuoka /hukuoka/ > 후쿠오카 /hukʰuokʰa/). Amanuma *et al.* (1978 : 54-55)[112] signalent que le nom et le prénom du premier auteur (天沼 寧 Amanuma Yasushi /amanuma jasusi/) est transcrit 아마누마 야스시 /amanuma jasɯsi/ en coréen.

Le F2 est en général supérieur à 1000 Hz (Figure 6, Sugitô 1995, Figure 8 et Figure 9, Mokhtari et Tanaka 2000), à la différence du /u/ du français métropolitain. Notons que la voyelle qui a le plus bas F2 parmi les 5 voyelles japonaises est le /o/. Les données de Mokhtari et Tanaka (2000), qui portent sur la partie stable des 5 voyelles contenues dans 22 mots isolés pour chacune des voyelles (= 110 mots « phonétiquement équilibrés » selon les auteurs), lus par 5 hommes japonophones, montrent bien que le F2 varie entre 1000 et 1500-1600 Hz[113], avec une variabilité intra-locuteur plus importante que /o/ (Figure 8 et Figure 9). Notons que la variabilité serait encore plus importante pour des corpus de lecture de texte et de parole spontanée. Keating et Huffman (1984)[114] ont analysé les 3 premiers formants des 5 voyelles japonaises prononcés par 7 hommes natifs du japonais de Tokyo. Les voyelles ont été inclues : 1) dans 15 mots isolés (3 mots cibles pour les 5 voyelles : /#h_bV#/, /#CVb_#/, /CVCVb_#/) lus deux fois ; 2) dans des textes lus, qui contenaient entre 59 et 100 occurrences de chaque voyelle effectivement analysées. Les résultats montrent que la dispersion de chaque voyelle dans l'espace F1-F2 est plus importante dans la lecture des textes que dans les mots isolés, résultant en des chevauchements des espaces des 5 voyelles phonémiquement distinctes (Figure 10). Nous pouvons y remarquer que le F2 du /u/ [ɯ] prononcé en lecture de textes s'élève jusqu'à presque 2000 Hz chez certains locuteurs. Le F2 de /o/ est inférieur à 1000 Hz dans des mots isolés, mais majoritairement supérieur à 1000 Hz en lecture de textes. Ces données illustrent non seulement l'effet de la coarticulation mais aussi du degré d'hyper/hypo-articulation.

---

[110] S. Shinohara. *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Thèse de doctorat, Université de la Sorbonne Nouvelle, 1997.

[111] Le symbole /ɨ/ est également utilisé pour transcrire ce phonème (ex. Lee *et al*. 1997).

I.-S. Lee, S.-O. Lee, W. Chai. *Kankokugo gaisetsu* [Aperçu de la langue coréenne], traduit en japonais par H. Umeda, T. Maeda. Tokyo, Taishûkan, 2004.

[112] Y. Amanuma, K. Ôtsubo, O. Mizutani. *Nihongo Onseigaku* [Phonétique du japonais]. Tokyo, Kuroshio, 1978.

[113] Le F2 de /o/ se trouve en dessous de 1000 Hz, celui de /u/ et /a/ entre 1000 Hz et 1500 Hz, et celui de /i/ et /a/ au-dessus de 1500 Hz (Figure 9).

[114] P. A. Keating, M. K. Huffman. Vowel variation in Japanese. *Phonetica* 41 : 191-207, 1984.

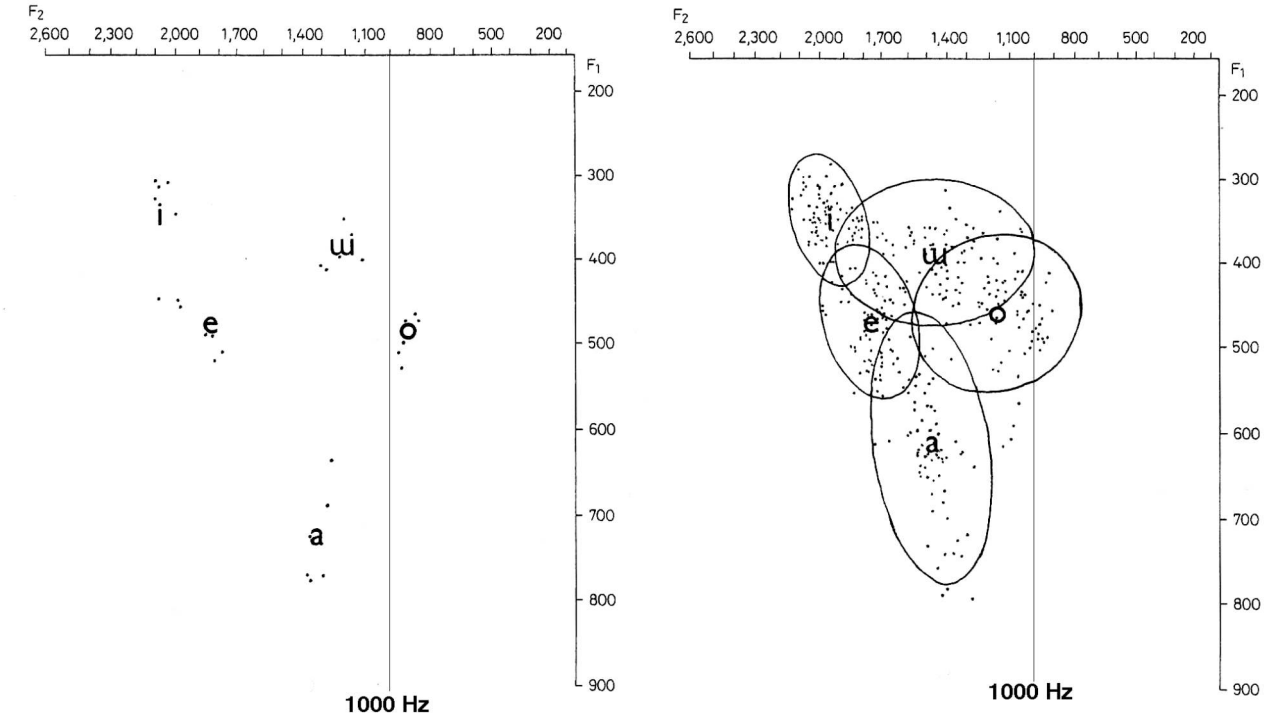



**Figure 10 :** Dispersion des 2 premiers formants des 5 voyelles du japonais de Tokyo prononcées ; i) dans 15 mots lus par un des 7 locuteurs examinés (2 répétitions : à gauche) : les valeurs de chaque production et la moyenne sont représentées (contexte : /#h_bV#/, /#CVb_#/, /CVCVb_#/) ; ii) dans des textes lus par un même locuteur (81 occurrences de chaque voyelle : à droite) : les valeurs de chaque production et la moyenne sont représentées (d'après Keating et Huffman 1984 : 198, 200).

Du point de vue articulatoire, le /u/ japonais est moins arrondi et plus antérieur que le [u] de l'API ou du français, comme le montrent les profils articulatoires d'Uemura (1990)[115] et de Wioland (1991)[116] à la Figure 11. Nous pouvons observer les caractéristiques suivantes du /u/ japonais, comparé au /u/ français :

1) la constriction labiale est moins étroite
2) la constriction vélaire est moins étroite
3) le volume de la cavité pharyngale est plus important

[115] 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]. Tokyo, Shûei shuppan, 1990.

[116] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

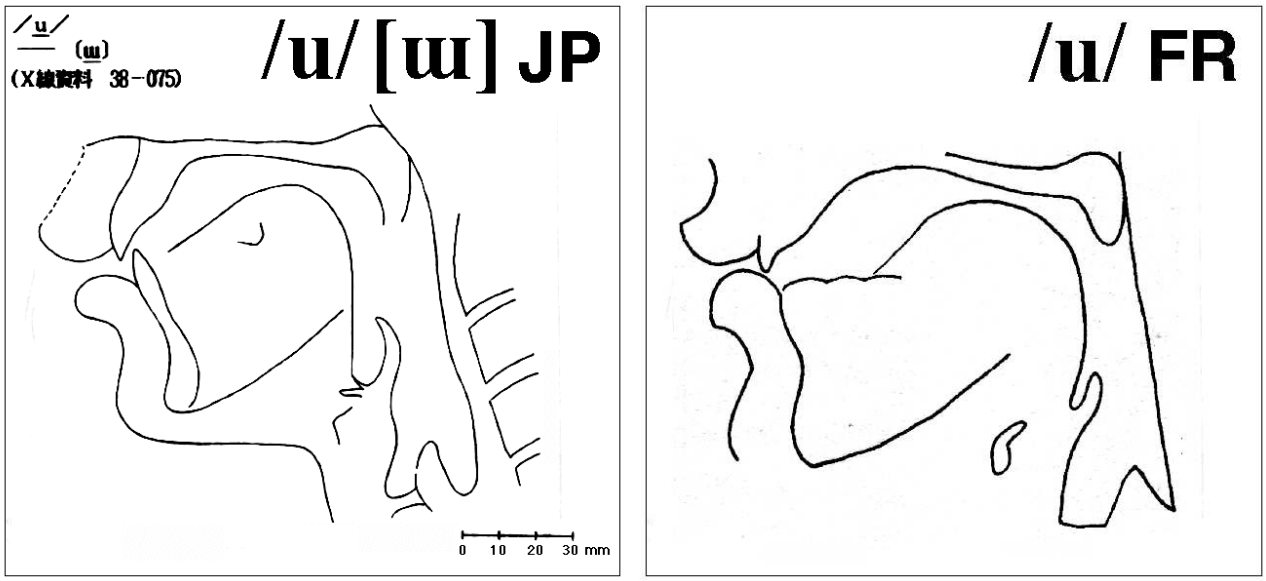



**Figure 11 :** Profil articulatoire du /u/ [ɯ] du japonais de Tokyo (d'après Uemura 1990 : 59, à gauche) et celui du /u/ français (d'après Wioland 1991 : 17, à droite).

Les nomogrammes de Fant (1960 : 76)[117] sont reproduits à la Figure 12. Nous pouvons y voir que quand la constriction se trouve au milieu du conduit vocal, F2 est fortement influencé par la configuration des lèvres : une protrusion / arrondissement des lèvres abaisse considérablement la valeur du F2, et légèrement celle du F1. Nous y observons également que le fait d'avancer la constriction a pour effet d'élever F2 dans cette zone. Cela est en accord avec les observations articulatoires et les valeurs élevées de F1 et F2 du /u/ japonais, dont la constriction est plus en avant et qui a un faible degré de labialité, par rapport au /u/ français.

[117] G. Fant. *Acoustic Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*. The Hague, Mouton, 1960.

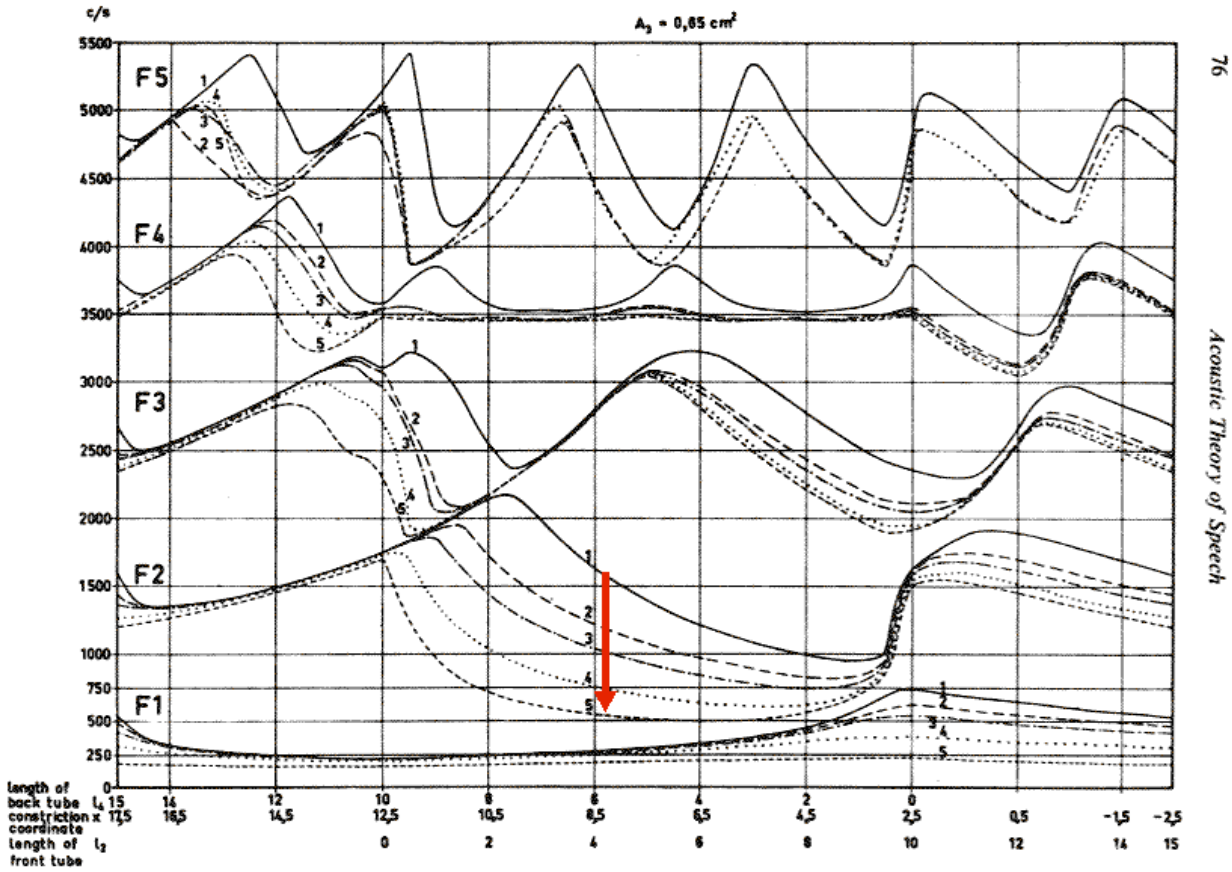



Fig. 1.4-9. Nomograms relating $F_1$, $F_2$, $F_3$, $F_4$, and $F_5$ of the three-parameter model *Fig. 1.4-8* to the location of the tongue

**Figure 12 :** Nomogrammes de Fant (d'après Fant 1960 : 78), montrant le lieu de constriction (sur l'axe horizontal : les lèvres à gauche, la glotte à droite), les valeurs de formants (sur l'axe vertical) estimées à partir d'un modèle à trois paramètres (position de la constriction, aire de la constriction, configuration des lèvres), et l'effet de la configuration des lèvres. Les courbes numérotées représentent de différents degrés de labialité, qui varie de 1 (moins) à 5 (plus). La flèche indique l'effet de l'aire et de la longueur des lèvres sur le deuxième formant quand la construction se trouve vers le milieu du conduit vocal.

Notons également que cette voyelle, malgré la transcription large courante [ɯ], n'est pas identique, comme Uwano (1993 : 224)[118] le signale, à la voyelle cardinale no. 16 (postérieure fermée non-arrondie), telle qu'elle est définie comme voyelle prononcée avec la position de la langue de la cardinale no. 8 [u], mais avec les lèvres étirées[119]. Takebayashi (1996 : 59)[120] et Vance (2008 : 56)[121] proposent la transcription étroite [ɯ̟] (avec le symbole diacritique d'antériorité) pour cette voyelle. Malgré tout, les données acoustiques, au moins celles des valeurs formantiques semblent être assez proches de celles de la cardinale no. 16 prononcée par des phonéticiens, dont Daniel Jones

[118] Z. Uwano "On no kouzou [Structure des sons de la parole]", in K. Kazama, *Gengogaku* [Linguistique]. Tokyo, Tokyo daigaku shuppankai [Presses de l'Université de Tokyo], 193-249, 1993.

[119] "… and No. 16 has the tongue position of No. 8, but said with spread lips." (Jones 1956 : 12)

D. Jones. *Cardinal Vowels*. London, Linguaphone Institute, 1956 (livret d'explication accompagné de 2 disques audio).

[120] S. Takebayashi. *Eigo onseigaku* [Phonétique de l'anglais]. Tokyo, Kenkyûsha, 1996.

[121] T. J. Vance. *The sounds of Japanese*. Cambridge, Cambridge University Press, 2008.

(Tableau 2). Le spectrogramme de la Figure 13 montre que les productions de Peter Ladefoged et de John Wells ressembleraient bien au /u/ japonais (chapitre 2.1. ; Figure 39, p. 120). En ce qui concerne la répartition des formants, c'est une voyelle acoustiquement centrale (Vaissière 2006[122] : F2 se trouve à mi-chemin entre F1 et F3, et il n'y a pas de groupement de formants).

**Tableau 2** : Valeurs moyennes des 4 premiers formants de la voyelle cardinale no. 16 [ɯ] prononcée par 4 phonéticiens (1 répétition, sauf 5 pour Daniel Jones), comparées avec celles du /u/ du japonais de Tokyo (2 locuteurs masculins, 12 répétitions au total : Figure 39, p. 120). Du haut en bas, les valeurs de Daniel Jones, de Peter Ladefoged, de John Wells, la moyenne de ces trois phonéticiens masculins, les valeurs de Jill House, les valeurs moyennes du /u/ japonais prononcés par deux locuteurs natifs. Source : Jones (1956, *op. cit*) pour Jones, site Web du laboratoire de phonétique de l'UCLA[123] pour Ladefoged, *The sounds of the IPA*[124] pour Wells et House.

| ***Voyelle cardinale No. 16*** | **F1 moyen** | **F2 moyen** | **F3 moyen** | **F4 moyen** |
|---|---|---|---|---|
| **Daniel Jones** | 308 | 1230 | 2485 | 3719 |
| **Peter Ladefoged** | 310 | 1305 | 2274 | 3254 |
| **John Wells** | 271 | 1108 | 2148 | 3490 |
| ***Moyenne DJ, PL, JW*** | 296 | 1214 | 2302 | 3488 |
| **Jill House** | 371 | 1476 | 3005 | 4377 |
| ***/u/ japonais (JP1, JP2)*** | 323 | 1125 | 2300 | 3508 |

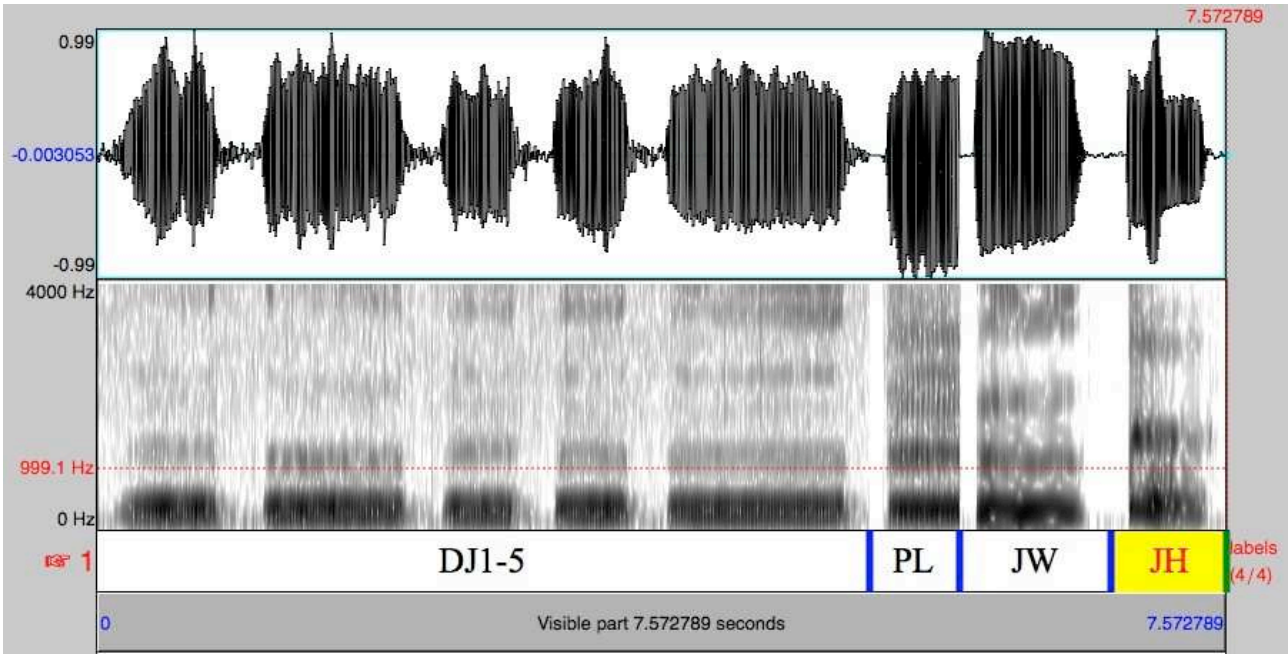



**Figure 13 :** Spectrogramme à bande large (Praat ; longueur de fenêtre : 5 millisecondes) de la voyelle cardinale no. 16 [ɯ] prononcée par Daniel Jones (5 répétitions), Peter Ladefoged, John Wells, et Jill House.

Dans la présente étude, nous utiliserons le symbole [ɯ] en transcription phonétique large pour représenter la réalisation phonétique de la voyelle fermée postérieure /u/ de cette langue.

---

[122] J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France, 2006.

[123] http://www.phonetics.ucla.edu/course/chapter1/secondary.aiff consulté le 6 septembre 2006.

[124] J. Wells, J. House. *The sounds of the IPA* (CD audio). London, University College London, 1995.

**Le dévoisement des voyelles en japonais de Tokyo**

Les voyelles fermées /i/ /u/ se réalisent en général sans voisement dans les deux contextes suivants (Labrune 2006 : 47[125], entre autres) :

1) lorsque la voyelle, accentuée ou non, est placée entre deux consonnes sourdes, y compris avant la première partie d'une géminée (ex. 草 *kusa* /kusa/[126] [kɯ̥sa] « herbe »).

2) lorsque la voyelle, inaccentuée, est placée après une consonne sourde et avant une pause (ex. 菓子 *kashi* /'kasi/ [kaɕi̥] « gâteau »).

Notons que la classe naturelle des voyelles fermées ([+ haut]) du français (/i y u/) subit également le dévoisement en français canadien (ex. poutine /putin/ [pu̥tsɪn] : Ostiguy *et al.* 1996 : 167[127], entre autres).

Le dévoisement est également observé, plus rarement, avec /a/ et /o/ sous certaines conditions. Selon Labrune (*op. cit.*), il s'agit alors d'une voyelle « inaccentuée, placée entre deux consonnes sourdes, et condition supplémentaire, une voyelle identique doit en outre figurer dans la more suivante ». Amanuma *et al.* (1978)[128] décrit le cas comme suit : des syllabes qui ont la structure « consonnes sourde + /o/ » ou « consonne sourde + /a/ » se succèdent, comme dans こころ／心 *kokoro* /koko'ro ~ ko'koro/ [ko̥koro] « cœur » ou かかし／案山子 *kakashi* /kakasi/ [kḁkaɕi] « épouvantail ». Ces auteurs mentionnent également le cas de dévoisement du /e/ que l'on peut observer dans せっかく／折角 *sekkaku* /seQkaku/ « avec beaucoup d'efforts » dans une prononciation emphatique. Par ailleurs, en étudiant les consonnes qui apparaissent avant et après les voyelles dévoisées dans les exemples cités par Labrune (2006) et Amanuma *et al.* (1978), nous nous rendons compte du fait qu'il s'agit des fricatives (/s h/) et de /k/ (ここ /ko̲ko/ « ici », かかる /ka̲'karu/ « s'accrocher », こころ /ko'koro ~ koko'ro/ « cœur », ほこり／埃 /ho̲kori/ « poussière », 墓 /ha̲'ka/ « tombe », せっかく／折角 /se̲Qkaku/ « avec beaucoup d'effort », 化粧 /ke̲'sjoR/ « maquillage »). Sawashima et Miyazaki (1973)[129], et Sawashima et Niimi (1974)[130] ont montré qu'en japonais de Tokyo l'ouverture de la glotte pendant /k/ était plus importante que celle des autres occlusives, et presque aussi importante que celle du /s/, au moins à l'initiale de mot. Étant donné que l'ouverture de la glotte empêche la vibration des plis vocaux, nous pourrions expliquer le fait que le dévoisement de voyelles est observé plus souvent dans des contextes où se trouvent les fricatives ainsi que le /k/ par l'ouverture plus importante de la glotte pour ces consonnes.

---

[125] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

[126] Sur la notation de l'accent lexical, voir ci-dessous (1.3.4. Prosodie).

[127] L. Ostiguy, R. Sarrasin, G. Irons. *Introduction à la phonétique comparée : les sons : le français et l'anglais nord-américains*. Sainte-Foy, Les Presses de l'Université Laval, 1996.

[128] Y. Amanuma, K. Ôtsubo, O. Mizutani. *Nihongo Onseigaku* [Phonétique du japonais]. Tokyo, Kuroshio, 1978.

[129] M. Sawashima, S. Miyazaki. Glottal Opening for Japanese Voiceless Consonants. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo* 7 : 1-10, 1973.

[130] M. Sawashima, S. Niimi. Laryngeal Conditions in Articulations of Japanese Voiceless Consonants. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo* 8 : 13-18, 1974.

Selon Labrune (2006 : 47), on trouve une description qui mentionnait probablement le phénomène de dévoisement dès le XVIIème siècle : « Collado (1632)[131] signale déjà que certains /i/ et /u/ japonais sont parfois inaudibles (mais dans des contextes qui ne sont toutefois pas typiquement ceux des dévocalisations modernes) ». Sur le plan synchronique, le phénomène est moins largement observé en japonais d'Osaka, et cette tendance serait due à des différences des gestes articulatoires du larynx : un seul geste d'ouverture et de fermeture de la glotte pour la séquence [CV̥C] pour le japonais de Tokyo, mais deux pour le japonais d'Osaka (Fujimoto *et al.* 2002[132], Fujimoto 2006[133]).

Nous présentons par la suite quelques exemples des conséquences du transfert du dévoisement du japonais à la production du français que nous avons observés auprès d'apprenants japonophones. Le mot « écouter » prononcé par des apprenants japonophones pourrait se réaliser [ekɯ̥te], car les apprenants ont tendance à dévoiser le /u/ qui se trouve entre deux consonnes sourdes. Les auditeurs francophones, notamment de France, auraient de la difficulté à percevoir /ekute/ à cause des deux écarts observés entre la forme prononcée par l'apprenant [ekɯ̥te] et la forme attendue [ekute] : la différence de la voyelle, et le dévoisement[134].

De la même manière, le /ø/, qui est confondu avec le /u/ comme nous le verrons dans les deuxième et troisième parties, au sens où il serait interprété comme /u/ [ɯ] japonais par les apprenants japonophones, pourrait se dévoiser chez eux dans un contexte sourd : « peut former » [pɯ̥foʀme] / [pɯ̥fɔʀme].

On observe également des cas où le [u] peut se dévoiser en finale de mots ou de groupes, ou tout au moins dans une position considérée comme telle par les apprenants. Dans « à tout moment » [a tɯ̥ mɔmɑ̃], et « beaucoup, beaucoup de … » [bokɯ̥ bokɯ̥ dɯ], le /u/ se dévoise comme si c'était la chute d'un « e caduc » dans « te » et « que ». L'écart est d'autant plus important dans ce cas, où la voyelle en question se trouve en fin de mot, car les voyelles sont en général bien

---

[131] D. Collado. *Ars Grammaticae Iaponicae Linguae*. Rome, 1632.

[132] M. Fujimoto, E. Murano, S. Niimi, S. Kiritani. Differences in glottal opening pattern between Tokyo and Osaka dialect speakers: factors contributing to vowel devoicing. *Folia Phoniatrica et Logopaedica* 54(3) : 133-143, 2002.

[133] M. Fujimoto. Is vowel devoicing in Japanese phonological or phonetic?: studies using PGG, MRI and EMG techniques. Poster présenté à la Dixième Conférence de Phonologie de Laboratoire, Paris, 29 juin – 1er juillet, 2006.

[134] Même si les francophones canadiens dévoisent cette voyelle [ekɯ̥te], le *F-pattern* et les résonances excitées durant [ɯ̥] (F2 et supérieurs) permettraient aux francophones de percevoir /u/, mais ce n'est pas le cas pour [ɯ̥], qui a un *F-pattern* et les résonances excitées différents.

Pour la notion de F-pattern, voir la définition de Vaissière (2007), qui a repris le terme de Fant (1960).

« *The filter function of the VT (vocal tract) without branching cavities contains only resonances (i.e., no zeros), labeled F1, F2, F3, F4 and so on. The term F-pattern (Fant 1960: 209) refers to these frequencies. The F-pattern can be calculated for any known VT shape. The F-pattern up to the first two, three, four and sometimes five formants is clearly visible on a spectrogram during oral vowels (with roughly one formant per kHz). The F-pattern for a consonant is determined by the place, degree and shape of the main oral constriction, and by the tongue and lip configuration required by the following vowel. For consonants, fewer clues to the formants are visible on spectrograms.* »

J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

G. Fant. *Acoustic Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*. The Hague, Mouton, 1960.

prononcées (sur le plan du timbre, de la durée, et de l'intensité sauf en fin de phrase déclarative ou d'énoncé) dans cette position (Wioland 1991[135]).

**La longueur vocalique en japonais**

Le japonais présente une opposition entre les voyelles brèves et longues. Contrairement à l'opposition en anglais entre les voyelles brèves (et surtout relâchées en même temps) et les voyelles longues (tendues), les voyelles correspondantes des deux séries (ex. /i/ bref et /i/ long) partagent des timbres similaires (voir la Figure 14 : Nishi *et al.* 2008[136]). Étant donnée la similitude de timbre entre les deux, les voyelles longues sont communément interprétées phonologiquement comme une séquence d'une voyelle brève suivie du phonème d'allongement vocalique /R/ (のばす音 *nobasu oto* ou 長音 *chouon*[137] « son allongé, son long ». Vance 2008[138] utilise le symbole /H/). Le phonème /R/ correspond donc à la deuxième more d'une voyelle longue, et fait partie des 3 phonèmes moraïques, avec /N/ (nasale moraïque, はねる音 *haneru oto* ou 撥音 *hatsuon* « son supprimé ») et /Q/ (première moitié de géminée, つまる音 *tsumaru oto* ou 促音 *sokuon* « son raccourci »). Le [o] long, ou [o:], et le [o] bref, par exemple, sont ainsi représentés phonémiquement /oR/ et /o/, comme dans 王 *ou* /oR/ « roi » et 尾 *o* /o/ « queue ».

[135] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

[136] K. Nishi, W. Strange, R. Akahane-Yamada, R. Kubo, S. A. Trent-Brown. Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America* 124(1) : 576-588, 2008.

[137] Les termes qui représentent les phonèmes moraïques sont donnés ici d'abord dans un mot Yamato (indigène), puis dans un *kango* (mot sino-japonais).

[138] T. J. Vance. *The sounds of Japanese*. Cambridge, Cambridge University Press, 2008.

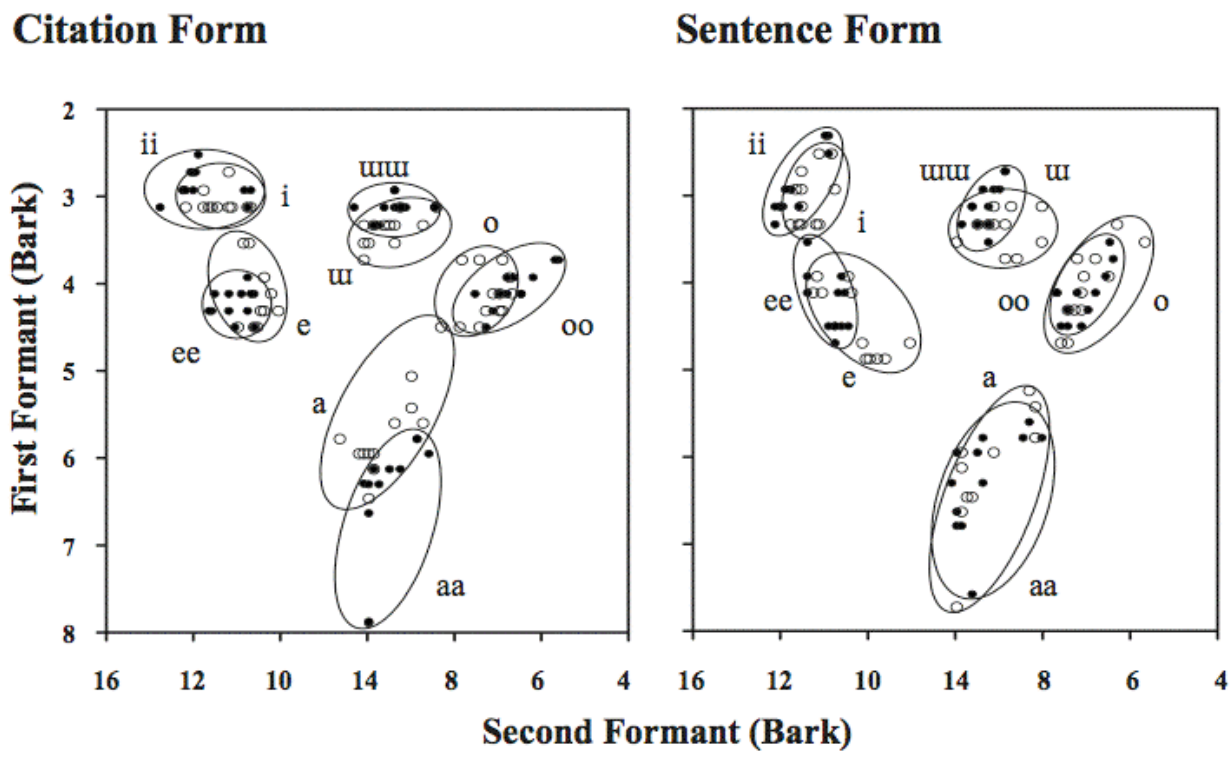
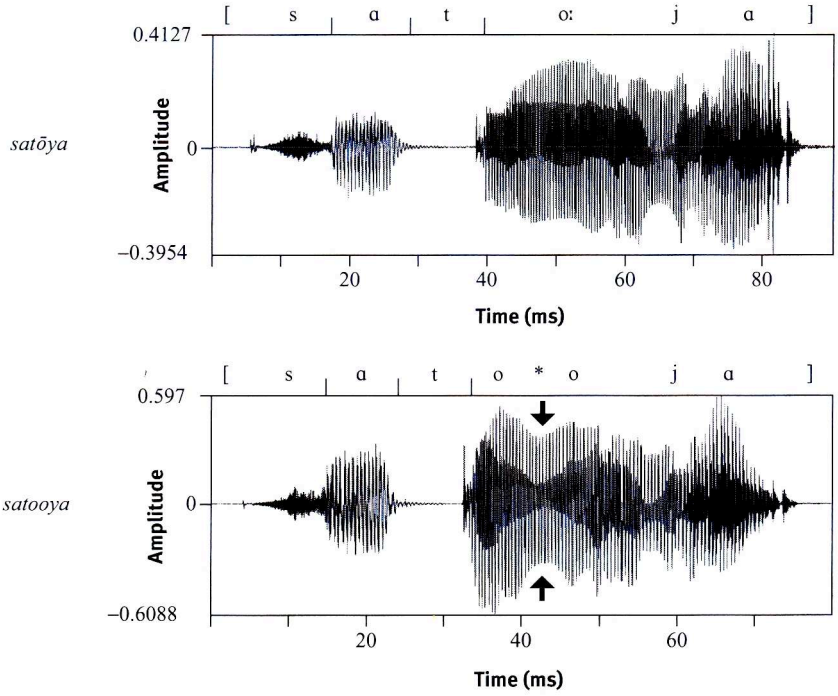



FIG. 1. Formant 1/formant 2 (bark) plots for ten Japanese vowels produced by four male speakers in [hVba] contexts in citation form (left panel) and sentence form (right panel): 12 tokens per vowel. Filled circles are for 2-mora vowels; open circles are for 1-mora vowels.

**Figure 14 :** Les deux premiers formants (en bark) des 5 voyelles brèves et longues du japonais de Tokyo produites par 4 locuteurs masculins dans [hVba] en citation (gauche) et en phrase (droite) : 12 occurrences par voyelle. Les cercles pleins représentent les voyelles longues (à 2 mores), les cercles vides, les voyelles brèves (à 1 more). Nishi *et al*. 2008 : 580.

Les voyelles longues peuvent se distinguer phonétiquement d'une suite de deux occurrences de la même voyelle à la frontière morphologique : ex. (砂糖 *satou* « sucre » + 屋 *ya* « magasin » >) 砂糖屋 *satouya* /satoRja/ [sato:ja] « marchand de sucre » vs. (里 *sato* « village » + 親 *oya* « parents » >) 里親 *satooya* /satooja/ [sato:ja] ou [sato$^{?}$oja] « parents nourriciers » (Labrune 2006[139], entre autres). Vance (2008 : 59) illustre la différence acoustique entre les deux mots par la présence ou l'absence d'une diminution d'amplitude au milieu de la séquence en question (Figure 15).

[139] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.




Waveforms of *satōya* (top) and *satooya* (bottom)

**Figure 15 :** Oscillogramme des mots *satouya* /satoRja/ « marchand de sucre » (en haut) et *satooya* /satooja/ « parents nourriciers » (en bas). La séquence /oo/ dans le dernier est marquée par une diminution d'intensité au milieu de la séquence en question (Vance 2008 : 59).

L'opposition entre les voyelles longues et brèves est neutralisée dans certains mots, par raccourcissement d'une voyelle longue (le raccourcissement est parfois noté dans la graphie) : ex. ほんとう／本当 /hoNtoR/ ou ほんと /hoNto/ « vrai », コンピューター /koN'pjuRtaR/ ou コンピュータ /koN'pjuRta/ « ordinateur » (< *computer*). Labrune (2006 : 61-62) résume les contextes dans lesquels ce phénomène d'« abrègement prosodique » est observé :

> *- Le phénomène […] se produit avant tout dans les emprunts occidentaux, et dans une moindre mesure dans les mots sino-japonais.*
>
> *- Il survient surtout en registre non formel, et frappe principalement les voyelles en position médiale ou finale dans les mores qui suivent une autre voyelle longue ou une nasale-more comme dans […]* hontou.
>
> *- […] l'abrègement ne se produit que rarement en position initiale, y compris dans ce qui peut être considéré comme un monosyllabe, par exemple* mou *« déjà ».*
>
> *- […] l'abrègement prosodique concernerait surtout les voyelles [- hautes] /a/, /e/ et /o/, mais beaucoup plus rarement /i/ et /u/.*

### *1.4.2. Les consonnes*

L'inventaire des phonèmes consonantiques du japonais est fourni dans le Tableau 3 ci-dessous (Labrune 2006 : 72[140]). À cette liste s'ajoutent les deux consonnes moraïques (autrement appelées, phonème moraïque モーラ音素 *môra onso*, phonème spécial 特殊音素 *tokushu onso*,

[140] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

phonème à more spéciale 特殊拍音素 *tokushuhaku onso*) /N/ (la nasale moraïque, はねる音 *haneru oto*, ou 撥音 *hatsuon*) et /Q/ (la première partie d'une géminée, つまる音 *tsumaru oto*, ou 促音 *sokuon*)[141]. Ces deux phonèmes n'occupent jamais la position d'attaque, et ils sont les seuls à occuper la coda. /N/ se réalise phonétiquement comme un phone nasal avec le lieu d'articulation et le degré de constriction qui résultent du passage de la voyelle précédente au segment suivant (il est uvulaire [ɴ] en fin d'énoncé). C'est ainsi que l'on trouve des transcriptions diverses selon les phonéticiens. La réalisation phonétique du /Q/ varie également selon le phonème suivant : occlusive non relâchée du lieu d'articulation du phonème suivant s'il est suivi d'une occlusive non-voisée (ex. 作家 *sakka* « écrivain » /saQka/ [sakˀka] [sakːa])[142] ; fricative du lieu d'articulation du phonème suivant s'il est suivi d'une fricative (ex. 喫茶店 *kissaten* « café (établissement) » /kiQsateN/ [kissateɴ] [kisːateɴ] ; /h/ et /r/ sont rares et limités aux mots mimétiques et emprunts récents).

Il existe également une série de consonnes palatalisées (拗音 *you-on*), communément interprétées comme des séquences d'une consonne suivie de /j/ (ex. 京都 *kyouto* « Kyoto » /ˈkjoRto/ [kʲoːto], et non [kioto]). Les consonnes palatalisées n'apparaissent jamais avant /i/, et peuvent apparaître avant /e/ dans des emprunts récents (シェフ *shefu* « chef cuisinier » /ˈsjehu/ [ɕeɸɯ/ɕeɸɯ̟]) ou dans des formes dialectales de certaines régions (ex. 九州 Kyûshû). Elles sont observées majoritairement dans le vocabulaire sino-japonais (pour les détails, voir Labrune 2006 : 105 ; voir *op. cit.* p. 27-36 pour les différentes strates lexicales du japonais).

[141] Il existe également le phonème moraïque non consonantique /R/ : la seconde partie d'une voyelle longue. Certains linguistes ajoutent un autre phonème moraïque non consonantique /J/ : la seconde partie d'une diphtongue se terminant par [i] (ex. 鯛 *tai* « daurade » /taJ/).

[142] Les occlusives voisées géminées sont possibles uniquement dans des mots d'emprunts récents : バッグ *baggu* « sac » /ˈbaQgu/ < *bag*, [gː] étant produit rarement comme occlusive entièrement voisée.

**Tableau 3** : Phonèmes consonantiques du japonais (Labrune 2006 : 72). Entre { } sont représentées les consonnes qui n'apparaissent que dans les emprunts récents, et entre ( ) les consonnes absentes selon certains phonologues. À cette liste s'ajoutent les deux consonnes moraïques /N/ (la nasale moraïque) et /Q/ (la première moitié d'une géminée). Le symbole /y/ représente la semi-consonne palatale (/j/ en API).

Tableau 2. Consonnes du japonais.

| | labiales | alvéolaires | palatales | vélaires | glottale |
|---|---|---|---|---|---|
| occlusives | p b | t d | | k g | |
| fricatives | {f v} | s z | (ɕ, ʑ) | | h |
| affriquées | | {ts} | (tɕ) | | |
| nasales | m | n | | (ŋ) | |
| glides | | | y | w | |
| liquide | | ɾ | | | |

Les phonèmes /s/ /t/ /z/ /d/ /g/ /h/ présentent des allophones transcrits communément avec des symboles différents. /s/ se réalise palatalisé [ɕ] (transcrit souvent [ʃ] dans une transcription large) devant les (semi-)voyelles palatales /i/ et /j/ (Tableau 4, gauche ; ex. 島 *shima* « île » /si'ma/ [ɕima]). /t/ subit non seulement la palatalisation devant les (semi-)voyelles palatales, mais aussi l'affrication devant les voyelles fermées : [t͡ɕ] (transcrit souvent [t͡ʃ] dans une transcription large) devant les (semi-)voyelles palatales /i/ et /j/, et [t͡s] devant /u/ (Tableau 4, droite ; ex. 地下鉄 *chikatetsu* « métro (chemin de fer sous-terrain) » /tikatetu/ [t͡ɕi̥katet͡sɯ̥ / t͡ɕi̥katet͡sɯ̥]).

**Tableau 4** : Allophones de /s/ (gauche) et de /t/ (droite).

/s/ -> | [s] / _ e, a, o, u
[ɕ] / _ i, j

/t/ -> | [t] / _ e, a, o
[t͡ɕ] / _ i, j
[t͡s] / _ u

De la même manière, /z/ se réalise palatalisé [ʑ] devant les (semi-)voyelles palatales /i/ et /j/, mais peut se réaliser également comme affriquée ([d͡z], et palatalisée [d͡ʑ] : Tableau 5, gauche. Ex. 時間 *jikan* « temps, heure » /zikaN/ [ʑikaɴ / d͡ʑikaɴ]). Les variantes affriquées sont typiquement observées en position initiale de mot et après /N/, les fricatives, à l'intervocalique. La palatalisation et l'affrication sont observées pour /d/ comme pour /t/. En revanche, l'opposition entre /d/ et /z/ se trouve neutralisée devant les voyelles fermées /i/ et /u/, les deux phonèmes se réalisant [ʑ] [d͡ʑ] devant /i/, et [z] [d͡z] devant /u/ (Tableau 5, droite).

**Tableau 5 :** Allophones de /z/ et de /d/.

| /z/ -> | | /d/ -> | |
|---|---|---|---|
| | [z] [d͡z] / _ e, a, o, u | | [d] / _ e, a, o |
| | [ʑ] [d͡ʑ] / _ i, j | | [ʑ] [d͡ʑ] / _ i, j |
| | | | [z] [d͡z] / _ u |

/h/ présente trois allophones en fonction de la (semi-)voyelle suivante : [h] avant /e a o/, [ç] avant /i j/, [ɸ] avant /u/ (Tableau 6 : ex. 花 *hana* « fleur » /ha'na/ [hana], 東 higashi « Est » /higasi/ [çiŋaɕi], 富士山 *Fuji-san* « Mont Fuji » /'huzisaN/ [ɸɯʑisaɴ]). Nous pouvons également décrire la réalisation phonétique de cette consonne comme un vocoïde produit sans voisement du lieu d'articulation de la voyelle suivante (/hi he ha ho hu/ [i̥i e̥e ḁa o̥o ɯ̥ɯ]), ce qui nous permettra de transcrire des cas de dévoisement [V̥ː] : ex. ふたり／二人 *futari* « deux personnes » /huta'ri/ [ɯ̥ːtaɾi].

**Tableau 6 :** Allophones de /h/.

| /h/ -> | |
|---|---|
| | [h] / _ e, a, o |
| | [ç] / _ i, j |
| | [ɸ] / _ u |

Le phonème /g/ se réalise en général comme occlusive en position initiale de mot, mais varie à l'intérieur de mot : nasale [ŋ], occlusive [g], ou fricative [ɣ]. Les locuteurs qui maintiennent la prononciation traditionnelle de la région de Tokyo produisent la nasale [ŋ][143], mais celle-ci est moins largement observée chez les jeunes locuteurs, au profit de la fricative ou de l'occlusive. Certains phonologues considèrent que la nasale /ŋ/ constitue un phonème à part, ce qui permettrait de distinguer 大ガラス *oogarasu* « grande vitre » /oRgarasu/ (< /oR/ + /garasu/) de 大烏 *oogarasu* « grand corbeau » /oRŋarasu/ (< /oR/ + /karasu/ : /k/ > /ŋ/ par 連濁 *rendaku* « voisement de composition »[144]), ou 毒蛾 *dokuga* « phalène venimeux » /dokuga/ de 毒牙 *dokuga* /dokuŋa/ « croc à venin » (Kindaichi 1942[145], Uwano 1993[146], Labrune 2006). Mais les paires minimales de ce type, qui ne sont pas nombreuses, semblent être limitées à des cas où la frontière morphologique est concernée[147]. C'est le phénomène de *rendaku* « voisement de composition »,

---

[143] Cet allophone est considéré comme correct par des organismes comme la NHK (*Nihon Hôsô Kyôkai* : radio-télévision publique) : les présentateurs de la NHK doivent apprendre à s'exprimer avec cet allophone.

[144] Traduit communément comme « *sequential voicing* » en anglais.

[145] H. Kindaichi. "Gagyô bion-ron [Essai sur les nasales de la ligne ga]" (1942), reproduit dans H. Kindaichi, *Nihongo on'in no kenkyuu* [Recherches sur la phonologie du japonais]. Tokyo, Tôkyôdô Shuppan, 168-197, 1972.

[146] Z. Uwano "On no kouzou [Structure des sons de la parole]", in K. Kazama, *Gengogaku* [Linguistique]. Tokyo, Tokyo daigaku shuppankai [Presses de l'Université de Tokyo], 193-249, 1993.

[147] Uwano (1993, *op. cit.*) reprend les exemples suivants de Kindaichi (1942, *op. cit.*) : 銃後 *juugo* [ʑɯːŋo] « arrière » (< 銃 /'zjuR/ « fusil » + 後 /'go/ « derrière »), 十五 *juugo* [ʑɯːgo] « quinze » (< 十 /'zjuR/ « dix » + 五 /'go/ « cinq »), et 十五夜 *juugoya* [ʑɯːŋoja] « nuit de la pleine lune ». Le mot 十五 *juugo*

observé à la frontière morphologique dans des mots composés, qui est en jeu dans la première paire minimale citée ci-dessus. Dans la dernière, 蛾 /ga/ « phalène » dans 毒蛾 *dokuga* « phalène venimeux » est un mot indépendant, tandis que 牙 /ŋa/ dans 毒牙 *dokuga* « croc à venin » n'est qu'un morphème qui apparaît dans des mots composés sino-japonais, ce qui fait que le premier serait susceptible d'être traité comme s'il était en début de mot, et le dernier, à l'intérieur.

**Tableau 7** : Allophones de /g/.

| | | |
|---|---|---|
| /g/ | -> | [g] / #_ |
| | | [ŋ] [g] [ɣ] / V_V |
| | | / N_V |

Certains des allophones présentés ci-dessus sont combinés avec d'autres voyelles dans des mots d'emprunts récents : ex. シェンゲン *Shengen* « Schengen » [ɕeɴgeɴ] ; チェック *chekku* « vérification » [t͡ɕekːɯ / t͡ɕekːɯ̟] (< *check*) ; ツァーリ *tsâri* « tsar » [t͡saːɾi] ; JR *Jêâru* « JR (groupe de compagnies anciennement nationales des chemin de fer japonais) » [dʑeːaːɾɯ / ʑeːaːɾɯ] ; ファースト *fâsuto* « première base (baseball) » [ɸaːsɯ̟to] (< *first*). C'est ainsi que Vance (2008)[148] considère les consonnes [ɕ], [t͡ɕ] (ou [c͡ɕ]), [t͡s], [ʑ / dʑ] (ou [ʒʑ]), et [ɸ] comme constituant des phonèmes à part entière (Tableau 8).

---

« quinze » s'oppose à 十四 *juushi* / *juuyon* « quatorze » ou 十六 *juuroku* « seize », et donc la coupure morphologique est ressentie entre 十 *juu* et 五 *go*, tandis que les mots 銃後 *juugo* [ʑɯːŋo] « arrière » et 十五夜 *juugoya* sont ressentis comme des mots tous faits.

Notons en passant que cet allophone n'est pas observé dans les dialectes du Kansaï (Kyoto, Osaka, Kobe, Nara, …), mais les annonces publiques y sont données en général avec l'accent de Tokyo, qui est considéré comme standard. Nous avons observé un cas d'hypercorrection à la gare JR (compagnie anciennement nationale des chemins de fer) de Kyoto : 45分発 *yonjuugofunhatsu* [jondʑɯːŋoɸɯ̃ɯ̃hat͡sɯ̟] « départ à (… heures) 45 minutes », qui serait normalement prononcé avec un [ɡ], ou éventuellement un [ɣ] par un(e) locuteur / locutrice de Tokyo possédant l'allophone [ŋ], car le même principe que le mot *juugo* « quinze » s'applique à tous les autres numéraux contenant la séquence *juugo* « quinze », et *go* « cinq » comme élément non initial (ex. *hyakugo* « cent cinq »).

[148] T. J. Vance. *The sounds of Japanese*. Cambridge, Cambridge University Press, 2008.

**Tableau 8** : Liste des phonèmes (consonnes) et allophones du japonais de Tokyo selon Vance (2008 : 237).

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| /p/ | [p] | | /d/ | [d] | | /š/ | [ɕ] | | /ǰ/ | [ʄ] | | /y/ | [j] |
| /b/ | [b] | | /f/ | [ɸ] | | /c/ | [ts] | | /m/ | [m] | | /w/ | [ɯ] |
| /t/ | [t] | | /s/ | [s] | | /č/ | [cɕ] | | /r/ | [ɾ] | | | |

| | | | | | |
|---|---|---|---|---|---|
| /k/ | [kʲ] | before /i/ or /y/ | /g/ | [gʲ] | before /i/ or /y/ |
| | [k] | elsewhere | | [g] | elsewhere |
| /h/ | [ç] | before /i/ or /y/ | /n/ | [ɲ] | before /i/ or /y/ |
| | [h] | elsewhere | | [n] | elsewhere |

| | | |
|---|---|---|
| /z/ | [dz] | word-initially or after /N/ or /Q/ |
| | [z] | elsewhere |

**Aspiration et voisement en japonais**

Nous avons constaté ci-dessus que les occlusives du japonais de Tokyo présentent une opposition entre les voisées et les non-voisées. Il a été montré que la réalisation phonétique de la série non-voisée est légèrement aspirée. Riney *et al.* (2007)[149] considèrent que le degré d'aspiration du japonais est intermédiaire entre les aspirées « prototypiques » (ex. anglais : VOT – *Voice Onset Time* : temps d'établissement de voisement – de 60-100 millisecondes, catégorisé comme « *long lag* » par Lisker et Abramson 1964[150]) et les non-aspirées « prototypiques » (ex. espagnol : VOT de 0-25 millisecondes, catégorisé comme « *short lag* » par Lisker et Abramson 1964). Les données de Shimizu (1996)[151] montrent que le VOT moyen des occlusives non-voisées /p t k/ au début de 22 mots monosyllabiques et de 16 onomatopées prononcés 3 fois dans une phrase-cadre (「これは～です」 *Kore wa … desu* « ceci est … ») par 6 locuteurs natifs du japonais (3 femmes et 3 hommes : leurs lieux d'origine ne sont pas mentionnés) était de 41, 30, et 66 millisecondes respectivement. En ce qui concerne la série voisée, les données de Shimizu (1996), recueillies avec la même méthode, montrent que le VOT de /b d g/ était négatif pour toutes les trois consonnes (-89, -75, et -75 millisecondes, respectivement), tandis que Kong et Beckman (2007)[152], et Kamiyama (2007)[153] présentent des occurrences des occlusives voisées sans pré-voisement.

En outre, il a été suggéré que la valeur du VOT varie en fonction de la position dans le mot : elle serait plus importante en position initiale qu'en milieu de mot. Dans Amino, Ji et

---

[149] T. J. Riney, N. Takagi, K. Ôta, Y. Uchida. The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35 : 439–443, 2007.

[150] L. Lisker, A. Abramson. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20 : 384–422, 1964.

[151] K. Shimizu. *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Tokyo, Seibidô Publishing, 1996.

[152] E.-J. Kong, M. Beckman. Fine-grained phonetics and developmental universals for glottal features. Travaux présentés à la Dixième Conférence de Phonologie de Laboratoire, Paris, 29 juin – 1er juillet, 2006.

[153] T. Kamiyama. Les occlusives sourdes et sonores en japonais et en français sont-elles phonétiquement similaires ? Communication orale présentée aux XXIèmes Journées de Linguistique d'Asie Orientale, EHESS, Paris, 28-30 juin 2007.

Shinohara (2007)[154], le VOT était de 32 et 14 millisecondes respectivement pour les deux occurrences de /p/ dans le logatome bisyllabique /papa/ prononcé isolément avec le patron d'accent lexical haut-bas 5 fois par 5 locuteurs natifs (5 hommes) de la variété de Tokyo. Les données de Kamiyama (2007) consistaient en des logatomes trisyllabiques /$C_iV_iC_iV_iC_iV_i$/, avec /p t k b d g/ pour consonne et /a i u/ pour voyelle (ex. /papapa/), prononcés 3 fois dans chacune des deux phrases-cadre différentes (「～と言いました」/… to iRmasita/ « (j'ai) dit … » ; 「それは～です」 /sorewa … desu/ « cela est … ») par 2 locuteurs natifs (1 femme et 1 homme) de la variété de Tokyo. Les résultats montrent que le VOT des occlusives sourdes était systématiquement plus long en position initiale qu'en position médiane (la deuxième et la troisième consonne) de mot pour les logatomes avec /a/ (ex. /papapa/, /tatata/, /kakaka/), tandis que la même tendance n'a pas toujours été observée avec les voyelles /i/ et /u/. Étant donné que l'ouverture de la glotte retarde le début du voisement, la valeur de VOT plus importante observée dans ces études corrobore les résultats de Sawashima et Niimi (1974)[155]. Ces auteurs ont mesuré, en utilisant l'endoscopie, l'ouverture de la glotte pendant la production de mots polysyllabiques qui contenaient [p], [t], [k], [s], et [ts] à l'initiale et au milieu de mot (ex. 計数 /keRsuR/, 撤兵 /teQpeR/, *etc.* : dans certains mots, la consonne médiane était géminée), prononcés par 3 locuteurs de Tokyo, et ont montré que la glotte était plus ouverte pour les occlusives et [ts] initiaux que pour ceux du milieu de mot.

Kamiyama (2007) a également comparé les données des 2 locuteurs natifs japonais (de Tokyo) avec celles de 4 locuteurs natifs francophones. Ces derniers ont montré un VOT bref pour les occlusives sourdes /p t k/ quelle que soit la position dans le mot, à la différence des locuteurs japonophones, en ce qui concerne les logatomes avec /a/, mais une fois de plus, cette tendance n'a pas toujours été observée avec /i/ et /u/.

**/Q/ : *sokuon*, ou *tsumaru oto* et voisement**

Le つまる音 *tsumaru oto* ou 促音 *sokuon*, /Q/ ne peut être suivi, comme nous l'avons vu ci-dessus, que par des obstruentes non-voisées (et rarement par /r/), sauf dans des mots d'emprunt récents, mais même dans ce dernier cas, les occlusives ne sont souvent pas entièrement voisées (Vance 2008 : 109-110)[156]. Cela limite les occurrences des géminées sonores, entièrement voisées. Il est connu que la tenue de l'occlusion et le maintien de la vibration des plis vocaux, les deux éléments qui constituent les occlusives sonores, sont difficilement compatibles sur le plan aérodynamique, étant donné que la vibration des plis vocaux nécessite un flux d'air au niveau glottique, résultant d'une différence entre les pressions sous-glottique et supra-glottique, d'une part, et que la présence d'une occlusion supra-glottique empêche le flux d'air venant de la glotte, d'autre part, et ce à moins d'agrandir la cavité buccale (en gonflant les joues, par exemple, pour l'occlusion labiale). C'est ainsi que l'on trouve plus de langues sans occlusive sonore que celles qui

---

[154] K. Amino, S.-R. Ji, S. Shinohara. Perception of Japanese plosives by Korean speakers. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1789-1792, 2007.

[155] M. Sawashima, S. Niimi. Laryngeal Conditions in Articulations of Japanese Voiceless Consonants. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo* 8 : 13-18, 1974.

[156] T. J. Vance. *The sounds of Japanese*. Cambridge, Cambridge University Press, 2008.

« *Speakers typically pronounce /Q/ plus a voiced obstruent with an extra-long stop that's voiced only at the beginning and not all the way through, […]* »

n'ont pas d'occlusive sourde (Ohala 1983 : 194-195[157]). En outre, plus la tenue de l'occlusion est longue, plus la possibilité est grande que le voisement s'arrête ; nous pouvons ainsi constater une forte tendance au dévoisement des occlusives longues (géminées) (Ohala *op. cit.*). Cette tendance est effectivement observée en japonais de Tokyo.

Il existe cependant des langues qui présentent une opposition systématique entre les consonnes simples et géminées, même pour la série voisée, comme l'italien. Des études acoustiques ont montré que la durée de l'occlusion et celle de la voyelle précédente permettaient de distinguer les géminées des simples, quel que soit le type du segment, ou du débit de la parole (Esposito et Benedetto 1999[158], Pickett *et al*. 1999[159], Payne 2005[160], entre autres). Le salentin, parlé dans la sous-région de Salento (partie méridionale de la région des Pouilles, Italie), présente un cas intéressant. Cette langue ne connaît pas d'opposition entre les occlusives voisées simples et géminées, à la différence des occlusives non-voisées. En revanche, la réalisation phonétique de l'occlusive bilabiale sonore /b/[161] implique une occlusion bilabiale longue, dont la durée est comparable à celle de la géminée correspondante /bb/ de l'italien. En outre, le voisement est entièrement maintenu. À ce phonème s'opposent /p/, qui se réalise [p]~[b]~[ß], et /pp/, qui reste encore à examiner dans une étude expérimentale, mais se réaliserait [pː] (Kamiyama et Gaillard-Corvaglia 2006a[162], 2006b[163], Gaillard-Corvaglia et Kamiyama 2008[164]).

Pour conclure, les géminées en japonais suivent les tendances générales concernant la tenue de l'occlusion et le maintien du voisement, bien que ce ne soit pas le cas dans toutes les langues.

### *1.4.3. Les structures syllabiques*

Les structures syllabiques possibles du japonais de Tokyo peuvent être résumées à la Figure 16 (Vance 2008 : 118). Pour l'attaque, qui est facultative, aucun branchement n'est autorisé sauf pour la semi-consonne /j/ (transcrit /y/ chez Vance), qui constitue avec une autre consonne la série palatalisée (ex. avec /k/, /kj/ comme dans 京都 *kyouto* « Kyoto » /'kjoRto/ [kʲoːto]). Le noyau, qui est un élément obligatoire, doit être constitué d'une voyelle. Celle-ci peut être ou non

---

[157] J. J. Ohala. "The origin of sound patterns in vocal tract constraints", in P. F. MacNeilage, *The production of speech*. New York, Springer-Verlag, 189-216, 1983.

[158] A. Esposito, M. G. Di Benedetto. Acoustic and Perceptual Study of Gemination in Italian Stops. *Journal of the Acoustical Society of America* 106(4 Pt 1) : 2051-2062, 1999.

[159] E. R. Pickett, S. E. Blumstein, M. W. Burton. Effects of Speaking Rate on the Singleton/Geminate Consonant Contrast in Italian. *Phonetica* 56(3-4) : 135-157, 1999.

[160] M. E. Payne. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2) : 153-181, 2005.

[161] Les occurrences du /d/ sont rares, et /g/ est inexistant, ce qui correspond à la difficulté relative du maintien du voisement durant les occlusives : l'occlusion bilabiale laisse plus d'espace entre la constriction buccale et le larynx que l'occlusion dentale, et encore moins pour la vélaire, ce qui permet de maintenir le voisement plus longuement pour la labiale.

[162] T. Kamiyama, A. Gaillard-Corvaglia. Le /b/ "fort" en salentin (sous-région des Pouilles, Italie du Sud) : une étude acoustique et perceptive. *Actes des 8èmes Rencontres Jeunes Chercheurs de l'École Doctorale 268 'Langage et langue'*, Université Paris III, 39-42, 2006.

[163] T. Kamiyama, A. Gaillard-Corvaglia. Le occlusive bilabiali in salentino (Puglia): uno studio acustico e percettivo. *Actes du second Congrès National de l'AISV* (Associazione Italiana di Scienze della Voce), Salerne, 683-694, 2006.

[164] A. Gaillard-Corvaglia, T. Kamiyama. La /b/ "forte" in salentino (Puglia): uno studio acustico, percettivo e fisiologico. *Actes du congrès international "la Comunicazione Parlata 2006"*, Nâples, 2008.

suivie par un des éléments suivants : i) une autre voyelle qui constitue une diphtongue avec la première (/ai/, /ui/, /ae/, /au/, … Voir Labrune 2006 : 67-71 pour les détails) ; ii) le phonème allongeant /R/ (transcrit /H/ chez Vance) ; iii) les phonèmes spéciaux /N/ /Q/. Dans le deuxième cas, l'élément qui suit la voyelle forme le noyau avec la première voyelle, tandis que dans le dernier cas, les phonèmes spéciaux /N/ /Q/ constituent la coda. /N/ et /Q/ sont donc les seules consonnes qui occupent la position de coda, ce qui fait que les syllabes ouvertes sont majoritaires dans la langue. Les syllabes qui se terminent par une voyelle (monophtongue) sans allongement sont des syllabes légères, qui comptent une more, et les autres, se terminant par un des trois éléments décrits ci-dessus, sont des syllabes lourdes, qui comptent deux mores (Figure 17). À ces deux types de syllabes s'ajoutent des syllabes extra-lourdes (3 mores), qui sont beaucoup plus rares (Labrune 2006 : 148) : ex. チェーン *chên* « chaîne » /'tjeRN/ (< *chain*) ; 東京っ子 *toukyoukko* « tokyoïte (personne d'origine de Tokyo) » /toRkjoRQko/.

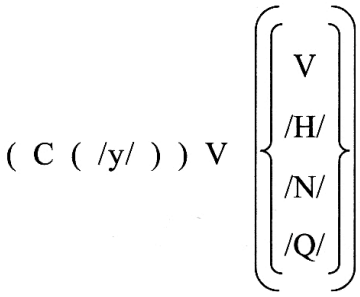



Figure 6-5 Japanese short/long syllable template

**Figure 16 :** Syllabes en japonais (Vance 2008 : 118). Le phonème allongeant /R/ est représenté par /H/, la semi-consonne /j/ par /y/.

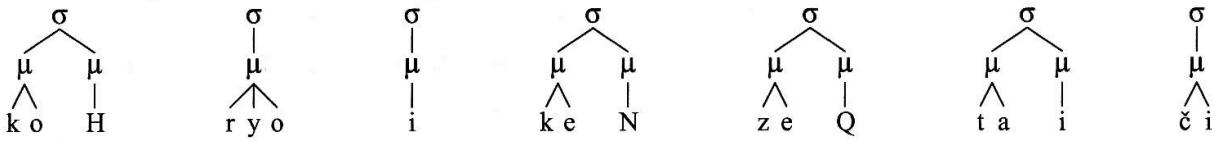



Figure 6-6 Japanese syllable and mora structure

**Figure 17 :** Types de syllabe (des syllabes lourdes qui comptent deux mores, et des syllabes légères qui n'en comptent qu'une) en japonais (Vance 2008 : 118). La première ligne représente des syllabes (σ), la deuxième, la more (*μ*). /H/ correspond au /R/ (phonème allongeant), /y/ au /j/ (semi-consonne), /či/ au /ti/ [tɕi].

La more est donc une unité sub-syllabique, qui est plus petite que la syllabe et plus grande que les segments. En revanche, selon Saitô (1997)[165], la structure syllabique du japonais n'est pas identique à celle des langues comme le français ou l'anglais (dans lesquelles la syllabe est constituée d'une attaque, d'un côté, et d'une rime, de l'autre, qui elle-même se décompose en un noyau et une coda) ou des langues chinoises (la partie initiale 聲母, qui correspond à l'attaque, et la

[165] Y. Saitô. *Nihongo onseigaku nyuumon* [Introduction à la phonétique du japonais]. Tokyo, Sanseidô, 1997.

finale 韻母, qui correspond à la rime : voyelle, consonne finale, ton). Cet auteur explique la différence dans un schéma simplifié (Figure 18), montrant l'unité de more en japonais.

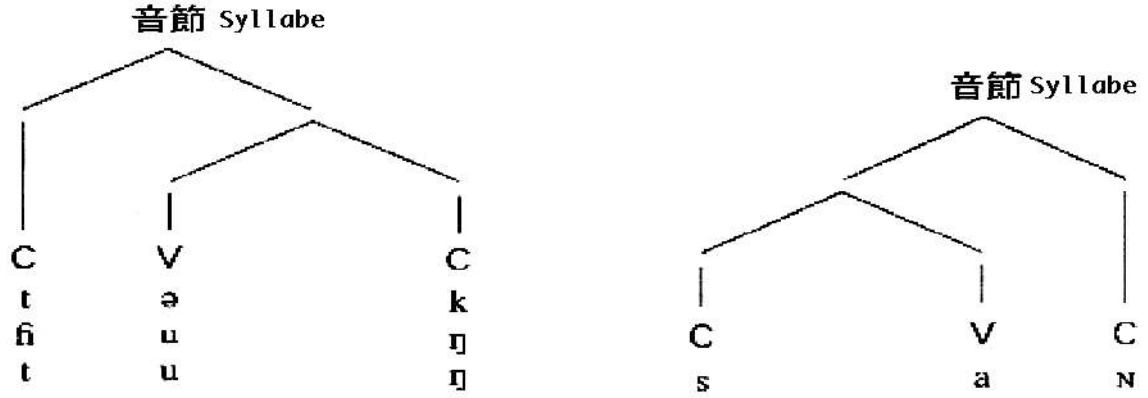



**Figure 18 :** Structure syllabique simplifiée du chinois et de l'anglais (gauche) et celle du japonais et du coréen (droite). Saitô (1997 : 101).

**More, syllabe et pied (*foot*)**

La more (ou 拍 *haku*) et la syllabe (ou 音節 *onsetsu*) sont toutes les deux des unités de base qui nous permettent d'expliquer différents phénomènes en japonais. La more se manifeste dans les cas suivants.

1. Dans la poésie

C'est le nombre de mores, et non de syllabes, qui compte pour former des vers dans la poésie traditionnelle. Le 俳句 *haiku* est constitué de trois vers de cinq, sept et cinq mores chacun, le 短歌 *tanka* de cinq vers de cinq, sept, cinq, sept et sept mores chacun (Figure 19).

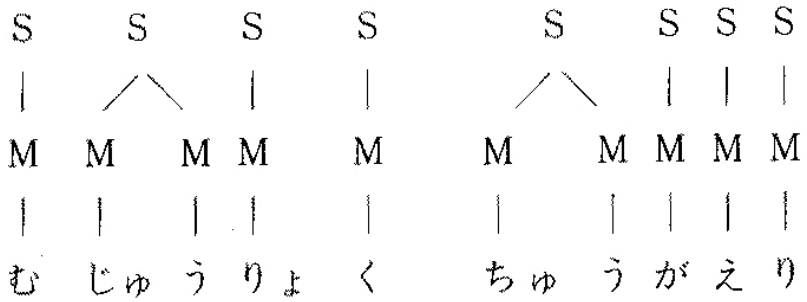



**Figure 19 :** Deux exemples de vers de cinq mores : むじゅうりょく／無重力 /mu 'zju R rjo ku/ [mɯ̟ʑɯːɾʲokɯ̟̥] « sans gravité » (gauche) et ちゅうがえり／宙返り /tju R 'ga e ri/ [t͡ɕɯːŋaeɾi] « saut périlleux » (droite). La première ligne représente les syllabes (S), la deuxième, les mores (M), et la troisième, l'orthographe en syllabaire *hiragana* (Kubozono 1999 : 144[166]).

2. L'écriture

[166] H. Kubozono. *Nihongo no onsei* [Système des sons du japonais]. Iwanami shoten, Tokyo, 1999.

Dans l'écriture en syllabaire かな／仮名 *kana* (*hiragana* et *katakana*), une lettre correspond à une more. Bien que l'origine du système remonte à plus de mille ans, la relation entre la graphie et la structure de mores n'aurait pas radicalement changé, à part l'ajout des lettres composées qui représentent les consonnes palatalisées (じゅ /zju/ et ちゅ /tju/ dans les deux exemples cités à la Figure 19)[167].

3. La description et les règles de l'accent lexical

La more est l'unité qui sert à décrire l'accent lexical qui se caractérise par une descente de la hauteur tonale après une more donnée. Elle nous permet également de rendre compte des règles d'attribution de l'accent des mots d'emprunt ou des néologismes (Shibata 1994[168], Akinaga 2001[169], Labrune 2006 *op. cit.*, entre autres). Elle n'explique cependant pas la totalité des phénomènes concernant l'accent lexical (voir ci-dessous pour le rôle de la syllabe).

4. L'unité d'isochronie rythmique et acoustique

La more est considérée comme unité de base qui constitue le rythme « *mora-timing* »[170]. Bien que Beckman (1982)[171] n'ait trouvé aucune preuve convaincante pour une véritable isochronie de mores sur le plan acoustique, Hoequist (1983)[172], Port *et al.* (1987)[173], et Sagisaka (1999)[174] soutiennent l'hypothèse du « *mora-timing* » même si la durée absolue des mores n'est pas invariable. Sagisaka (1999), par exemple, affirme qu'il existe une corrélation négative entre les durées des segments environnants, et que la corrélation entre la consonne précédente et la voyelle suivante est particulièrement forte. Cette tendance est illustrée à la Figure 20 (Kaiki *et al*. 1992[175], reprise dans Kurematsu 1997[176]). Nous y voyons que la durée des consonnes (ex. les fricatives sont

---

[167] Ces séquences de sons se rencontrent rarement dans les mots Yamato (indigènes). Labrune (2006 : 105).

L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

[168] T. Shibata. "Gairaigo ni okeru akusento-kaku no ichi [Position du noyau d'accent dans les mots d'emprunt]", in K. Satô, *Gendaigo, hougen no kenkyuu* [Recherches sur le japonais contemporain et des dialectes]. Tokyo, Meiji shoin, 388-418, 1994.

[169] K. Akinaga. "Toukyou akusento no shuutoku housoku [Règles pour l'acquisition de l'accent de Tokyo]", in H. Kindaichi, K. Akinaga, *Shin-Meikai nihongo akusento jiten* [Dictionnaire Shin-Meikai de l'accent lexical du japonais]. Tokyo, Sanseidô, pages annexes 10-106, 2001.

[170] Pour l'origine des termes « *stress-timing* », « *syllable-timing* », « *mora-timing* », voir Pike (1945), Ladefoged (1993).

K. L. Pike. *The Intonation of American English*. Ann Arbor, University of Michigan Press, 1945.

P. Ladefoged. *A Course in Phonetics* (troisième édition). Fort Worth, Texas, Harcourt Brace Jovanovich, 1993.

[171] M. Beckman. Segment Duration and the 'Mora' in Japanese. *Phonetica* 39 : 113-135, 1982.

[172] C. J. Hoequist. Syllable Duration in Stress-, Syllable- and Mora-Timed Languages. *Phonetica* 40 : 203-237, 1983.

[173] R. F. Port, J. Dalby, M. O'Dell. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81 : 1574-1585, 1987.

[174] Y. Sagisaka. Nihongo on'in no jikanchou seigyo to chikaku [Contrôle de durée et perception des phonèmes en japonais]. *Gengo* [Langage] 28(9) : 51-56, 1999.

[175] N. Kaiki, K. Takeda, Y. Sagisaka. Gengojouhou o riyou shita boin keizoku jikanchou no seigyo [Contrôle de la durée des voyelles par des informations linguistiques]. *Denshi jouhou tsuushin gakkai ronbun-shi* J-75A(3) : 467-473, 1992.

[176] A. Kurematsu. "Inritsu jouhou no jouhou kagaku eno tekiyou [Application des informations prosodiques aux sciences de l'information]", in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 :*

plus longues que les occlusives) est compensée par celle de la voyelle suivante, et que la durée intrinsèque des voyelles (ex. les voyelles ouvertes sont plus longues que les voyelles fermées) est compensée, ne serait-ce que légèrement, par la durée de la consonne précédente, ce qui contribue à équilibrer la durée de la more. Hoequist (1983) suggère également que le début (*onset*) de l'unité d'isochronie se situerait au « *perceptual center (P-center)* », c'est-à-dire le début de la voyelle.

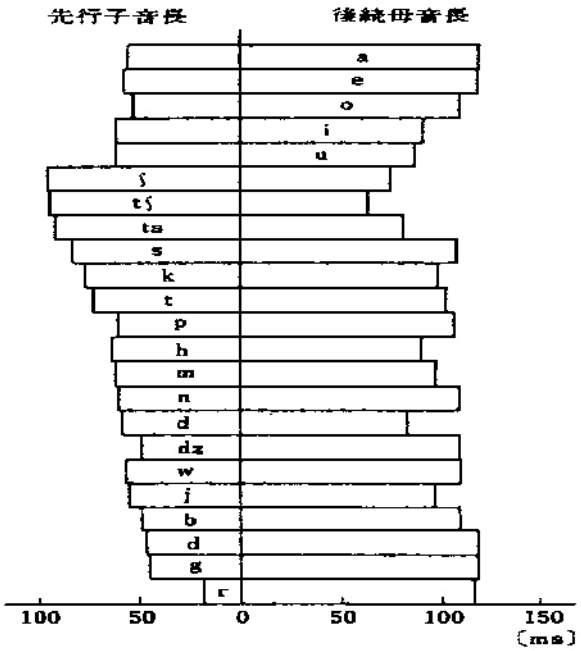



**Figure 20 :** Durée moyenne des mores qui contiennent les 5 voyelles du japonais (les cinq premières lignes : la durée de la consonne précédente est représentée à gauche) et les consonnes (la durée de la voyelle suivante est représentée à droite). Kaiki *et al.* (1992), repris dans Kurematsu (1997).

5. L'unité de perception

La more sert également d'unité de perception chez les locuteurs natifs du japonais. Les résultats d'une série d'expériences de détection de la cible (CV et CVN) dans un stimulus, effectuées par Ôtake *et al*. (1993)[177], suggèrent que les japonophones perçoivent les stimuli sur la base de la more.

6. L'unité psychologique (représentation mentale)

Tsujimura (1996[178] : en prenant les exemples de Kubozono 1985[179]) et Kubozono (1999 : 163-167)[180] analysent les lapsus (*speech errors*) chez les japonophones, et concluent que les erreurs

*akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 303-318, 1997.

[177] T. Ôtake, G. Hatano, A. Cutler, J. Mehler. Mora or Syllable? Speech Segmentation in Japanese. *Journal of Memory and Language* 32(2) : 258-278, 1993.

[178] N. Tsujimura. *An Introduction to Japanese Linguistics*. Oxford, Blackwell, 1996.

[179] H. Kubozono. Speech errors and syllable structure. *Linguistics and Philology* 6 : 220-243, 1985.

[180] H. Kubozono. *Nihongo no onsei* [Système des sons du japonais]. Iwanami shoten, Tokyo, 1999.

(remplacements de segments par exemple) se produisent au niveau de la more[181], ce qui suggère que la more est l'unité psychologique de base chez les locuteurs du japonais. Kubozono (1999) présente également d'autres preuves : le bégaiement se produit le plus souvent au niveau de la more ; quand les Japonais créent des mots valises (混成語、かばん語 : *blend*) à partir de deux entrées, consciemment (formation de néologismes) ou non (erreurs de fusionnement), ou encore, dans des conditions expérimentales, ils ont tendance à segmenter à partir de l'unité CV, ce qui correspond à la more.

Itô et Kagawa (2000)[182] ont mené une étude sur l'acquisition et le développement de l'unité de more chez des enfants japonophones. Dans leur expérience, ils ont demandé aux enfants japonophones de segmenter des mots qu'ils avaient entendus. Les résultats montrent que, chez ceux qui n'avaient jamais appris le système orthographique : 1) la more est une unité phonologique consciente chez les enfants de trois ans dans le cas le plus précoce, mais elle n'est pas aussi bien établie que chez les enfants qui avaient appris l'orthographe ; 2) la syllabe est une unité phonologique consciente plus élémentaire que la more ; 3) à partir de l'âge de trois ans, la more s'établit progressivement comme unité phonologique consciente.

Le terme 音節 *onsetsu* (syllabe) a souvent été utilisé dans la phonologie traditionnelle japonaise pour désigner la notion qui correspond à la more (拍 *haku*). L'exemple suivant illustre cet usage : « 音節 *onsetsu* est également appelé 拍 *haku*, モーラ *môra*, ou モラ *mora* » (Amanuma *et al.* 1978[183] : 165, traduit par nous).

La notion de syllabe telle qu'elle est acceptée en général de nos jours, bien distincte de la more, ne s'était pas vue attribuer d'importance dans la phonologie traditionnelle. Cependant, il a été montré que la syllabe nous permet de mieux rendre compte de certains phénomènes concernant l'accent lexical tels que sa position dans les mots d'emprunt (surtout quand une syllabe lourde est concernée), l'accent lexical des mots composés, *etc*. (Kubozono et Ôta 1998 : chapitre 2[184], Kubozono 1996[185], 1999 : chapitre 7[186], entre autres).

---

[181] Ex. /ketyaQpu/ « ketchup » -> /tyakeQpu/, /iN o humaseru/ « faire des rimes » -> /huN o imaseru/.

[182] T. Itô, A. Kagawa. Moji kakutoku-mae no on'in tan'i eno jikakuteki bunsetsuka no hattatsu [Développement de la segmentation consciente en unités phonologiques avant l'acquisition de la langue écrite]. *Rapport annuel du projet "Kokoro no hattatsu: ninchiteki seichou no kikou* [Développement mental : organisation du développement cognitif]" (année 1999-2000), 141-148, 2000.

[183] Y. Amanuma, K. Ôtsubo, O. Mizutani. *Nihongo Onseigaku* [Phonétique du japonais]. Tokyo, Kuroshio, 1978.

[184] H. Kubozono, S. Ôta. *On'in kouzou to akusento* [Structure phonologique et accent]. Tokyo, Kenkyûsha, 1998.

[185] H. Kubozono. *Nihongo no onsei* [Système des sons du japonais]. Iwanami shoten, Tokyo, 1999.

[186] H. Kubozono. Syllable and Accent in Japanese: Evidence from Loanword Accentuation. *Onseigakkai kaihou / Journal of the Phonetic Society of Japan* 211 : 71-82, 1996.

Il reste une autre unité importante : le pied (フット *foot*) bimoraïque, constitué de deux mores. En ce qui concerne la versification, selon Poser (1990 : 80)[187], c'est Kawakami (1973)[188] qui a proposé une unité qui regroupe deux mores. Nakamichi (1980)[189], quant à lui, propose le terme « タクト *takuto* » (*tact*), qui est constitué de deux mores, comme unité perceptive de reconnaissance de la parole et unité rythmique qui forme le « よい口調 *yoi kuchou* (bon ton, rythme) ». Poser (1990) étend la notion du pied bimoraïque (*bimoraic foot*) pour rendre compte de divers phénomènes morphologiques. Kubozono (1998)[190] présente le rôle du pied comme unité minimale des mots formés (ex. mots abrégés d'emprunt, comme スト /'suto/ (< ストライキ /suto'raiki/ < *strike*), et non *ス */su/) à partir de mots existants (la minimalité). En outre, la forme abrégée des mots d'emprunt est souvent composée de deux pieds bimoraïques (Labrune 2006 : 168, entre autres. Ex. コマタレ *komatare* < コマーシャル *komâsharu* + タレント *tarento* ; *コマータレン **komâtaren*). Le pied bimoraïque permet également d'expliquer la position d'accent « par défaut » du japonais, reflétée sur celle de mots d'emprunt (Shinohara 2000[191]).

### *1.4.4. La prosodie : l'accent lexical et l'intonation*

La prosodie peut être définie de différentes manières, comme le signale Vaissière (2006 : chapitre X)[192]. En première approximation, on peut la définir comme étant la variation de paramètres acoustiques (F0, durée, intensité, timbre des voyelles, qualité de voix) non prévisible à partir des informations segmentales. Les valeurs intrinsèques et co-intrinsèques sont considérées comme prévisibles (Vaissière 2006, *op. cit.*). Nous traiterons dans cette section des variations de la F0.

Pour cela, nous considérons deux composantes : l'accent lexical et l'intonation. Ces deux éléments sont d'ailleurs ceux qui sont utilisés dans le modèle de Fujisaki (Fujisaki et Hirose 1984[193] : Figure 21) : dans ce modèle de génération de la courbe de F0 en japonais, le résultat est obtenu en superposant les « commandes d'accent » et les « commandes de *phrase* (syntagme) ».

---

[187] W. J. Poser. Evidence for Foot Structure in Japanese. *Language* 66(1) : 78-105, 1990.

[188] S. Kawakami. "Nihongo onsuuritsu no saikentou / Reexamination of the rhythm of Japanese verse", in *Ônishi hakushi kiju kinen onseigaku sekai ronbunshuu / Worldwide anthology of papers in phonetics in honor of Dr. Onishi's seventy-seventh birthday*. Tokyo, Phonetic Society of Japan, 665-671, 1973.

[189] M. Nakamichi. Nihongo no rizumu no tan'i ni tsuite [Sur l'unité de rythme en japonais]. *Onsei, gengo no kenkyuu* [Recherches en sons de la parole et langage], Laboratoire de phonétique de l'Université des langues étrangères de Tokyo 1 : 15-22, 1980.

[190] H. Kubozono. "On'inron [Phonologie]", in *Onsei* [Sons de la parole]. Tokyo, Iwanami shoten, 53-91, 1998.

[191] S. Shinohara. Default Accentuation and Foot Structure in Japanese: Evidence from Japanese Adaptations of French Words. *Journal of East Asian Linguistics* 9 : 55-96, 2000.

[192] J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France, 2006.

[193] H. Fujisaki, K. Hirose. Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustic Society of Japan* 5(4) : 233–242, 1984.

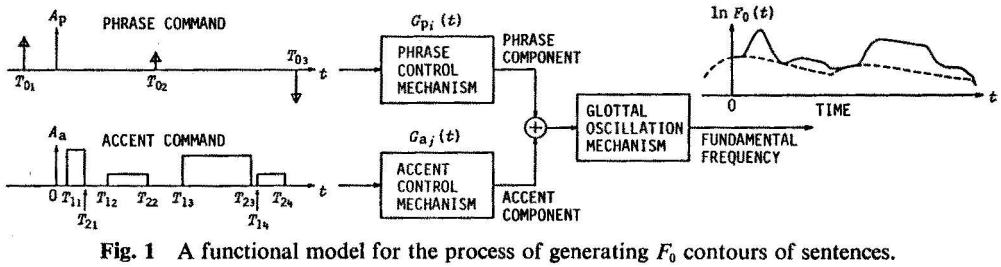



**Fig. 1** A functional model for the process of generating $F_0$ contours of sentences.

**Figure 21** : Modèle de génération de la courbe de F0 de Fujisaki (Fujisaki et Hirose 1984).

Nous commencerons par l'accent lexical (1.4.4.1.), suivi de l'intonation au niveau de syntagme et de phrase (1.4.4.2.).

### 1.3.4.1. L'accent lexical

Le japonais de Tokyo est communément classé comme une langue à accent mélodique. L'accent lexical a deux fonctions en japonais : les fonctions démarcative (il n'existe qu'un seul endroit par mot, ou aucun, où la mélodie descend) et distinctive. Tanaka et Kubozono (1999)[194] indiquent que la première est plus importante que la dernière, étant donné le nombre limité des paires minimales homophones que la différence d'accent permet de distinguer.

La hauteur tonale de chaque more est communément décrite sur une échelle à deux niveaux, haut et bas[195], mais sur le plan phonologique, les patrons d'accent du japonais de Tokyo peuvent être décrits par la position du noyau d'accent (アクセント核 *akusento-kaku*)[196], après lequel la mélodie descend. Cette description est conforme à celle de nombreux chercheurs contemporains, y compris Kawakami (1973 : 7-10)[197], Hattori (1984 : 157)[198], Uwano (1989)[199], Shibatani (1990 : 177-180)[200], Kubozono (1993)[201], Tsujimura (1996 : Section 3.4)[202], Labrune (2006)[203], entre autres.

---

[194] S. Tanaka, H. Kubozono. *Nihongo no hatsuon kyoushitsu* [Classe de prononciation du japonais]. Tokyo, Kuroshio, 1999.

[195] Il existe également un autre point de vue selon lequel trois niveaux de hauteur devraient être considérés. Cette position, d'après Sugitô (1997), a été proposée par SAKUMA Kanae en 1929, et ensuite soutenue par JIMBO Kaku, ce qui a renforcé son influence. À l'époque, un grand nombre de chercheurs ont suivi la pratique, mais la description à deux niveaux est dominante de nos jours.

M. Sugitô. "Nihongo no akusento to intonêshon: toukyou hougen no 'hana' to 'hana' no soui [Accent et intonation en japonais : différence entre '*hana*' (fleur) et '*hana*' (nez) en dialecte de Tokyo]", in M. Sugitô, *"Hana" to "hana"* ["Hana" (fleur) et "hana" (nez)]. Osaka, Izumi shoin : 3-22, 1997.

[196] Le terme « accent » tout court est aussi souvent utilisé, mais cela risque de donner une fausse impression que la more « accentuée » soit proéminente par rapport aux autres. Étant donné que les locuteurs ne sont pas conscients de la position du noyau en général, nous employons ici le terme « noyau d'accent », ou la more « marquée ».

[197] S. Kawakami. *Nihongo akusento-hou* [Méthode de l'accent en japonais]. Tokyo, Gaku shobô shuppan, 1973.

[198] S. Hattori. *Onseigaku* [Phonétique]. Tokyo, Iwanami shoten, 1984.

[199] Z. Uwano. "Nihongo no akusento [Accent en japonais]", in *Kouza nihongo to nihongo-kyouiku 2 : Nihongo no onsei, on'in I* [Série : Langue japonaise et enseignement du japonais 2 : Phonétique et phonologie du japonais, tome I]. Tokyo, Meiji Shoin : 178-205, 1989.

[200] M. Shibatani. *The Languages of Japan*. Cambridge, Cambridge University Press, 1990.

[201] H. Kubozono. *The Organization of Japanese Prosody*. Tokyo, Kuroshio, 1993.

[202] N. Tsujimura. *An Introduction to Japanese Linguistics*. Oxford, Blackwell, 1996.

Toutefois, certains chercheurs favorisent une description simple avec les tons haut et bas, comme Yukawa (1999)[204], notamment pour les dialectes du Kansai (voir ci-dessous). Il est d'ailleurs intéressant de noter qu'il est spécialiste des langues bantoues. Selon Hyman (2009)[205], le japonais serait parmi les langues à tons avec des contraintes, comme d'autres langues classées communément comme « langues à accent mélodique ».

Le noyau peut tomber sur n'importe quelle more (y compris la dernière), sauf sur la deuxième more d'une syllabe lourde. C'est d'ailleurs pour cette raison qu'il est commode de dire que la syllabe, plutôt que la more, est l'unité porteuse du noyau accent. Il existe également des mots sans noyau (« non-marqués », ou « non-accentués »), qui comptent environ la moitié du vocabulaire japonais, toutes catégories confondues (selon Kubozono 1999 : 205)[206]. Les mots non-marqués et ceux qui ont le noyau sur la dernière more ne se distinguent pas en contexte isolé[207], mais la descente due au noyau d'accent des mots marqués sur la dernière more se réalise quand ils sont suivis d'une particule, qui est prononcée avec une hauteur tonale basse.

Parmi les transcriptions du noyau d'accent largement utilisées, nous en présentons deux :

1) le symbole ˺, qui est identique au symbole de l'API (Association Phonétique Internationale) utilisé pour représenter l'absence de relâchement audible des occlusives, placé après les segments de la more marquée (ex. 箸 /ha˺si/ « baguette », 橋 /hasi˺/ « pont »).

2) le symbole ', placé avant la more marquée, comme pour la syllabe qui porte l'accent lexical (stress) en anglais (ex. 箸 /'hasi/ « baguette », 橋 /ha'si/ « pont »).

Dans la présente étude, nous utiliserons le second système.

La grande majorité des locuteurs non linguistes, cependant, ne sont pas conscients de la position précise du noyau d'accent, même s'ils discriminent la même séquence prononcée avec et sans noyau, et avec des positions différentes de noyau d'accent[208]. La position du noyau d'accent ou son absence n'est pas transcrite dans l'orthographe, et ne figure pas non plus dans les dictionnaires généraux (il existe des dictionnaires d'accent). Ce cas de figure contraste avec celui de la position de l'accent lexical (stress) en anglais.

Les locuteurs japonais auraient une meilleure conscience de la différence de patron mélodique entre les mots avec noyau d'accent, d'un côté, et sans, de l'autre, le dernier groupe étant

---

[203] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

[204] Y. Yukawa. *Gengogaku* [Linguistique]. Tokyo, Hituzi syobô, 1999.

[205] L. M. Hyman. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31(2-3) : 213-238, 2009.

[206] H. Kubozono. Nihongo no onsei [Système des sons du japonais]. Iwanami shoten, Tokyo, 1999.

[207] La F0 pourrait être plus élevée par anticipation de la descente virtuelle, qui ne se réalise pas, pour les mots marqués sur la dernière more (voir Sugitô 1997).

M. Sugitô. "Nihongo no akusento to intonêshon: toukyou hougen no 'hana' to 'hana' no soui [Accent et intonation en japonais : différence entre "hana" (fleur) et "hana" (nez) en dialecte de Tokyo]", in M. Sugitô, *"Hana" to "hana"* ["Hana" (fleur) et "hana" (nez)]. Osaka, Izumi shoin : 3-22, 1997.

[208] Ils ressentent en général un accent régional si le noyau d'accent est déplacé par rapport à la forme attendue.

perçu comme plat : c'est d'ailleurs le terme « 平板型 *heiban-gata* » (patron de la « planche plate » ou du plateau) qui est employé dans la tradition japonaise de la phonologie. La Figure 22, qui représente un extrait d'un manga, pourrait être considérée comme une illustration de la représentation mentale de l'accent lexical chez les japonophones. Les segments パンツ /paNtu/ [pantsɯ / pantsɯ̟] signifient « le slip » s'ils sont prononcés avec le noyau sur la première more, « le pantalon » si prononcés sans noyau. Les deux mots sont donc représentés comme /'paNtu/ et /paNtu/ respectivement[209].

[209] Les deux mots sont une sorte de doublets. /'paNtu/ vient de l'usage britannique du mot anglais *pants*, tandis que /paNtu/ vient de l'usage américain du même mot. Ce qui est important ici est la différence de l'époque d'emprunt en plus de la langue d'origine. C'est le premier qui a d'abord été emprunté, et le mot a récemment acquis, parmi les spécialistes (dans le domaine de la mode) et les jeunes en particulier, une nouvelle signification fondée sur l'usage américain (et *pantsu* avec le nouveau sens, sans noyau, a remplacé le mot *zubon* chez certains locuteurs). Il existe quelques paires minimales comme celle-ci, dont l'un des membres a une nouvelle signification chez les jeunes et les spécialistes d'un domaine concerné. En général, le mot avec le sens originel suit la règle d'accentuation des mots d'emprunt (noyau sur la syllabe qui contient la more antépénultième), et l'autre, avec la nouvelle signification, est prononcé sans noyau. Certains affirment que le rôle distinctif de l'accent est ainsi productif, mais d'autres disent que c'est un phénomène transitoire. Le fait que le personnage du manga cité ci-dessous ait mal compris le mot relève de l'usage nouveau de /paNtsu/ sans accent, mais suggère, de l'autre côté, que l'accent n'ait pas véritablement un rôle distinctif. D'ailleurs, l'accent lexical varie d'une région à l'autre, et pourtant, les japonophones de différentes régions peuvent communiquer entre eux.




**Figure 22 :** Exemple de représentation des patrons mélodiques des mots パンツ /'paNtu/ « slip » (HBB, à gauche) et パンツ /paNtu/ « pantalon » (sans accent, BHH ~ HHH, à droite) par un non-linguiste. Notons que l'orthographe est identique pour les deux mots. UEDA Masashi, *Kobo-chan* 7112, journal *Yomiuri shimbun* 45284 du dimanche 28 avril 2002, p. 31.

Quant au patron mélodique du début de mot isolé, les deux première mores sont prononcées avec des hauteurs tonales différentes. Si la première more est marquée, ce sera HB (haut – bas)[210], et dans les autres cas, ce sera BH (bas – haut). Cette montée est ainsi prévisible, et donc considérée en général comme appartenant au niveau de l'intonation de syntagme, plutôt qu'à

[210] Malgré le fait que la première more soit communément décrite H, on observe, au niveau acoustique, une montée qui s'achève en général durant cette more (voir Kubozono 1993 : 91, Figures 6, 9 et 10).

H. Kubozono. *The Organization of Japanese Prosody*. Tokyo, Kuroshio, 1993.

celui de l'accent lexical (Kawakami 1956[211], Uwano 1989[212]). Il peut cependant y avoir une montée légère même à l'intérieur de syntagmes, et c'est en partie pour cette raison que cette montée est considérée par certains chercheurs comme faisant partie de l'accent lexical, qui ne se manifeste pas clairement dans certains environnements phrastiques (Kôri 1997)[213].

La réalisation de cette montée initiale dépend également de la structure syllabique. Si les deux premières mores constituent une syllabe lourde, avec /R/, /N/ ou le deuxième élément d'une diphtongue comme deuxième more, la montée s'effectue plus tôt (Kawakami 1973 : 22[214], Saitô 1997 : 130[215], Pierrehumbert et Beckman 1988 : 127-128[216]) et la montée est plus faible (Kamiyama 2003[217]). Dans le cas de la syllabe lourde avec /Q/, elle est reportée d'une more à droite (Kawakami 1973 : 21-22, Saitô 1997 : 130). Ces phénomènes pourraient être considérés comme relevant du niveau phonétique et physiologique, mais une réalisation précoce ou tardive de la montée peut également jouer un rôle émotionnel ou attitudinal : surprise pour une montée précoce, doute pour une montée tardive (Kawakami 1956, Saitô 1997 : 129-130).

L'accentuation lexicale des mots composés en japonais de Tokyo et d'autres dialectes est un sujet qui a été abordé par de nombreux chercheurs. Les mots composés n'ont qu'un seul noyau d'accent, ou aucun (pour les mots non-marqués), mais sa position ne correspond pas nécessairement à celle des mots d'origine (ex. 雪山 *yukiyama* /jukijama/ « montagne enneigée » < 雪 *yuki* /ju'ki/ « neige » + 山 *yama* /ja'ma/ « montagne » ; 雪国（ゆきぐに） *yukiguni* /ju'kiguni/ « pays enneigé » < 雪 *yuki* /ju'ki/ + 国（くに） *kuni* /kuni/ « pays »). Comme Labrune (2006 : 199)[218] le signale, « le schéma accentuel d'un composé […] semble très généralement prévisible à partir de ses constituants », mais « les facteurs qui le déterminent sont nombreux et divers, et leurs interactions difficiles à saisir ». Cette auteure compte les éléments suivants parmi les paramètres à considérer pour analyser l'accentuation lexicale des mots composés (Labrune 2006 : 200, basé sur Kindaichi et Akinaga 2001[219]) : 1) la nature de la relation morphologique et syntaxique entre les deux constituants ; 2) la taille du composé ; 3) la taille de chacun des constituants ; 4) la catégorie

---

[211] S. Kawakami. Buntou no intonêshon [Intonation du début de phrase]. *Kokugogaku* [Philologie japonaise] 25 : 21-30, 1956.

[212] Z. Uwano. "Nihongo no akusento [Accent en japonais]", in *Kouza nihongo to nihongo-kyouiku 2 : Nihongo no onsei, on'in I* [Série : Langue japonaise et enseignement du japonais 2 : Phonétique et phonologie du japonais, tome I]. Tokyo, Meiji Shoin : 178-205, 1989.

[213] S. Kôri. "Nihongo no intonêshon: kata to kinou [Intonation du japonais: ses patrons et fonctions]" in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 169-202, 1997.

[214] S. Kawakami. *Nihongo akusento-hou* [Méthode de l'accent en japonais]. Tokyo, Gaku shobô shuppan, 1973.

[215] Y. Saitô. *Nihongo onseigaku nyuumon* [Introduction à la phonétique du japonais]. Tokyo, Sanseidô, 1997.

[216] J. B. Pierrehumbert, M. E. Beckman. *Japanese Tone Structure*. Cambridge, Massachusetts ; London, England, The MIT Press, 1988.

[217] T. Kamiyama. Initial pitch in words beginning with a CVV syllable with a long vowel in Tokyo Japanese. *Proceedings of the 16th International Congress of Phonetic Sciences*, 543-546, 2003.

[218] L. Labrune. *La phonologie du japonais*. Leuven ; Paris, Peeters, 2006.

[219] K. Akinaga. "Toukyou akusento no shuutoku housoku [Règles pour l'acquisition de l'accent de Tokyo]", in H. Kindaichi, K. Akinaga, *Shin-Meikai nihongo akusento jiten* [Dictionnaire Shin-Meikai de l'accent lexical du japonais]. Tokyo, Sanseidô, pages annexes 10-106, 2001.

grammaticale des constituants (nom, verbe, *etc.*) ; 5) l'accent intrinsèque de chacun des constituants ; 6) la strate lexicale des constituants.

En revanche, ce phénomène n'est pas observé dans tous les dialectes japonais. Dans le japonais d'Osaka, pour lequel l'accent lexical d'un mot est défini non seulement par la position du noyau mais aussi par le ton du début de mot (haut ou bas), ce dernier élément, plus précisément, l'opposition entre le début bas et le début haut des composants est maintenue dans les mots composés. C'est ainsi que certains mots composés présentent plusieurs montées et descentes dans le dialecte d'Osaka, tandis que l'on n'observe qu'une seule montée (initiale) et une descente éventuelle dans celui de Tokyo (ex. 移動性高気圧 *idousei-koukiatsu* « anticyclone migratoire » : /idoRseRkoRkiatu/ BBBBHBBHBB pour Osaka *vs.* /idoRseRkoR'kiatu/, soit BHHHHHHHBB pour Tokyo. Sugitô 2001 : 201[220]). Sugitô (2001) mentionne ce phénomène pour dire que le japonais d'Osaka présente certaines caractéristiques d'une langue à accent, car le début de chaque mot est prononcé haut ou bas selon son schéma d'accent lexical, quelle que soit la structure grammaticale de la phrase ; en revanche, le japonais de Tokyo est caractérisé plutôt comme une langue à intonation, car son intonation est influencée par la structure grammaticale de la phrase, le ton du début de mot étant non pertinent à la distinction des patrons accentuels. Kamiyama (2004)[221] a testé la perception de ces caractéristiques auprès de 22 auditeurs natifs du japonais de Tokyo, qui ont eu pour tâche de dire si ce qu'ils ont entendu était du japonais de Tokyo ou non. Les stimuli, créés en utilisant Mbrola[222], étaient constitués de 9 paires de phrases délexicalisées[223] des dialectes de Tokyo et d'Osaka, mais gardant la durée segmentale et la courbe de F0 des phrases d'origine. Les résultats suggèrent que les auditeurs japonophones peuvent détecter la prosodie qui ne relève pas du dialecte de Tokyo, du moins pour des phrases relativement longues.

**1.3.4.2. L'intonation au niveau du syntagme et de la phrase**

Comme nous l'avons déjà vu, le japonais est communément catégorisé comme langue à accent mélodique. Les langues à tons, telles que les langues chinoises ou le thaï, utilisent la F0 au niveau lexical pour distinguer des mots qui ont des tons différents. Dans ces langues, la F0 ne joue qu'un rôle relativement peu important (Saitô 1997 : 134)[224]. En cantonais (de Hong Kong), par exemple, plusieurs types d'intonation sont observés (Matthews et Yip 1994 : 28, 318[225] ; Yoshikawa

---

[220] M. Sugitô. "Bunpou to nihongo no akusento oyobi intonêshon: Toukyou to Ôsaka no baai" [Grammaire, et accent et intonation du japonais : les cas de Tokyo et d'Osaka]", in Onsei bunpou kenkyuukai / Spoken Language Working Group, *Bunpou to onsei III / Speech and grammar III*. Tokyo, Kuroshio, 197-210, 2001.

[221] T. Kamiyama. Tokyo and Osaka Japanese: is it possible to distinguish them by prosody alone? *Proceedings of the Workshop MIDL*, Paris, France, 167-172, 2004.

[222] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken. The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of the International Conference on Spoken Language Processing (ICSLP) '96*, 1393-1396, 1996.

[223] Des énoncés composées des mêmes structures syllabiques que les phrases d'origine, mais qui n'ont pas de sens : ex. Osaka /oRsaka/ -> /aRsata/.

[224] Y. Saitô. *Nihongo onseigaku nyuumon* [Introduction à la phonétique du japonais]. Tokyo, Sanseidô, 1997.

[225] S. Matthews, V. Yip. *Cantonese : A Comprehensive Grammar*. London, Routledge, 1994.

2001 : 94[226]), mais ils s'appliquent tous sur la fin de phrase ou de syntagme, ou encore sur une réponse monosyllabique.

En japonais, la F0 est également utilisée au niveau lexical pour réaliser la descente liée au noyau d'accent (ou non-descente pour les mots non-marqués). La position de la descente ou son absence est respectée en principe dans des phrases et des énoncés longs[227]. La phrase interrogative, par exemple, est accompagnée d'une montée sur la dernière more, précédée par la descente due au noyau d'accent s'il y en a un (Kôri 1997 : 191[228], entre autres).

Concernant l'intonation des questions totales (« *yes-no question* »), de nombreuses langues semblent manifester une tendance à l'annulation de la ligne de déclinaison : en général, on observe dans beaucoup de langues une descente globale de F0 au fur et à mesure que la phrase progresse (Maeda 1976[229] pour l'anglais américain, Vaissière 1983[230] pour la tendance générale), mais cette tendance est souvent absente dans les phrases interrogatives (Delattre 1966[231], Léon et Léon 1997[232], Vaissière 2006[233] pour le français, Thorsen[234] pour le danois, Yuan *et al*. 2002[235] pour le chinois mandarin, par exemple). En japonais, cependant, l'annulation de la ligne de déclinaison serait absente. Shôchi *et al.* (à paraître)[236] ont effectué une expérience de perception du type *gating* sur deux phrases prononcées par un locuteur japonais dans 12 attitudes différentes (y compris la déclarative et la question totale sans marque d'interrogation *ka*) : les dix auditeurs japonophones qui y ont participé ont perçu correctement l'interrogation dans moins de 10% des cas seulement quand ils ont entendu les questions totales sans la dernière more. Ces résultats suggèrent que les phrases déclaratives et les questions totales sont perceptivement ressemblantes jusqu'à l'avant-

---

[226] M. Yoshikawa. *Honkon etsugo: hatsuon / An introduction to Hong Kong Cantonese: Pronunciation*. Tokyo, Hakuteisha, 2001.

[227] Sadanobu (2005) mentionne des cas où la descente accentuelle est annulée au profit de certains styles d'intonation (ex. présentateurs décrivant en direct des actions importantes dans une partie sportive).

T. Sadanobu. *Sasayaku koibito, rikimu repôtâ: kuchi no naka no bunka* [Murmures des amoureux, voix pressée des reporters : culture dans la bouche]. Tokyo, Iwanami, 2005.

[228] S. Kôri. "Nihongo no intonêshon: kata to kinou [Intonation du japonais: ses patrons et fonctions]" in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 169-202, 1997.

[229] S. Maeda. *A characterization of American English intonation*. Ph.D. dissertation, Massachusetts Institute of Technology, 1976.

[230] J. Vaissière. "Language Independent Prosodic Features", in A. Cutler, R. Ladd, *Prosody: Models and Measurements*. Berlin, Springer Verlag, 53-65, 1983.

[231] P. Delattre. Les dix intonations de base du français. *The French Review* 40(1) : 1-14, 1966.

[232] M. Léon, P. Léon. La prononciation du français. Paris, Nathan Université, 1997.

[233] J. Vaissière. *La Phonétique*. Paris, Presses Universitaires de France, 2006.

[234] N. G. Thorsen. A study of the perception of sentence intonation – evidence from Danish. *Journal of the Acoustical Society of America 67* : 1014-1030, 1980.

[235] La courbe globale de F0 est plus élevée dans les interrogatives que dans les déclaratives.

J. Yuan, C. Shih, G. P. Kochanski. Comparison of declarative and interrogative intonation in Chinese. *Proceedings of the International Congress: Speech Prosody 2002*, 711-714, 2002.

[236] T. Shôchi, V. Aubergé, A. Rilliard. Nihongo bogo-washa wa hatsuwa no ichibu o kikudakede taido o chikaku dekirunoka? – Gating paradaimu niyoru jikken o tôshite [Peut-on percevoir les attitudes japonaises avant la fin de phrase ? – test perceptif selon le paradigme de "gating"]. *Furansu Nihongo Kyouiku* [Enseignement du Japonais en France] 4, Association des enseignants de japonais en France, 9 pages, à paraître.

dernière more. Kamiyama et Sakamoto (en préparation)[237] ont montré en effet que les apprenants japonophones du français langue étrangère avaient tendance à ne pas réaliser l'annulation de la ligne de déclinaison quand ils lisaient des questions totales en français, et que ce patron mélodique, superposé à des voix françaises (en utilisant Mbrola[238]), ont été souvent perçues comme une continuative plutôt qu'une question.

La continuation n'est marquée ni par allongement final ni par montée de F0 dans un style normal en japonais de Tokyo : l'influence de cette tendance sur la production du français langue étrangère a été étudiée dans Kamiyama (2002[239], 2004[240], 2005[241]). C'est plutôt la montée initiale (et le rehaussement de la ligne de déclinaison) qui marque le début de syntagme et de phrase. Il existe néanmoins des locuteurs, surtout parmi les jeunes, qui utilisent, sur la dernière more de syntagme, un contour montant, mais suivi d'une portion descendante. Ce type d'intonation, appelé « 尻上がりイントネーション *shiriagari intonêshon* »[242], est accompagné également d'un allongement de la voyelle. L'usage de cette intonation est cependant fortement limité[243], et ne peut pas être considéré comme représentatif de la tendance générale de la langue.

Parmi d'autres fonctions de l'intonation, le focus et son influence sur l'accent lexical du mot sous focus et ceux des autres mots ont été le sujet de nombreuses études (Pierrehumbert et Beckman 1988[244], Kubozono 1993[245], entre autres).

L'intonation transmet également diverses attitudes ainsi que la politesse. Ce qui complique la tâche d'apprentissage des langues étrangères et secondes est la présence des cas de correspondance différente entre le patron prosodique et l'attitude transmise dans la langue cible et

---

[237] T. Kamiyama, M. Sakamoto. Intonation of yes-no questions in French: is it really enough to raise pitch at the end? En préparation.

[238] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken. The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of the International Conference on Spoken Language Processing (ICSLP) '96*, 1393-1396, 1996.

[239] T. Kamiyama. *Les aspects suprasegmentaux des phrases déclaratives courtes en français lues par les apprenants japanophones*. Mémoire de DEA (Diplôme d'Études Approfondies) de phonétique, Université de la Sorbonne Nouvelle, 2002.

[240] T. Kamiyama. L1 nihongo gakushuusha no furansugo hatsuon ni okeru tsuyosa to takasa: tanbun yomiage kara mita sono tokuchou [Des caractéristiques sur l'intensité et la F0 dans de courtes phrases en français lues par des apprenants japonophones]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2003 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2003-2004], Graduate School of Arts and Sciences, Université de Tokyo, 69-84, 2004.

[241] T. Kamiyama. L1 nihongo washa no furansugo hatsuon ni okeru purosodjî: sono keikou no shikaku-ka to suuchi-ka no kokoromi [La prosodie de phrases lues en français par des apprenants japonophones : essai de représentation graphique et de quantification de quelques tendances observées]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2004 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2004-2005], Graduate School of Arts and Sciences, Université de Tokyo, 72-79, 2005.

[242] « Montée à la queue ». Cette appellation a commencé à s'employer parmi des non-linguistes, ce qui suggère que la portion montante soit plus importante du point de vue perceptif.

[243] Inoue (1997) traite de l'aspect non seulement phonétique mais aussi social de cette intonation.

F. Inoue. "Intonêshon no shakaisei [Aspect social de l'intonation]", in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 143-168, 1997.

[244] J. B. Pierrehumbert, M. E. Beckman. *Japanese Tone Structure*. Cambridge, Massachusetts ; London, England, The MIT Press, 1988.

[245] H. Kubozono. *The Organization of Japanese Prosody*. Tokyo, Kuroshio, 1993.

la langue source. Shôchi *et al*. (2006)[246] ont montré que certaines attitudes du japonais étaient difficiles à apprendre à percevoir par les apprenants francophones, car certaines sont des « faux amis » prosodiques et attitudinaux. Kamiyama et Yamamoto (2007)[247] ont illustré la difficulté d'apprentissage des expressions de demande en japonais par des apprenants français, et présenté la pratique de classe traitant de cette tâche. Certaines formules de demande se terminent par la particule か *ka*, utilisée également pour former des questions. Comme cette particule est accompagnée d'une intonation montante dans une question, les apprenants ont tendance à appliquer le même patron intonatif aux expressions de la demande. Or, les locuteurs natifs du japonais manifestent en général une montée moins importante que pour les questions, et un patron plat ou légèrement descendant dans « 「～していただけませんでしょうか」… *shite itadakemasendeshouka* » (serait-il possible de vous demander …). Les demandes prononcées par des apprenants avec une montée importante et sans diminution d'intensité risquent d'être prises comme arrogantes par certains auditeurs. Cet article mentionne également l'usage attitudinal de la qualité de voix « りきみ／力み *rikimi* » (voix pressée : voir Sadanobu 2004[248]), qui exprime l'attitude de 恐縮 *kyoushuku* en japonais (selon Sadanobu 2004, « L'émotion / attitude *kyoushuku* correspond très approximativement à un mélange de honte et d'embarras, provenant de la conscience du locuteur du fait que sa demande impose une charge à son interlocuteur »[249]).

Le mouvement de F0 en japonais est caractérisé ainsi par une tendance plutôt descendante due aux noyaux d'accent et de l'absence de montée de continuation, à la différence de la tendance montante du français. Cette caractéristique est observée dès le bas âge. Hallé *et al*. (1991)[250] ont observé des vocalisations bisyllabiques de quatre jeunes enfants japonais de 18 mois environ et quatre enfants français du même âge. Ils ont trouvé plus de contours descendants (F0 de la première syllabe est supérieure à celui de la seconde) que montants chez les jeunes enfants japonais, et la tendance était inverse chez les enfants français. Par ailleurs, les enfants français ont produit la seconde syllabe plus longue que la première dans la plupart du temps, ce qui n'était pas le cas chez les enfants japonais.

### *1.4.5. Convention de transcription*

Nous adoptons le système Hepburn pour la transcription orthographique du japonais dans la présente étude (sur les différents systèmes de transcription orthographique en alphabet latin,

---

[246] T. Shôchi, V. Aubergé, A. Rilliard. Comment les attitudes prosodiques sont parfois de « faux-amis » : les affects sociaux du japonais vs. français. *Actes des XXVIèmes Journées d'études sur la parole*, Dinard, 259-262, 2006.

[247] T. Kamiyama, Y. Yamamoto. Visual representation of prosody for tactful communication skills: the case of request in Japanese as a Foreign Language taught to French university students. *Proceedings of the Phonetics Teaching and Learning Conference 2007*, University College London, Paper ID 37 (4 pages), 2007.

[248] T. Sadanobu. A natural history of Japanese pressed voice. *Journal of the Phonetic Society of Japan* 8(1) : 29-44, 2004.

[249] « *The emotion/attitude kyooshuku, very roughly corresponding to a mixture of suffering ashamedness and embarrassment, comes from the speaker's consciousness of the fact that his/her utterance of request imposes a burden to the hearer* » (Sadanobu 2004 : 34).

[250] P. A. Hallé, B. de Boysson-Bardies, M. M. Vihman. Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* 34 : 299-318, 1991.

voir Labrune 2000[251]). Mais おう <ou> prononcé /oR/ [o:], et えい <ei> prononcé /eR/ [e:] seront notés respectivement <ou> et <ei>, ainsi que les mêmes séquences de voyelles avec une consonne d'attaque (ex. こう <kou> prononcé /koR/ [ko:] sera noté <kou>), à l'exception des emprunts récents écrits en *katakana* (ex. イントネーション *intonêshon*[252]), ainsi que des noms d'auteurs (ex. 杉藤 Sugitô, au lieu de Sugitou) et des maisons d'édition (ex. 三省堂 Sanseidô, au lieu de Sanseidou) par souci d'uniformité avec la transcription courante qui ne marque pas la longueur vocalique (Sugito / Sanseido). Les toponymes japonais plus ou moins fréquemment utilisés en français (Tokyo, Osaka, Yokohama, *etc.*) seront transcrits tels quels, sauf dans les exemples et les titres d'ouvrage.

[251] L. Labrune. Transcrire le japonais. *Daruma Revue d'Études Japonaises* 6/7 (printemps 2000), Philippe Picquier : 340-356, 2000.

[252] Notons que l'allongement /R/ est noté par le caractère qui représente un allongement (ー) dans les mots en *katakana*.

**Tableau 9** : Syllabaires *kana* (*hiragana* à gauche, et *katakana* à droite de chaque case), la transcription phonémique, et la transcription orthographique en alphabet latin (*rômaji*) selon le système Hepburn. /Q/ (première moitié d'une géminée) est représenté en redoublant la consonne suivante, /N/ (nasale moraïque) par <n>. Les voyelles longues (/VR/) sont représentées par un ˆ (ex. <î> pour /iR/).

| あ ア | い イ | う ウ | え エ | お オ | | | | |
|---|---|---|---|---|---|---|---|---|
| /a/ | /i/ | /u/ | /e/ | /o/ | | | | |
| a | i | u | e | o | | | | |
| か カ | き キ | く ク | け ケ | こ コ | | きゃ キャ | きゅ キュ | きょ キョ |
| /ka/ | /ki/ | /ku/ | /ke/ | /ko/ | | /kja/ | /kju/ | /kjo/ |
| ka | ki | ku | ke | ko | | kya | kyu | kyo |
| が ガ | ぎ ギ | ぐ グ | げ ゲ | ご ゴ | | ぎゃ ギャ | ぎゅ ギュ | ぎょ ギョ |
| /ga/ | /gi/ | /gu/ | /ge/ | /go/ | | /gja/ | /gju/ | /gjo/ |
| ga | gi | gu | ge | go | | gya | gyu | gyo |
| さ サ | し シ | す ス | せ セ | そ ソ | | しゃ シャ | しゅ シュ | しょ ショ |
| /sa/ | /si/ | /su/ | /se/ | /so/ | | /sja/ | /sju/ | /sjo/ |
| sa | shi | su | se | so | | sha | shu | sho |
| ざ ザ | じ ジ | ず ズ | ぜ ゼ | ぞ ゾ | | じゃ ジャ | じゅ ジュ | じょ ジョ |
| /za/ | /zi/ | /zu/ | /ze/ | /zo/ | | /zja/ | /zju/ | /zjo/ |
| za | ji | zu | ze | zo | | ja | ju | jo |
| た タ | ち チ | つ ツ | て テ | と ト | | ちゃ チャ | ちゅ チュ | ちょ チョ |
| /ta/ | /ti/ | /tu/ | /te/ | /to/ | | /tja/ | /tju/ | /tjo/ |
| ta | chi | tsu | te | to | | cha | chu | cho |
| だ ダ | ぢ ヂ | づ ヅ | で デ | ど ド | | ぢゃ ヂャ | ぢゅ ヂュ | ぢょ ヂョ |
| /da/ | /zi/ | /zu/ | /de/ | /do/ | | /zja/ | /zju/ | /zjo/ |
| da | ji | zu | de | do | | ja | ju | jo |
| な ナ | に ニ | ぬ ヌ | ね ネ | の ノ | | にゃ ニャ | にゅ ニュ | にょ ニョ |
| /na/ | /ni/ | /nu/ | /ne/ | /no/ | | /nja/ | /nju/ | /njo/ |
| na | ni | nu | ne | no | | nya | nyu | nyo |
| は ハ | ひ ヒ | ふ フ | へ ヘ | ほ ホ | | ひゃ ヒャ | ひゅ ヒュ | ひょ ヒョ |
| /ha/ | /hi/ | /hu/ | /he/ | /ho/ | | /hja/ | /hju/ | /hjo/ |
| ha | hi | fu | he | ho | | hya | hyu | hyo |
| ぱ パ | ぴ ピ | ぷ プ | ぺ ペ | ぽ ポ | | ぴゃ ピャ | ぴゅ ピュ | ぴょ ピョ |
| /pa/ | /pi/ | /pu/ | /pe/ | /po/ | | /pja/ | /pju/ | /pjo/ |
| pa | pi | pu | pe | po | | pya | pyu | pyo |
| ば バ | び ビ | ぶ ブ | べ ベ | ぼ ボ | | びゃ ビャ | びゅ ビュ | びょ ビョ |
| /ba/ | /bi/ | /bu/ | /be/ | /bo/ | | /bja/ | /bju/ | /bjo/ |
| ba | bi | bu | be | bo | | bya | byu | byo |
| ま マ | み ミ | む ム | め メ | も モ | | みゃ ミャ | みゅ ミュ | みょ ミョ |
| /ma/ | /mi/ | /mu/ | /me/ | /mo/ | | /mja/ | /mju/ | /mjo/ |
| ma | mi | mu | me | mo | | mya | myu | myo |
| や ヤ | | ゆ ユ | | よ ヨ | | | | |
| /ja/ | | /ju/ | | /jo/ | | | | |
| ya | | yu | | yo | | | | |
| ら ラ | り リ | る ル | れ レ | ろ ロ | | りゃ リャ | りゅ リュ | りょ リョ |
| /ra/ | /ri/ | /ru/ | /re/ | /ro/ | | /ɾja/ | /ɾju/ | /ɾjo/ |
| ra | ri | ru | re | ro | | rya | ryu | ryo |
| わ ワ | | | | を ヲ | | | | |
| /wa/ | | | | /o/ | | | | |
| wa | | | | o (wo) | | | | |

### 1.5. PHONÉTIQUE ET PHONOLOGIE DU FRANÇAIS PARISIEN

Nous présenterons dans ce qui suit ce que l'on appelle communément le français parisien de nos jours. Au début du XXIème siècle, la variété d'accent qui est probablement la plus fréquemment observée dans la région Île-de-France et qui pourrait être considérée comme proche de celle qui est utilisée dans les média, n'a pas de particularité régionale, et est perçue comme « standard », à la différence de l' « accent parisien », qui était perçu comme un accent particulier de

la région au cours du XXème siècle. Nous définirons ici le terme « français parisien » comme reflétant très approximativement les accents de la moitié nord de la France, qui ne sont marqués ni de l'accent du sud, ni de l'est.

### *1.5.1. Les voyelles : généralité*

En français, comme c'est souvent le cas dans de nombreuses langues du monde, le nombre de consonnes ne varie pas d'une région à l'autre, alors que celui de voyelles est variable (Martinet 1969[253], Vaissière 2006 : 12-13[254]). Pour le français parisien de nos jours, nous pouvons compter 10 voyelles orales et 3 nasales : les voyelles antérieures non arrondies /i e ɛ/, antérieures arrondies /y ø œ/, postérieures[255] /u o ɔ/, basse /a/, et nasales /ɛ̃ ɑ̃ õ/ (Wioland 1991[256], Lauret 2007[257], entre autres).

Nous utiliserons ici les symboles /ɑ̃/ pour la voyelle du mot 'cent', et /õ/ pour la voyelle du mot 'onze'. Ces choix sont conformes à ceux de Lauret (1998)[258] et de Léon (2000)[259], bien que les phonéticiens français de actuels tendent à utiliser /ɔ̃/ au lieu de /õ/ suivant la convention traditionnelle. Considérant les différences entre le français de France et le français canadien, tel qu'il est décrit par Ostiguy *et al*. (1996)[260] avec les symboles /ã ɔ̃/ qui reflètent l'antériorité de la voyelle du mot 'cent' et une ouverture plus grande de la voyelle du mot 'onze' que celle de son équivalent français, ces choix nous semblent justifiés. Les résultats de l'étude de Montagu (2004)[261] corroborent également cette position.

Sur le plan acoustique, Vaissière (2006, *op. cit*., entre autres) distingue 3 séries de voyelles, définies sur une base uniquement acoustique : 1) antérieures étirées /i e ɛ y/, dans lesquelles la distance F1-F2 est supérieure à celle de F2-F3 ; 2) labio-postérieures /u o ɔ/[262], dans lesquelles la distance F1-F2 est inférieure à celle de F2-F3 ; 3) acoustiquement centrales /ø œ a/, dans lesquelles le F2 se situe à mi-chemin entre F1 et F3 (Figure 23).

---

[253] A. Martinet. *Le Français sans Fard*. Paris, Presses Universitaires de France, 1969.

[254] J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France, 2006.

[255] Toutes les trois étant arrondies, les voyelles postérieures du français ne connaissent pas d'opposition phonémique arrondi / non-arrondi.

[256] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

[257] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[258] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français*. Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[259] P. R. Léon. *Phonétisme et prononciations du français* (4ème édition). Paris, Nathan, 2000.

[260] L. Ostiguy, R. Sarrasin, G. Irons. *Introduction à la phonétique comparée : les sons : le français et l'anglais nord-américains*. Sainte-Foy, Les Presses de l'Université Laval, 1996.

[261] J. Montagu. Les sons sous-jacents aux voyelles nasales en français parisien : indices perceptifs des changements. *Actes des XXIVèmes Journées d'Étude sur la Parole*, 385-388, 2004.

[262] L'auteure y rajoute le /ɑ/.

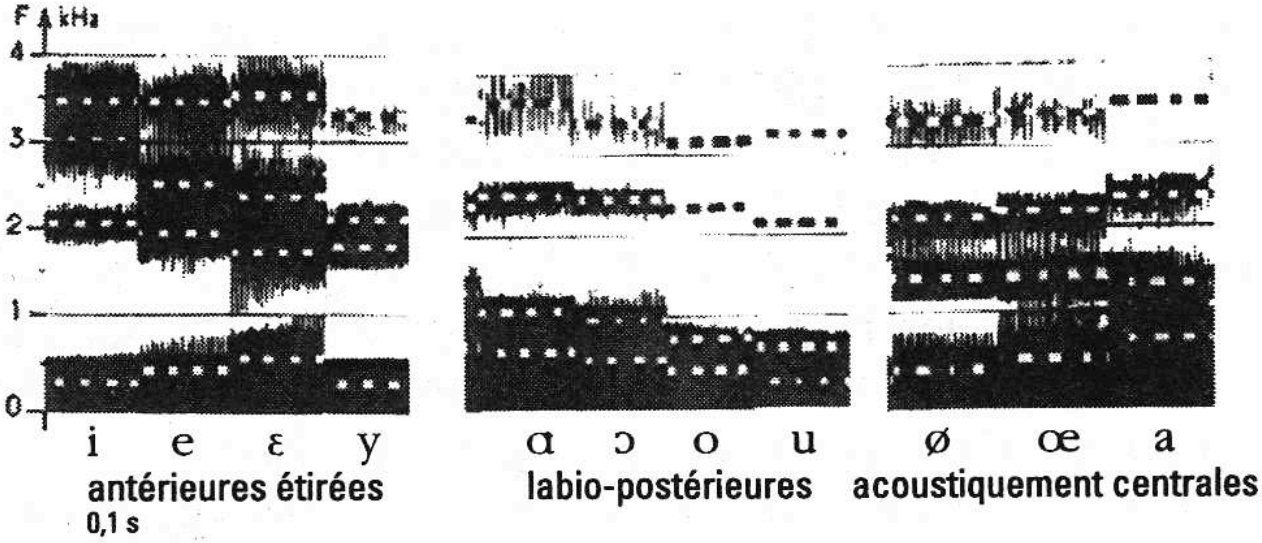



Fig. 10. – Spectrogrammes de voyelles typiques en français (extrait d'un livre de J.-S. Liénard)

**Figure 23 :** Spectrogrammes des voyelles antérieures étirées (gauche), labio-postérieures (milieu), et acoustiquement centrales (droite), extraits de Liénard (1977)[263]. Vaissière (2006 : 72).

Il existe également deux autres phones vocaliques qui étaient autrefois des phonèmes indéniablement distincts, mais plus ou moins neutralisés de nos jours : le /ɑ/ qui s'oppose au /a/, d'une part, et le /œ̃/ qui s'oppose au /ɛ̃/, de l'autre. Comme Lauret (1998 : 21) le signale, l'opposition /a/-/ɑ/ existerait probablement dans la compétence de nombreux locuteurs, se manifestant souvent par une différence de durée (/ɑ/ long et /a/ bref), mais uniquement quand les paires minimales comme pâte /pɑt/ et patte /pat/ sont présentées aux locuteurs de manière explicite. Notons en passant que ces deux oppositions sont maintenues dans certaines autres variétés comme le français canadien (Ostiguy *et al*. 1996).

À cette liste se rajoute le /ə/ (Fougeron et Smith 1999[264], entre autres), qui apparaît uniquement en syllabe inaccentuée (c'est-à-dire, non finale), et est susceptible de disparaître. La distinction entre /ə/ (et /ø œ/), d'un côté, et /e ɛ/, de l'autre, est importante dans l'enseignement et l'apprentissage du français langue étrangère et seconde, notamment du point de vue de la graphie-phonie (Wioland 2001[265], entre autres). Néanmoins, comme nous nous intéressons essentiellement aux voyelles isolées, qui peuvent apparaître dans les syllabes ouvertes finales, nous ne détaillerons pas cette voyelle.

L'opposition entre les voyelles moyennes (mi-fermées et mi-ouvertes) /e ɛ/, /ø œ/, et /o ɔ/ est partiellement neutralisée. Leur distribution est ainsi semi-complémentaire (« loi de

---

[263] J.-S. Liénard. *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole*. Paris, Masson, 1977.

[264] C. Fougeron, C. L. Smith. "Illustrations of the IPA: French", in International Phonetic Association, *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge, Cambridge University Press, 78–81, 1999.

[265] F. Wioland. Que faire de la graphie « e » dans le cadre de la didactique de la prononciation ? *Le français dans le monde* 318 : 31-33, 2001.

position ») : les voyelles mi-fermées tendent à apparaître dans les syllabes ouvertes, et les voyelles mi-ouvertes dans les syllabes fermées. Il existe des paires minimales qui permettent d'opposer /e/-/ɛ/ dans les syllabes ouvertes ('épée' /epe/ *vs*. 'épais' /epɛ/), /o/-/ɔ/ dans les syllabes fermées ('paume' /pom/ *vs*. 'pomme' /pɔm/), /ø/-/œ/ dans les syllabes fermées ('jeûne' /ʒøn/ *vs*. 'jeune' /ʒœn/), mais le /e/ n'apparaît jamais dans les syllabes fermées, et /œ/ et /ɔ/ n'apparaissent jamais dans les syllabes ouvertes. Par ailleurs, le nombre des paires minimales est très limité, sauf pour la paire /e/-/ɛ/ dans des terminaisons verbales (futur /e/ *vs*. conditionnel /ɛ/). C'est ainsi que l'on observe une tendance vers le respect de la « loi de position » : Lauret (1998 : 21) signale une « *certaine confusion entre* /e/ *et* /ɛ/ *en syllabe ouverte* » ; Cécile Fougeron (communication personnelle) a observé la production [ʁɔz] ('rose' : au lieu de la forme attendue [ʁoz]) chez un enfant vivant dans la région parisienne, qui semble respecter la tendance générale de distribution plutôt que la forme définie au niveau lexical.

Quant aux syllabes non-accentuées (non-finales), l'opposition entre les voyelles mi-fermées et mi-ouvertes serait neutralisée, et subirait une harmonie vocalique, c'est-à-dire, l'influence du timbre de la syllabe accentuée (finale) : 'épée' [epe], 'épais' [ɛpɛ].

Notons en passant que cette distribution diffère d'une variété à l'autre. D'un côté, dans beaucoup de variétés du Midi, la distribution des voyelles mi-fermées et mi-ouvertes est parfaitement complémentaire. De l'autre côté, en Franche-Comté, par exemple, toutes les voyelles moyennes (mi-fermées et mi-ouvertes) sont attestées dans les syllabes fermées et ouvertes, bien que /e/ en syllabe fermée (ex. 'vieille' /viej/) et /œ/ en syllabe ouverte (ex. 'peut' /pœ/ : mot régional courant signifiant « vilain ») soient rares. On atteste ainsi des paires minimales telles que 'peau' /po/ *vs*. 'pot' /pɔ/ ; 'un peut' /œ̃pœ/ « vilain » *vs*. 'un peu' /œ̃pø/[266] (Bôle-Richard, communication personnelle)[267]. Le Tableau 10 ci-dessous résume la distribution des voyelles moyennes dans les trois variétés mentionnées.

**Tableau 10** : Distribution des voyelles moyennes en français parisien (standard), en français comptois, et en français méridional. [e] en syllabe fermée et [œ] en syllabe ouvertes sont rares en français comptois. Bôle-Richard (communication personnelle).

| | Français Standard | Français Comtois | Français Méridional |
|---|---|---|---|
| Syllabe ouverte | [e] [ø] [o]<br>[ɛ] | [e] [ø] [o]<br>[ɛ] ([œ]) [ɔ] | [e] [ø] [o] |
| Syllabe fermée | [ø] [o]<br>[ɛ] [œ] [ɔ] | ([e]) [ø] [o]<br>[ɛ] [œ] [ɔ] | [ɛ] [œ] [ɔ] |

[266] Notons que quatres voyelles nasales /ɛ̃ ɑ̃ ɔ̃ œ̃/ sont distinguées dans cette région.

[267] R. Bôle-Richard. Note sur les voyelles du français local de Franche-Comté. 2009.

La même distribution est observée en dialecte (ou « patois ») du Saugeais (Haut-Doubs), par exemple : [eːtr] « être », [feːr] « faire » ; [ɔ] « os », [sabɔ] « sabot », [vɔ] « sabot » (Bôle-Richard 2009).

R. Bôle-Richard. *Les mots du Saugeais : dictionnaire du patois sauget, d'après les travaux de Joseph Bobillier*. Besançon, Cêtre, 2009.

C'est pour les raisons décrites ci-dessus que certains phonéticiens et phonologues postulent pour le français parisien les archiphonèmes /E/, /Œ/, et /O/, qui englobent respectivement /e ɛ/, /ø œ/, et /o ɔ/. Dans l'enseignement de la prononciation du français langue étrangère et seconde, c'est le respect de la « loi de position » qui devrait être travaillé en priorité, avant de perfectionner l'opposition entre les voyelles mi-fermées et mi-ouvertes (Wioland 1991, Lauret 2007).

Les voyelles du français sont adaptées au japonais dans des mots d'emprunt et des noms propres en générale de la manière représentée au Tableau 11 (voir Shinohara 1997[268] pour une étude approfondie).

**Tableau 11** : Adaptation des voyelles françaises au japonais dans des mots d'emprunt et des noms propores.

| Français | Japonais | *Katakana* | Exemple | Français | Japonais | *Katakana* |
|---|---|---|---|---|---|---|
| /i/ | /i/ | イ | Paris | /paʀi/ | /pari/ | パリ |
| /e/ | /e/ | エ | épée | /epe/ | /epe/ | エペ |
| /ɛ/ | /e/ | エ | baguette | /bagɛt/ | /bageQto/ | バゲット |
| /a/ | /a/ | ア | chocolat | /ʃokola/ | /sjokora/ | ショコラ |
| /ɔ/ | /o/ | オ | Languedoc | /lãgdɔk/ | /raNgudoQku/ | ラングドック |
| /o/ | /o/ | オ | Saint-Malo | /sɛ̃malo/ | /saNmaro/ | サン・マロ |
| /u/ | /u/ | ウ | Tours | /tuʀ/ | /t(w)uRru/ | トゥール（ツール） |
| /y/ | /ju/ | ユ | Camus | /kamy/ | /kamju/ | カミュ |
| /ø/ | /u/ | ウ | Dreux | /dʀø/ | /doruR/ | ドルー |
| /œ/ | /u/ | ウ | Sacré Cœur | /sakʀekœʀ/ | /sakurekuRru/ | サクレ・クール |
| /ɛ̃/ | /aN/ | アン | Benjamin | /bɛ̃ʒamɛ̃/ | /baNzjamaN/ | バンジャマン |
| /ã/ | /aN/ | アン | Caen | /kã/ | /kaN/ | カン |
| /õ/ | /oN/ | オン | Besançon | /bəzãsõ/ | /buzaNsoN/ | ブザンソン |

### *1.5.2. Les voyelles : des données acoustiques*

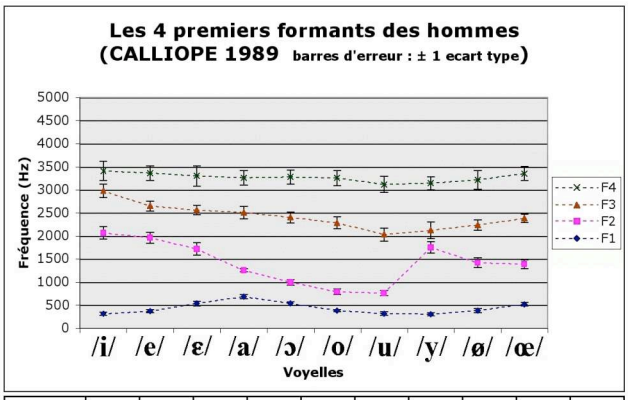
Examinons maintenant des données acoustiques des voyelles. Concernant les valeurs formantiques des voyelles orales du français, nous disposons de celles de CALLIOPE (1989)[269], ainsi que celles de Gendrot et Adda (2004)[270], entre autres.

Les données de CALLIOPE (Figure 24) portent sur 2 répétitions des 10 voyelles orales françaises prononcées par 10 locuteurs masculins et 9 locuteurs féminins. Les voyelles /e o u y ø/ étaient dans le contexte /p_/, tandis que /i ɛ a ɔ œ/ étaient dans le contexte /p_ʀ/. Nous devrions donc considérer, comme les auteurs l'indiquent (p. 84), l'effet de coarticulation du /ʀ/ pour le dernier groupe de voyelles.

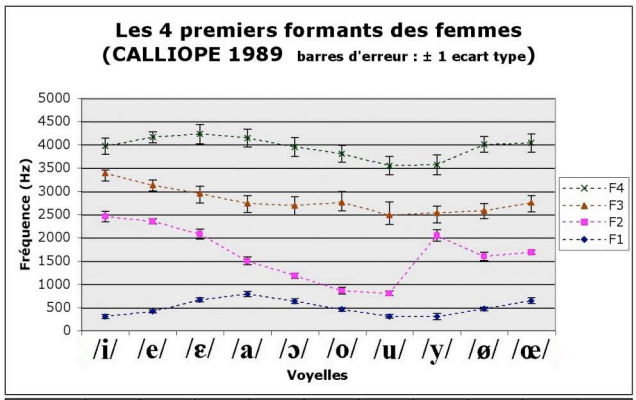
[268] S. Shinohara. *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Thèse de doctorat, Université de la Sorbonne Nouvelle, 1997.

[269] CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

[270] C. Gendrot, M. Adda-Decker. Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande. *Proceedings of the Workshop MIDL*, Paris, France, 7-12, 2004.



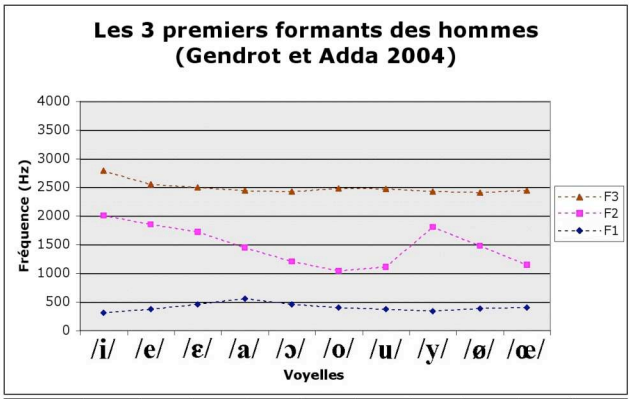

| Hommes | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3407 | 3362 | 3300 | 3262 | 3278 | 3256 | 3118 | 3145 | 3215 | 3353 |
| F3 | 2976 | 2644 | 2558 | 2503 | 2399 | 2283 | 2027 | 2120 | 2235 | 2379 |
| F2 | 2064 | 1961 | 1718 | 1256 | 998 | 793 | 764 | 1750 | 1417 | 1391 |
| F1 | 308 | 365 | 530 | 684 | 531 | 383 | 315 | 300 | 381 | 517 |



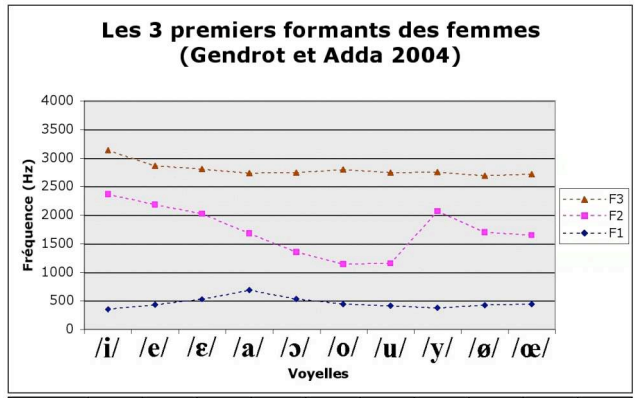

| Femmes | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3966 | 4161 | 4231 | 4143 | 3950 | 3805 | 3550 | 3570 | 4005 | 4038 |
| F3 | 3389 | 3128 | 2954 | 2737 | 2690 | 2756 | 2485 | 2535 | 2581 | 2753 |
| F2 | 2456 | 2351 | 2080 | 1503 | 1180 | 855 | 804 | 2046 | 1605 | 1690 |
| F1 | 306 | 417 | 660 | 788 | 634 | 461 | 311 | 305 | 469 | 647 |

**Figure 24** : Valeurs médianes en Hertz des 4 premiers formants de 2 répétitions des voyelles orales du français chez 10 hommes (à gauche) et 9 femmes (à droite). Contexte : /p_/ pour /e o u y ø/, /p_ʀ/ pour /i ɛ a ɔ œ/ (CALLIOPE 1989 : 84). Les barres d'erreur représentent l'écart-type.

Celles de Gendrot et Adda (Figure 25) ont été recueillies sur une grande base de données, qui consistent d'enregistrements radio- et télédiffusés d'émissions journalistiques. Le corpus, extrait pour la majeure partie d'émissions de France Inter enregistrées en 1998, correspond à 2 heures environ de parole produite par 15 hommes et 15 femmes. Comme les auteurs l'indiquent (p. 7), il s'agit de la parole préparée, et non pas de la parole spontanée, mais l'on attend un effet plus fort de coarticulation que dans les données de CALLIOPE, celles-ci ayant été recueillies dans une condition de laboratoire. Le second formant du /u/, par exemple, est nettement supérieur à celui de CALLIOPE, aussi bien pour les hommes que pour les femmes.




| Hommes | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F3 | 2785 | 2545 | 2491 | 2438 | 2420 | 2477 | 2470 | 2425 | 2405 | 2441 |
| F2 | 2004 | 1849 | 1717 | 1444 | 1203 | 1041 | 1105 | 1803 | 1474 | 1145 |
| F1 | 310 | 370 | 456 | 557 | 456 | 396 | 371 | 336 | 384 | 399 |




| Femmes | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F3 | 3129 | 2861 | 2801 | 2733 | 2743 | 2791 | 2742 | 2746 | 2687 | 2714 |
| F2 | 2365 | 2177 | 2016 | 1677 | 1347 | 1140 | 1153 | 2063 | 1693 | 1643 |
| F1 | 348 | 423 | 526 | 685 | 528 | 438 | 404 | 371 | 419 | 436 |

**Figure 25 :** Valeurs en Hertz des 3 premiers formants des voyelles orales du français dans une base de données qui consistent d'enregistrements radio- et télédiffusés d'émissions journalistiques (Gendrot et Adda 2004). Les auteurs n'ont malheureusement pas mesuré le quatrième formant.

Malgré les différences de corpus décrites ci-dessus, nous observons des caractéristiques communes à ces deux séries de données :

**1) /i/** : F3 est élevé. Le rapprochement de F3 et F4 caractériserait le /i/ français, à la différence de la réalisation typique du /i/ de l'anglais britannique (Vaissière 2007[271], Gendrot *et al*. 2008[272]). C'est pour cette raison qu'il est indispensable de considérer F4 quand on étudie les voyelles du français. Notons également que la valeur de F2 est assez similaire pour /i/ et /e/ dans les données de CALLIOPE (2064 Hz et 1961 Hz respectivement chez les hommes, 2456 Hz et 2351 Hz chez les femmes), mais que ces deux voyelles sont bien distinguées par F3.

**2) /u/** : F1 et F2 sont proches, comme les deux autres voyelles postérieures /o ɔ/, et très bas. Cette tendance est moins nette dans Gendrot et Adda (2004), car tous les contextes consonantiques et tous les débits sont confondus dans ces données. F2 de /u/ (1105 Hz chez les hommes, 1153 Hz chez les femmes) est même plus élevé que celui de /o/ (1041 Hz chez les hommes, 1140 Hz chez les femmes). Notons que la cible F2 n'est souvent pas atteinte dans le contexte dental (ex. doute /dut/), notamment avec un débit élevé (CALLIOPE 1989 : 87-89).

**3) /y/** : F2 et F3 sont proches dans les environs de 1900 Hz pour les hommes, de 2300 Hz pour les femmes. Une fois de plus, le rapprochement est moins net dans Gendrot et Adda (2004), probablement pour la même raison.

**4) /ø/ et /œ/** : F2 se situe ni haut ni bas (1300-1500 Hz pour les hommes, 1600-1700 Hz pour les femmes), à mi-chemin entre F1 et F3.

### *1.5.3. Les consonnes*

Le français parisien compte 16 consonnes : les occlusives /p t k b d g/, les fricatives /f s ʃ v z ʒ/, et les sonantes /m n l ʁ/. À cette liste se rajoutent les 3 semi-consonnes /j ɥ w/, et éventuellement 2 sonantes /ɲ ŋ/ selon certains chercheurs : /ɲ/ pourrait être également interprété comme /nj/ pour les locuteurs qui prononcent la même consonne pour « Agnès » /aɲɛs/ et « nièce » /njɛs/, et dans ce cas-là, il n'est pas nécessaire de considérer /ɲ/ comme phonème à part entière ; l'occurrence du /ŋ/, qui se réalise [ŋg] chez certains locuteurs, est limitée aux mots d'emprunt (ex. « parking » /paʁkiŋ/).

Selon Wioland (1985)[273], la fréquence d'occurrence des phonèmes (consonnes et voyelles) dans le discours est comme indiquée dans le Tableau 12 ci-dessous.

---

[271] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

[272] C. Gendrot, M. Adda-Decker, J. Vaissière. Les voyelles /i/ et /y/ du français : focalisation et variations formantiques. *XXVIIèmes Journées d'Étude sur la Parole*, juin 2008, Avignon, 205-208, 2008.

[273] F. Wioland. *Les structures syllabiques du français : fréquence et distribution des phonèmes consonantiques ; contraintes idiomatiques dans les séquences consonantiques*. Genève, Slatkine / Paris, Champion, 1985.

**Tableau 12** : Fréquence d'occurrence des phonèmes du français à partir d'un échantillon de 200.000 phonèmes dans le français parlé (conversation et théâtre). Wioland (1985 : 14).

C – La fréquence d'occurrence des phonèmes

1. Phonèmes consonantiques d'une part et phonèmes vocaliques d'autre part

| Phonèmes consonantiques | | Phonèmes vocaliques | | | |
|---|---|---|---|---|---|
| 56,55 % | | 43,45 % | | | |
| 1 - /R/ | 7,25 | | | | |
| 2 - /s/ | 6,00 | 1 - /E/ | 10,60 | /ɛ/ | 5,58 |
| 3 - /l/ | 5,63 | | | /e/ | 5,02 |
| 4 - /t/ | 5,335 | 2 - /a/ | 8,55 | | |
| 5 - /k/ | 4,06 | 3 - /i/ | 5,115 | | |
| 6 - /d/ | 4,035 | 4 - /OE/ | 4,31 | graphie e | 3,26 |
| 7 - /m/ | 3,845 | | | graphie eu | 1,05 |
| 8 - /p/ | 3,715 | 5 - /O/ | 3,36 | /o/ | 1,965 |
| 9 - /n/ | 3,095 | | | /ɔ/ | 1,395 |
| 10 - /v/ | 2,755 | 6 - /ɑ̃/ | 3,09 | | |
| 11 - /j/ | 2,00 | 7 - /u/ | 2,425 | | |
| 12 - /ʒ/ | 1,66 | 8 - /ɔ̃/ | 2,255 | | |
| 13 - /z/ | 1,535 | 9 - /y/ | 1,90 | | |
| 14 - /f/ | 1,40 | 10 - /ɛ̃/ | 1,845 | /ɛ̃/ | 1,395 |
| 15 - /w/ | 1,40 | | | /œ̃/ | 0,45 |
| 16 - /b/ | 1,31 | | | | |
| 17 - /ʃ/ | 0,535 | | | | |
| 18 - /ɥ/ | 0,515 | | | | |
| 19 - /g/ | 0,475 | | | | |

L'effet de la coarticulation consonantique est susceptible d'influencer la perception et la production des voyelles par les apprenants, comme l'ont montré Levy et Strange (2008)[274], mais nous ne le détaillerons pas dans notre étude, qui traite essentiellement des voyelles isolées.

### *1.5.4. Les structures syllabiques*

En français, le noyau syllabique est toujours constitué d'une voyelle, le nombre de consonnes d'attaque et de coda variant de zéro à trois au maximum (la troisième est toujours une liquide /l ʀ/), mais le type de syllabe le plus fréquent est CV. Léon (2000 : 95)[275] présente le recensement « *de tous les types de syllabes du français* » effectué par Georges Gougenheim[276], qui montre que « *les plus fréquents sont constitués par la séquence consonne + voyelle (CV)* ».

Delattre (1965)[277] a mené une étude sur la distribution des types de syllabes dans quatre langues (l'anglais, le français, l'allemand et l'espagnol) dans un corpus de 2000 syllabes chacun, 1000 tirées d'un texte narratif, 1000 d'un passage dramatique. Le résultat confirme le recensement de Gougenheim. La structure CV occupe une place majoritaire en français. Sa fréquence est

[274] E. S. Levy, W. Strange. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36 : 141–157, 2008.

[275] P. R. Léon. *Phonétisme et prononciations du français* (4ème édition). Paris, Nathan, 2000.

[276] G. Gougenheim. *Éléments de phonologie française*. Paris, Belles Lettres, 1935.

[277] P. Delattre. "Comparing the Prosodic Features of English, German, Spanish and French", in P. Delattre, *Comparing the Phonetic Features of English, French, German and Spanish: an Interim Report*. Heidelberg, Julius Groos Verlag, 23-43, 1965.

presque le double de celle de l'anglais. CCV aussi est considérablement plus fréquent qu'en anglais. Tout cela fait que les syllabes ouvertes sont dominantes en français, tandis qu'elles ne le sont ni en allemand ni en anglais (Tableau 13).

**Tableau 13** : Distribution (en pourcentage) des quatre types de syllabes les plus fréquents (gauche), et des syllabes fermées et ouvertes (droite). Delattre (1965 : 41-42).

| | CVC | VC | CV | CCV |
|---|---|---|---|---|
| Allemand | 38,1 | 9,8 | 28,7 | 3,3 |
| Anglais | 31,8 | 11,9 | 27,6 | 4,0 |
| Espagnol | 19,8 | 3,1 | 55,6 | 10,2 |
| **Français** | **17,1** | **1,9** | **54,9** | **14,2** |

| | Syllabes fermées | Syllabes ouvertes |
|---|---|---|
| Allemand | 63 | 37 |
| Anglais | 60 | 40 |
| Espagnol | 28 | 72 |
| **Français** | **24** | **76** |

Les données plus récentes de Wioland (1985)[278], qui porte sur un corpus constitué de 20 échantillons de 10.000 phonèmes, soit 200.000 au total, dans des dialogues enregistrés à la télévision de mars à octobre 1973, confirment ces tendances : 80,36% des syllabes étudiées étaient des syllabes ouvertes, et 55,61% avaient la structure CV (Tableau 14, Tableau 15).

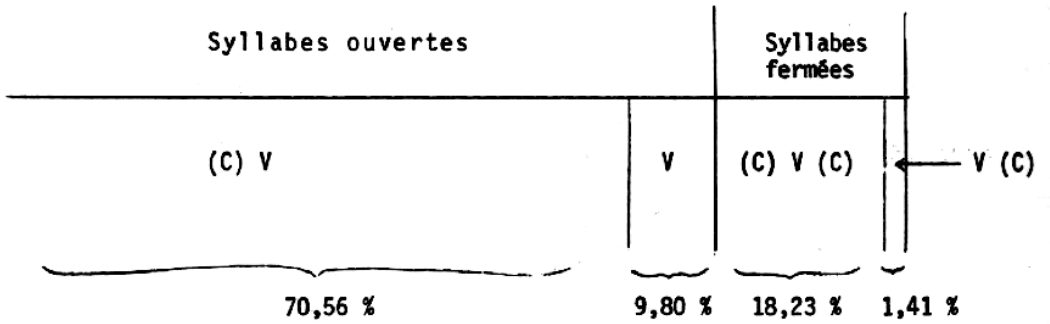
**Tableau 14** : Structures syllabiques observées dans un corpus de dialogues à la télévision (mars – octobre 1973) qui contiennent 200.000 phonèmes (Wioland 1985 : 260).

Les 16 structures syllabiques observées classées par ordre de fréquence décroissante sont les suivantes :

| | Structures syllabiques | Occurrences | % | Total partiel |
|---|---|---|---|---|
| 1 | CV | 48 347 | 55,61 % | |
| 2 | CCV | 12 088 | 13,90 % | 69,51 % |
| 3 | CVC | 11 777 | 13,55 % | 83,06 % |
| 4 | V | 8 519 | 9,80 % | 92,86 % |
| 5 | CCVC | 2 301 | 2,65 % | 95,51 % |
| 6 | CVCC | 1 303 | 1,50 % | 97,01 % |
| 7 | VC | 1 151 | 1,32 % | 98,33 % |
| 8 | CCCV | 880 | 1,01 % | 99,34 % |
| 9 | CCVCC | 253 | 0,29 % | |
| 10 | CCCVC | 155 | 0,18 % | |
| 11 | VCC | 73 | 0,08 % | |
| 12 | CVCCC | 42 | 0,05 % | |
| 13 | CCCCV | 36 | 0,04 % | |
| 14 | CCCVCC | 6 | 0,01 % | |
| 15 | CCCCVC | 2 | 0,002 % | |
| 16 | VCCC | 1 | 0,001 % | |
| | | 86 934 | | |

[278] F. Wioland. *Les structures syllabiques du français : fréquence et distribution des phonèmes consonantiques ; contraintes idiomatiques dans les séquences consonantiques*. Genève, Slatkine / Paris, Champion, 1985.

**Tableau 15** : Structures syllabiques observées dans un corpus de dialogues à la télévision (mars – octobre 1973) qui contiennent 200.000 phonèmes (Wioland 1985 : 264).




Nous pouvons constater ainsi que les syllabes ouvertes, notamment du type CV, dominent en français. Notons que cette tendance est favorisée en partie par l'enchaînement, qui est un processus de réorganisation de la structure syllabique au profil des syllabes ouvertes, et par la liaison, qui permet d'éviter des hiatus, au profil des syllabes du type CV plutôt que V.

Il résulte de cette dominance des syllabes ouvertes en français que le japonais et le français se ressemblent dans leur préférence pour les syllabes ouvertes, à la différence des langues comme l'anglais ou l'allemand.

### *1.5.5. La prosodie*

#### 1.5.5.1. Le rythme

Le français est communément classé comme « *syllable-timed* ». Ce classement relève de l'idée que la syllabe en français, tout comme la more en japonais, est l'unité d'isochronie. Selon l'échelle proposée par Dauer (1983)[279], qui est fondée sur les différences de structures syllabiques, la réduction des voyelles, et la réalisation phonétique du stress (accent lexical) et son influence sur le système linguistique[280], le français se trouve à proximité du japonais (Figure 26).

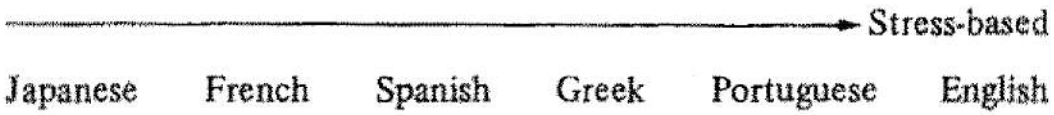



**Figure 26 :** Échelle de degré de « *stress-based* » selon Dauer (1983 : 60).

Cependant, l'existence de l'isochronie dans son sens strict est en général niée dans les études acoustiques. Hoequist (1983)[281] cite l'exemple de Wenk et Wioland (1982)[282] qui ont montré que les durées syllabiques en français varient largement (en particulier à cause du phénomène

[279] R. M. Dauer. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11 : 51-62, 1983.

[280] « *differences in syllable structure, vowel reduction, and the phonetic realization of stress and its influence on the linguistic system* » (Dauer 1983 : 51).

[281] C. J. Hoequist. Syllable Duration in Stress-, Syllable- and Mora-Timed Languages. *Phonetica* 40 : 203-237, 1983.

[282] B. J. Wenk, F. Wioland. Is French really syllable-timed? *Journal of Phonetics* 10 : 193-216, 1982.

d'allongement final) et conclu que les caractéristiques basées sur la durée ont peu de connexion avec les catégories rythmiques fondées sur l'impression perceptive.

À la différence de l'isochronie supposée des syllabes non-finales, il a été observé que le français présente un allongement final au niveau du groupe rythmique. Delattre (1966)[283] distingue les syllabes inaccentuées, désaccentuées (fin de mot non-final de groupe), et accentuées, et affirme que les proportions entre les voyelles de ces positions, devant consonnes simples, sont « très approximativement comme les chiffres 4, 6, 9. » Léon (2000)[284] constate qu'« en français standard une syllabe accentuée est en moyenne deux fois plus longue qu'une syllabe inaccentuée ». O'Shaughnessy (1981)[285] a mené une étude sur un corpus de 285 mots mis dans une phrase-cadre et lus par un locuteur canadien, et conclu qu'« il existe une tendance vers les syllabes d'une durée égale, sauf que les syllabes qui précèdent immédiatement une pause sont significativement plus longues »[286]. Vaissière (1997)[287] montre une hiérarchie d'importance de l'allongement en fonction de la position prosodique. Elle distingue six degrés d'allongement, du plus fort au plus faible (Figure 27) :

> *« 1) l'allongement final en fin de groupe de souffle, marqué par une montée de continuation avant une pause non finale,*
> *2) l'allongement en fin de phrase,*
> *3) l'allongement en fin de syntagme,*
> *4) l'allongement en fin de mot,*
> *5) la durée de référence des syllabes médiales de mot et des syllabes finales non allongées,*
> *6) la durée raccourcie des mots grammaticaux. »*

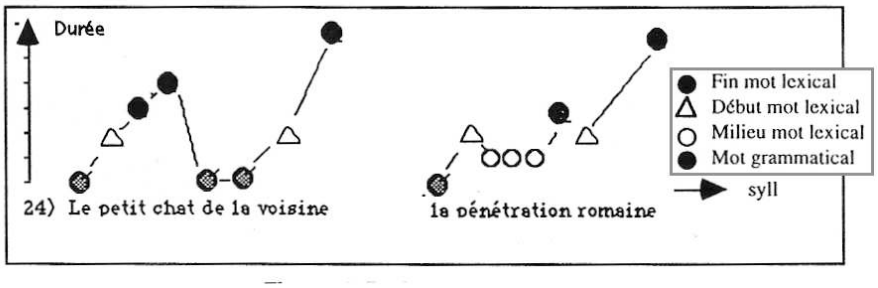



**Figure 27 :** Durée syllabique selon Vaissière (1997 : 69) : 6 niveaux de durée sont représentés.

Nous avons vu dans la section 1.4.5. que les japonophones ne manifestent pas d'allongement final en général, et que cette tendance est retrouvée dans la production des

---

[283] P. Delattre. "Accent de mot et accent de groupe", in P. Delattre, *Studies in French and Comparative Phonetics*. The Hague, Mouton & co., 69-72, 1966.

[284] P. R. Léon. *Phonétisme et prononciations du français* (4ème édition). Paris, Nathan, 2000.

[285] D. O'Shaughnessy. A study of French vowel and consonant durations. *Journal of Phonetics* 9 : 385-406, 1981.

[286] « *there is a tendency toward equal duration syllables, except that syllables immediately prior to a pause are significantly longer*. »

[287] J. Vaissière. Langues, prosodies et syntaxe. *Revue Traitement Automatique des Langues, ATALA* 38(1) : 53-82, 1997.

apprenants japonophones parlant français (Kamiyama 2002[288], 2004[289], 2005[290]). Rappelons également que les francophones présentent un allongement final dès leur plus jeune âge (18 mois dans l'étude de Hallé *et al*. 1991[291]).

Un autre élément du rythme du français est ce que Lauret (2007)[292] appelle la « dynamique syllabique », qu'il définit comme la « répartition de l'énergie dans la syllabe » (Lauret 2007 : 57). Les syllabes sont prononcées de manière lisse et continue en français, tandis que les apprenants qui parlent comme langue native le vietnamien ou le chinois, par exemple, ont tendance à produire un rythme « saccadé » ou « haché » (Lauret, communication personnelle). Wioland (2007)[293] décrit très astucieusement cette particularité du rythme du français : « seul le frottement de l'archet sur les cordes du violon peut évoquer et traduire le rythme du français parlé, et non pas les battements d'un instrument à percussion quel qu'il soit, encore moins ceux d'une batterie ! » (Wioland 2007 : 5).

**1.5.5.2. La f0**

Le français est communément classé comme une langue à frontières (*boundary language* : Vaissière 1991[294]). On observe principalement un allongement de la fin de groupe rythmique (Delattre 1966[295], Léon 2000[296], entre autres), et aucun pattern de F0 est déterminé intrinsèquement au niveau lexical. C'est l'intonation, tout comme en anglais, qui donne le contour mélodique à chaque réalisation du mot, en s'alignant essentiellement sur les début et fin de mot.

Le début de groupe peut être marqué par ce qui est appelé communément « l'accent initial » ou l'accent d'insistance, qui se réalise comme une montée de F0, une augmentation

---

[288] T. Kamiyama. *Les aspects suprasegmentaux des phrases déclaratives courtes en français lues par les apprenants japanophones*. Mémoire de DEA (Diplôme d'Études Approfondies) de phonétique, Université de la Sorbonne Nouvelle, 2002.

[289] T. Kamiyama. L1 nihongo gakushuusha no furansugo hatsuon ni okeru tsuyosa to takasa: tanbun yomiage kara mita sono tokuchou [Des caractéristiques sur l'intensité et la F0 dans de courtes phrases en français lues par des apprenants japonophones]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2003 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2003-2004], Graduate School of Arts and Sciences, Université de Tokyo, 69-84, 2004.

[290] T. Kamiyama. L1 nihongo washa no furansugo hatsuon ni okeru purosodjî: sono keikou no shikaku-ka to suuchi-ka no kokoromi [La prosodie de phrases lues en français par des apprenants japonophones : essai de représentation graphique et de quantification de quelques tendances observées]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2004 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2004-2005], Graduate School of Arts and Sciences, Université de Tokyo, 72-79, 2005.

[291] P. A. Hallé, B. de Boysson-Bardies, M. M. Vihman. Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* 34 : 299-318, 1991.

[292] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[293] F. Wioland. "Préface", in B. Lauret, *Enseigner la prononciation du français : questions et outils*. Hachette, Paris, 2007.

[294] J. Vaissière. "Rhythm, accentuation and final lengthening in French", J. Sundberg, L. Nord, R. Carlson, *Music, Language, Speech and Brain*. London, Macmillan Press, 108-120, 1991.

[295] P. Delattre. "Accent de mot et accent de groupe", in P. Delattre, *Studies in French and Comparative Phonetics*. The Hague, Mouton & co., 69-72, 1966.

[296] P. R. Léon. *Phonétisme et prononciations du français* (4ème édition). Paris, Nathan, 2000.

d'intensité, et une augmentation de la durée de la première consonne. Cet accent permettrait de faciliter la détection du début de mot (Welby 2007[297] pour le rôle de la montée de F0).

La fin de groupes mélodiques ou intonatifs non-finaux est marquée par une montée de F0 qui exprime la continuation. Delattre (1966)[298] distingue deux sortes de continuation, à savoir la continuation majeure et la continuation mineure. Martin (1982)[299] emploie trois paramètres binaires [± *Extreme*], [± *Rising*], et [± *Ample*] pour tenir en compte le degré plus ou moins important de continuation. Selon Vaissière (2002)[300], « toutes choses égales par ailleurs, plus la montée est importante, plus le mot est indépendant »[301] de celui qui le suit. Dans une phrase déclarative typique, une ou plusieurs occurrences de montée de continuation sont suivies par une descente qui marque la finalité, ce qui produit un contraste de pente entre les contours montant(s) et descendant. La Figure 28 illustre le phénomène. Nous y observons également que les mots grammaticaux sont associés à un ton bas[302].

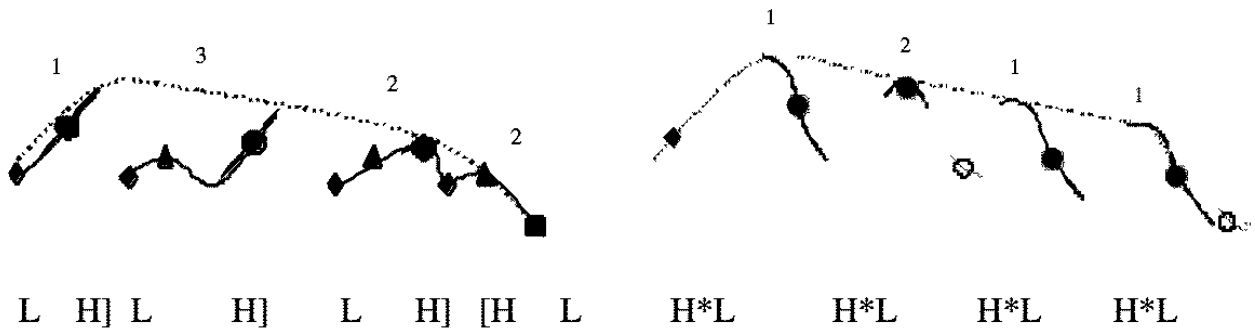



**Figure 1: A typical declarative pattern in French (left) and in English (right). Plain circles correspond to word-final syllables in French and to stressed syllables in English; triangles to lexical word-initial syllables, and lozenges to the function words of French. Unfilled circles correspond to unstressed final English syllables. The numbers indicate the number of syllables in the lexical word.**

**Figure 28** : Contour typique d'une phrase déclarative en français (gauche) et en anglais (droite). Les cercles remplis correspondent aux syllabes finales de mots en français, et aux syllabes accentuées lexicalement (*stressed*) en anglais ; les triangles aux syllabes initiales de mots lexicaux, et les losanges aux mots grammaticaux en français et en anglais. Les cercles non remplis correspondent aux syllabes finales inaccentuées en anglais. Les nombres indiquent le nombre de syllabes dans le mot lexical. Vaissière (2002 : 152).

---

[297] P. Welby. The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication* 49 : 28–48, 2007.

[298] P. Delattre. Les Dix Intonations de base du français. *French Review* 40(1) : 1-14, 1966.

[299] Ph. Martin. Phonetic Realisations of Prosodic Contours in French. *Speech Communication* 1 : 283-294, 1982.

[300] J. Vaissière. "Cross-Linguistic Prosodic Transcription: French versus English", in N. B. Volskaya, N. D. Svetozarova, P. A. Skrelin, *Problems and Methods of experimental Phonetics: in honour of the 70th anniversary of Prof. L. V. Bondarko*. St.-Petersburg, St.-Petersburg State University, 147-164, 2002.

[301] « *all other things being equal, the greater the rise, the more independent the word is* ».

[302] « *French contrasts between lexical words, associated with an H, and function words, associated with an L* » (Vaissière 2002).

La tendance selon laquelle les mots grammaticaux sont prononcés avec une F0 basse, et une durée courte distinguera prosodiquement des paires comme le numéral « deux » et le partitif « du », tels qu'ils sont utilisés dans des phrases comme « nous avons commandé deux chocolats », et « nous avons commandé du chocolat ». Kamiyama (2005)[303] a montré que les tendances prosodiques attendues (« deux » plus long et plus haut que « du ») ont effectivement été observées dans des phrases de ces deux types prononcées par des locuteurs natifs du français, et que la connaissance explicite sur ces différences prosodiques pourrait faciliter les apprenants japonophones à mieux distinguer cette paire de déterminants, pour qui l'opposition /y/-/ø/ est difficile à distinguer perceptivement.

La tendance montante du français, caractérisée par la montée de continuation, s'avère évidente non seulement en comparaison avec le japonais, mais présente aussi un contraste très marqué avec l'anglais. La montée de continuation est dominante en français tandis que la continuation est plutôt descendante (Delattre 1965[304], 1966[305]), ou seulement facultativement montante (Vaissière 2002) en anglais américain.

Il y existe toutefois un phénomène à noter : l'utilisation du contour terminal haut-montant (HRT : *high-rise terminal*) pour la continuation en anglais, qui a été observée d'abord en Australie, puis en Nouvelle-Zélande, l'Amérique du Nord, et finalement en Angleterre (Cruttenden 1997 : 129-131)[306]. Tout comme le contour montant-descendant en japonais, son utilisation est limitée sur le plan social (surtout les jeunes, la classe ouvrière, les femmes), et aussi pragmatique (employé particulièrement en narrative).

La tendance montante du français, et descendante du japonais et de l'anglais semble se manifester très tôt dans la vie, tout comme nous l'avons vu sur l'allongement (voir la section précédente 1.5.5.1). Hallé *et al.* (1991)[307] ont observé plus de contours descendants que montants chez les enfants japonais de 18 mois, et la tendance est inverse chez les enfants français du même âge. Quant à la comparaison entre le français et l'anglais, les résultats de Levitt et Wang (1991)[308] présentent une fréquence plus élevée d'occurrences de l'allongement final et du contour montant de F0 de la syllabe finale chez les enfants français que chez les enfants anglophones américains. Ils ont analysé des formes de babillage consistant en syllabes répétées produites par des enfants français et américains de sept à onze mois d'âge. L'allongement final et la F0 montante sur la

---

[303] T. Kamiyama. Does Explicit Knowledge Help L2 Comprehension? The Case of Determiners "*du*" and "*deux*" in French Learned by Japanese-speaking Learners. *Proceedings of the Phonetics Teaching and Learning Conference 2005*, Department of Phonetics and Linguistics, University College London, July 2005, Paper ID 05 (4 pages), 2005.

[304] P. Delattre. "Comparing the Prosodic Features of English, German, Spanish and French", in P. Delattre, *Comparing the Phonetic Features of English, French, German and Spanish: an Interim Report*. Heidelberg, Julius Groos Verlag, 23-43, 1965.

[305] P. Delattre. "La leçon d'intonation de Simone de Beauvoir : étude d'intonation déclarative comparée", in P. Delattre, *Studies in French and Comparative Phonetics*. The Hague, Mouton & co., 75-82, 1966.

[306] A. Cruttenden. *Intonation* (2nd edition). Cambridge, Cambridge University Press, 1997.

[307] P. A. Hallé, B. de Boysson-Bardies, M. M. Vihman. Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* 34 : 299-318, 1991.

[308] A. Levitt, Q. Wang. Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech* 34 : 235-249, 1991.

syllabe finale ont été observés plus souvent chez les enfants français (54 %) que chez les enfants américains (24 %).

### 1.6. PHONÉTIQUE ET PHONOLOGIE DE L'ANGLAIS AMÉRICAIN (*GENERAL AMERICAN*)

Nous présenterons brièvement le système phonémique et la réalisation phonétique de l'anglais américain (GA : *General American*), en particulier son système vocalique. Le GA correspond très approximativement aux accents qui ne sont marqués ni par l'accent de l'Est (*eastern* : accents non-rhotiques de Boston, de l'est de la Nouvelle Angleterre, et de la ville de New York), ni par celui du sud (*southern*), et qui n'ont pas de caractéristique régionale particulière. Il est également mentionné comme « *Network English* », car cette variété est la plus acceptable sur les chaînes de télévision qui couvrent l'ensemble des États-Unis (Wells 1982 : 470)[309].

Il nous semble indispensable de considérer le système vocalique de cette variété de langue, à laquelle une grande majorité des apprenants japonophones du français sont exposés davantage, et qui peut ainsi être considérée comme la deuxième langue pour eux.

#### *1.6.1. Le système vocalique de l'anglais américain (GA :* **General American***)*

Wells (1982) présente le système vocalique de l'anglais américain GA (Figure 29). Nous utiliserons dans la présente étude le système de transcription de Wells (2000)[310], qui permet de décrire également l'anglais britannique de manière uniforme[311].

---

[309] J. C. Wells. *Accents of English*. Cambridge, Cambridge University Press, 1982.

[310] J. C. Wells. *Longman Pronunciation Dictionary* (Second edition). Harlow, Pearson Education Limited, 2000.

Ce dictionnaire compte, avec Jones (2006), qui a pratiquement la même convention de transcription, parmi les dictionnaires de prononciation de l'anglais les plus utilisés de nos jours. Selon Ladefoged (1993), « *Everyone seriously interested in English pronunciation should be using this dictionary* [tous ceux qui s'intéressent sérieusement à la prononciation de l'anglais devraient utiliser ce dictionnaire] ». L'auteur a rajouté *Cambridge English Pronouncing Dictionary* pour les éditions ultérieures.

D. Jones, P. Roach, J. Hartman, J. Setter. *Cambridge English Pronouncing Dictionary* (17th edition). Cambridge, Cambridge University Press, 2006.

P. Ladefoged. *A Course in Phonetics* (troisième édition). Fort Worth, Texas, Harcourt Brace Jovanovich, 1993.

[311] Pour d'autres systèmes de transcription de l'anglais largement utilisés, voir Ladefoged (1993 : Table 4.1.), entre autres.

P. Ladefoged. *A Course in Phonetics* (troisième édition). Fort Worth, Texas, Harcourt Brace Jovanovich, 1993.

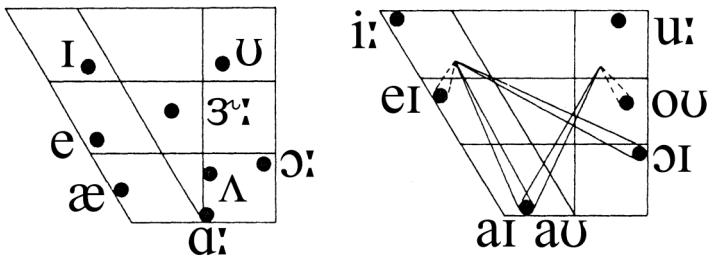



**Figure 29 :** Voyelles de l'anglais américain (figure fondée sur Wells 1982 : 486, avec le système de transcription adopté dans Wells 2000).

Nous y comptons ainsi 15 voyelles : *hit* /hɪt/, *head* /hed/, *had* /hæd/, *hot* /hɑːt/, *hut* /hʌt/, *law* /lɔː/, *hurt* /hɜːt/, *hood* /hʊd/, *heat* /hiːt/, *hate* /heɪt/, *height* /haɪt/, *how* /haʊ/, *boy* /bɔɪ/, *hole* /hoʊl/, *who* /huː/. Cependant, l'opposition entre /ɔː/ (dans *thought, law*) et /ɑː/ (dans *lot, hot*) est perdue chez certains locuteurs (Wells 1982 : 473-476). Un sous-ensemble de ces 15 voyelles, /ɪ e ɑː ɔː ʊ aɪ aʊ/, peuvent être suivies par /r/ et former des séquences vocaliques rhotiques : *hear* /hɪr/, *hare hair* /her/, *heart* /hɑːrt/, *course* /kɔːrs/, *poor* /pʊr/, *hire* /haɪr/, hour /aʊr/. Les voyelles relâchées (*lax vowels*) /ɪ e æ ʌ ʊ/ ne peuvent apparaître en position finale de mot.

Notons que les voyelles de l'anglais ont tendance à se diphtonguer, y compris les voyelles phonémiquement monophtongues. Selon Wells (1982 : 487), les voyelles fermées /iː/ et /uː/, par exemple, sont généralement soit monophtongues soit très légèrement diphtonguées [iˑ ~ i̯i], [uˑ ~ u̯u].

Il est également important de noter que la série antérieure arrondie n'existe pas sur le plan phonémique en anglais américain (ni dans d'autres variétés principales de l'anglais), même si un vocoïde antérieur arrondi peut être observé comme allophone (ex. /uː/ dans un contexte antérieur comme dans « *dude* » /duːd/). Les données de Kataoka (2009)[312] suggèrent que la réalisation antériorisée du /uː/ dans un contexte alvéolaire avec un F2 élevé serait un allophone « phonologisé », qui serait utilisé comme cible articulatoire dans le contexte alvéolaire quel que soit le débit de la parole, et non le résultat de la cible articulatoire non atteinte (*undershoot*).

La réalisation phonétique du /uː/ de l'anglais américain de nos jours n'est d'ailleurs pas complètement postérieure même dans d'autres contextes consonantiques, à la différence du /u/ français (Ostiguy *et al*. 1996[313], entre autres), qui est caractérisé par sa nature quantique (rapprochement des deux premiers formants sur les basses fréquences) sur le plan acoustique, comme nous l'avons déjà vu. Le F2 du début du /uː/ américain est typiquement assez élevé, et descend vers la fin, comme illustré à la Figure 30. Nous y voyons que le F2 du mot *who* /huː/ se trouve séparé du F1, au-dessus de 1000 Hz, au début de la voyelle, et descend progressivement vers la fin de la voyelle, où les deux premiers formants sont situés en dessous de 1000 Hz. Le mot

[312] R. Kataoka. A production study on phonologization of /u/-fronting in alveolar context. *Journal of the Acoustical Society of America* 125(4) : 2570, 2009.

[313] Ces auteurs signalent une « faute courante » de /uː/ chez les francophones (nord-américains) comme suit : « [locuteur francophone] *produit une voyelle trop brève, trop arrondie et trop postérieure lorsqu'elle survient en syllabe ouverte* » (Ostiguy *et al*. 1996 : 82).

L. Ostiguy, R. Sarrasin, G. Irons. *Introduction à la phonétique comparée : les sons : le français et l'anglais nord-américains*. Sainte-Foy, Les Presses de l'Université Laval, 1996.

*dude* /duːd/, dans lequel la voyelle /uː/ est entourée de deux consonnes alvéolaires, présente un F2 beaucoup plus élevé dès le début jusqu'à la fin : il se situe autour de 2000 Hz au début, descend progressivement, mais remonte à la fin sans atteindre 1000 Hz. Notons qu'il s'agit d'une prononciation soignée, faite pour un dictionnaire de prononciation.

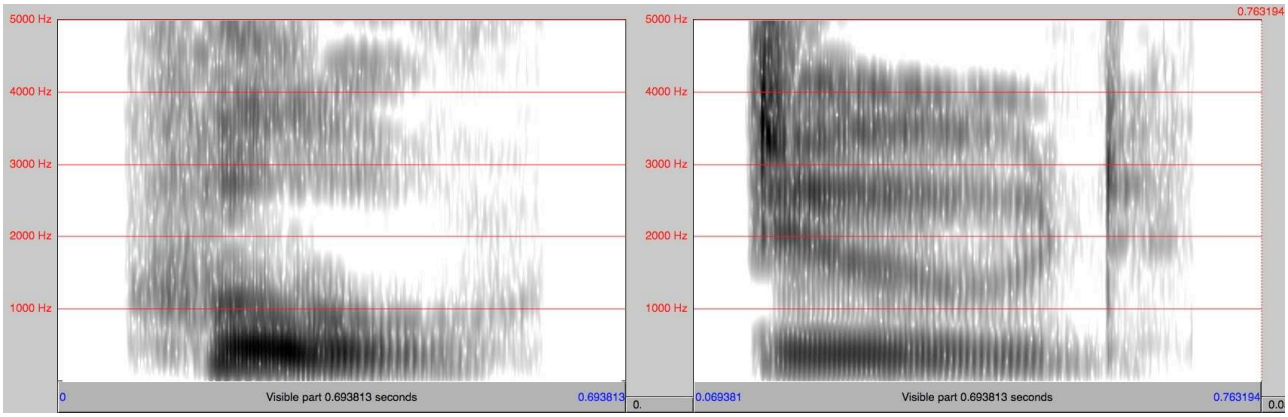



**Figure 30** : Spectrogrammes à bande large (Praat ; longueur de fenêtre : 5 millisecondes) des mots *who* /huː/ (à gauche) et *dude* /duːd/ (à droite) prononcés par des hommes dans une prononciation américaine. Source : Jones *et al*. (2006)[314], CD-ROM.

Les données acoustiques de Hillenbrand *et al*. (1995)[315] confirment cette tendance. Ces auteurs ont enregistré 12 voyelles de l'anglais américain /iː, ɪ, eɪ, e, æ, ɑː, ɔː, oʊ, ʊ, uː, ʌ, ɝː/ prononcées dans des listes randomisées de mots contenant ces voyelles dans le contexte /h_d/ (« *heed* », « *hid* », « *hayed* », « *head* », « *had* », « *hod* », « *hawed* », « *hoed* », « *hood* », « *who'd* », « *hud* », et « *heard* »). Les listes de mots ont été lues par 45 hommes adultes, 48 femmes adultes, et 46 enfants (27 garçons et 19 filles) âgés de 10 à 12 ans, majoritairement (87%) d'origine de l'État de Michigan, et ils maintenaient tous l'opposition /ɔː/-/ɑː/. Les valeurs des deux premiers formants sont représentées à la Figure 31.

Nous pouvons y observer que la moyenne de F2 du /uː/ des hommes se trouve autour de 1000 Hz (997 Hz), même si l'on trouve de nombreuses occurrences de cette voyelle où le F2 est plus bas. Notons également que les hommes et les femmes montrent des valeurs différentes pour cette voyelle : 378 Hz et 459 Hz respectivement pour F1, 997 Hz et 1105 Hz pour F2. Par contraste, le /u/ français, qui se réalise typiquement avec deux constrictions très fortes, labiale et vélaire, qui donnent lieu à une forme de conduit vocal modélisée avec deux cavités de la résonance de Helmholtz, a ses deux premiers formants semblables pour les hommes et les femmes. Les données de CALLIOPE, obtenues dans le contexte /p_/ pour /u/, présentent effectivement des zones de dispersion F1-F2 de /u/ très similaires pour les hommes et les femmes, avec des valeurs moyennes semblables : 315 Hz et 764 Hz chez les hommes, 311 Hz et 804 Hz chez les femmes (CALLIOPE 1989 : 84-85 ; le graphique des zones de dispersion reproduit à la Figure 67, p. 185, les valeurs

[314] D. Jones, P. Roach, J. Hartman, J. Setter. *Cambridge English Pronouncing Dictionary* (17th edition). Cambridge, Cambridge University Press, 2006.

[315] J. Hillenbrand, L. A. Getty, M. J. Clark, K. Wheeler. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5) : 3099-3111, 1995.

moyennes représentées à la Figure 24, p. 81)[316]. La présence ou l'absence de différence entre les deux premiers formants des hommes et des femmes peut donc être interprétée comme une des preuves de la différence articulatoire entre le /uː/ de l'anglais américain et le /u/ français.

Par ailleurs, le petit nombre d'occurrences du /uː/ américain où les deux premiers formants étaient trop proches pour être mesurés séparément montre également la nature non quantique de cette voyelle : 1,3% seulement pour /uː/, contre 10,0% pour F2 et F3 de /iː/, et 15,3% pour F2 et F3 de /ɜː/[317] (Hillenbrand *et al*. 1995 : 3101). Par contraste, en ce qui concerne les voyelles orales du français, c'est avec les deux premiers formants de /u/ que l'on rencontre très souvent cette difficulté technique.

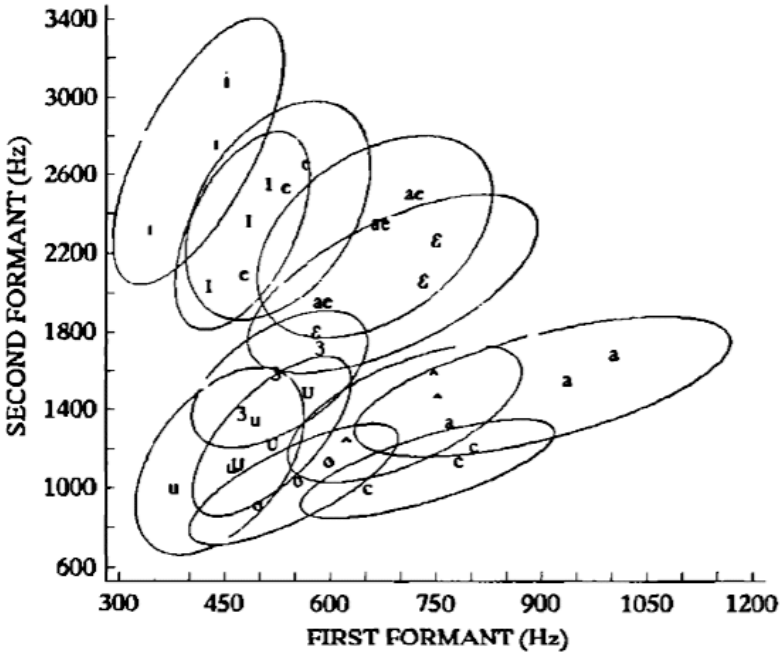



FIG. 3. Average values of *F*1 and *F*2 for men, women, and child talkers for 12 vowels with ellipses fit to the data ("ae"=/æ/, "a"=/ɑ/, "c"=/ɔ/, "ʌ"=/ʌ/, "ɜ"=/ɜː/).

[316] CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

[317] Notons que le F2 du /iː/ de l'anglais américain est souvent très élevé et proche de F3 (à la différence du /i/ français avec F3/F4 proches), et que le /ɜː/ est caractérisé par sa nature quantique avec un F3 extra-bas, en dessous de 2000 Hz et proche de F2.

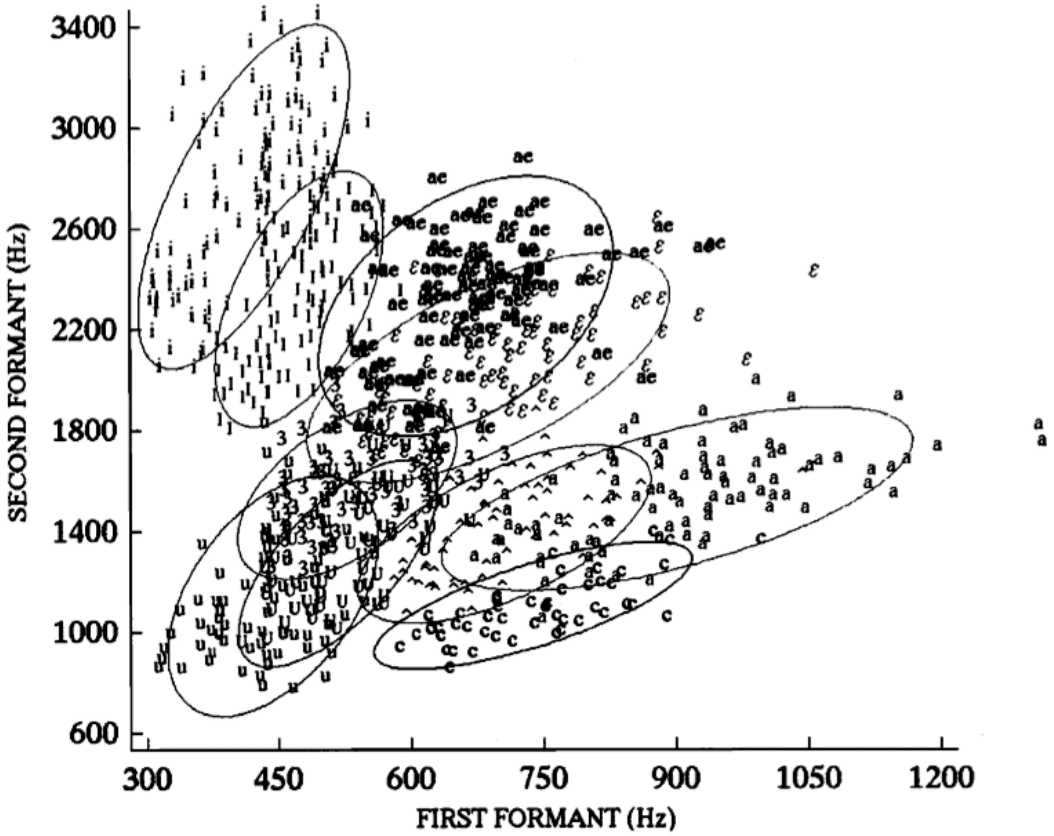



FIG. 4. Values of *F*1 and *F*2 for 46 men, 48 women, and 46 children for 10 vowels with ellipses fit to the data ("ae"=/æ/, "a"=/ɑ/, "c"=/ɔ/, "ʌ"=/ʌ/, "ɜ"=/ɝ/). Measurements for /e/ and /o/ have been omitted, and the data have been thinned of redundant data points.

**Figure 31 :** Les deux premiers formants (F1 en abscisse, F2 en ordonnée) de 12 voyelles de l'anglais américain /iː, ɪ, eɪ, e, æ, ɑː, ɔː, oʊ, ʊ, uː, ʌ, ɝː/ dans les mots « *heed* », « *hid* », « *hayed* », « *head* », « *had* », « *hod* », « *hawed* », « *hoed* », « *hood* », « *who'd* », « *hud* », et « *heard* » (contexte /h_d/) prononcés par 45 hommes[318], 48 femmes et 46 enfants de 10 à 12 ans, majoritairement originaires de l'État de Michigan. Les moyennes de chacun des 3 groupes de locuteurs (hommes, femmes, enfants) en haut, et les valeurs de chaque occurrence (en bas). Les auteurs utilisent les symboles suivants : i pour /iː/ (« *heed* »), e pour /eɪ/ (« *hayed* »), ɛ pour /e/ (« *head* »), ae pour /æ/ (« *had* »), a pour /ɑː/ (« *hod* »), c pour /ɔː/ (« *hawed* »), o pour /oʊ/ (« *hoed* »), U pour /ʊ/ (« *hood* »), u pour /uː/ (« *who'd* »), ɜ pour /ɝː/ (« *heard* »). Les données de /eɪ/ et /oʊ/ ainsi que des points redondants ont été éliminées de la figure d'en bas (Hillenbrand *et al*. 1995 : 3103-3104).

Néanmoins, cette valeur élevée de F2 de /uː/ serait un phénomène relativement récent. Les mêmes auteurs ont comparé les données présentées dans leur article de 1995 avec des années comparables (10 voyelles /iː, ɪ, eɪ, e, æ, ɑː, ɔː, oʊ, ʊ, uː, ʌ, ɝː/ prononcées dans le même contexte /h_d/) publiées en 1952 par Peterson et Barney[319] (Figure 32). Nous y voyons que le F2 ainsi que le F1 de /uː/ était moins élevé dans les données de Peterson et Barney (1952), suggérant que cette voyelle aurait eu autrefois une nature quantique, avec les deux premiers formants proches.

[318] Dans la légende originale du graphique figure le nombre 46, mais d'après le texte rédigé, il s'agit de 45 hommes.

[319] G. E. Peterson, H. L. Barney. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24(2) : 175-184, 1952.

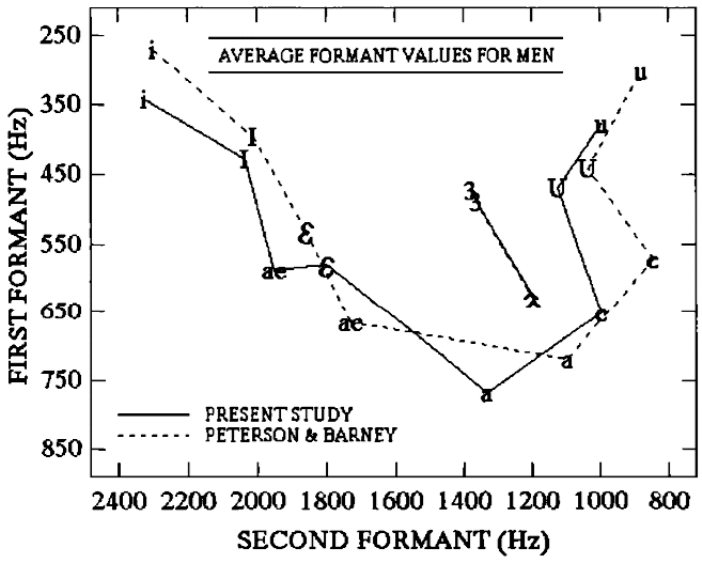



FIG. 5. Acoustic vowel diagrams showing average formant frequencies for men from the present study and from Peterson and Barney ("ae"=/æ/, "a"=/ɑ/, "c"=/ɔ/, "ʌ"=/ʌ/, "ɜ"=/ɝ/).

**Figure 32** : Les deux premiers formants de 10 voyelles de l'anglais américain /iː, ɪ, e, æ, ɑː, ɔː, ʊ, uː, ʌ, ɝː/ dans les mots « *heed* », « *hid* », « *head* », « *had* », « *hod* », « *hawed* », « *hood* », « *who'd* », « *hud* », et « *heard* » (contexte /h_d/) observés dans les données de Hillenbrand *et al*. (1995 : traits continus) et des données comparables de Peterson et Barney (1952 : traits pointillés). Les auteurs utilisent les symboles suivants : i pour /iː/ (« *heed* »), ɛ pour /e/ (« *head* »), ae pour /æ/ (« *had* »), a pour /ɑː/ (« *hod* »), c pour /ɔː/ (« *hawed* »), U pour /ʊ/ (« *hood* »), u pour /uː/ (« *who'd* »), ɜ pour /ɝː/ (« *heard* »). Hillenbrand *et al*. (1995 : 3105).

La montée de F2 du /uː/ est observée également en anglais britannique, et ce, de manière plus marquée. Le mot *who* /huː/ est illustré à la Figure 33 dans une prononciation britannique. À la différence de l'illustration de la prononciation américaine (Figure 30, p. 92), le F2 reste au-dessus de 1000 Hz, assez écarté de F1 jusqu'à la fin.

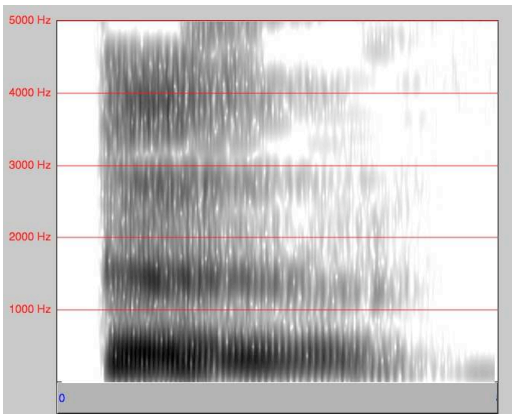



**Figure 33** : Spectrogramme à bande large (Praat ; longueur de fenêtre : 5 millisecondes) du mot *who* /huː/ prononcé par un homme dans une prononciation britannique. Source : Jones *et al*. (2006)[320], CD-ROM.

Selon Cruttenden (2001 : 81-83)[321], l'avancement (de la langue) du /uː/ to [ʉː] est un changement bien établi en RP (*Received Pronunciation*), alors que la délabialisation (*unrounding*) de /uː/ (et de /ʊ/) est considérée comme une innovation récente. Notons que ces deux phénomènes contribuent en effet à élever le F2, comme nous le verrons dans l'Expérience 9 de notre étude. La prononciation complètement postérieure de /uː/, par contraste, est stéréotypiquement associée à la U-RP (*Upper-crust RP* : RP de la couche supérieure ; voir Wells 1982 : 280-283), avec d'autres caractéristiques comme [ɐː] pour /ɜː/, et [ɛæ] pour /æ/ (Hawkins et Midgley 2005 : 184)[322]. Nous pouvons observer la réalisation complètement postérieure de /uː/ chez Daniel Jones (né en 1881), par exemple. Le mot *two* /tuː/ en anglais qu'il a prononcé dans son disque d'illustration des voyelles cardinales (Jones 1956)[323] est montré à la Figure 34. Nous y voyons que le F2 baisse en dessous de 1000 Hz et rejoint le F1 dans la première moitié de la voyelle pour rester à ce niveau jusqu'à la fin.

[320] D. Jones, P. Roach, J. Hartman, J. Setter. *Cambridge English Pronouncing Dictionary* (17th edition). Cambridge, Cambridge University Press, 2006.

[321] A. Cruttenden. *Gimson's Pronunciation of English: 6th edn. of A. C. Gimson, An Introduction to the Pronunciation of English. London*, Edward Arnold, 2001.

[322] S. Hawkins, J. Midgley. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35 : 183-199, 2005.

[323] D. Jones. *Cardinal Vowels*. London, Linguaphone Institute, 1956 (livret d'explication accompagné de 2 disques audio).

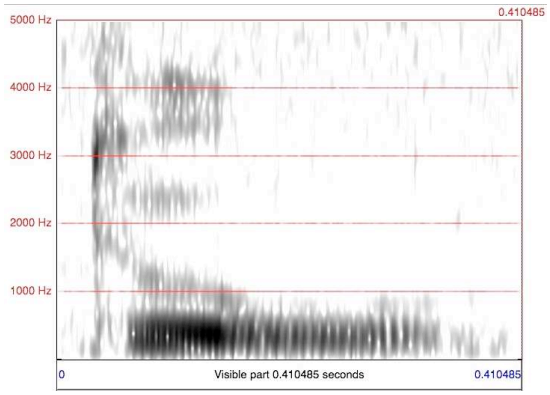



**Figure 34** : Spectrogramme à bande large (Praat ; longueur de fenêtre : 5 millisecondes) du mot *two* /tuː/ en anglais prononcé dans la séquence « *number two* » par Daniel Jones dans son disque d'illustration des voyelles cardinales (Jones 1956).

Les données de Hawkins et Midgley (2005, *op. cit*.) confirment la montée de F2 au cours du XXème siècle. Ils ont mesuré les deux premiers formants de la partie stable de 11 voyelles monophtongues de la RP (*Received Pronunciation*) /iː ɪ e æ ɑː ɒ ɔː ʊ uː ɜː ʌ/ prononcées dans le contexte /h_d/ (les mots *heed, hid, head, had, hard, hod, hoard, hood, who'd, herd, hud*) par 20 locuteurs (hommes) de la RP. Les locuteurs appartenaient chacun à l'un des 4 groupes d'âge suivants à raison de cinq locuteurs par groupe d'âge : 65-73 ans (nés en 1928-1936), 50-55 ans (nés en 1946-1951), 35-40 ans (nés en 1961-1966), et 20-25 ans (nés en 1976-1981). Notons qu'en anglais RP, le mot *hod* est prononcé /hɒd/, et non /hɑːd/ comme en anglais américain GA, et que /ɜː/ rhotique de la GA correspond à /ɜː/ en RP. Le résultat est représenté à la Figure 35. Nous y observons, parmi d'autres tendances de changement comme la montée de F1 et la baisse de F2 pour /æ/, une montée remarquable de F2 pour /uː/ : autour de 1000 Hz (994 Hz) pour les locuteurs au-dessus de 65 ans, contre plus de 1600 Hz (1616 Hz) chez les plus jeunes de 20-25 ans. La valeur élevée de F2 du /uː/ est également observée dans les données de Deterding (2006)[324].

[324] D. Deterding. The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association* 36 : 187-196, 2006.

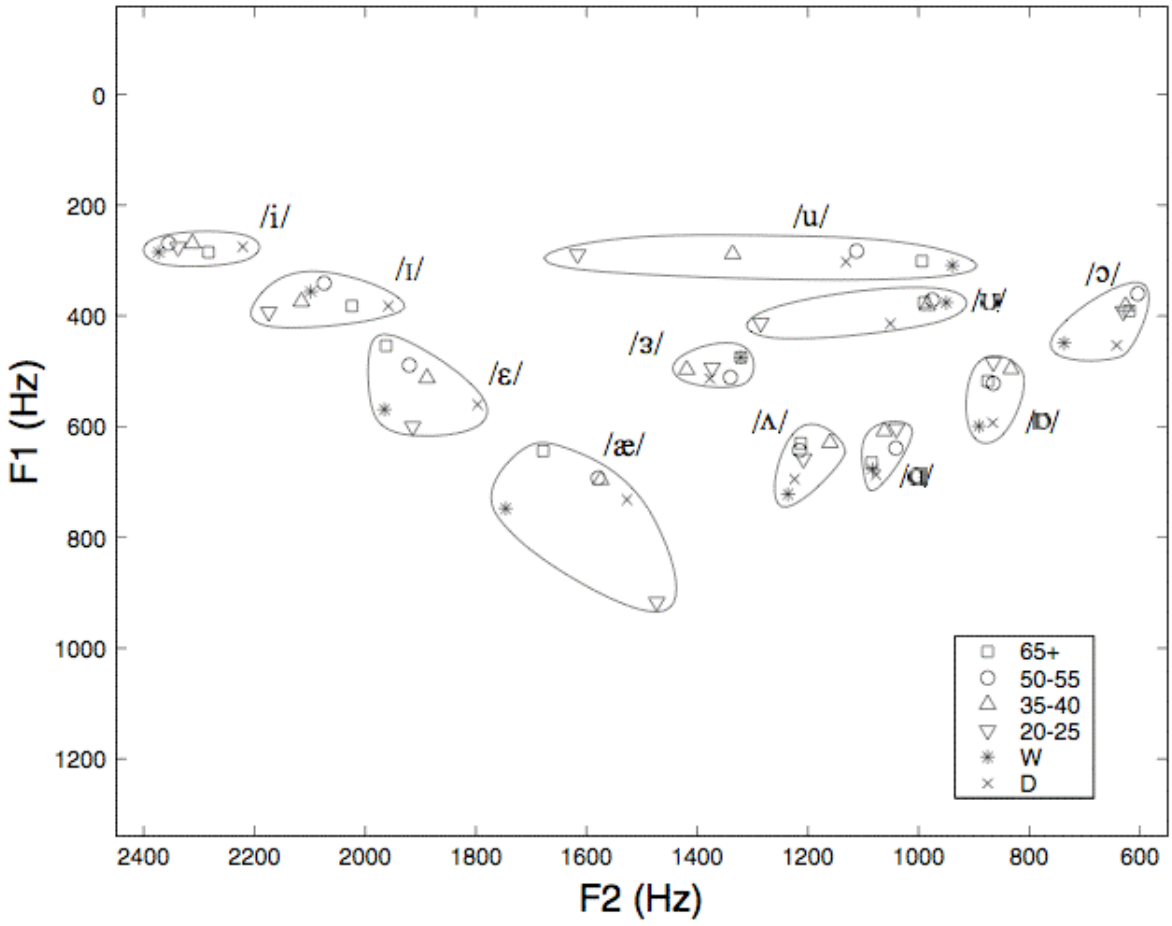



**Figure 1** Mean frequencies (Hz) of the first and second formants of each of the eleven monophthongs, for each age group of the present study, together with those of Wells (1962) and Deterding (1990). The lines surrounding each vowel phoneme separate the vowel distributions but have no statistical status.

**Figure 35** : Fréquences moyennes des deux premiers formants de 11 monophtongues /iː ɪ e æ ɑː ɒ ɔː ʊ uː ɜː ʌ/ dans le contexte /h_d/ pour 4 groupes d'âge différents (20-25 ans, 35-40 ans, 50-55 ans, 65 ans ou plus) de 5 hommes chacun, et les données de Wells (1962)[325] et Deterding (1990)[326]. Les lignes entourant chaque phonème vocalique séparent les distributions de voyelles, mais ne correspondent à aucune analyse statistique. Les auteurs utilisent les symboles suivants : i pour /iː/ (« *heed* »), ɛ pour /e/ (« *head* »), ɑ pour /ɑː/ (« *hard* »), ɔ pour /ɔː/ (« *hoard* »), u pour /uː/ (« *who'd* »). Hawkins et Midgley (2005 : 186).

La valeur élevée de F2 du /uː/ de l'anglais américain et du /u/ [ɯ] du japonais, à la différence du F2 bas du /u/ français, fait que ces deux voyelles-là se ressemblent acoustiquement

[325] J. C. Wells. *A Study of the Formants of the Pure Vowels of British English*. MA dissertation, University College London, 1962. <http://www.phon.ucl.ac.uk/home/wells/formants/index.htm>

[326] D. H. Deterding. *Speaker Normalisation for Automatic Speech Recognition*. Ph.D. dissertation, University of Cambridge, 1990.

et perceptivement. Nishi *et al*. (2008)[327] ont trouvé une distance acoustique faible (fondée sur les deux premiers formants) entre le /uː/ GA et le /u/ [ɯ] (et la séquence /uR/ [ɯː], ou [ɯɯ]) japonais dans une analyse acoustique sur les données des deux langues, et leur ressemblance perceptive dans un test de perception effectué auprès de locuteurs natifs de l'anglais américain. Ces auteurs ont mesuré les deux premiers formants des voyelles brèves (/V/) et longues (/VR/ [Vː] ou [VV]) du japonais prononcées 3 fois par 4 locuteurs (hommes) du japonais de Tokyo dans le contexte /h_ba/ dans une liste de logatomes (« citation ») et dans la phrase-cadre 「これは/h_ba/ ですね」 *kore wa* /h_ba/ *desune* /kore wa h_ba desune/ « ceci est /h_ba/, n'est-ce pas ? ». Les formants de ces voyelles ont été comparés avec ceux de 11 voyelles /iː ɪ eɪ e æ ɑː ɔː oʊ ʊ uː ʌ/ de l'anglais américain prononcées par 4 locuteurs natifs (hommes) dans des contextes comparables. Les résultats, représentés à la Figure 36 montrent que la voyelle de l'anglais américain la plus proche du /u/ [ɯ] japonais est effectivement le /uː/ et que les zones de dispersion (sur le plan F1-F2 uniquement) de ces deux voyelles coïncident très largement.

Le Tableau 16 montre les résultats de l'expérience d'assimilation perceptive effectuée par les mêmes auteurs. Dans cette expérience, 12 auditeurs natifs de l'anglais américain ont écouté 240 stimuli des voyelles du japonais (5 voyelles x 2 longueurs x 4 locuteurs x 2 contextes x 3 répétitions) produits pour l'étude acoustique présentée ci-dessus, et ils devaient choisir laquelle des 11 voyelles de l'anglais américain /iː ɪ eɪ e æ ɑː ɔː oʊ ʊ uː ʌ/ ils ont entendue. Ils ont choisi /uː/ de l'anglais américain pour /u/ ([ɯ]) et /uR/ ([ɯː], ou [ɯɯ]) japonais dans 89-96% des cas.

[327] K. Nishi, W. Strange, R. Akahane-Yamada, R. Kubo, S. A. Trent-Brown. Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America* 124(1) : 576-588, 2008.

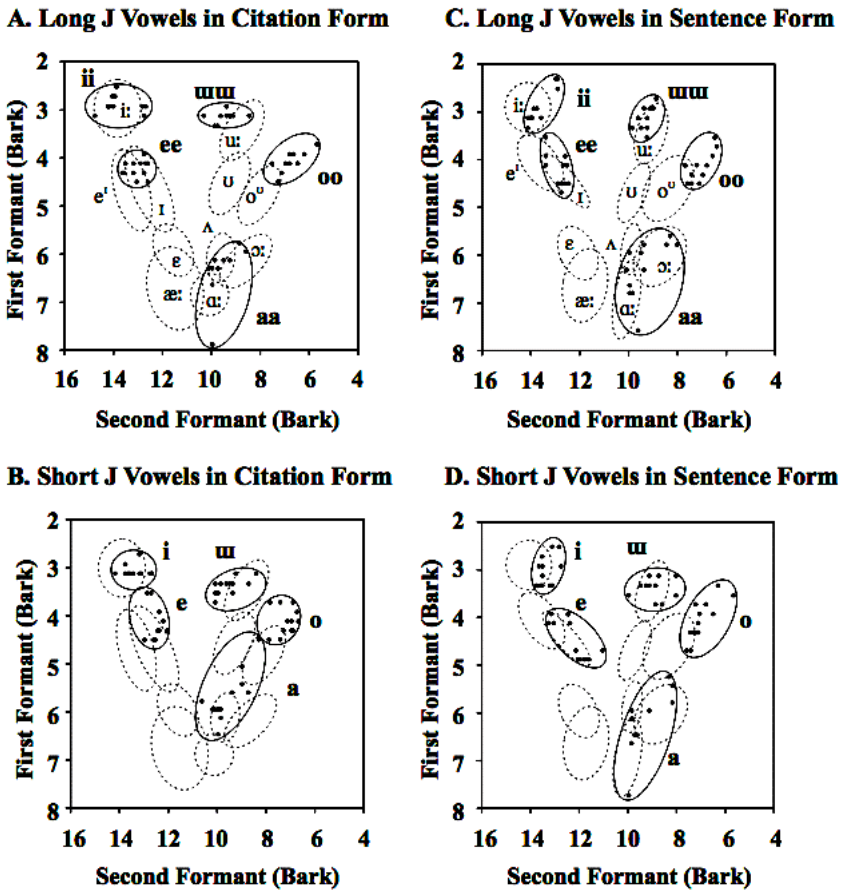



FIG. 2. Formant 1/formant 2 (bark) plots for 2-mora (upper panels) and 1-mora Japanese vowels (lower panels) produced by four male speakers in [hVba] in citation form (left panels) and sentence form (right panels) superimposed on 11 AE vowels (dotted ellipses). Solid ellipses encircle all 12 tokens of a Japanese vowel; individual AE tokens are not plotted.

**Figure 36** : Les deux premiers formants (sur l'échelle de Bark) des 5 voyelles à une more (/V/ : en haut) et à deux mores (/VR/ : en bas) du japonais produites dans le contexte /h_ba/ 3 fois par 4 locuteurs natifs (hommes) du japonais de Tokyo sous forme de citation (à gauche) et dans une phrase (à droite), et ceux de 11 voyelles /iː ɪ eɪ e æ ɑː ɔː oʊ ʊ uː ʌ/ de l'anglais américain (les ellipses pointillées) superposés. Les ellipses continues entourent toutes les 12 occurrences (3 répétitions x 4 locuteurs) de chaque voyelle japonaise ; les occurrences individuelles ne sont pas représentées pour l'anglais américain. Les auteurs utilisent les symboles suivants : eˈ pour /eɪ/ (comme dans « *hay* »), ɛ pour /e/ (comme dans « *head* »), æː pour /æ/ (comme dans « *had* »), oᵛ pour /oʊ/ (comme dans « *boat* »). Nishi *et al*. (2008 : 581).

**Tableau 16** : Patterns d'assimilation perceptive : les réponses de catégorisation exprimées en pourcentage de réponses totales additionnées sur les locuteurs et les auditeurs pour la forme de citation (A) et de phrase (B). Les nombres en gras représentent les cas de classification perceptive modale, les nombres encadrés, les cas de classification acoustique modale. Nishi *et al*. 2008 : 584.

TABLE IV. Perceptual assimilation patterns: Categorization responses, expressed as percentages of total responses summed over speakers and listeners for citation-form (A) and sentence-form (B) materials. Bold =modal perceptual classification; boxed=modal acoustic classification.

| | AE response categories | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | iː | ɪ | eᴵ | ɛ | æː | ɑː–ɔː | ʌ | oᵁ | ʊ | uː |
| **(A) Citation** | | | | | | | | | | |
| **2-mora** | | | | | | | | | | |
| ii | **99** | 1 | | | | | | | | |
| ee | | 2 | **94** | 5 | | | | | | |
| aa | | | | | 2 | **89** | 9 | | | |
| oo | | | | | | 1 | | **99** | | 1 |
| ɯɯ | | | | | | | 1 | 2 | 5 | **92** |
| **1-mora** | | | | | | | | | | |
| i | **95** | 4 | | 1 | | | | | | |
| e | | 16 | **76** | 8 | | | | | | |
| a | | | | | 3 | **57** | 39 | | | |
| o | | | | | | | 1 | **95** | 1 | 2 |
| ɯ | | | | | | | 3 | 1 | 5 | **91** |
| **(B) Sentence** | | | | | | | | | | |
| **2-mora** | | | | | | | | | | |
| ii | **99** | | | | | | | | | |
| ee | | 1 | **97** | 2 | | | | | | |
| aa | | | | | 1 | **96** | 3 | | | |
| oo | | | | | | | | **98** | | 1 |
| ɯɯ | | | | | | | | 2 | 2 | **96** |
| **1-mora** | | | | | | | | | | |
| i | **98** | 2 | | | | | | | | |
| e | 1 | 23 | **48** | 28 | | | | | | |
| a | | | | | 2 | **77** | 21 | | | |
| o | | | | | | | | **95** | | 4 |
| ɯ | | | | | | | 2 | 1 | 7 | **89** |

Les études présentées ci-dessus illustrent la nature non quantique (pas de regroupement des deux premiers formants) du /uː/ de l'anglais américain, ce qui fait que cette voyelle ressemble au /u/ [ɯ] japonais et diffère du /u/ français.

Les voyelles de l'anglais sont adaptées au japonais dans des mots d'emprunt et des noms propres en générale de la manière représentée au Tableau 17 (voir Shinohara 1997[328], entre autres, pour une étude approfondie).

[328] S. Shinohara. *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Thèse de doctorat, Université de la Sorbonne Nouvelle, 1997.

**Tableau 17** : Adaptation des voyelles de l'anglais au japonais dans des mots d'emprunt.

| Anglais | Japonais | *Katakana* | Exemple | Anglais | Japonais | *Katakana* |
|---|---|---|---|---|---|---|
| /iː/ | /iR/ | イー | beat | /biːt/ | /biRto/ | ビート |
| /ɪ/ | /i/ | イ | bit | /bɪt/ | /biQto/ | ビット |
| /ɪr/ (RP /ɪə/) | /i(j)a(R)/ | イア／ヤ（ー） | year | /jɪr/ | /ijaR/ | イヤー |
| /eɪ/ | /eR/ | エー | gate | /geɪt/ | /geRto/ | ゲート |
| /e/ | /e/ | エ | get | /get/ | /geQto/ | ゲット |
| /er/ (RP /eə/) | /ea/ | エア（ー） | air | /er/ | /ea(R)/ | エア（ー） |
| /æ/ | /a/ | ア | bag | /bæg/ | /baQgu/ | バッグ |
| N.B. /kæ/ /gæ/ | /ja/ | キャ／ギャ | cat | /kæt/ | /kjaQto/ | キャット |
| /ʌ/ | /a/ | ア | cut | /kʌt/ | /kaQto/ | カット |
| /ɑː/ | /a(R)/ | ア（ー） | father | /fɑːðə/ | /hwa(R)zaR/ | ファ（ー）ザー |
| /ɑː/ (RP /ɒ/) | /o/ | オ | hot | /hɑːt/ | /hoQto/ | ホット |
| /ɑːr/ (RP /ɑː/) | /aR/ | アー | car | /kɑːr/ | /kaR/ | カー |
| /ɜː/ (RP /ɜː/) | /aR/ | アー | girl | /gɜːl/ | /gaRru/ | ガール |
| /ɔː/ | /oR/ | オー | call | /kɔːl/ | /koRru/ | コール |
| /ɔːr/ (RP /ɔː/) | /oR/ | オー | court | /kɔːrt/ | /koRto/ | コート |
| /oʊ/ (RP /əʊ) | /oR/ | オー | coat | /koʊt/ | /koRto/ | コート |
| /uː/ | /uR/ | ウー | boot(s) | /buːts/ | /buRtu/ | ブーツ |
| /ʊ/ | /u/ | ウ | book | /bʊk/ | /buQku/ | ブック |
| /ʊr/ (RP /ʊə ~ ɔː/) | /ua(R)/ | ウア（ー） | poor | /pʊr/ | /pua/ | プア |
| /aɪ/ | /ai/ | アイ | eye | /aɪ/ | /ai/ | アイ |
| /aɪr/ (RP /aɪə ~ aə/) | /aia(R)/ | アイア／ヤ（ー） | fire | /faɪr/ | /fai(j)a(R)/ | ファイア／ヤ（ー） |
| /aʊ/ | /au/ | アウ | out | /aʊt/ | /auto/ | アウト |
| /aʊr/ (RP /aʊə ~ aə/) | /awa(R)/ | アワ（ー） | hour | /aʊr/ | /awaR/ | アワー |
| /ɔɪ/ | /o(R)i/ | オ（ー）イ | boy | /bɔɪ/ | /boRi/ | ボーイ |

### *1.6.2. D'autres caractéristiques à noter*

Parmi les phénomènes qui caractérisent l'anglais américain par rapport au japonais et au français parisien, nous pouvons retenir les suivants :

**Consonnes** : l'opposition entre /l/ et /r/ est plus difficile à distinguer perceptivement que l'opposition entre /l/ et /ʀ/ français (Yamasaki et Hallé 1999[329], Ooigawa 2008[330]) ; l'aspiration des occlusives sourdes (Lisker et Abramson 1964[331], entre autres) ; /t/ communément considéré comme caractérisé par son lieu d'articulation alvéolaire par opposition au lieu d'articulation dental du /t/ français, mais la différence essentielle entre les deux serait celle de la forme de la langue (Straka 1965 : 31[332], Vaissière 2006[333]) ; allophone vélarisé [ɬ] du /l/ en position post-vocalique,

[329] H. Yamasaki, P. A. Hallé. How do native speakers of Japanese discriminate and categorize French /r/ and /l/? *Proceedings of the XIVth International Congress of Phonetic Sciences*, 909-912, 1999.

[330] T. Ooigawa. Nihon-jin, Amerika-jin, Furansu-jin choushusha ni yoru eigo, futsugo, seigo ni okeru boinkan no ryuuon no benbetsu : yobi kenkyuu [Discrimination de liquides intervocaliques en anglais, en français, et en espagnol chez des auditeurs japonais, américains et français : étude préliminaire]. *Sophia University Working Papers in Phonetics* 2008, Sophia University Phonetics Laboratory, Tokyo : 2-20, 2008.

[331] L. Lisker, A. Abramson. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20 : 384–422, 1964.

[332] G. Straka. *Album phonétique*. Sainte-Foy, Les Presses de l'Université Laval, 1965.

relativement vélarisé aussi en position pré-vocalique en anglais américain (Ladefoged 1993[334], Takebayashi 1996[335], entre autres).

**Structures syllabiques** : les syllabes fermées sont nombreuses, comme nous l'avons vu dans les données de Delattre (1965 : 41-42[336], tableau reproduit dans le Tableau 13, p. 84) ; l'ambisyllabicité perceptive en anglais (Cutler, Mehler, *et al*. 1986)[337], à la différence de la syllabicité perceptive du français (Mehler, Dommergues, *et al*. 1981)[338], et à la perception moraïque du japonais (Ôtake, Hatano, *et al*. 1993[339]), montrée dans des expérience de détection de cible (syllabe ou non).

**Rythme** : l'anglais est caractérisé comme « *stress-timed* », ou comme ayant un fort degré de « *stress-based* » (Dauer 1983 : 60, figure reproduite à la Figure 26, p. 85), à la différence du japonais (« *mora-timed* » ou faible degré de « *stress-based* ») et du français (« *syllable-timed* » ou faible degré de « *stress-based* »).

**Prosodie** : l'allongement final est moins marqué en anglais qu'en français, et la F0 présente une tendance générale descendante par rapport au français, comme nous l'avons vu dans la chapitre précédent (1.5.).

## 1.7. COMPARAISON DES TROIS LANGUES CONCERNÉES : LE JAPONAIS DE TOKYO (L1), L'ANGLAIS AMÉRICAIN (L2), ET LE FRANÇAIS PARISIEN (L3)

La comparaison entre les caractéristiques phonémiques et phonétiques présentées dans les trois derniers chapitres nous permettra de prédire des possibilités de transferts positifs de L1 (japonais) et de L2 (anglais américain GA) vers L3 (français), qui pourraient avoir lieu lors de l'apprentissage du français langue étrangère par des apprenants japonophones. Ces possibilités sont présentées dans le Tableau 18 ci-dessous.

---

[333] J. Vaissière. Using an articulatory model as an integrated tool for a better understanding of the combined articulatory, acoustic and perceptual aspects of speech. Communication orale présentée au *INAE Workshop on Image and Speech Processing*, Chennai, Inde, 2006.

[334] P. Ladefoged. *A Course in Phonetics* (troisième édition). Fort Worth, Texas, Harcourt Brace Jovanovich, 1993.

[335] S. Takebayashi. *Eigo onseigaku* [Phonétique de l'anglais]. Tokyo, Kenkyûsha, 1996.

[336] P. Delattre. "Comparing the Prosodic Features of English, German, Spanish and French", in P. Delattre, *Comparing the Phonetic Features of English, French, German and Spanish: an Interim Report*. Heidelberg, Julius Groos Verlag, 23-43, 1965.

[337] A. Cutler, J. Mehler, D. Norris, J. Segui. The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25(4) : 385-400, 1986.

[338] J. Mehler, J.-Y. Dommergues, U. Frauenfelder, J. Segui. The Syllable's Role in Speech Segmentation. *Journal of Verbal Learning and Verbal Behavior* 20 : 298-305, 1981.

[339] T. Ôtake, G. Hatano, A. Cutler, J. Mehler. Mora or Syllable? Speech Segmentation in Japanese. *Journal of Memory and Language* 32(2) : 258-278, 1993.

**Tableau 18** : Comparaison de principales caractéristiques phonémiques et phonétiques du japonais, de l'anglais et du français, et possibilités de transferts positifs de L1 japonais et de L2 anglais américain (GA) vers L3 français : 1) voyelles ; 2) consonnes ; 3) suprasegmental[340].

1) voyelles

| *Voyelles* | **Grand inventaire** | **Diphtongaison** | **Voyelles phonémiquement antérieures arrondies** | **/u/ quantique (F1/F2 groupés)** | **Plusieurs voyelles ouvertes** | **Voyelles phonémiquement nasales** | **Opposition tendue-relachée** |
|---|---|---|---|---|---|---|---|
| **L1 japonais** | Non | Non | Non | Non | Non | Non | Non |
| **L2 anglais** | Oui | Oui | Non | Non | Oui | Non | Oui |
| **L3 français** | Oui | Non | Oui | Oui | Non (?) | Oui | Non |
| **Transfert positif possible ?** | | | | | | | |
| **de L1 vers L3** | Non | Oui | Non | Non | Oui (?) | Non | Oui |
| **de L2 vers L3** | Oui ? | Non | Non | Non | Non (?) | Non | Non |

2) consonnes

| *Consonnes* | **Labio-dentales /f/ /v/** | **Deux liquides ?** | **Arrondissement des lèvres pour /ʃ/** | **Aspiration** |
|---|---|---|---|---|
| **L1 japonais** | Non | Non | Non | Non (?) |
| **L2 anglais** | Oui | Oui | Oui | Oui |
| **L3 français** | Oui | Oui | Oui | Non |
| **Transfert positif possible ?** | | | | |
| **de L1 vers L3** | Non | Non | Non | Oui (?) |
| **de L2 vers L3** | Oui | Oui | Oui | Non |

3) suprasegmental

| *Suprasegmental* | **Rythme** | **Isochronie (perçue)** | **Structures syllabiques** | **Groupes de consonnes (clusters)** | **Tendance de f0** |
|---|---|---|---|---|---|
| **L1 japonais** | *Mora-timed* | Oui | majoritairement ouvertes | Non | Descendante |
| **L2 anglais** | *Stress-timed* | Non | majoritairement fermées | Oui | Descendante |
| **L3 français** | *Syllable-timed* | Oui | majoritairement ouvertes | Oui | Montante |
| **Transfert positif possible ?** | | | | | |
| **de L1 vers L3** | Oui ? | Oui ? | Oui (?) | Non | Non |
| **de L2 vers L3** | Non | Non | Non | Oui | Non |

Nous y voyons des cas où un phénomène donné ne se trouve ni dans L1 ni dans L2, et il est attendu donc que son acquisition soit difficile. Parmi les voyelles, les cas des voyelles phonémiquement antérieures arrondies, de la nature quantique (regroupement de formants) du /u/, et des voyelles nasales entrent dans cette catégorie. C'est pour cette raison qu'il nous paraît particulièrement intéressant de les examiner. Il existe en effet de nombreuses études qui portent,

[340] Une version antérieure de ces tableaux a été présentée dans la communication suivante :

T. Kamiyama. Acquisition of French vowels by Japanese-speaking learners: close and close-mid rounded vowels. Phonetics and Phonology in Third Language Acquisition (réunion satellite du XVI$^{ème}$ Congrès International des Sciences Phonétiques), Université de Fribourg (en Brisgau), Allemagne, 3-4 août 2007.

entre autres, sur l'acquisition des voyelles antérieures arrondies par des locuteurs anglophones, comme nous le verrons dans le chapitre suivant.

## 1.8. L'APPRENTISSAGE DES VOYELLES L2 CONCERNANT LE FRANÇAIS, L'ANGLAIS, ET LE JAPONAIS

Nous présenterons dans ce chapitre divers types d'expériences de perception et de production réalisés dans les études inter-langues en général (1.8.1), et des études inter-langues antérieures concernant les voyelles du français, de l'anglais et du japonais (1.8.2).

### *1.8.1. Types d'expérience de perception et de production*

Dans les études qui portent sur la perception et la production des langues non natives, divers types d'expériences sont réalisées.

Les tableaux ci-dessous résument les différents types d'expérience.

**A. Perception**

| -- catégorisation (réponse / transcription libre, choix forcé)<br>-- détection de phonème, de mot<br>-- discrimination de type AX, AXB, ABX, | (cible) | (contexte) | (mesure) |
|---|---|---|---|
| | -- consonnes et voyelles<br>-- syllabes<br>-- accent lexical<br>-- mots | -- contexte vide<br>-- syllabes<br>-- mots / logatomes<br>-- syntagmes<br>-- phrases<br>-- discours | -- réponse donnée par l'auditeur<br>-- temps de réponse |

**B. Production**

| -- lecture<br>--- texte écrit en orthographe<br>--- texte écrit en orthographe accompagné de transcription phonémique<br><br>-- répétition<br>--- répétition immédiate (*shadowing*)<br>--- répétition différée<br><br>-- parole semi-spontanée (production guidée : dénomination d'image, *map-task*, …)<br>-- parole spontanée | (objet de l'analyse) | (mesure) |
|---|---|---|
| | -- consonnes et voyelles<br>-- syllabes<br>-- mots / logatomes<br>-- syntagme<br>-- phrase<br>-- discours | -- analyse acoustique (durée, F0, formants, intensité, …)<br>-- analyse articulatoire (débit d'air, ouverture / fermeture de la glotte, position et forme de la langue, …)<br>-- temps de réponse |

**C. perception de la production non native par les auditeurs natifs**

-- identification (et évaluation)

-- évaluation globale

**D. comparaison inter-langue de la perception par des auditeurs de différentes langues maternelles**

-- identification (aux catégories de la langue native) et évaluation de degré de bon exemplaire

-- discrimination (AX, ABX, AXB, *etc.*)

**A. Perception**

**Catégorisation** Dans une tâche d'identification (catégorisation) d'un segment d'une langue non native, les sujets-auditeurs sont amenés à écouter le stimulus et à le classer dans une des catégories. S'il s'agit des catégories de la langue non native, cette tâche nécessite des connaissances explicites sur le système phonologique (ou au moins orthographique) de la langue testée de la part des sujets-auditeurs, ou il existe une phase d'entraînement pendant laquelle les sujets-auditeurs apprennent les catégories phonémiques testées (avec *feedback*). Les auditeurs peuvent être amenés également à identifier des catégories natives et à donner une note reflétant dans quelle mesure les stimuli est un « bon » exemplaire (« assimilation perceptive »). La réponse peut être donnée soit librement (les auditeurs écrivent ce qu'ils ont entendu), soit au choix forcé (les catégories).

**Détection** Dans une expérience de détection de phonème ou de mot, les tâches des auditeurs sont d'écouter des séquences de sons non natifs, et de réagir (en appuyant sur un bouton, par exemple) quand ils ont perçu la cible demandée (phonème, syllabe, mot, *etc.*). Ce type d'expérience demande également des connaissances sur les catégories non natives concernant la cible à détecter, sauf si les stimuli présentés contiennent une cible ressemblant à un son natif (ex. auditeurs japonophones qui détectent la consonne /r/ japonais dans un énoncé en français).

**Discrimination** Dans un test de discrimination AX, les auditeurs écoutent les sons et jugent si les deux sons présentés sont identiques ou différents. Dans un test ABX, des triplets de sons sont présentés : les deux premiers sont des repères, et les auditeurs jugent si le troisième est identique au premier ou au deuxième. Le type AXB fonctionne de la même manière, sauf que c'est le premier et le dernier qui sont les repères, le deuxième étant identique soit au premier, soit au dernier. Dans AXB, le stimulus "test" se trouve à distance égale des deux stimuli repères. Que ce soit de AX, ABX, ou AXB, les tests de discrimination peuvent être réalisés auprès des non apprenants de la langue testée, car il s'agit de comparer les stimuli présentés, et non de reconnaître les stimuli comme membre d'une catégorie non native.

**Différence de tâche** Les résultats des expériences de perception peuvent varier selon les types de tâche, par exemple, entre les tâches d'identification et de discrimination (Flege et Hillenbrand 1984). La tâche d'identification demande une classification fondée davantage sur les catégories phonémiques (donc linguistiques) que la tâche de discrimination, qui relève davantage des différences acoustiques (donc physiques) qui se trouvent entre les stimuli, même si les stimuli d'une même paire ou d'un même triplet ne se succèdent pas immédiatement (l'intervalle entre les stimuli était de 500 ms dans le test AXB de Levy et Strange 2008, 1 seconde dans le nôtre).

À titre d'exemple[341], Gottfried (1984) a effectué un test d'identification de 11 voyelles orales du français (y compris le /ɑ/ postérieur) soutenues et isolées auprès de 16 auditeurs natifs du français parisien ; le taux d'erreurs des auditeurs francophones est de 20% environ. Par contraste, les résultats du test de discrimination AXB de Levy et Strange (2008), dans lequel 6 paires de voyelles françaises (/u/-/y/, /i/-/y/, /u/-/œ/, /y/-/œ/, /i/-/u/, /i/-/œ/) ont été présentées dans des logatomes /ʁab_p/ et /ʁad_t/, montrent que les auditeurs francophones natifs ont présenté 2-3% d'erreurs en moyenne pour /y/-/œ/, /i/-/u/, et 1% environ pour /i/-/œ/, et presque zéro pour /u/-/y/, /i/-/y/, /u/-/œ/. Dans notre étude, les 13 voyelles (orales et nasales) ont été identifiées par cinq auditeurs francophones natifs avec un taux moyen d'erreur de 17% (Expérience 3, Tableau 20, p. 127), alors que six paires de voyelles (/u/-/ø/, /y/-/ø/, /u/-/y/, /o/-/u/, /i/-/e/) ont été discriminées par cinq auditeurs francophones natifs dans un test de discrimination AXB (utilisant un sous-ensemble des stimuli présentés dans le test d'identification ci-dessus) avec un taux d'errer inférieur à 1% (Expérience 6, section 3.1.5.3.).

**Cible** Les unités qui peuvent être la cible incluent des consonnes, des voyelles, des syllabes, des mots, l'accent lexical, l'accent de syntagme ou de phrase.

**Mesures** Les mesures peuvent porter sur les réponses données par les auditeurs et sur le temps de réaction.

**B. Production**

**Lecture** Dans une tâche de lecture, le corpus est présenté à l'écritsous forme orthographique ou phonétique. Il est donc demandé aux sujets-locuteurs de posséder des connaissances sur l'orthographe ou un autre système de transcription (transcription phonémique, ou translittération en alphabet latin pour les langues écrites dans d'autres systèmes d'écriture) pour qu'ils puissent décoder le corpus. Cette tâche ne peut être effectuée par les non apprenants, à moins d'avoir une séance d'entraînement. La production est susceptible d'être influencée par le rapport graphie-phonie de la langue native, ou des autres langues acquises ou apprises (ex. la romanisation du japonais et l'orthographe de l'anglais pour les japonophones apprenant le français).

**Répétition** Dans une tâche de répétition (immédiate ou différée), le corpus est présenté oralement (par un locuteur en présence, ou des enregistrements). Les sujets-locuteurs doivent donc décoder perceptivement ce qu'ils entendent pour qu'ils puissent produire le son cible. La production est ainsi influencée par la perception des sujets-locuteurs. La répétition différée relève davantage d'un traitement phonologique par rapport à la répétition immédiate, qui dépend plus d'un traitement acoustique. Les non apprenants peuvent participer à une tâche de répétition, mais la tâche serait plus difficile pour la répétition différée.

**Production semi-spontanée** Les apprenants peuvent être invités à produire de la parole semi-spontanée dans une tâche de production guidée. La production peut être guidée par une

[341] Il est plus difficile de comparer les résultats de différentes études qui portent sur les auditeurs natifs à cause de différence de niveau de compétence linguistique.

image (dénomination d'objet), ou dans une tâche utilisant un plan (*map task*). La production ne subit pas une influence directe ni des connaissances sur l'orthographe ni de la perception, mais il est demandé aux sujets-locuteurs de posséder une compétence linguistique suffisante pour effectuer la tâche. Les non apprenants ne peuvent donc pas être testés.

**Production spontanée** Une tâche de production spontanée permet d'étudier ce que les apprenants produisent quand ils se concentrent sur la transmission de message plutôt que sur l'exactitude de la forme (phonétique, grammaticale) d'une langue non native. En revanche, il est difficile de contrôler la forme produite.

**Objet de l'analyse** Comme pour les tâches de perception, divers niveaux d'unités peuvent être étudiés, des voyelles isolées, syllabes, mots et logatomes, phrases, *etc*. Notons que la difficulté de la tâche augmente en général en fonction de la longueur de la production. Dans une tâche de lecture de phrases ou de discours, les sujets-locuteurs doivent effectuer une analyse syntaxique et pragmatique pour produire correctement le corpus. Une tâche de répétition d'une séquence longue demande une mémorisation, aidée ou non de manière *top-down* (du haut en bas) par les connaissances lexicales, morphologiques, syntaxiques, pragmatiques, *etc*.

**Mesures** Les mesures peuvent porter sur des analyses acoustiques (durée, F0, formants, intensité, …), des mesures articulatoires (débit d'air, ouverture / fermeture de la glotte, position et forme de la langue, …), et le temps de réaction.

**C. Perception de la production non native par les auditeurs natifs**

La production des locuteurs non natifs peut être soumise à une expérience de perception auprès d'auditeurs natifs. Ces derniers ont à identifier ou à évaluer le degré de « bon » exemplaire des productions des apprenants (voyelles isolées, syllabes, mots, phrases, ...) ..

**D. Comparaison de la perception des stimuli synthétisés auprès de plusieurs groupes d'auditeurs parlant différentes langues natives**

Des stimuli synthétisés peuvent être utilisés pour comparer les comportements perceptifs de plusieurs groupes d'auditeurs parlant différentes langues natives (ex. perception des francophones non parlant le japonais et des japonophones non parlant le français pour étudier le français appris par des apprenants japonophones, ou le japonais appris par des apprenants francophones). Des tâches d'identification aux catégories de la langue native des auditeurs et d'évaluation de degré de bon exemplaire sont envisageable, ainsi que celles de discrimination (AX, ABX, AXB, ...).

### *1.8.2. Études inter-langues antérieures concernant les voyelles du français, de l'anglais, et du japonais*

Nous présentons par la suite des études antérieures sur l'acquisition et l'apprentissage, ainsi que sur la perception non-native des voyelles, notamment des voyelles antérieures arrondies, concernant le français, l'anglais et le japonais.

En ce qui concerne les comportements des apprenants japonophones apprenant l'anglais langue étrangère (ALE), Lambacher *et al.* (2000)[342] ont montré que les japonophones apprenant l'ALE avaient une certaine difficulté perceptive quant à l'identification des voyelles postérieures /ʌ/ et /ɔː/, et des voyelles ouvertes /æ/ et /ɑː/ de l'anglais américain (GA). Notons que ces voyelles sont en général perçues comme proches du /a/ japonais.

Strange, Akahane, *et al.* (1998)[343] ont mené une étude d'assimilation perceptive dans laquelle 24 auditeurs japonophones ont écouté 11 voyelles de l'anglais américain /iː ɪ eɪ e æ ɑː ʌ ɔː oʊ ʊ uː/ prononcées dans les contextes /h_b/ et /ˈh_ˌbɑ/ par 4 locuteurs natifs de l'anglais américain, pour choisir la voyelle du japonais la plus proche de ce qu'ils ont entendu. Les auteurs concluent que les voyelles extrêmes (/iː/, /ɑː/, et /uː/) ont été assimilées aux trois voyelles correspondantes du japonais (/i/, /a/, et /u/ [ɯ]) d'une manière relativement stable, mais moins pour les autres. Les voyelles longues (et diphtongues) ont été perçues comme similaires aux voyelles longues du japonais, mais seulement dans le contexte /h_b/ prononcé et présenté dans une phrase-cadre (« *I say* /h_b/ *on the tape* »), et non dans des logatomes isolés (/ˈh_ˌbɑ/). Nous avons vu dans le chapitre 1.6. les résultats de Nishi *et al*. (2008), qui ont montré que les auditeurs anglophones américains ont perçu le /u/ [ɯ] japonais comme proche du /uː/ américain (tableau des résultats reproduit dans le Tableau 16, p. 101). Ceux de Strange, Akahane, *et al*. (1998) montrent ainsi qu'une tendance similaire est observée dans l'autre sens : les auditeurs japonophones considèrent le /uː/ de l'anglais américain comme proche du /u/ [ɯ] japonais.

Les autres tendances d'assimilation perceptive incluent les suivantes (Tableau 19) :

- Les voyelles fermées tendues et relâchées, /i ɪ/ et /uː ʊ/, sont assimilées aux voyelles /i/ et /u/ (ou à /iR/ et uR/) du japonais, mais le taux d'assimilation du /ɪ/ américain au /i/ japonais est plus faible (59%) que celui du /ʊ/ américain au /u/ japonais (97%) ;

- Les voyelles moyennes diphtongues et monophtongues, /eɪ e/ et /oʊ ɔː/ sont assimilées aux voyelles /e/ et /o/ (ou à /eR/ et /oR/) du japonais ;

- Les trois voyelles ouvertes /æ ɑː ʌ/ sont assimilées à /a/ (ou à /aR/, /ja/, /jaR/) du japonais.

Notons que la réponse modale de chaque voyelle correspond aux patterns d'adaptation observés dans des mots d'emprunt et des noms propres de l'anglais en japonais (Tableau 17, p. 102), sauf /ɑː/ : /ɑː/ dans des mots tels que *hot* ou *lot* est adapté comme /o/ en japonais, qui correspond plutôt à la forme de la RP (*Received Pronunciation*) /ɒ/.

**Tableau 19** : Pattern d'assimilation perceptive de 11 voyelles de l'anglais américain prononcées dans des logatomes disyllabiques /ˈh_ˌbɑ/ isolés (A) et dans /h_b/ inséré dans la phrase cadre « *I say* /h_b/ *on the tape* » (B) par 4 locuteurs de l'anglais américain et perçues par 24 auditeurs japonophones. Réponses de catégorisation représentées en pourcentage de toutes les réponses additionnées sur les locuteurs et les auditeurs pour les conditions disyllabe (A : en haut), et phrase

[342] S. Lambacher, W. Martens, G. Molholt. Comparison of identification of American English vowels by native speakers of Japanese and English. *Proceedings of the Phonetic Society of Japan 2000*, 213-217, 2000.

[343] W. Strange, R. Akahane-Yamada, R. Kubo, S. A. Trent, K. Nishi, J. J. Jenkins. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics* 26 : 311-344, 1998.

(B : en bas). La réponse modale de chaque stimulus de voyelle est donnée en caractères gras. Les voyelles de l'anglais américain présentées en ordonnée, les catégories de réponses en japonais en abscisse. Les auteurs utilisent les symboles ɛ pour /e/ anglais, VV pour /VR/ en japonais (ex. aa pour /aR/). Strage, Akahane *et al*. (1998 : 322).

TABLE I. Spectral Assimilation Patterns. Categorization responses, expressed as percentages of total responses summed over speakers and listeners within Disyllable (A) and Sentence (B) conditions

| | Japanese response categories | | | | |
|---|---|---|---|---|---|
| A. *Disyllable* | High front<br>i, ii | Mid front<br>e, ee, ei | Low<br>a, aa, ʲa, ʲaa | Mid back<br>o, oo, ou | High back<br>u, uu, ʲu |
| iː | **99** | < 1 | | | < 1 |
| ɪ | **59** | 41 | | | |
| e͡ɪ | 9 | **82** | 3 | 2 | 4 |
| ɛ | 1 | **85** | 14 | | |
| æː | < 1 | 38 | **61** | < 1 | |
| ɑː | | | **99** | | < 1 |
| ʌ | | | **68** | 13 | 19 |
| ɔː | | | 29 | **70** | 1 |
| o͡ʊ | | | | **94** | 6 |
| ʊ | | < 1 | < 1 | 2 | **97** |
| uː | | < 1 | | | **99** |

| B. *Sentence* | High front<br>i, ii | Mid front<br>e, ee, ei | Low<br>a, aa, ʲa, ʲaa | Mid back<br>o, oo, ou | High back<br>u, uu, ʲu |
|---|---|---|---|---|---|
| iː | **100** | | | | |
| ɪ | **81** | 18 | | | 1 |
| e͡ɪ | 19 | **81** | | | |
| ɛ | 2 | **63** | 35 | | |
| æː | | 29 | **71** | | |
| ɑː | 1 | 1 | **93** | 3 | 2 |
| ʌ | | 1 | **69** | 13 | 17 |
| ɔː | | | 28 | **72** | |
| o͡ʊ | | | | **76** | 24 |
| ʊ | | | 3 | 2 | **95** |
| uː | | | | | **100** |

Il existe de nombreuses études qui portent sur la perception et la production des voyelles antérieures arrondies du français par des locuteurs natifs de l'anglais.

Gottfried (1984)[344] a montré, dans une expérience d'identification perceptive, la difficulté que les auditeurs anglophones américains rencontrent face aux voyelles antérieures arrondies, qui n'existent pas dans le système phonémique de l'anglais.

Levy et ses collègues ont examiné la perception des voyelles antérieures arrondies du français et de l'allemand par les locuteurs natifs de l'anglais américain, et éclairci particulièrement l'effet du contexte consonantique et de l'expérience d'apprentissage (Levy et Strange 2002[345],

[344] T. L. Gottfried. Effects of consonant context on the perception of French vowels. *Journal of Phonetics* 12 : 91-114, 1984.

[345] E. S. Levy, W. Strange. Effects of consonantal context on perception of French rounded vowels by American English adults with and without French language experience. *Journal of the Acoustical Society of America* 111(5) : 2361-2362, 2002.

Strange, Levy, *et al.* 2004[346], Levy 2006[347], Levy et Strange 2008[348], Levy 2009[349]). Dans leurs expériences d'assimilation perceptive (Levy *et al.* 2004, Levy 2009), les auteurs ont montré que la série antérieure arrondie était perçue comme plus similaire aux voyelles postérieures qu'antérieures de l'anglais. Levy et Strange (2008) ont effectué une expérience de discrimination du type AXB en utilisant les voyelles /u y œ i/ prononcées par 3 locuteurs natifs du français parisien dans des logatomes /rab_p/ et /rad_t/. La perception de deux groupes de 10 auditeurs de l'anglais américain (ceux qui ne parlaient pas français – non expérimentés – et ceux qui avaient appris le français de manière extensive en commençant après l'âge de 13 ans – expérimentés –) a été comparée. Le groupe expérimenté a répondu plus correctement que le groupe non expérimenté sur les paires /u/-/œ/, /i/-/y/, et /y/-/œ/ (taux moyen d'erreur : expérimenté 5% *vs.* non expérimenté 24% pour /y/-/œ/), alors que les deux groupes ont présenté peu de différence pour /u/-/y/ (expérimenté 30% *vs.* non expérimenté 24% : Figure 37). Il se peut qu'un des facteurs qui pourraient expliquer la mauvaise discrimination de la paire /u/-/y/ par le groupe expérimenté (apprenants de français) soit leur connaissance métalinguistique sur l'orthographe : le graphème <u> correspond non pas au phonème /u/ mais à /y/ (/u/ correspond à la séquence <ou>).

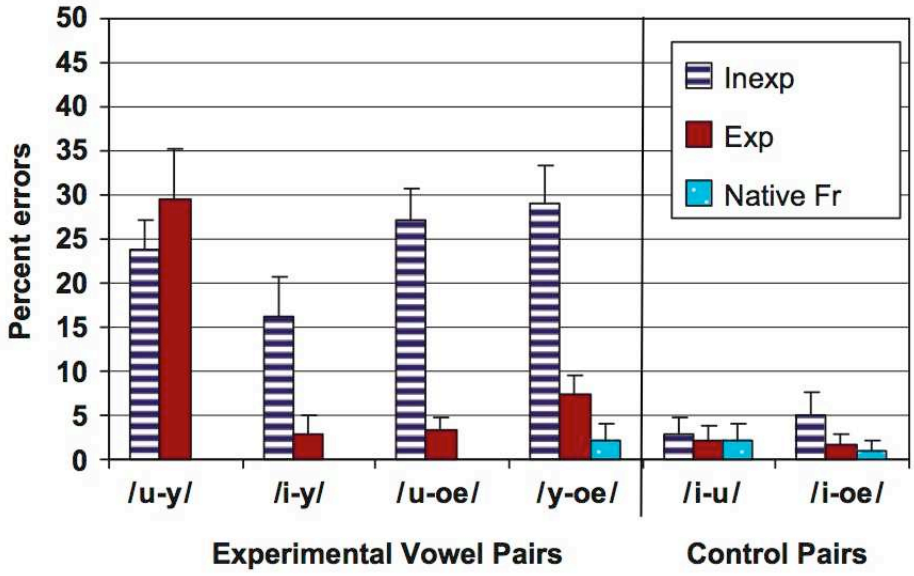



Fig. 2. Overall errors on experimental and control contrasts. Performance (percent errors) for Inexperienced (dark and light stripes) and Experienced (dark solid) and Native French (light dotted) groups for experimental and control vowel pairs (collapsed over context). Error bars represent standard errors.

**Figure 37** : Taux d'erreur (en pourcentage) du test de discrimination AXB pour les groupes non expérimenté (« *Inexp* »), expérimenté (« *Exp* »), et natif du français (« *Native Fr* ») dans les contextes /rab_p/ et /rad_t/ confondus. Les barres d'erreur représentent une erreur type. Levy et Strange (2008 : 149)

---

[346] W. Strange, E. S. Levy, R. Lehnholf. Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. *Journal of the Acoustical Society of America* 115(5) : 2606, 2004.

[347] E. S. Levy. Consonantal context and experience effects on American L2-learners' perception of French vowels. *Journal of the Acoustical Society of America* 119(5) : 3424, 2006.

[348] E. S. Levy, W. Strange. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36 : 141–157, 2008.

[349] E. S. Levy. Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America* 125(2) : 1138-1152, 2009.

Levy (2009) a montré les patterns d'assimilation perceptive des voyelles /y, œ, u, o, i, ɛ, a/ dans les contextes /rab_p/ et /rad_t/ par 3 groupes de 13 auditeurs de l'anglais américain (1. non apprenants – non expérimenté – ; 2. ceux qui ont reçu de l'instruction formelle – moyennement expérimenté – ; 3. ceux qui ont reçu de l'instruction formelle et de l'immersion – hautement expérimenté –). Parmi les résultats de cette expérience, notons les points suivants : 1) /u/ a été assimilé majoritairement à /uː/ dans tous les cas (75-91%), mais dans le groupe hautement expérimenté, le taux de réponse pour cette voyelle était plus élevé dans /rad_t/ (91%) que dans /rab_p/ (78%) ; 2) /y/ a été assimilé majoritairement à /juː/ (61-85%), mais plus dans /rab_p/ (72-85%) que dans /rad_t/ (61-71%) ; 3) /œ/ a été assimilé majoritairement à la voyelle /ʊ/ de l'anglais par le groupe hautement expérimenté dans les deux contextes (60% dans /rab_p/, 61% dans /rad_t/), alors que les réponses /uː/ étaient nombreuses dans le groupe non expérimenté (34% dans /rab_p/, 59% dans /rad_t/). Le point 1) suggère qu'il serait difficile de discerner les différences phonétiques entre le /u/ français et le /uː/ américain, et par conséquent, qu'il serait difficile d'établir la catégorie /u/, et cela notamment dans le contexte alvéolaire. La tendance observée dans le point 2) serait due au fait que la séquence /djuː/ n'existe pas en anglais américain à cause de la neutralisation de l'opposition /uː/-/juː/ (chute de /j/) après /d/, ce qui n'est pas le cas pour le contexte /b_/. Le point 3) suggère que les auditeurs anglophones apprennent à différencier /u/ (voyelle perçue comme proche de /uː/) et /œ/ (voyelle perçue comme proche de /ʊ/) avec l'apprentissage, ce qui corrobore le fait que le groupe expérimenté de Levy et Strange (2008) ont distingué /u/-/œ/ et /y/-/œ/ beaucoup plus correctement que le groupe non expérimenté (voir la Figure 37).

Quant à la production, Flege et Hillenbrand (1984)[350] ont analysé les formants des voyelles /u/ et /y/ prononcées deux fois dans un contexte dental (7 paires de syntagmes telles que « Tous les prêtres » /tu le pʀɛtʀ/ et « Tu les montres » /ty le mõtʀ/) par 7 locutrices natives du français et deux groupes de 7 locutrices natives de l'anglais américain parlant français (7 étudiantes qui avaient passé l'année universitaire précédente à Paris, et 7 enseignantes de français dans une université américaine). Ils ont trouvé que les locutrices américaines produisaient /u/ avec un F2 significativement plus élevé que les locutrices francophones, mais produisaient /y/ avec un F2 comparable à celui des locutrices natives. Le /y/ prononcé par les deux groupes de locutrices américaines a été identifié correctement en tant que tel dans une tâche de choix binaire forcé (/tu/ ou /ty/) par 7 locutrices natives du français dans 70% des cas environ, alors que le /u/ prononcé par le groupe plus expérimenté a été beaucoup mieux identifié (82,9%) que le /u/ prononcé par le groupe moins expérimenté (44,8%).

Rochet (1995)[351] a montré que le /y/ prononcé par 10 anglophones canadiens dans des monosyllabes qui contenaient différentes consonnes a été jugé par 3 locuteurs natifs du français comme [u] ou un vocoïde ressemblant à [u] (*[u]-like vowel*), ou un vocoïde entre [y] et [u] dans 92% des cas (8% seulement comme un vocoïde ressemblant à [i], ou un vocoïde entre [y] et [i] : notons

[350] J. E. Flege, J. Hillenbrand. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76(3) : 708-721, 1984.

[351] B. L. Rochet. "Perception and production of second-language speech sounds by adults", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 379-410, 1995.

que les locuteurs de certaines langues comme le portugais tendent à produire un vocoïde qui ressemble plutôt à [i]).

Lauret (1998)[352] a trouvé que les locuteurs natifs de l'anglais américain apprenant le français avaient tendance à produire pour /y/ une séquence similaire à /ju:/ de l'anglais américain, où le F2 descendait graduellement pendant la voyelle (Figure 38).

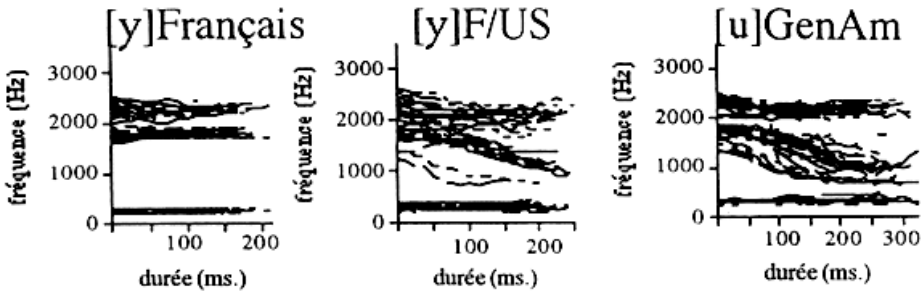



**Figure 38** : Mouvements des trois premiers formants de la voyelle /y/ dans des mots CV (C = consonne dentale) en français par 6 locuteurs francophones natifs (à gauche), en français par 6 locuteurs natifs de l'anglais américain (au milieu), et de la voyelle de l'anglais américain GA (*General American*) [u:] par 6 locuteurs natifs de l'anglais américain (à droite) ; 6 locuteurs x 3 occurrences = 18 réalisations pour chaque graphique (Lauret 1998 : 176). L'auteur utilise le symbole [u] pour la voyelle [u:] du GA.

Flege (1987)[353] a montré, comme dans l'une de ses études précédentes avec Hillenbrand (1984), que le F2 de la voyelle /y/ prononcée dans les mêmes contextes (ex. « Tu les montres » /ty le mõtʀ/) par des locuteurs natifs de l'anglais américain, expérimentés en français, n'était pas différent de celui du /y/ prononcé par les Français natifs. En revanche, le F2 du /u/ français prononcé par tous les locuteurs de l'anglais américain, quelle que soit l'expérience d'apprentissage, était significativement plus élevé que celui des francophones natifs[354]. En se fondant sur ces résultats, il distingue deux cas de figure différents : les phones « similaires » et les phones « nouveaux ».

Les phones L2 « similaires » diffèrent systématiquement d'un équivalent L1 facilement identifiable comme tel (Flege 1987 : 48)[355]. C'est le cas pour le /u/ du français (F2 bas, inférieur à

---

[352] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français*. Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[353] J. E. Flege. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15 : 47-65, 1987.

[354] Il faut néanmoins signaler que les deux voyelles n'ont pas été mises dans un contexte prosodique identique. /u/ était placé dans une syllabe qui aurait pu recevoir un accent d'insistance (ex. « tous les prêtres » /tu le pʀɛtʀ/), alors que /y/ était dans un mot grammatical, donc moins susceptible d'être mis en avant (ex. « tu les montres » /ty le mõtʀ/).

Une autre particularité du corpus de cette étude consiste en le fait que les deux voyelles /u/ et /y/ étaient placées après /t/. Comme nous l'avons vu ci-dessus (Lévy et Strange 2008), c'est ce contexte consonantique qui rend difficile la tâche de discrimination perceptive de cette paire de voyelles, par rapport au contexte labial.

[355] « *"Similar" L2 phones […] differ systematically from an easily identifiable counterpart in L1* ». (Flege 1987 : 48)

1000 Hz, regroupé avec F1) et le /uː/ de l'anglais (F2 plus élevé, et descendant). L'auteur donne également l'exemple de /t/ français et anglais dans le même article : le VOT (*voice onset time* : temps d'établissement de voisement) plus long en anglais qu'en français[356].

Les phones L2 « nouveaux », en revanche, n'ont pas d'équivalent en L1, et donc, par définition, diffèrent acoustiquement des phones trouvés en L1. C'est le cas du /y/ français pour les anglophones. La catégorie /y/ n'existe pas en anglais, bien que le phone [y] puisse apparaître comme allophone du /uː/ anglais (réalisation antériorisée dans un contexte alvéolaire). Les locuteurs natifs de l'anglais américain pourraient identifier, au début de l'apprentissage, le /y/ français comme /u/ (français) malgré les grandes différences acoustiques entre les réalisations des deux phonèmes. Cependant, le résultat de l'expérience de Flege (1987) suggère que les anglophones apprendraient à reconnaître que [y] n'est ni une réalisation d'une catégorie de l'anglais, ni une réalisation de la voyelle /u/ du français[357].

La comparaison de ces deux cas, représentés par ceux de l'acquisition du /u/ (phone « similaire ») et du /y/ (phone « nouveau ») par les locuteurs-auditeurs de l'anglais américain apprenant le français langue étragnère (FLE) nous permet de prédire qu'il sera plus difficile de produire d'une manière « authentique » (comme les natifs) les phones qui sont similaires mais en même temps différents par rapport aux équivalents dans le système phonémique de la L1 (ex. /u/), que les sons nouveaux, qui n'ont pas d'équivalent dans le système de la L1 (ex. /y/). Cette prédiction correspond aux hypothèses H2 et H3 du SLM (*Speech Learning Model* : modèle d'apprentissage de la parole) postulées par le même auteur (Flege 1995 : 239)[358].

> *« H3 Plus la dissimilarité phonétique perçue entre un son L2 et le son le plus proche de la L1 est grande, plus il est probable que les différences phonétiques entre les sons seront discernées. »*[359]

Nous pouvons appliquer cette hypothèse au cas de /u/ et /y/ français appris par des anglophones américains de la manière suivante : la dissimilarité phonétique entre le /y/ du français et le /uː/ de l'anglais américain (la distance F1/F2 et F2/F3, entre autres) est plus grande que celle qui se trouve entre le /u/ du français et le /uː/ de l'anglais américain ; par conséquent, il est plus probable que les différences phonétiques entre le /y/ français et le /uː/ de l'anglais

---

[356] Notons qu'il existe également des différences de « *F-pattern* » (qui se manifestent dans les formants, notamment le F2, durant la transition vers la voyelle suivante) dues à la différence de forme de la langue, comme nous l'avons vu dans la section 1.6.2. (Straka 1965, Vaissière 2006)

G. Straka. *Album phonétique*. Sainte-Foy, Les Presses de l'Université Laval, 1965.

J. Vaissière. Using an articulatory model as an integrated tool for a better understanding of the combined articulatory, acoustic and perceptual aspects of speech. Communication orale présentée au INAE Workshop on Image and Speech Processing, Chennai, Inde, 2006.

[357] « *"New" L2 phones have no counterpart in the L1 and so, by definition, differ acoustically from phones found in L1. For example, realizations of French /y/ are "new" phones for native speakers of English. Although [y] phones may sometimes occur on the phonetic surface of American English as an allophone of /u/ (as in [mjyzɪk] for "music"), English has no /y/ category. Native English speakers may at first identify French /y/ as /u/, despite the fact that it differs considerably in terms of second formant (F2) frequency from English /u/ (and to an even greater extent from French /u/ […]). However, an important assumption made here is that they will eventually come to recognize that [y] is not the realization of an English category (or French /u/).* » (Flege 1987 : 48)

[358] J. E. Flege. "Second language speech learning: Theory, findings, and problems". in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 233-277, 1995.

[359] « *H3 The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.* » (Flege 1995 : 239)

américain soient discernées que celles qui se trouvent entre le /u/ français et le /uː/ de l'anglais américain.

> *« H2 Une nouvelle catégorie phonétique peut être établie pour un son L2 qui diffère phonétiquement du son de la L1 le plus proche si les bilingues discernent au moins quelques-unes des différences phonétiques entre les sons de la L1 et de la L2. »*[360]

Ce qui relève de cette hypothèse est le suivant : /u/ et /y/ français diffèrent phonétiquement du /uː/ de l'anglais américain, et le /y/ français diffère plus du /uː/ américain que le /u/ français ; il est plus probable qu'une nouvelle catégorie /y/ est établie que celle du /u/ français, car les locuteurs anglophones discerneront plus probablement au moins quelques-unes des différences phonétiques entre le /y/ français et le /uː/ américain que celles qui se trouvent entre /u/ français et /uː/ américain.

Quant aux locuteurs japonophones apprenant le FLE, nous pouvons trouver des similitudes avec le cas des locuteurs anglophones américains concernant les voyelles /u/ et /y/ du français. Comme nous l'avons déjà vu, le /u/ [ɯ] japonais est une voyelle acoustiquement centrale, avec le F2 entre 1000 et 1500 Hz, et est donc phonétiquement différent des voyelles /u/ et /y/ du français, mais plus proche de ces deux voyelles que les autres voyelles japonaises /i e a o/. Parmi les deux voyelles françaises /u/ et /y/, c'est /y/ qui peut être considéré comme plus éloigné du /u/ [ɯ] japonais. Dans ce cas-là, les hypothèses du SLM de Flege prédisent, comme pour les locuteurs de l'anglais américain, plus de chance pour que la catégorie /y/ soit établie plutôt que la catégorie /u/.

Outre /u/ et /y/, nous devrons considérer le cas de /ø/ (et de /œ/). LeBel (1990)[361] définit 7 voyelles du français (/i E a O u y Œ/[362]) qui doivent être distinguées même pour assurer une communication minimale en français. Nous y trouvons non seulement /u/ et /y/ mais aussi /ø/ (/Œ/), ce qui nous amène à supposer que la distinction entre ces trois voyelles est indispensable pour produire des énoncés compréhensibles, quel que soit le but ultérieur de l'apprentissage. Étant donné que les autres voyelles /i E a O/ ne semblent pas présenter des différences majeures par rapport aux voyelles correspondantes du japonais (/i e a o/ : voir la section 1.4.1.), il nous conviendra de nous intéresser plus particulièrement aux voyelles /u/, /y/, et /ø/ (et /œ/)/.

Par ailleurs, il nous paraît important de considérer les ressemblances acoustiques entre le /ø/ français et le /u/ [ɯ] japonais dans le sens où les deux sont des voyelles acoustiquement centrales. Que se passera-t-il dans l'apprentissage de /ø/ par les apprenants japonophones ?

---

[360] « *H2 A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.* » (Flege 1995 : 239)

[361] J.-G. LeBel. *Traité de correction phonétique ponctuelle*. Québec, Centre international de Recherche en aménagement linguistique, 1990.

[362] /E O Œ/ sont des archi-phonèmes qui englobent /e ɛ/, /o ɔ/, et /ø œ/ respectivement, qui présentent une distribution semi-complémentaire (voir la section 1.4.).

Polivanov (1931 : 86)[363] a en effet observé la confusion des voyelles françaises /u y ø œ/ chez des japonophones, ainsi que chez des Tagals (locuteurs du tagalog ou tagal aux Philippines) :

> *« Les Tagals (peuple malais habitant l'île Luçon), parlant français – et certains possèdent cette langue d'une façon tout à fait courante – substituent leur représentation phonologique* u *à quatre phonèmes voyelles qualitativement différents du français :* u *(dans loup, coup),* y *(dans tu, lune),* ø *(dans deux),* œ *(dans peur). La même substitution a lieu dans la prononciation et les emprunts japonais de mots français. Ce fait s'explique par le système pauvre des vocalismes en cause (tagal et japonais), qui ne distinguent, au point de vue qualitatif, que les voyelles :* a, i, u, e, o. *» (Polivanov 1931 : 86)*

Cependant, en nous fondant sur les études antérieures présentées dans le présent chapitre, nous pouvons supposer que les phonèmes /u y ø (et œ)/ seront certes confondus par les apprenants japonophones, mais la difficulté de leur apprentissage et acquisition ne sera pas identique pour les trois. À notre connaissance, nous ne disposons pas de données obtenues dans des études expérimentales qui portent sur la production et la perception des voyelles orales du français par les apprenants japonophones, et qui puissent nous permettre d'éclaircir cette question.

### 1.9. HYPOTHÈSE DE NOTRE ÉTUDE

En nous fondant sur les études antérieures présentées ci-dessus, nous pouvons nous poser la question de savoir si les apprenants japonophones apprendront à prononcer le /y/ plus facilement que le /u/ français comme les locuteurs natifs de l'anglais américain, étant donné que la réalisation phonétique du /u/ [ɯ] japonais, équivalent phonémique du /u/ français, est assez éloignée de celle du /u/ français, et que le /y/ français n'a pas de phonème correspondant en japonais. Qu'en est-il du /ø/, qui serait considéré phonémiquement nouveau dans le sens où le système japonais n'a pas de voyelles antérieures arrondies, mais phonétiquement ressemblant au /u/ [ɯ] japonais ?

Nous pouvons postuler l'hypothèse suivante : les voyelles antérieures arrondies du français ainsi que l'opposition entre les voyelles mi-ouvertes et les voyelles mi-fermées seraient difficiles à percevoir et à produire pour les apprenants japonophones (au moins pour ceux qui sont au niveau élémentaire), et ce serait encore plus le cas concernant la voyelle /u/. Les voyelles « similaires » sont-elles véritablement plus difficiles à acquérir en production et en perception que les voyelles « nouvelles » ? Afin de vérifier cette hypothèse, une série d'expériences de production et de perception a été effectuée.

---

[363] E. Polivanov. La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague* 4 : 79-96, 1931.

# Deuxième partie : des études préliminaires sur l'apprentissage des dix voyelles orales isolées du français parisien par des apprenants japonophones.


## RÉSUMÉ

Le but de la série d'expériences présentées dans cette partie est de cerner les tendances générales de difficultés de perception et de production des voyelles orales du français par les apprenants japonophones.

1) Quatre locuteurs natifs du japonais de tokyo (deux hommes et deux femmes) ont prononcé six fois les cinq voyelles du japonais dans la phrase cadre « *sorewa* /V/ *to iimasu* » /sorewa ... to iRmasu/. Les quatre premiers formants mesurés présentent des tendances similaires à celles que l'on trouve dans les études précédentes (Mokhtari et Tanaka 2000, entre autres).

2) Quatre locuteurs natifs du français (deux hommes et deux femmes) ont prononcé trois fois les dix voyelles orales et les trois voyelles nasales du français dans la phrase cadre « *je dis* /V/ *comme dans …* ». Les quatre premiers formants mesurés présentent des tendances similaires à celles que l'on trouve dans les études précédentes (CALLIOPE 1989, entre autres).

3) Des apprenants japonophones ont passé une expérience d'identification (cinq auditeurs) sur les 13 voyelles isolées (y compris les 3 nasales) prononcées par 4 francophones natifs (deux hommes et deux femmes) dans une phrase-cadre du type « Je dis /u/ comme dans *loup* ». Les résultats du test d'identification montrent un taux de confusion élevé (41% de toutes les réponses, contre 17% auprès de 5 auditeurs francophones natifs). On observe des confusions concernant les voyelles nasales (33% de cas confusion) et entre les voyelles moyennes (28% de cas de confusion), mais aussi entre /u/, /y/, et /ø/ (17% de cas de confusion). Parmi les trois dernières voyelles, /y/ et /u/ (généralement moins arrondies quand elles sont prononcées par des locuteurs japonophones) ont rarement été confondus (1% seulement de cas de confusion). Notons que les locuteurs natifs de nombreuses langues qui n'ont pas de distinction entre les voyelles arrondies antérieures et postérieures ont tendance à confondre /u/ (généralement arrondi dans leur langue maternelle) avec /y/ (Lauret 2007), contrairement aux japonophones.

4) Afin d'étudier la production de 3 apprenants japonophones, les valeurs formantiques de 10 voyelles orales françaises isolées ont été mesurées (en lecture et en répétition) et elles ont été comparées à celles de 4 locuteurs francophones natifs (2 hommes et 2 femmes) mesurées de la même manière. Les déviations les plus importantes sont les suivantes ; 1) le F2 du /u/ est plus élevé (1000-1300 Hz pour les deux apprenants japonophones, 1300-1400 Hz pour l'autre apprenante), ce qui résulte en une voyelle acoustiquement proche du /ø/. 2) Deux apprenants ont produit des F2 et F3 proches pour le /y/, mais avec un F2 élevé (> 2000 Hz pour les hommes), tandis que l'apprenante a produit une diphtongue (baisse du F2 au cours de la voyelle) proche de la séquence /ju/ du japonais. 3) Comme attendu, les différences entre les mi-fermées et les mi-ouvertes sont moins bien marquées que chez les locuteurs natifs.

5) Les 10 voyelles orales du français prononcées par les 3 apprenants japonais ont été identifiées et évaluées par 26 francophones natifs. La voyelle /u/ prononcée en lecture par les 3 apprenants japonais a été reconnue comme /u/ dans 43 % des cas seulement (39 % comme /ø/, 13 % comme /y/), tandis que le /y/ a été reconnu en tant que tel dans 69 % des cas (14 % comme /u/, 7 % comme /ø/).

Ces expériences préliminaires de perception et de production illustrent les trois types de voyelles qui posent des difficultés aux apprenants japonophones : 1) les voyelles fermées arrondies /u/, /y/ et antérieure mi-fermée arrondie /ø/ ; 2) les voyelles mi-fermées et mi-ouvertes ; 3) les voyelles nasales. Ces résultats correspondent aux différences phonémiques et phonétiques qui existent entre les voyelles du français et du japonais.

Dans cette partie, nous présenterons cinq expériences préliminaires qui ont été effectuées afin de confirmer les tendances générales des difficultés de perception et de production des voyelles du français rencontrées par les japonophones apprenant l'ALE (L2) et le FLE (L3) :

1) Expérience 1 : Production des voyelles du japonais par des locuteurs japonophones natifs

2) Expérience 2 : Production des voyelles orales du français par des locuteurs francophones natifs

3) Expérience 3 : Identification des voyelles isolées du français par des japonophones apprenant le FLE

4) Expérience 4 : Production des voyelles orales du français par des japonophones apprenant le FLE

Tâche 1 : lecture
Tâche 2 : répétition immédiate

5) Expérience 5 : Identification et évaluation des voyelles orales produites pour l'Expérience 4 auprès d'auditeurs francophones natifs.

## 2.1. LES VOYELLES DU JAPONAIS DE TOKYO ET DU FRANÇAIS PARISIEN

### *2.1.1. Les 5 voyelles isolées du japonais prononcées par des locuteurs natifs du japonais de Tokyo (Expérience 1).*

Dans la section 1.4.1., nous avons présenté les données formantiques concernant les 5 voyelles du japonais d'après des études antérieures. Afin de vérifier les tendances observées, nous avons enregistré et analysé ces 5 voyelles du japonais.

Les 5 voyelles du japonais ont été prononcées par 4 locuteurs natifs du japonais de Tokyo (2 femmes et 2 hommes : âgés de 33 à 64 ans) dans le contexte d'une phrase-cadre (「それは～といいます」 « *sorewa* /V/ *to iimasu* » /sorewa ... to iRmasu/). Les phrases ont été présentées sous forme conventionnelle de l'écriture japonaise (la voyelle a été représentée par la lettre *katakana* correspondante, entre des signes de citation : ex. 「ウ」 /u/), une par une sur un écran d'ordinateur. Aux 5 phrases contenant les 5 voyelles isolées, 3 comportant des séquences avec /j/ (/ja/ /ju/ /jo/) ont été rajoutées, et les 8 phrases ont été arrangées dans un ordre semi-aléatoire préétabli (/jo/ /a/ /u/ /i/ /ju/ /e/ /o/ /ja/). Les locuteurs ont été invités à répéter la liste des phrases 6 fois. La production sonore a été enregistrée à la fréquence d'échantillonnage de 44,1 kHz (16 bits), et reconvertie ultérieurement à 22,05 kHz. Les données ont été segmentées manuellement, à l'aide d'un oscillogramme et d'un spectrogramme à bande large (la largeur de la fenêtre était fixée à 5 millisecondes) sous Praat[364] : les portions retenues comme voyelles sont celles pour lesquelles les formants supérieurs (notamment le F2) sont clairement visibles. Pour chacune des 5 voyelles cibles enregistrées, les 4 premiers formants ont été mesurés sur 3 points (déterminés manuellement : vers le début, au milieu, vers la fin) de la durée de la voyelle sous Praat, en

[364] P. Boersma, D. Weenink. *Praat: doing phonetics by computer* (logiciel). http://www.praat.org/

enregistrant les données sur un « *Log file* »[365] : en ce faisant, nous pouvons vérifier à la main les mesures des formants détectés automatiquement. Les points de mesure choisis pour le début étaient situés à 13,3% en moyenne du début de la voyelle (120 mesures ; l'écart-type était de 3,0%), ceux du milieu étaient situés à 49,9% en moyenne du début de la voyelle (120 mesures ; l'écart-type était de 0,6%), ceux de la fin, à 85,7% (120 mesures ; l'écart-type était de 3,5%). La valeur moyenne de ces 3 mesures et des 6 répétitions (18 mesures) a été retenue comme valeur formantique de chacun des 4 locuteurs. Notons que les voyelles isolées /i e a o u/ n'étaient diphtonguées dans aucune des occurrences.

Les résultats sont indiqués dans la Figure 39. Nous y remarquerons les points détaillés ci-dessous :

**/u/** : comme nous l'avons constaté dans les études antérieures (les données de Sugitô 1995[366], représentées à la Figure 6, et celles de Mokhtari et Tanaka 2000[367] dans les Figure 8 et Figure 9), les 4 locuteurs présentent tous une réalisation acoustiquement centrale, avec le F2 situé entre 1000 et 1600 Hz, sauf le locuteur JP1 (F2 : 990 Hz). Ils ont tous une valeur de F2 supérieure à celle de /o/, cette dernière étant inférieure à 1000 Hz chez les 4 locuteurs. Notons que le F2 du /o/ français est supérieur à celui du /u/ français.

**/o/** : la voyelle /o/ est ainsi celle qui présente le F2 le plus bas dans le système vocalique du japonais. On peut l'expliquer par un lieu de constriction plus postérieur et un degré plus important de labialisation que pour /u/ (voir les données d'Uemura 1990[368], représentées à la Figure 7). Il s'agit d'une voyelle focale, avec un regroupement des deux premiers formants.

**/a/** : le F2 est relativement bas chez le locuteur JP1, relativement élevé chez la locutrice JP3 (même en tenant compte du fait qu'il s'agit d'une femme (la distance entre F1 et F2 est grande en conséquence), intermédiaire chez JP2 et JP4. Ces valeurs de F2 suggérent que la langue se trouve plutôt dans une position postérieure chez JP1, antérieure chez JP3. Il s'agit d'une voyelle focale, avec un regroupement des deux premiers formants (plutôt [ɑ] que [a]), sauf chez JP3.

**/i/** : Vaissière (2007)[369] montre qu'il existe deux types différents de réalisation pour les voyelles antérieures fermées non-arrondies, selon le degré de rapprochement de la langue vers le palais. Les conséquences acoustiques seraient un regroupement de F3 et F4 (avec le rapprochement palatal : réalisation typique du /i/ français), d'un côté, et un F2 maximalement élevé (observé souvent avec le /i/ en anglais), de l'autre. Parmi nos 4 locuteurs japonais, JP1 et JP2 montrent un rapprochement de F3 et F4, tandis que F2 est très élevé chez JP4. Il est intéressant de noter que

---

[365] Pour les détails, voir le site : http://www.fon.hum.uva.nl/praat/manual/Log_files.html

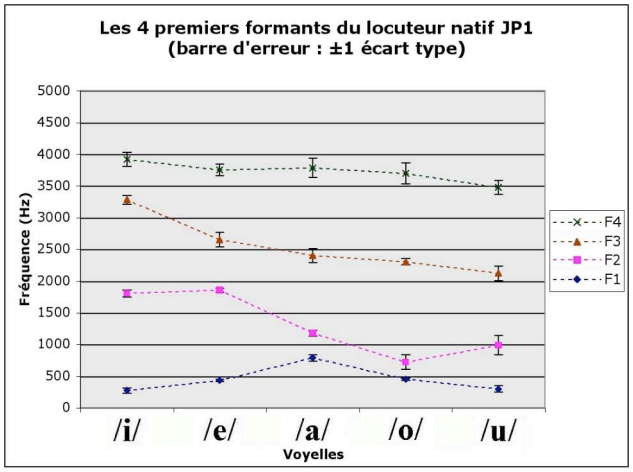
[366] M. Sugitô. *Ôsaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [Dictionnaire CD-ROM sonore d'accent des japonais d'Osaka et de Tokyo : commentaires]. Tokyo, Maruzen, 1995.

[367] P. Mokhtari, K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66, 2000.

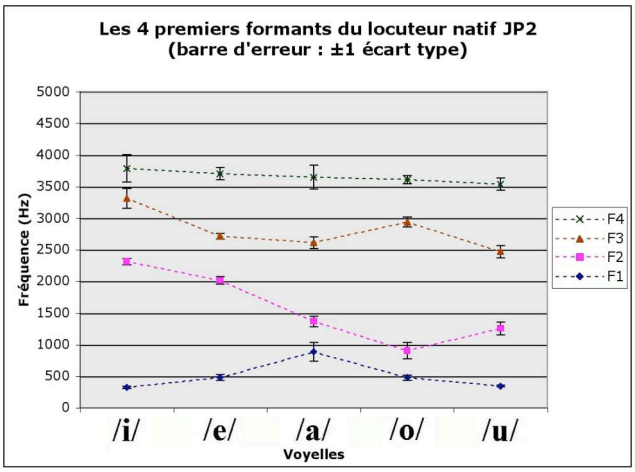
[368] Y. Uemura (Kokuritsu Kokugo Kenkyûjo [Institut national de la langue japonaise]). *Nihongo no boin, shiin, onsetsu: chouon undou no jikken-onseigakuteki kenkyuu* [Voyelles, consonnes et syllabes en japonais : étude en phonétique expérimentale sur les mouvements articulatoires]. Tokyo, Shûei shuppan, 1990.

[369] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

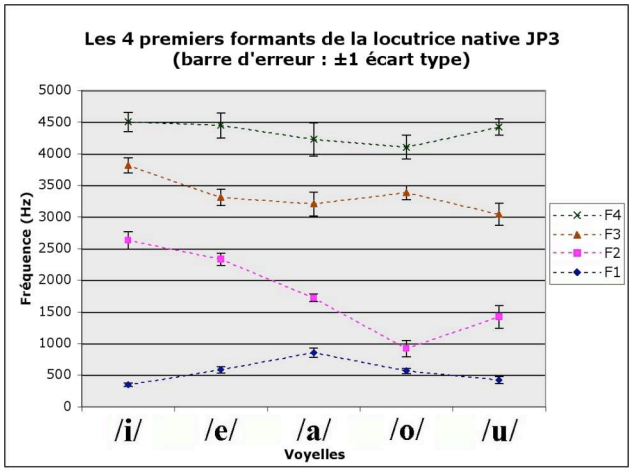
pour la réalisation de JP1, F2 se situe dans la même zone de fréquence (1800-1900 Hz) pour /i/ et /e/, et que ce sont les valeurs de F1 et de F3 (son rapprochement avec F4) qui différencient ces deux voyelles. En revanche, dans le cas de la locutrice JP4, la valeur de F2 est remarquablement différente pour /i/ et /e/.



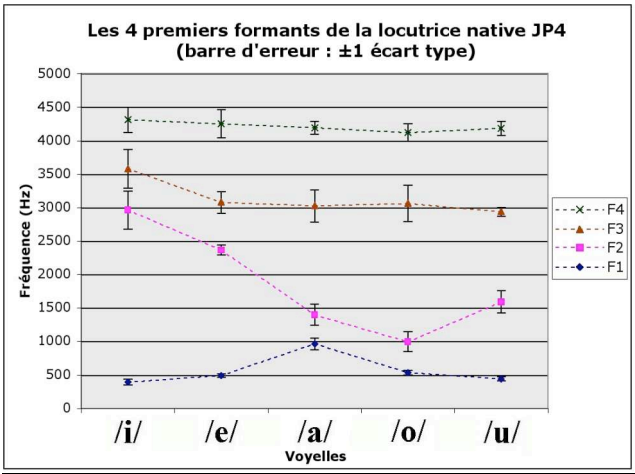

| JP1(H) | /i/ | /e/ | /a/ | /o/ | /u/ |
|---|---|---|---|---|---|
| F4 | 3915 | 3750 | 3781 | 3695 | 3476 |
| F3 | 3280 | 2653 | 2401 | 2304 | 2125 |
| F2 | 1805 | 1857 | 1182 | 721 | 990 |
| F1 | 276 | 433 | 788 | 453 | 299 |




| JP2(H) | /i/ | /e/ | /a/ | /o/ | /u/ |
|---|---|---|---|---|---|
| F4 | 3789 | 3705 | 3649 | 3610 | 3539 |
| F3 | 3315 | 2715 | 2613 | 2939 | 2474 |
| F2 | 2317 | 2017 | 1371 | 906 | 1259 |
| F1 | 326 | 481 | 887 | 478 | 346 |




| JP3(F) | /i/ | /e/ | /a/ | /o/ | /u/ |
|---|---|---|---|---|---|
| F4 | 4499 | 4446 | 4223 | 4101 | 4417 |
| F3 | 3811 | 3306 | 3204 | 3380 | 3039 |
| F2 | 2627 | 2329 | 1718 | 917 | 1419 |
| F1 | 347 | 582 | 856 | 566 | 422 |




| JP4(F) | /i/ | /e/ | /a/ | /o/ | /u/ |
|---|---|---|---|---|---|
| F4 | 4309 | 4250 | 4191 | 4118 | 4181 |
| F3 | 3577 | 3075 | 3022 | 3057 | 2937 |
| F2 | 2962 | 2364 | 1397 | 995 | 1590 |
| F1 | 389 | 488 | 962 | 534 | 438 |

**Figure 39 :** Les quatre premiers formants des 4 locuteurs natifs (2 hommes : JP1 et JP2, en haut, à gauche et à droite ; et 2 femmes : JP3 et JP4, en bas, à gauche et à droite) du japonais de Tokyo (valeur moyenne de 3 mesures x 6 répétitions). Les barres d'erreur représentent l'écart-type.

### *2.1.2. Les voyelles isolées du français prononcées par des locuteurs natifs (Expérience 2).*

Afin de vérifier les tendances observées dans les études antérieures présentées dans le chapitre 1.5., une étude de production a été effectuée. Dans cette étude, 4 locuteurs natifs du français du

nord de France (2 hommes et 2 femmes ; 3 de la région parisienne, 1 de Nantes ; âgés de 27 à 31 ans) ont prononcé les 10 voyelles orales françaises dans une phrase-cadre (« Je dis /V/ comme dans … » /ʒə di … kɔm dɑ̃ …/). L'enregistrement a été effectué avec un micro-casque dans la salle d'enregistrement de l'ILPGA (Institut de Linguistique et Phonétique Générales et Appliquées). Le taux d'échantillonnage était de 16000 Hz (16 bits). Pour chacune des voyelles cibles enregistrées, les 4 premiers formants ont été mesurés sur 5 sections (1 : du début jusqu'à un cinquième de la durée de la voyelle, 2 : le deuxième cinquième, … 5 : jusqu'à la fin) sur Praat (Figure 40). Parmi ces 5 valeurs, celles qui sont discontinues et aberrantes ont été éliminées après vérification du spectrogramme et/ou du spectre. Le taux de rejet était de 0,8% (20 valeurs sur 2400, dont 12 pour /u/ et 6 pour /i/ : notons que le /u/ français est caractérisé par un rapprochement de F1/F2, et le /i/, par un rappochement de F3/F4).

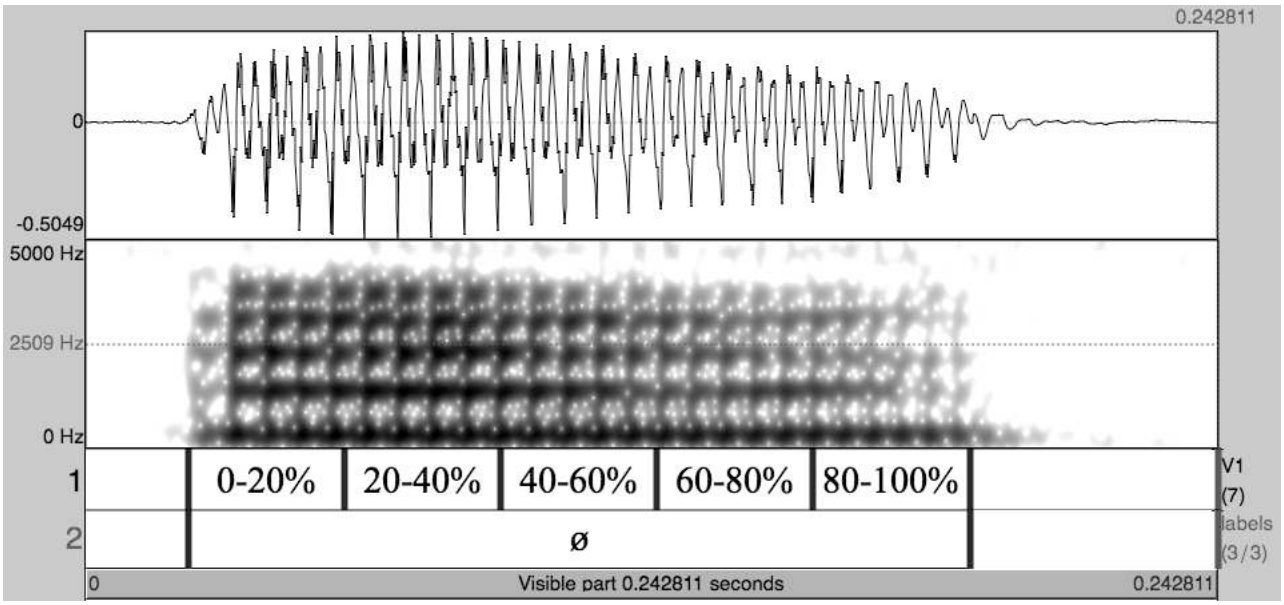



**Figure 40** : Exemple de cinq zones de mesure durant la voyelle : 1) du début jusqu'à un cinquième de la durée de la voyelle, 2) le deuxième cinquième, … 5) jusqu'à la fin. /ø/ prononcé par le locuteur natif FR1. Oscillogramme et spectrogramme à bande large (Praat ; largeur de fenêtre : 5 millisecondes).

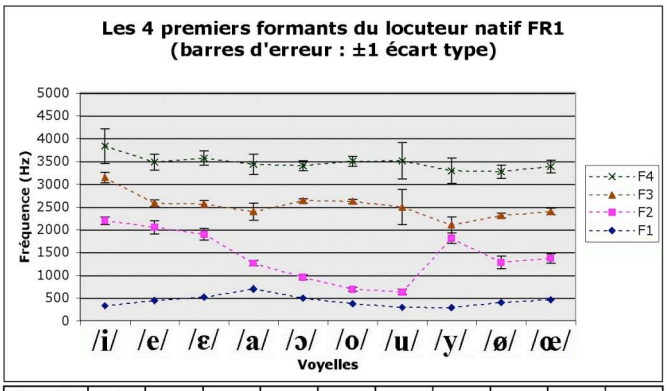
Comme on pouvait s'y attendre pour les voyelles du français, aucune n'était diphtonguée. La Figure 41 ci-dessous indique la valeur moyenne des quatre premiers formants pour les 3 répétitions de chaque locuteur, et les Figure 42 et Figure 43 représentent la moyenne des trois premiers formants des deux locuteurs masculins (FR1 et FR2) et des deux locutrices féminines (FR3 et FR4). Nous pouvons effectivement y retrouver les tendances décrites dans la section 1.5.2. (Figure 24, p. 81 ; Figure 25 ; p. 81) :

- regroupement de F1/F2 pour les voyelles postérieures /u o ɔ/ : les deux premiers formants sont les plus bas pour /u/.

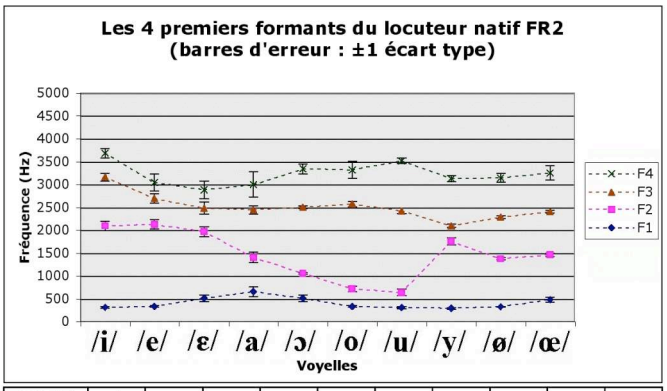
- regroupement de F2/F3 pour /y/.

- regroupement de F3/F4 pour /i/. Notons que les valeurs de F1 et F2 sont très similaires pour /i/ et /e/ chez le locuteur FR2. C'est F3 et F4 qui permettent de distinguer ces deux voyelles chez ce locuteur.

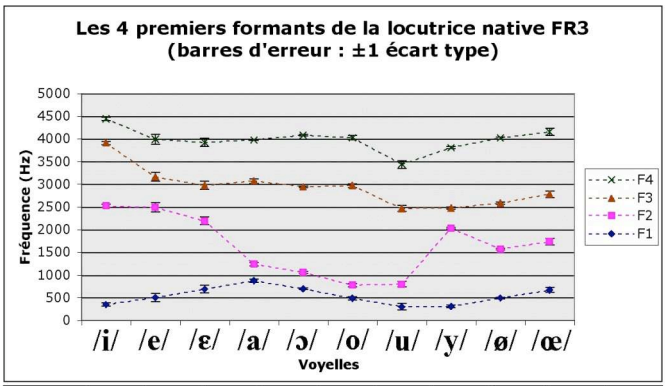
- distribution équilibrée de formants pour les voyelles « acoustiquement centrales » /ø œ/.



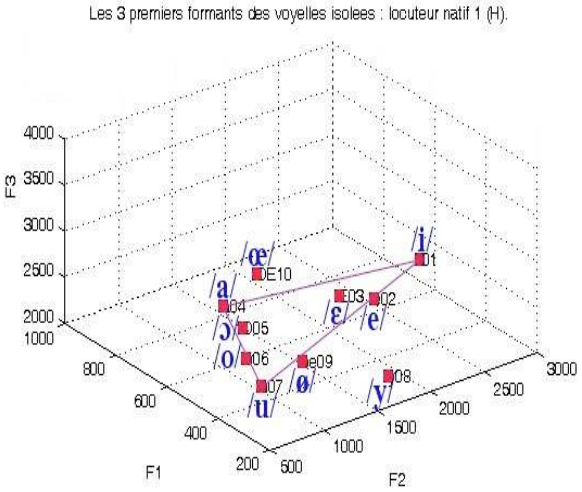

| FR1 (H) | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3829 | 3479 | 3569 | 3434 | 3401 | 3496 | 3503 | 3290 | 3271 | 3386 |
| F3 | 3140 | 2576 | 2556 | 2392 | 2629 | 2618 | 2489 | 2100 | 2304 | 2390 |
| F2 | 2194 | 2048 | 1891 | 1261 | 951 | 688 | 631 | 1806 | 1276 | 1366 |
| F1 | 323 | 441 | 512 | 699 | 495 | 372 | 289 | 284 | 396 | 463 |

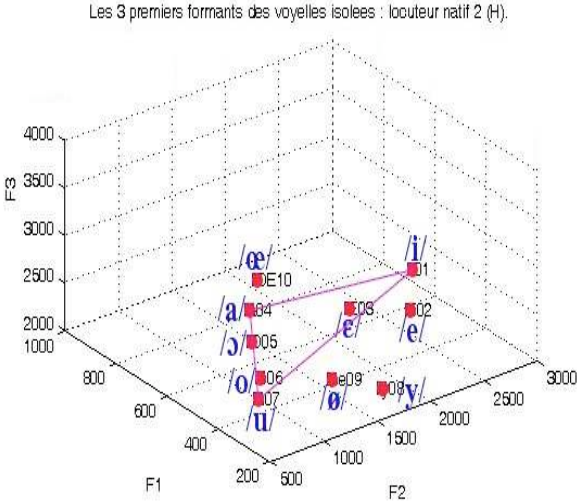


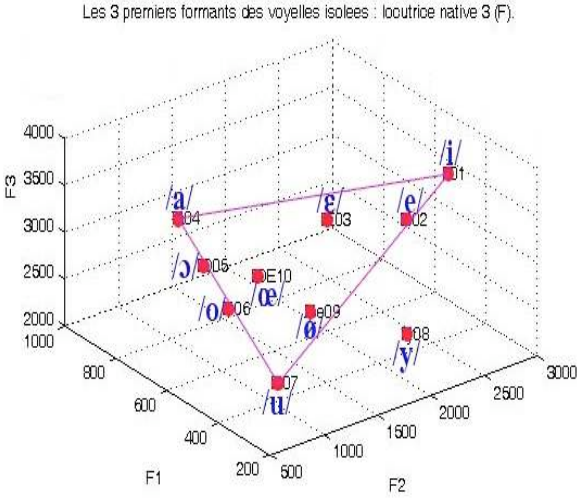

| FR2 (H) | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3679 | 3041 | 2879 | 2993 | 3336 | 3317 | 3514 | 3128 | 3144 | 3251 |
| F3 | 3160 | 2691 | 2478 | 2444 | 2501 | 2565 | 2416 | 2089 | 2282 | 2401 |
| F2 | 2091 | 2123 | 1964 | 1396 | 1053 | 715 | 637 | 1749 | 1372 | 1461 |
| F1 | 312 | 332 | 505 | 654 | 505 | 330 | 304 | 288 | 324 | 479 |

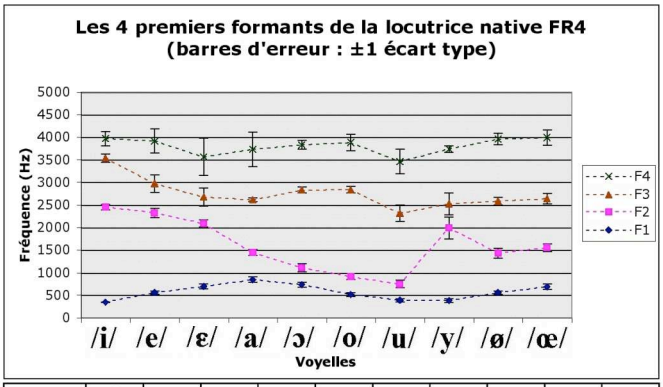



| FR3 (F) | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 4439 | 3985 | 3917 | 3977 | 4078 | 4024 | 3435 | 3804 | 4012 | 4149 |
| F3 | 3903 | 3162 | 2972 | 3071 | 2940 | 2968 | 2462 | 2471 | 2583 | 2778 |
| F2 | 2519 | 2488 | 2190 | 1238 | 1051 | 776 | 798 | 2032 | 1570 | 1733 |
| F1 | 347 | 500 | 686 | 869 | 694 | 481 | 302 | 306 | 490 | 669 |






| FR4 (F) | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3967 | 3914 | 3562 | 3731 | 3827 | 3878 | 3457 | 3733 | 3954 | 3987 |
| F3 | 3536 | 2966 | 2667 | 2612 | 2826 | 2835 | 2311 | 2522 | 2583 | 2635 |
| F2 | 2451 | 2321 | 2088 | 1441 | 1108 | 912 | 744 | 1989 | 1431 | 1551 |
| F1 | 343 | 558 | 696 | 845 | 735 | 515 | 385 | 378 | 556 | 689 |

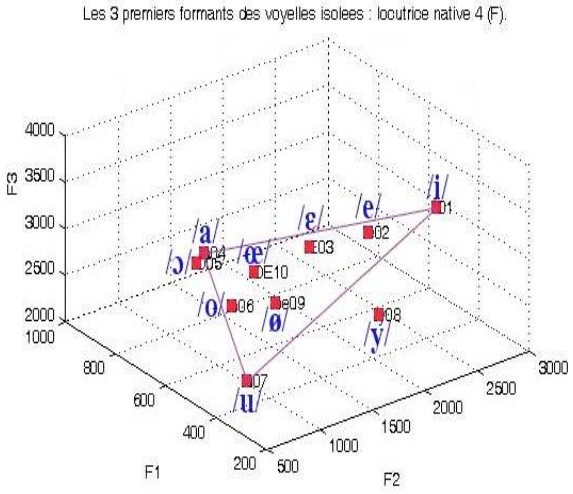



**Figure 41 :** Formants des voyelles orales du français en Hertz (5 mesures par voyelle x 3 répétitions). Locuteurs FR1-4 (FR1-2 : hommes ; FR3-4 : femmes) du haut en bas. Les quatre premiers formants (à gauche : les barres d'erreur représentent l'écart-type) et les trois premiers formants dans un triangle vocalique à trois dimensions (à droite).

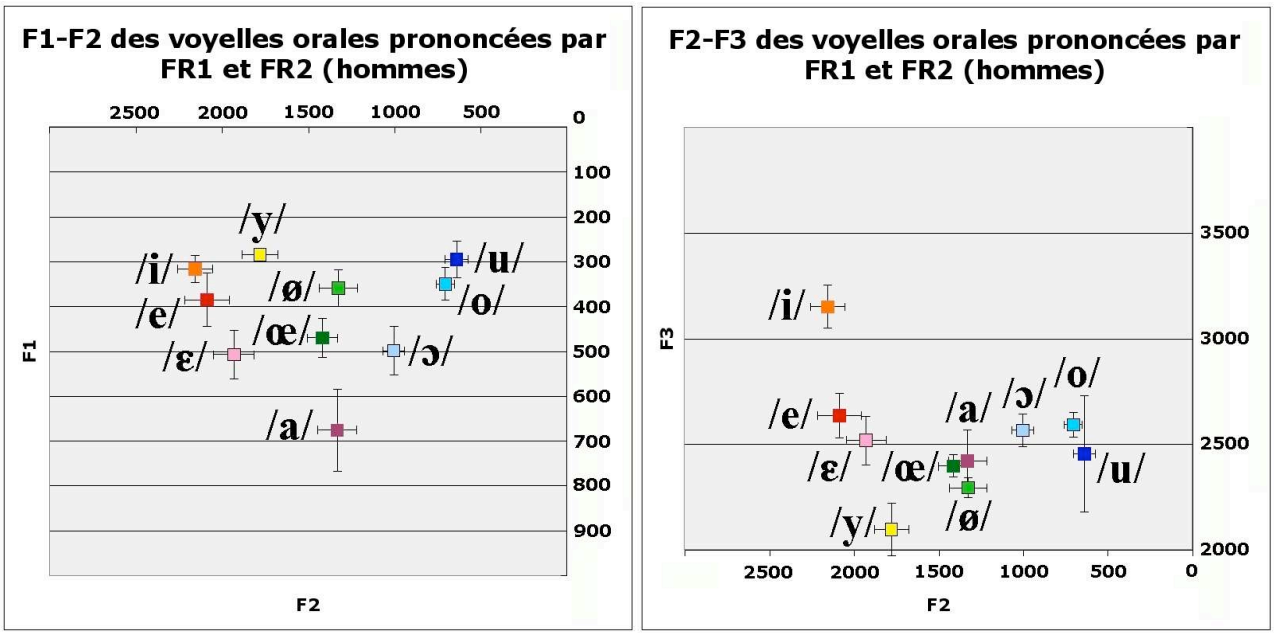



**Figure 42** : Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par les locuteurs francophones natifs FR1 et FR2 (hommes) dans la phrase cadre « *je dis* /V/ *comme dans* … ». F1 sur l'axe vertical et F2 sur l'axe horizontal (à gauche) ; F2 sur l'axe horizontal et F3 sur l'axe vertical (à droite). 5 mesures par voyelle x 3 répétitions x 2 locuteurs. Les barres d'erreur représentent l'écart-type.

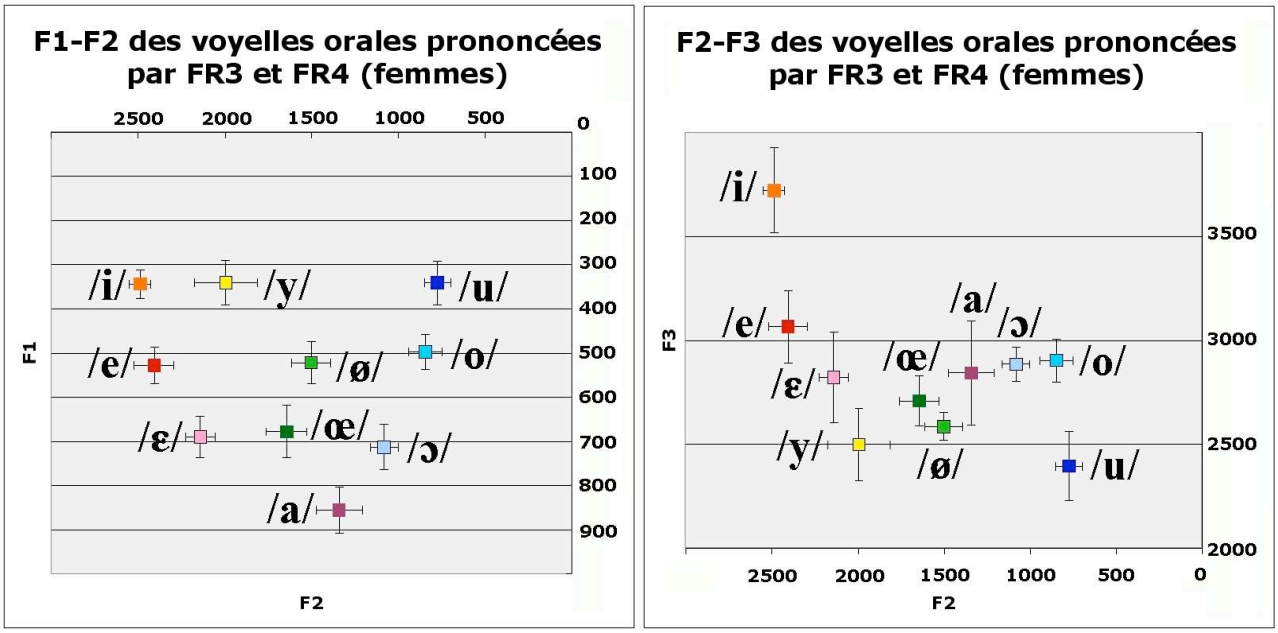



**Figure 43** : Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par les locuteurs francophones natifs FR3 et FR4 (femmes) dans la phrase cadre « *je dis* /V/ *comme dans* … ». F1 sur l'axe vertical et F2 sur l'axe horizontal (à gauche) ; F2 sur l'axe horizontal et F3 sur l'axe vertical (à droite). 5 mesures par voyelle x 3 répétitions x 2 locuteurs. Les barres d'erreur représentent l'écart-type.

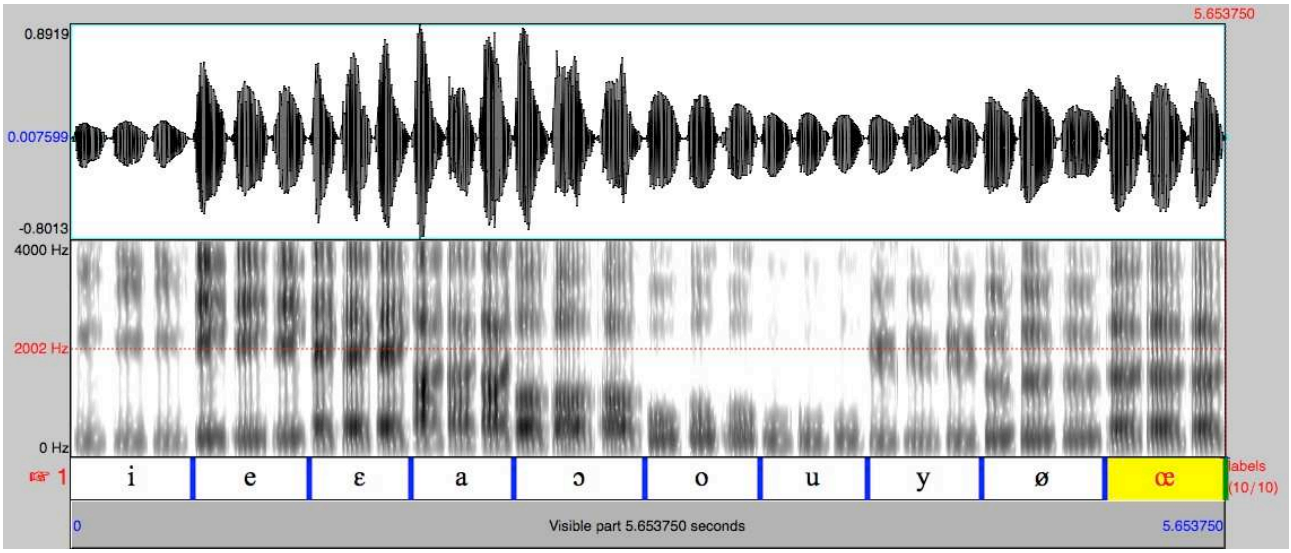



**Figure 44 :** Spectrogramme à bande large (Praat ; largeur de fenêtre : 5 millisecondes) des dix voyelles orales du français prononcées par le locuteur natif FR1 dans la phrase-cadre « *Je dis* /V/ *comme dans* … » (3 répétitions). Notons également l'intensité intrinsèque : les voyelles ouvertes sont plus intenses que les voyelles fermées.

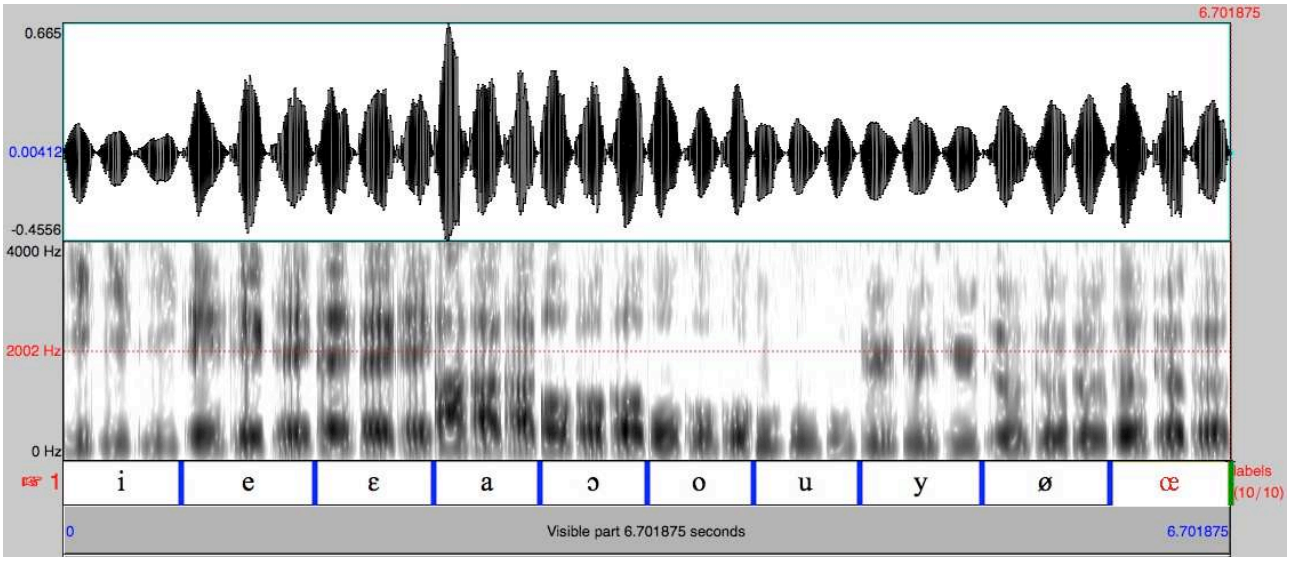



**Figure 45 :** Spectrogramme à bande large (Praat ; largeur de fenêtre : 5 millisecondes) des dix voyelles orales du français prononcées par le locuteur natif FR2 dans la phrase-cadre « *Je dis /V/ comme dans …* » (3 répétitions).

## 2.2. ANALYSE PERCEPTIVE : PERCEPTION DES VOYELLES DU FRANÇAIS PAR DES APPRENANTS JAPONOPHONES : IDENTIFICATION DES VOYELLES ISOLÉES (EXPÉRIENCE 3)[370]

Après avoir éclairci les caractéristiques acoustiques des voyelles orales du français à travers la littérature et nos expériences personnelles, nous avons utilisé les enregistrements obtenus dans la première partie (Expérience 2) pour examiner la perception de ces voyelles par des apprenants japonophones. Cette expérience de perception consistait en un test d'identification de toutes les voyelles isolées, dans lequel les apprenants devaient identifier chaque voyelle présentée isolément. Elle a été effectuée auprès d'un nombre de sujets relativement limité afin de cerner la tendance générale de confusion. Une autre expérience de perception (discrimination AXB), présentée dans la troisième partie, examinera des paires de voyelles.

### *2.2.1. Les stimuli*

Les stimuli consistaient en 13 voyelles françaises isolées (10 orales + 3 nasales /ɑ̃ ɔ̃ ɛ̃/) extraites des phrases-cadre « Je dis /V/ comme dans … », produites par les 4 locuteurs francophones natifs pour l'Expérience 2 (Figure 41 à la page 123, Figure 44 et Figure 45 à la page 125). Nous avons extrait 2 répétitions de chacune des 13 voyelles cibles (ex. /i/ isolé dans « Je dis /i/ comme dans lit »). Comme chacun des 4 locuteurs avaient mis une courte pause avant et après la voyelle cible, le début et la fin du fichier extrait ont été fixés durant ces pauses. Nous avons ainsi obtenu 104 stimuli (13 voyelles x 4 locuteurs x 2 répétitions). La durée moyenne des voyelles était de 160 ms (Locuteur 1), 178 ms (Locuteur 2), 256 ms (Locutrice 3), et 134 ms (Locutrice 4).

[370] Une partie des résultats de cette expérience a été publiée dans :

T. Kamiyama. La production et la perception des voyelles orales françaises par les apprenants japonophones. *Actes des XXVIèmes Journées d'Étude sur la Parole*, 49-52, 2006.

### *2.2.2. Les procédures*

L'expérience comportait trois étapes. Tout d'abord, les sujets ont bénéficié d'une période d'habituation. Ils ont eu la possibilité d'écouter des mots qui contenaient les 13 voyelles concernées (« lit », « les », « lait », « la », « lors », « l'eau », « loup », « lu », « leu », « leur », « lin », « lent », et « long »). Ces mots ont été prononcés par la locutrice FR4 de l'Expérience 2 (Figure 41, p. 123). Chaque mot correspondait à un bouton sur l'écran sur lequel étaient notées la transcription phonémique de la voyelle et l'orthographe du mot. Les sujets pouvaient écouter les mots autant de fois qu'ils le voulaient, en cliquant sur les boutons (Figure 46, à gauche : Microsoft PowerPoint a été utilisé pour présenter l'image et les fichiers son).

Deuxièmement, ils sont passés à l'entraînement. Ils ont écouté les 13 voyelles isolées prononcées par la même locutrice (FR4), dans un ordre semi-aléatoire préalablement établi. Il s'agissait des stimuli qui n'étaient pas inclus dans le test. Ils ont écouté les voyelles isolées précédées par un bip sonore (440 Hz, -6dB par rapport au plafond de l'enregistrement), dont la durée était de 50 ms, placé au moins 500 ms avant chaque stimulus), et choisi la voyelle qu'ils pensaient avoir perçue en cliquant sur la case correspondant à la réponse (Figure 46, à droite).

Troisièmement, ils ont passé le test. À cette étape ainsi que dans l'entraînement, la tâche des auditeurs consistait à écouter la voyelle (précédée par un bip sonore, comme dans l'entraînement) et à répondre à la question « Quelle voyelle entendez-vous » en cliquant sur la case correspondant à la réponse (Figure 46, à droite). Les stimuli ont été présentés dans un ordre semi-aléatoire pour que la même voyelle, ou les stimuli produits par le même locuteur ne se succèdent pas. Les auditeurs ont eu une pause au milieu du test. L'interface utilisée pour l'entraînement et le test est celle de Praat (version 3 de « Experiment MFC »).

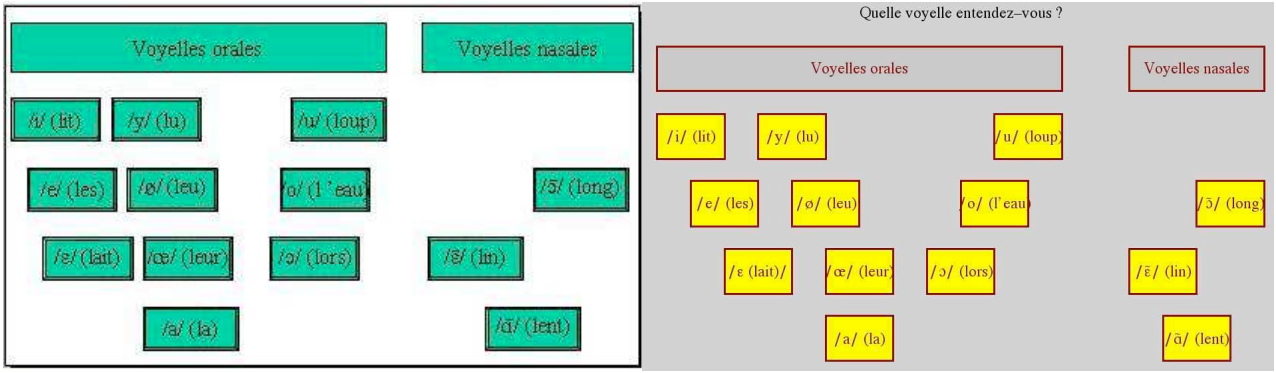



**Figure 46 :** Copie des écrans présentés aux sujets-auditeurs pendant l'habituation (à gauche), et pendant l'entraînement et le test (à droite).

### *2.2.3. La perception d'auditeurs natifs*

Dans un premier temps, nous avons effectué cette expérience auprès de 5 auditrices natives du français afin de vérifier la perception des voyelles isolées chez les natifs. Il est possible que les auditeurs natifs ne perçoivent pas nécessairement ce que d'autres locuteurs natifs ont pensé produire, car il ne s'agit que d'une voyelle isolée prononcées dans une phrase cadre (N.B. la

phrase-cadre était « je dis /V/ comme dans … »[371]). Les résultats de cette expérience nous serviront ainsi comme références quand nous analyserons les comportements des apprenants.

Cinq locutrices natives du français ont participé à l'expérience. Les auditeurs sont originaires de la France métropolitaine.

Le Tableau 20 ci-dessous montre la matrice de confusion. Le taux de confusion était de 17%. Ce taux de confusion peut paraître étonnant, mais considérant le résultat de Gottfried (1984)[372], nos données semblent raisonnables. Cet auteur a effectué un test d'identification de 11 voyelles orales du français (y compris le /ɑ/ postérieur) auprès de 16 auditeurs natifs du français parisien et 8 auditeurs anglophones américains. La figure 1 de son article (Gottfried 1984 : 95) montre que le taux d'erreurs des auditeurs francophones était de 20% environ pour le contexte isolé, ce qui est inférieur à celui des auditeurs anglophones (c. 33-34%).

Parmi les cas de confusion, les stimuli des voyelles nasales étaient responsables de 38% de toutes les confusions, tandis que des confusions entre les voyelles moyennes (/e/-/ɛ/, /o/-/ɔ/, /ø/-/œ/) représentaient 24%.

**Tableau 20** : Résultats du test d'identification des voyelles isolées auprès de 5 auditeurs natifs du français, exprimés en nombre de réponses d'identification et en pourcentage entre parenthèses. 40 réponses = 8 stimuli par voyelle (2 hommes et 2 femmes x 2 répétitions) x 5 auditeurs francophones natifs). Les stimuli sont représentés sur l'axe vertical, les réponses, sur l'axe horizontal.

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ | ɛ̃ | ɑ̃ | õ | total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 40 (100) | | | | | | | | | | | | | 40 (100) |
| e | 4 (10) | 31 (78) | 5 (13) | | | | | | | | | | | 40 (100) |
| ɛ | | 1 (3) | 39 (98) | | | | | | | | | | | 40 (100) |
| a | | | | 32 (80) | | | | | 2 (5) | 4 (10) | | 1 (3) | 1 (3) | 40 (100) |
| ɔ | | | | 5 (13) | 27 (68) | 1 (3) | | | | 4 (10) | | 2 (5) | 1 (3) | 40 (100) |
| o | | | | | | 35 (88) | 5 (13) | | | | | | | 40 (100) |
| u | | | | | | | 39 (98) | | | | | | 1 (3) | 40 (100) |
| y | | | | | | | 1 (3) | 39 (98) | | | | | | 40 (100) |
| ø | | | | | | | | | 40 (100) | | | | | 40 (100) |
| œ | | 1 (3) | | | | | | | 14 (35) | 24 (60) | | 1 (3) | | 40 (100) |
| ɛ̃ | | | | 1 (3) | 2 (5) | 1 (3) | | | 2 (5) | | 31 (78) | 3 (8) | | 40 (100) |
| ɑ̃ | | | | | 7 (18) | 3 (8) | | | | | | 27 (68) | 3 (8) | 40 (100) |
| õ | | | | | 1 (3) | 9 (23) | 1 (3) | | | | | | 29 (73) | 40 (100) |
| total_réponses | 44 (110) | 33 (83) | 44 (110) | 38 (95) | 37 (93) | 49 (123) | 46 (115) | 39 (98) | 58 (145) | 32 (80) | 31 (78) | 34 (85) | 35 (88) | |

[371] Gottfried et Strange (1980) ont montré que 10 voyelles de l'anglais américain ont été identifiées plus correctement dans des contextes consonantiques divers (/pVp/, /bVp/, /kVk/, /kV/, /Vk/) qu'en contexte isolé. Notons au passage que les voyelles relâchées (*lax*) de l'anglais ne peuvent apparaître en syllabes ouvertes (et donc en contexte isolé).

T. L. Gottfried, W. Strange. Identification of coarticulated vowels. *Journal of the Acoustical Society of America* 68 : 1626-1635, 1980.

[372] T. L. Gottfried. Effects of consonant context on the perception of French vowels. *Journal of Phonetics 12* : 91-114, 1984.

En ce qui concerne les confusions entre les voyelles moyennes, notons que, dans la langue, /ɔ/ et /œ/ ne se prononcent jamais seuls, mais uniquement suivies d'une ou plusieurs consonnes (Wioland 2005 :108 [373], entre autres). Cette contrainte nous permettra d'expliquer les taux relativement faibles d'identification de ces deux voyelles : la catégorisation des auditrices francophones aurait été orientée vers les voyelles qui peuvent apparaître dans le contexte isolé (/a/ et /o/ au lieu de /ɔ/, /ø/ au lieu de /œ/, sauf les 4 réponses /œ/ pour des stimuli /ɔ/).

Les 4 réponses de /œ/ pour /ɔ/ pourraient s'expliquer par l'antériorisation du /ɔ/, un phénomène remarqué par André Martinet dès les années 1950 et décrit dans un article dont le titre l'illustre astucieusement (« C'est jeuli le Mareuc ! »)[374]. Cette tendance a été confirmée dans des données plus récentes recueillies dans le cadre du projet « Phonologie du Français Contemporain[375] » : le F2 du /O/ dans le mot « connaît », par exemple, est de 1500 Hz dans le nord contre 1100 Hz dans le sud (une phrase lue par 3 hommes et 3 femmes de 3 générations différentes de chacune de 3 points d'enquête dans le nord[376] et le sud[377], respectivement : Woehrling et Boula de Mareüil 2006 : 94[378]), ce qui suggère une antériorisation, sans exclure la possibilité de la délabialisation[379].

---

[373] F. Wioland. *La vie sociale des sons du français*. Paris, l'Harmattan, 2005.

[374] A. Martinet. C'est jeuli le Mareuc ! *Romance Philology* 11 : 345-355, 1958. (Réédité dans A. Martinet, *Le Français sans Fard*. Paris, Presses Universitaires de France, 1969.)

Notons en passant que le /ʌ/ long du coréen est prononcé à un lieu d'articulation antérieur [ə:] par rapport à son équivalent bref [ʌ] dans la prononciation traditionnelle de Séoul.

[375] Projet coordonné par Jacques Durand, Chantal Lyche, et Bernard Laks.

J. Durand, B. Laks, C. Lyche. Le projet 'Phonologie du français contemporain'. *La tribune internationale des langues vivantes* 33 : 3-9, 2003.

http://www.projet-pfc.net/

[376] Brécey (Normandie), Treize-vents (Vendée), le canton de Vaud (Suisse romande).

[377] Biarritz (Pays Basque), Douzens (Languedoc) et Marseille (Provence).

[378] C. Woehrling, P. Boula de Mareüil. Identification d'accents régionaux en français : perception et catégorisation. *Bulletin PFC (Phonologie du Français Contemporain : Usages, Variétés et Structure) numéro 6 (Prosodie du français contemporain)* : 89-102, 2006.

[379] Les voyelles antérieures arrondies et les voyelles postérieures non-arrondies sont des voyelles « acoustiquement centrales », caractérisées par un second formant situé vers 1500 Hz (Vaissière 2006 : 73). Cette tendance a été constatée dans l'Expérience 9 de notre étude, présentée dans la troisième partie, et dans Kamiyama (2009).

J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France. 2006.

T. Kamiyama. Acoustic similarities between front rounded and back unrounded vowels as evidenced by French /ø/ and /u/ produced by Japanese-speaking learners. *The Journal of the Acoustical Society of America* 125(4) : 2756, 2009.

**Tableau 21** : Résultats du test d'identification auprès de 5 auditeurs francophones natifs, exprimés en nombre de réponses d'identification et en pourcentage entre parenthèses. Les voyelles mi-fermées et mi-ouvertes ont été regroupées (80 réponses pour les voyelles mi-fermées et mi-ouvertes, 40 pour les autres). Les stimuli sont représentés sur l'axe vertical, les réponses, sur l'axe horizontal.

| stimuli\réponses | i | e/ɛ | a | ɔ/o | u | y | ø/œ | ɛ̃ | ɑ̃ | õ | total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **i** | 40 (100) | | | | | | | | | | 40 (100) |
| **e/ɛ** | 4 (5) | 76 (95) | | | | | | | | | 80 (100) |
| **a** | | | 32 (80) | | | | 6 (15) | | 1 (3) | 1 (3) | 40 (100) |
| **ɔ/o** | | | 5 (6) | 63 (79) | 5 (6) | | 4 (5) | | 2 (3) | 1 (1) | 80 (100) |
| **u** | | | | | 39 (98) | | | | | 1 (3) | 40 (100) |
| **y** | | | | | 1 (3) | 39 (98) | | | | | 40 (100) |
| **ø/œ** | | 1 (1) | | | | | 78 (98) | | 1 (1) | | 80 (100) |
| **ɛ̃** | | | 1 (3) | 3 (8) | | | 2 (5) | 31 (78) | 3 (8) | | 40 (100) |
| **ɑ̃** | | | | 10 (25) | | | | | 27 (68) | 3 (8) | 40 (100) |
| **õ** | | | | 10 (25) | 1 (3) | | | | | 29 (73) | 40 (100) |
| total_réponses | 44 | 77 | 38 | 86 | 46 | 39 | 90 | 31 | 34 | 35 | |

Le Tableau 21 montre le résultat après avoir regroupé les stimuli et les réponses des voyelles moyennes (le résultat est représenté en pourcentage). Nous y voyons maintenant que les confusions sont concentrées principalement sur les stimuli des voyelles nasales, et des voyelles /a/, /ɔ/ et /o/. Outre les nasales, il ne reste qu'un seul cas où le taux de confusion était supérieur à 10 % (/a/ perçu comme /ø/ ou /œ/).

### *2.2.4. Les apprenants japonophones*

Par la suite, nous avons effectué la même expérience auprès de 5 auditeurs natifs du japonais apprenant le français langue étrangère dans un contexte exolingue (à l'Université de Tokyo, Japon). Leur expérience d'apprentissage variait de 3 mois à 2 ans, mais tous avaient appris le français au Japon auprès d'enseignants natifs et non-natifs (japonophones), et par conséquent, ils n'avaient pas eu beaucoup « d'input » sonore dans leur vie quotidienne.

**Tableau 22** : Résultats du test d'identification des voyelles isolées, exprimés en nombre de réponses d'identification et en pourcentage entre parenthèses.
40 réponses : 8 stimuli par voyelle (2 hommes et 2 femmes x 2 répétitions) x 5 auditeurs japonophones apprenant le français). Les stimuli sont représentés sur l'axe vertical, les réponses, sur l'axe horizontal.

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ | ɛ̃ | ɑ̃ | õ | total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 39 (98) | 1 (3) | | | | | | | | | | | | 40 (100) |
| e | 5 (13) | 16 (40) | 16 (40) | | | | | | 1 (3) | 1 (3) | 1 (3) | | | 40 (100) |
| ɛ | | 15 (38) | 24 (60) | | | | | | | 1 (3) | | | | 40 (100) |
| a | | | | 34 (85) | | | | | 1 (3) | 2 (5) | | 3 (8) | | 40 (100) |
| ɔ | | | | 1 (3) | 24 (60) | 6 (15) | | | 2 (5) | 3 (8) | 1 (3) | 3 (8) | | 40 (100) |
| o | | | | | 6 (15) | 23 (58) | 8 (20) | | 3 (8) | | | | | 40 (100) |
| u | | | | | | | 22 (55) | 2 (5) | 12 (30) | | | | 4 (10) | 40 (100) |
| y | | | | | | | | 33 (83) | 6 (15) | 1 (3) | | | | 40 (100) |
| ø | | | | | | | 12 (30) | 5 (13) | 17 (43) | 6 (15) | | | | 40 (100) |
| œ | | | | | | | 1 (3) | 1 (3) | 11 (28) | 24 (60) | 1 (3) | 2 (5) | | 40 (100) |
| ɛ̃ | | | | 5 (13) | | | | | | 2 (5) | 21 (53) | 12 (30) | | 40 (100) |
| ɑ̃ | | | | | 8 (20) | 2 (5) | | | | | 3 (8) | 14 (35) | 13 (33) | 40 (100) |
| õ | | | | | 3 (8) | 5 (13) | | | 3 (8) | 1 (3) | 5 (13) | 8 (20) | 15 (38) | 40 (100) |
| total_réponses | 44 (110) | 32 (80) | 40 (100) | 40 (100) | 41 (103) | 36 (90) | 43 (108) | 41 (103) | 56 (140) | 41 (103) | 32 (80) | 42 (105) | 32 (80) | |

Les résultats sont présentés dans la matrice de confusion ci-dessus (Tableau 22). Le taux global de confusion était de 41% de toutes les réponses (contre 17% auprès des 5 auditeurs francophones natifs). Nous pouvons y observer trois tendances principales :

**1)** Comme attendu, les voyelles nasales ont été largement confondues : les stimuli des 3 voyelles nasales /ɛ̃ ɑ̃ õ/ ont été responsables de 33% des confusions.

**2)** Les voyelles mi-ouvertes et mi-fermées (/e/-/ɛ/, /o/-/ɔ/, /ø/-/œ/) ont été également confondues presque autant (28% des confusions). /œ/ a été perçu comme /ø/ plus souvent (11 cas sur 40) que /ø/ a été perçu comme /œ/ (6 cas sur 40)[380]. Cette tendance est commune aux auditeurs francophones (14 confusions sur 40 pour /œ/ perçu comme /ø/, mais aucune dans le sens inverse), mais /o/ et /ɔ/ ont été confondus dans les deux sens (6 cas sur 40 pour les deux sens), contrairement aux auditeurs francophones. En regroupant les stimuli et les réponses des voyelles moyennes, nous obtenons le Tableau 23 ci-dessous.

**3)** La confusion entre les voyelles arrondies fermées /u y/ et mi-fermée /ø/ représente 17% de toutes les confusions. /u/ et /ø/ ont été confondus dans les deux sens (12 cas sur 40), et nous y observons une tendance similaire mais moins marquée entre /y/ et /ø/ (6 sur 40 pour /y/->/ø/, 5 sur 40 pour /ø/->/y/). En revanche, /u/ et /y/ ont été rarement confondus (2 sur 40 pour /u/->/y/ ; /y/ n'a jamais été perçu comme /u/).

[380] Sur l'asymétrie perceptive des consonnes, voir Chang *et al.* (2001).

S. Chang, M. C. Plauché, J. J. Ohala. "Markedness and consonant confusion asymmetries", in E. Hume, K. Johnson, *The role of speech perception in phonology*. San Diego, Academic Press, 79-101, 2001.

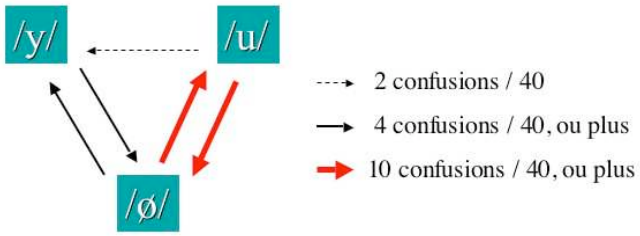



**Figure 47 :** Pattern de confusion entre les voyelles /u y ø/.

Cette dernière tendance (c'est-à-dire la confusion entre les voyelles arrondies mi-fermée /ø/ et fermées /u y/) sera traitée en détail dans la troisième partie, et cela, en relation avec la production.

Nous pouvons également constater les phénomènes suivants :

**4)** Parmi les 3 paires de voyelles moyennes (Tableau 21), le taux de confusion était plus élevé pour /e/-/ɛ/ (16 cas sur 40 pour /e/->/ɛ/, 15 cas sur 40 pour /ɛ/->/e/) que pour les deux autres (6 sur 40 pour /o/-/ɔ/ dans les deux sens ; 6 sur 40 pour /ø/->/œ/ et 11 sur 40 pour /œ/->/ø/).

**5)** /o/ et /ɔ/ ont été perçus comme /ø/ ou /œ/ dans 10% des cas (Tableau 23). Rappelons que les auditeurs natifs ont présenté une tendance similaire (/ɔ/ perçu comme /œ/ dans 4 cas sur 40 : Tableau 20, p. 127).

**6)** Nous trouvons également des cas de confusion moins fréquents entre les séries fermée et mi-fermée (/i/-/e/ et /u/-/o/, la paire /y/-/ø/ étant déjà mentionnée). Ce qui est intéressant avec ces paires est que ce sont toujours les stimuli des voyelles mi-fermées qui ont été perçus comme voyelles fermées. Cela serait dû au timbre intermédiaire (entre mi-fermé et mi-ouvert) des voyelles /e/ et /o/ du japonais (rappelons que c'est un système à 5 voyelles et à 3 degrés d'aperture), telles qu'elles sont décrites dans la littérature (Chapitre 1.4. ; Takebayashi 1996[381] : 58 les transcrit [e̞] et [o̞] ; Saitô[382] 1997 : 84 propose [e̞] ou [ɛ̞], et [o̞ ]), et mesurées acoustiquement (figures de Mokhtari et Tanaka 2000[383] reproduites dans les Figure 8, p. 41 et Figure 9, p. 42, nos données présentées dans la présente partie : Expérience 1, Figure 39).

**7)** Il se trouve également un petit nombre de cas où les apprenants ont perçu des voyelles orales au lieu de nasales. Ils ont montré une tendance à entendre /a/, /ɔ/, et /o/ respectivement pour /ɛ̃/, /ɑ̃/, et /õ/. On remarquera que les confusions /a/-/ɛ̃/ et /ɔ/-/ɑ̃/ sont inattendues ; on s'attendrait plutôt à des confusions entre les deux membres des paires /a/-/ɑ̃/ et /ɔ o/-/õ/. Cette tendance nous évoque l'étude de Montagu (2004)[384], qui a montré que des auditeurs francophones natifs ont perçu majoritairement ces mêmes voyelles /a/, /ɔ/, et /o/ quand ils ont entendu le

[381] S. Takebayashi. *Eigo onseigaku* [Phonétique de l'anglais]. Tokyo, Kenkyûsha, 1996.

[382] Y. Saitô. *Nihongo onseigaku nyuumon* [Introduction à la phonétique du japonais]. Tokyo, Sanseidô, 1997.

[383] P. Mokhtari, K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66, 2000.

[384] J. Montagu. Les sons sous-jacents aux voyelles nasales en français parisien : indices perceptifs des changements. *Actes des XXIVèmes Journées d'Étude sur la Parole*, 385-388, 2004.

début (les 30 premières millisecondes) de chacune de ces voyelles nasales (/ɛ̃/, /ɑ̃/, et /ɔ̃/). Elle est également conforme aux comportements des 5 auditeurs francophones natifs testés dans la section précédente de notre étude (sauf pour le /ɛ̃/ : Expérience 3, Tableau 20).

**Tableau 23** : Résultats du test d'identification auprès de 5 apprenants japonophones, exprimés en nombre de réponses d'identification et en pourcentage entre parenthèses. Les voyelles mi-fermées et mi-ouvertes ont été regroupées (80 réponses pour les voyelles mi-fermées et mi-ouvertes, 40 pour les autres). Les stimuli sont représentés sur l'axe vertical, les réponses, sur l'axe horizontal.

| stimuli\réponses | i | e/ɛ | a | ɔ/o | u | y | ø/œ | ɛ̃ | ɑ̃ | ɔ̃ | total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 39 (98) | 1 (3) | | | | | | | | | 40 (100) |
| e/ɛ | 5 (6) | 71 (89) | | | | | 3 (4) | 1 (1) | | | 80 (100) |
| a | | | 34 (85) | | | | 3 (8) | | 3 (8) | | 40 (100) |
| ɔ/o | | | 1 (1) | 59 (74) | 8 (10) | | 8 (10) | 1 (1) | 3 (4) | | 80 (100) |
| u | | | | | 22(55) | 2 (5) | 12 (30) | | | 4 (10) | 40 (100) |
| y | | | | | | 33 (83) | 7 (18) | | | | 40 (100) |
| ø/œ | | | | | 13 (16) | 6 (8) | 58 (73) | 1 (1) | 2 (3) | | 80 (100) |
| ɛ̃ | | | 5 (13) | | | | 2 (5) | 21 (53) | 12 (30) | | 40 (100) |
| ɑ̃ | | | | 10 (25) | | | | 3 (8) | 14 (35) | 13 (33) | 40 (100) |
| ɔ̃ | | | | 8 (20) | | | 4 (10) | 5 (13) | 8 (20) | 15 (38) | 40 (100) |
| total_réponses | 44 | 72 | 40 | 77 | 43 | 41 | 97 | 32 | 42 | 32 | |

Le test d'identification présenté dans ce chapitre nous permet de repérer ainsi les trois sources majeures de difficulté de perception que les apprenants japonophones rencontrent : 1) les voyelles nasales ; 2) les voyelles moyennes ; 3) les voyelles arrondies mi-fermée /ø/ et fermées /u y/.

### 2.3. ANALYSE ACOUSTIQUE : LA PRODUCTION DES VOYELLES ORALES DU FRANÇAIS PAR DES APPRENANTS JAPONOPHONES : LECTURE ET RÉPÉTITION (EXPÉRIENCE 4)[385]

Les difficultés de la production reflètent-elles celles que nous avons décrites du point de vue perceptif ?

**Deux tâches différentes : lecture et répétition immédiate** Pour cela, nous avons recueilli des données de production en utilisant deux différents types de tâche : la lecture (section 2.3.1.) et la répétition immédiate après un modèle (section 2.3.2.). Lors de l'interprétation de la première tâche, nous devrons considérer l'effet de l'écriture orthographique sur la représentation mentale des phonèmes chez les apprenants, même si nous avons présenté les voyelles cibles sous forme de mots exemples écrits en orthographe et en transcription phonémique en même temps. En

[385] Une partie des résultats de cette expérience a été publiée dans :

T. Kamiyama. La production et la perception des voyelles orales françaises par les apprenants japonophones. *Actes des XXVIèmes Journées d'étude sur la Parole*, 49-52, 2006.

revanche, lors de la répétition, la production est contrainte par la perception du modèle. Nous examinerons si les mêmes tendances se retrouvent pour la production et la perception (le tableau récapitulatif se trouve au Tableau 25, p. 150).

### *2.3.1. Lecture*

#### 2.3.1.1. Le corpus et les procédures

Le corpus utilisé pour la tâche de lecture est le même que celui qui a été utilisé pour les locuteurs natifs : la phrase-cadre « *Je dis /V/ comme dans …* ». Ces phrases ont été imprimées sur une feuille (Tableau 22)[386]. Même si les locuteurs japonophones produisent assez rarement une montée de continuation en japonais (voir la Section 1.4.4.), nous avons répété la première phrase à la fin de la liste pour que l'avant-dernière phrase soit prononcée avec des schémas intonatifs similaires aux phrases précédentes.

Je dis [i] comme dans « lit ».
Je dis [e] comme dans « les ».
Je dis [ɛ] comme dans « lait ».
Je dis [a] comme dans « la ».
Je dis [ɔ] comme dans « lors ».
Je dis [o] comme dans « l'eau ».
Je dis [u] comme dans « loup ».
Je dis [y] comme dans « lu ».
Je dis [ø] comme dans « leu ».
Je dis [œ] comme dans « leur ».
Je dis [ɛ̃] comme dans « lin ».
Je dis [ɑ̃] comme dans « lent ».
Je dis [ɔ̃] comme dans « long ».
Je dis [i] comme dans « lit ».

**Figure 48 :** Feuille présentée aux apprenants lors de la tâche de lecture.

**Les procédures** Les apprenants ont lu la liste de phrases deux fois après entraînement. Ils étaient équipés d'un micro-casque. L'enregistrement a été effectué dans une cabine d'enregistrement. Les données ont été enregistrées au taux d'échantillonnage de 16 kHz, 16 bits.

**La segmentation et la détection de formants** La segmentation a été effectuée en consultant l'oscillogramme et le spectrogramme sous Praat. La partie où les 4 premiers formants (notamment le F2) étaient clairement visibles a été sélectionnée pour l'analyse formantique. Les 4 premiers formants ont été mesurés sous Praat, en utilisant un script écrit par Cédric Gendrot[387] et modifié par nous. Quand les valeurs détectées étaient aberrantes, nous avons vérifié le

[386] Nous avons utilisé des crochets [ ] pour la transcription phonémique, car les apprenants sont en général plus familiers avec cette notation en tant que transcription de la prononciation.

[387] Disponible sur :
http://www.personnels.univ-paris3.fr/users/cgendrot/pub/download/analyse1.zip

spectrogramme et changé des paramètres (le nombre de formants à détecter entre 0 et n Hz) quand c'était nécessaire, afin d'obtenir des valeurs correctes. Les valeurs qui étaient toujours aberrantes après modification de paramètres ont été rejetées (3 valeurs sur 3200, soit 0,09%, pour les deux types de tâche confondus).

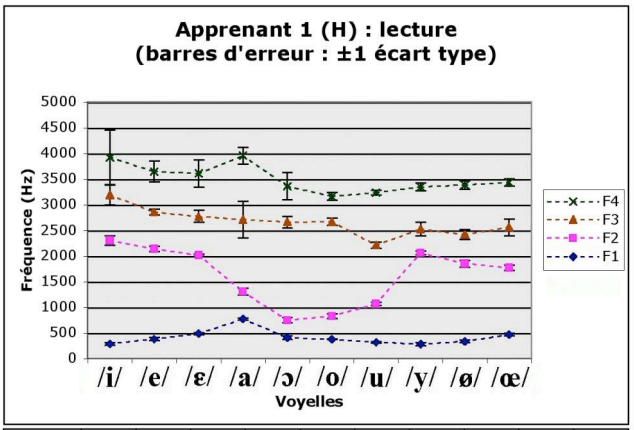
**Les mesures de formants** Les formants ont été mesurés dans cinq endroits de la durée la voyelle (Figure 40, p. 121) : la moyenne du premier cinquième de la voyelle (0%-20% de la durée), celle du deuxième cinquième (20%-40%), ainsi de suite jusqu'au dernier cinquième (80%-100%).

#### 2.3.1.2. Les locuteurs

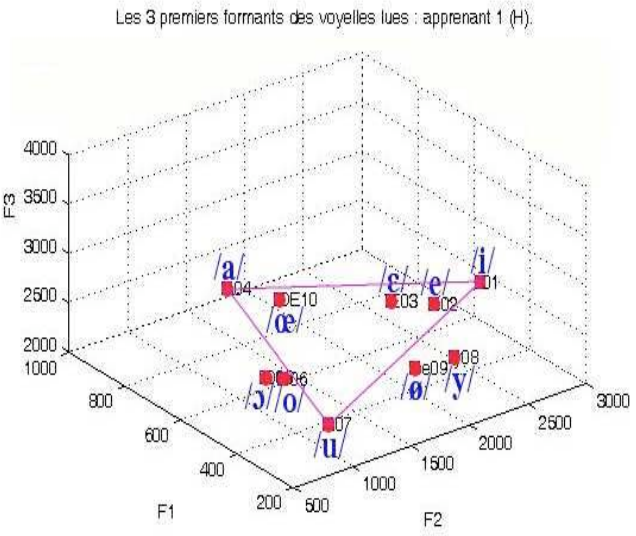
Trois apprenants (2 hommes et 1 femme) ont participé aux enregistrements. Ils étaient tous des étudiants en première année de l'Université de Tokyo lors de l'expérience (effectuée en juillet 2005). Ils avaient choisi le français comme deuxième langue étrangère (une langue au choix est obligatoire en tant que deuxième langue étrangère). Ils l'avaient appris depuis 3 mois, à raison de 4,5-7,5 heures par semaine, avec des enseignants japonophones et francophones.

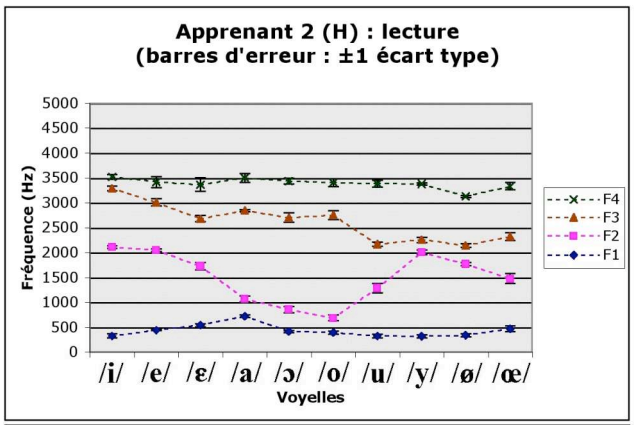
#### 2.3.1.3. Les résultats

Les valeurs moyennes des cinq mesures des quatre premiers formants (pour chacune des deux répétitions) sont représentées à la Figure 49, et une représentation à deux dimensions (F1-F2 et F2-F3) des trois premiers formants en comparaison avec les valeurs des locuteurs natifs (Figure 42 et Figure 43, p. 123) est proposée aux Figure 50-Figure 52.

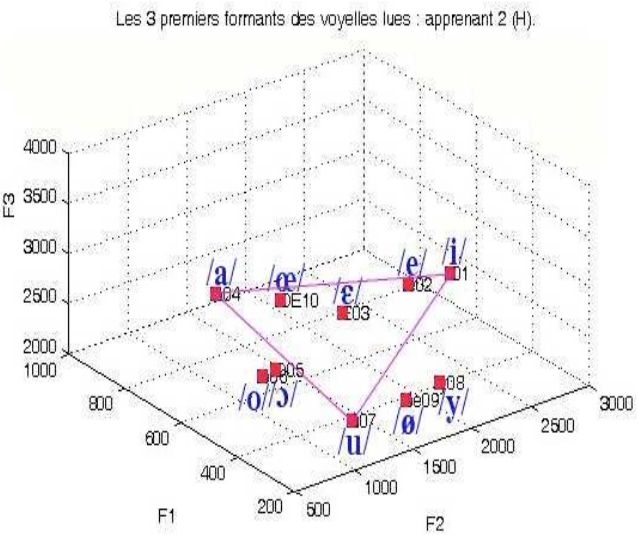


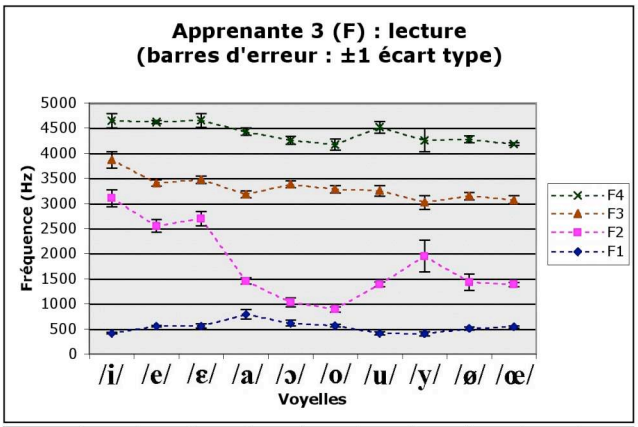

| Appr. 1 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3923 | 3643 | 3606 | 3955 | 3360 | 3164 | 3234 | 3346 | 3383 | 3433 |
| F3 | 3188 | 2855 | 2770 | 2709 | 2656 | 2665 | 2212 | 2523 | 2414 | 2555 |
| F2 | 2297 | 2135 | 2015 | 1304 | 749 | 830 | 1065 | 2062 | 1850 | 1770 |
| F1 | 285 | 382 | 486 | 772 | 407 | 376 | 316 | 281 | 334 | 465 |




| Appr. 2 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3523 | 3420 | 3361 | 3502 | 3434 | 3403 | 3384 | 3375 | 3126 | 3332 |
| F3 | 3290 | 3006 | 2680 | 2843 | 2703 | 2751 | 2168 | 2256 | 2144 | 2318 |
| F2 | 2114 | 2045 | 1729 | 1073 | 858 | 685 | 1283 | 2005 | 1767 | 1474 |
| F1 | 324 | 444 | 543 | 721 | 414 | 391 | 326 | 317 | 336 | 468 |




| Appr. 3 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 4647 | 4621 | 4653 | 4426 | 4255 | 4172 | 4514 | 4257 | 4276 | 4181 |
| F3 | 3865 | 3405 | 3471 | 3175 | 3380 | 3282 | 3250 | 3018 | 3146 | 3068 |
| F2 | 3101 | 2547 | 2693 | 1452 | 1027 | 888 | 1388 | 1945 | 1429 | 1382 |
| F1 | 411 | 551 | 553 | 785 | 610 | 558 | 406 | 399 | 504 | 538 |

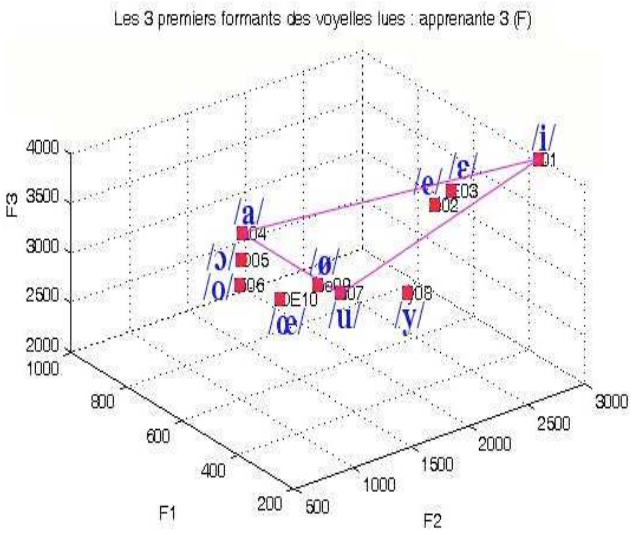



**Figure 49** : Formants (en Hertz) des 10 voyelles orales du français lues par 3 apprenants japonais (moyenne de 5 mesures par voyelle x 2 répétitions). A gauche : représentation des quatre premiers

formants (les barres d'erreur représentent l'écart-type) ; à droite : représentation des trois premiers formants dans un triangle vocalique à trois dimensions (à droite).

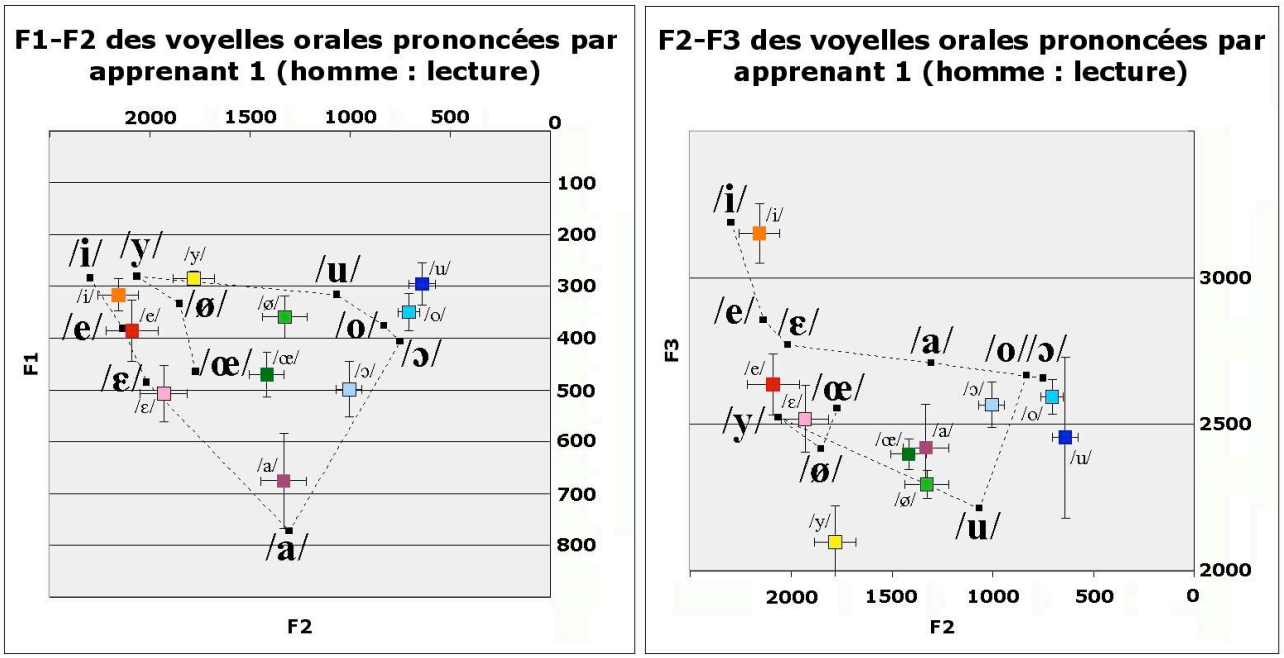



**Figure 50** : Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par Apprenant 1 (homme) dans la phrase cadre « *je dis* /V/ *comme dans* … » (petits carreaux avec transcription en grand caractère : 5 mesures par voyelle x 2 répétitions), en comparaison avec ceux de la moyenne des deux locuteurs natifs FR1 et FR2 (Figure 42, p. 123 : grands carreaux avec des barres d'erreur ; 5 mesures par voyelle x 3 répétitions x 2 locuteurs ; les barres d'erreur représentent l'écart-type). F1 sur l'axe vertical et F2 sur l'axe horizontal (à gauche) ; F2 sur l'axe horizontal et F3 sur l'axe vertical (à droite).

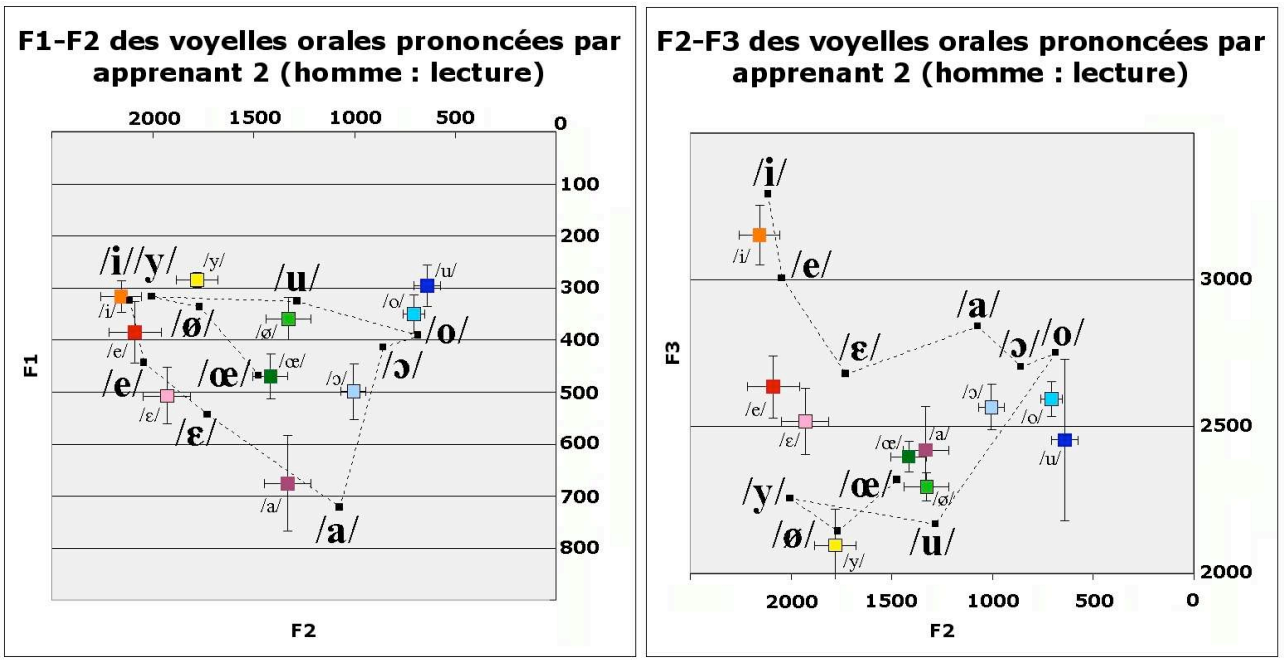



**Figure 51** : Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par Apprenant 2 (homme) dans la phrase cadre « *je dis* /V/ *comme dans* … » (petits carreaux avec

transcription en grand caractère : 5 mesures par voyelle x 2 répétitions), en comparaison avec ceux de la moyenne des deux locuteurs natifs FR1 et FR2 (Figure 42, p. 123 : grands carreaux avec des barres d'erreur ; 5 mesures par voyelle x 3 répétitions x 2 locuteurs ; les barres d'erreur représentent l'écart-type). F1 sur l'axe vertical et F2 sur l'axe horizontal (à gauche) ; F2 sur l'axe horizontal et F3 sur l'axe vertical (à droite).

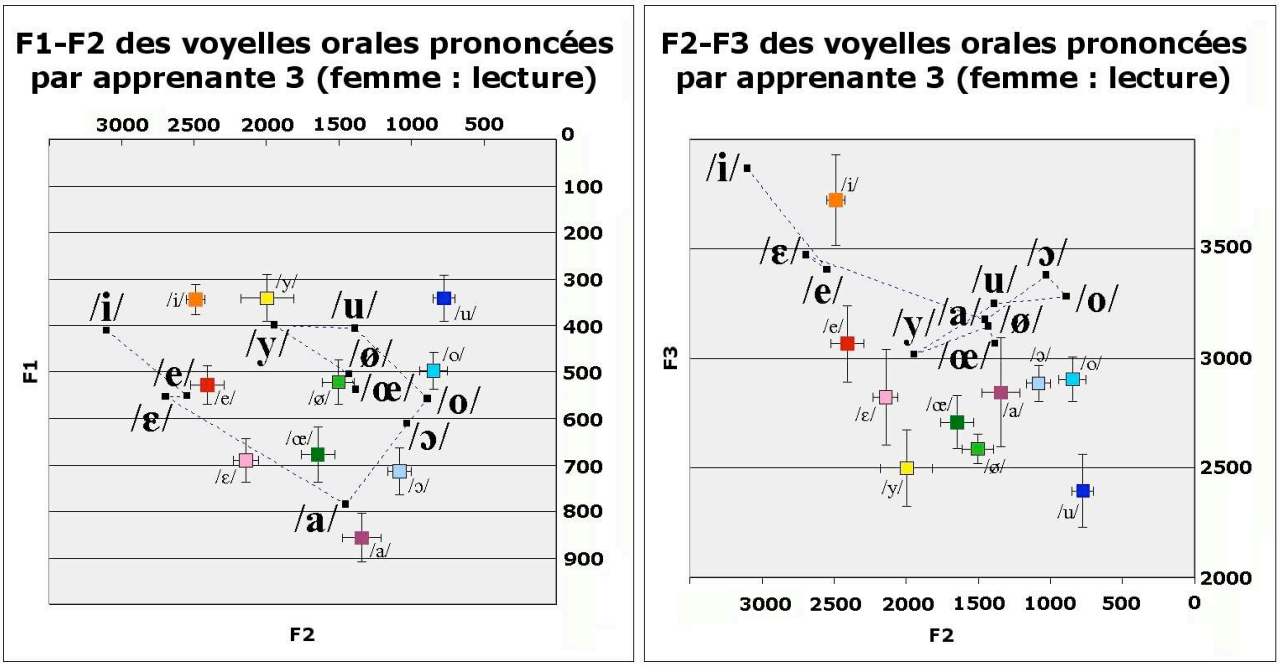



**Figure 52** : Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par Apprenante 3 (femme) dans la phrase cadre « *je dis* /V/ *comme dans* … » (petits carreaux avec transcription en grand caractère : 5 mesures par voyelle x 2 répétitions), en comparaison avec ceux de la moyenne des deux locutrices natives FR3 et FR4 (Figure 43, p. 124 : grands carreaux avec des barres d'erreur ; 5 mesures par voyelle x 3 répétitions x 2 locuteurs ; les barres d'erreur représentent l'écart-type). F1 sur l'axe vertical et F2 sur l'axe horizontal (à gauche) ; F2 sur l'axe horizontal et F3 sur l'axe vertical (à droite).

Nous y observons les tendances suivantes :

**1) /y/** Les apprenants 1 et 2 ont prononcé /y/ avec ses F2 et F3 proches, malgré un F2 relativement élevé (2062 Hz et 2005 Hz en moyenne, respectivement), notamment dans la première répétition de l'apprenant 1 (premier spectrogramme de la Figure 53). En revanche, l'apprenante 3 a produit une diphtongue (Figure 54 et Figure 55 : nous pouvons constater que le F2 descend progressivement et que le F3 monte légèrement ; l'écart-type important que l'on observe à la Figure 49 suggère également cette tendance), qui ressemble à la syllabe /ju/ du japonais (Figure 132, p. 259). Ce phénomène sera traité en détail dans la discussion de la troisième partie (3.5.).

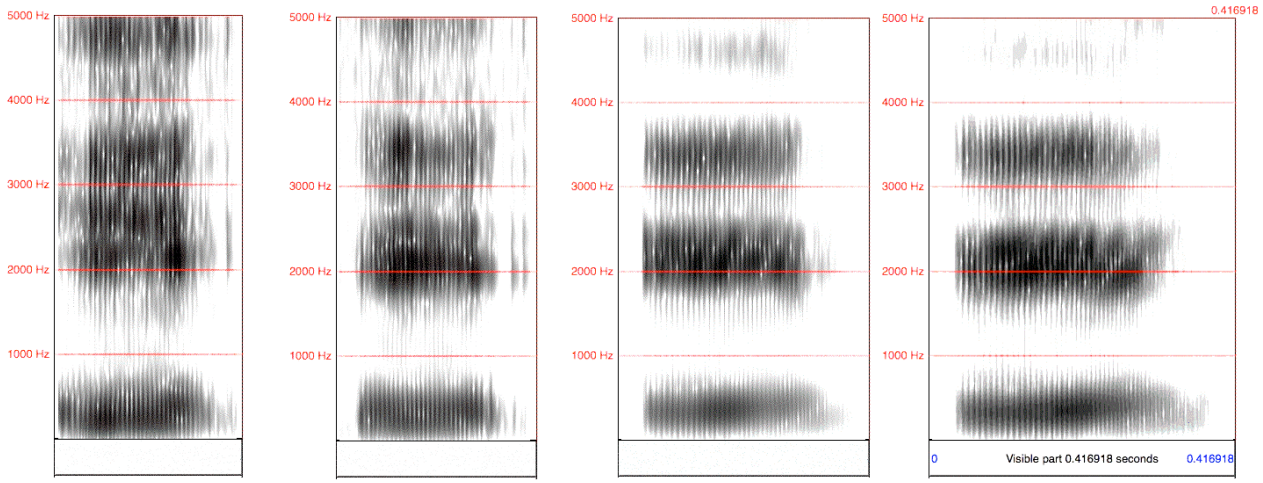



**Figure 53 :** Spectrogrammes à bande large (Praat ; longueur de fenêtre : 5 millisecondes) du /y/ prononcé par l'apprenant 1 (les deux à gauche) et l'apprenant 2 (les deux à droite).

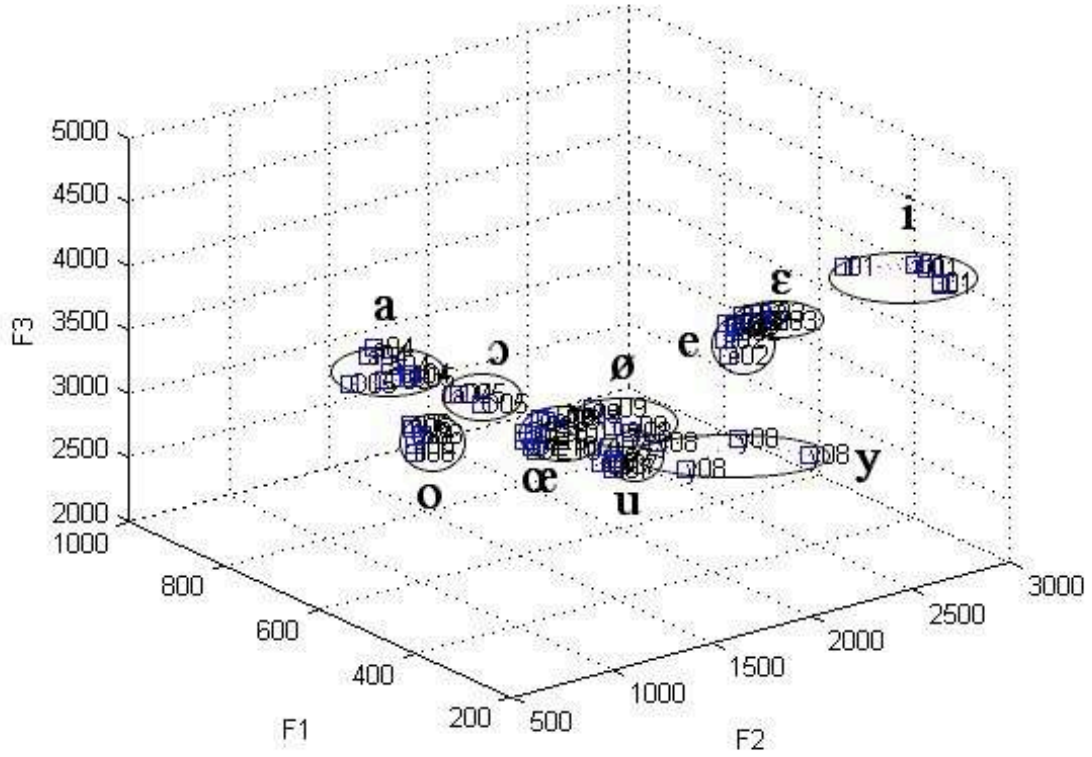



**Figure 54** : Les trois premiers formants (en Hertz) de l'apprenante 3 mesurés dans 5 zones de la durée de la voyelle (une répétition x 5 mesures de la durée de la voyelle).

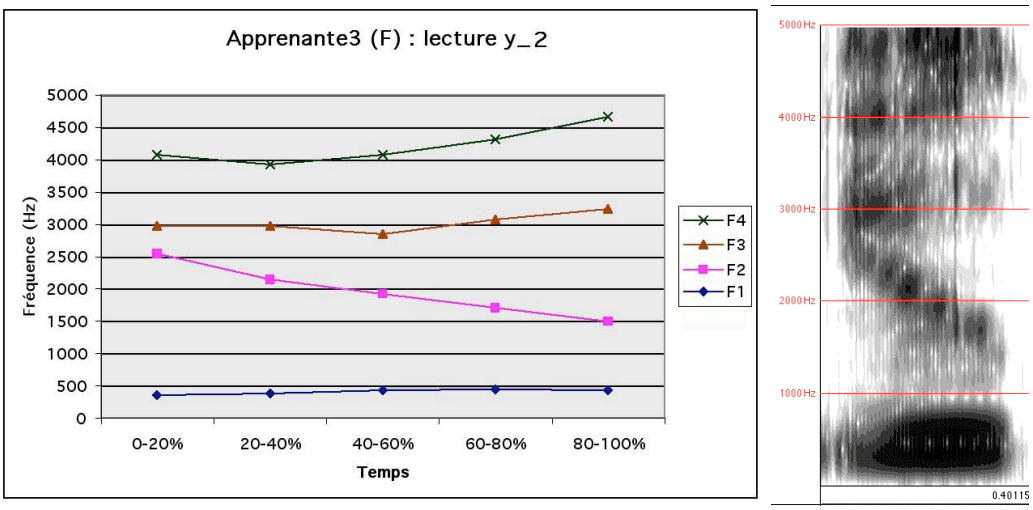



**Figure 55 :** Changement des valeurs formantiques (gauche) et le spectrogramme (droite) du /y/ prononcé dans la tâche de lecture par l'apprenante 3 (une des deux répétitions).

**2) /u/** Les trois apprenants ont prononcé la voyelle /u/ avec un F2 plus élevé (1065 Hz, 1283 Hz, et 1388 Hz en moyenne) que celui de /o/ (830 Hz, 685 Hz, et 888 Hz) et de /ɔ/ (749 Hz, 858 Hz, et 1027 Hz). Notons qu'ils ont produit ces deux voyelles /o/ et /ɔ/ avec une valeur de F2 qui n'était pas très différente de celle des locuteurs natifs (Figure 41, p. 123 : 688 Hz, 715 Hz, 776 Hz, 912 Hz, inférieure donc à 1000 Hz pour /o/ ; 951 Hz, 1053 Hz, 1051 Hz, 1108 Hz pour /ɔ/). Par contraste, la valeur élevée du F2 du /u/ des apprenants japonophones se rapproche à celle du /ø/ français (Figure 41, p. 123 : 1276 Hz, 1372 Hz, 1570 Hz, 1431 Hz), notamment dans la première répétition de l'apprenant 2 (le troisième spectrogramme de la Figure 56).

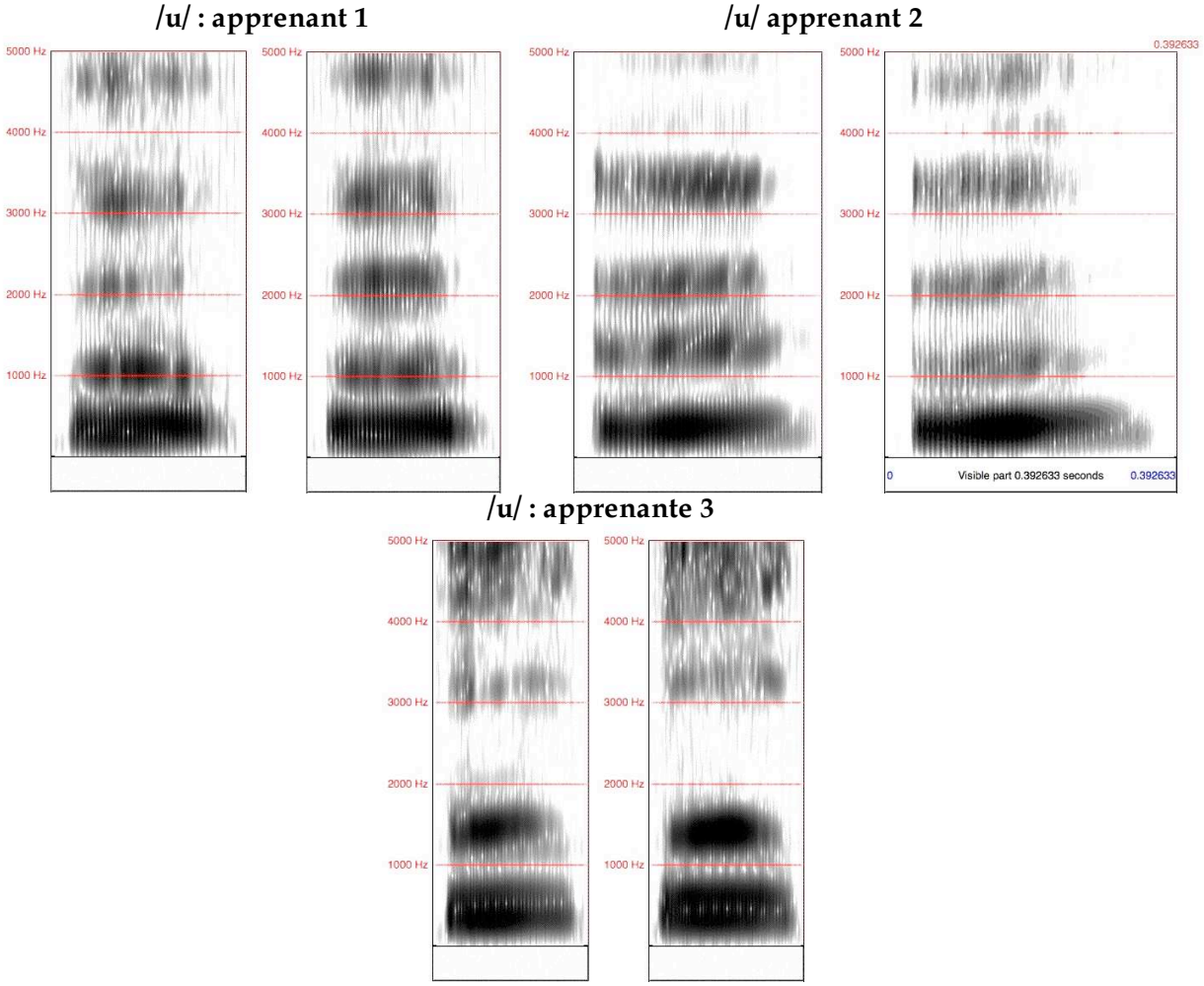



**Figure 56 :** Spectrogrammes à bande large (longueur de fenêtre : 5 millisecondes) du /u/ prononcé par les 3 apprenants : les deux répétitions de l'apprenant 1 (en haut, à gauche), de l'apprenant 2 (en haut, à droite), et de l'apprenante 3 (en bas).

Par ailleurs, l'intensité relative des formants supérieurs est plus forte que dans le /u/ des locuteurs français. Une comparaison d'énergie de 3 différentes zones de fréquence (0-1 kHz, 1-2 kHz, et 2-3 kHz) confirme cette tendance de manière quantitative (Figure 57). Le /u/ prononcé par les 3 apprenants a plus d'énergie sur les moyennes fréquences (-18dB dans 1-2 kHz, -33dB dans 2-3 kHz, par rapport à l'intensité de la zone 0-1 kHz) que le /u/ prononcé par les 4 locuteurs natifs francophones (-37dB dans 1-2 kHz, -41dB dans 2-3 kHz, par rapport à l'intensité de la zone 0-1 kHz).

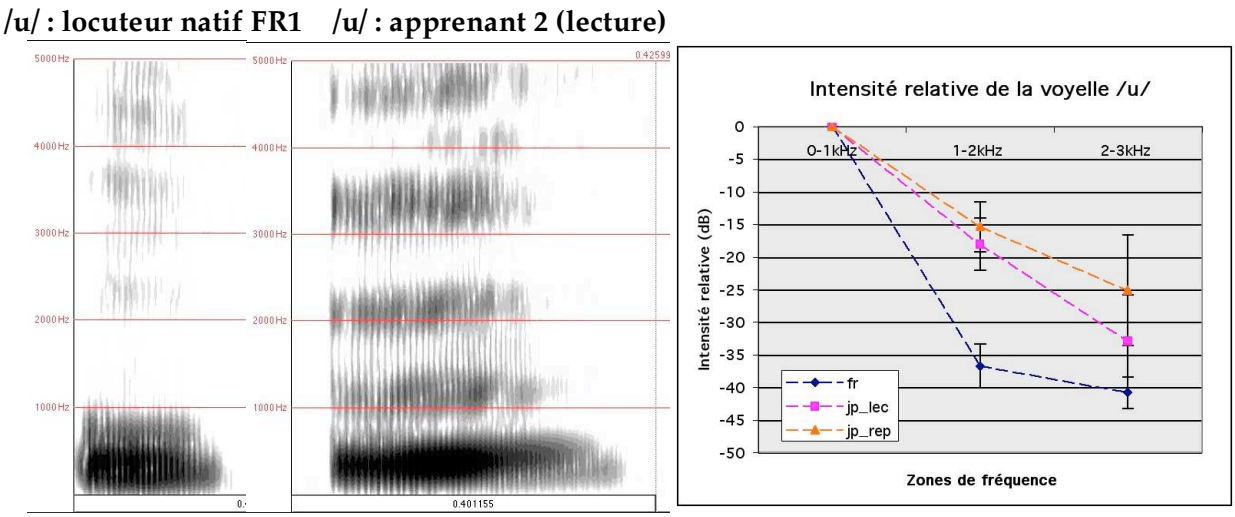



**Figure 57 :** Spectrogrammes à bande large (longueur de fenêtre : 5 millisecondes) de la voyelle /u/ prononcée par le locuteur natif FR1 (à gauche) et par l'apprenant 2 (au milieu). L'intensité relative sur 3 zones de fréquence (0-1 kHz, 1-2 kHz, 2-3 kHz) du /u/ prononcé par les 4 locuteurs natifs (12 occurrences : fr), lu (6 : jp_lec) et répété par les 3 apprenants (10 : jp_rep) : les barres d'erreur représentent ±1 écart-type.

**/u/ et /ø œ/** L'apprenante 3 a produit des valeurs de F2 et F3 similaires pour /u/, /ø/ et /œ/, même si le F1 de son /u/ (406 Hz) était moins élevé que celui de /ø/ (504 Hz) et de /œ/ (538 Hz) : les 3 premiers formants étaient de 406, 1388, et 3250 Hz pour /u/, 504, 1429 et 3146 Hz pour /ø/, 538, 1382, et 3068 Hz pour /œ/ (spectrogramme à la Figure 58).

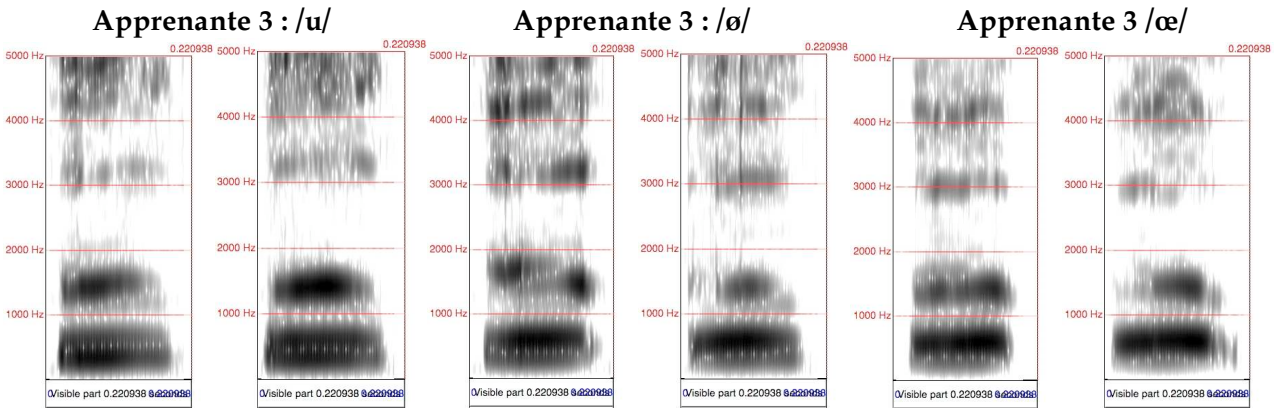



**Figure 58** : Spectrogrammes à bande large (longueur de fenêtre : 5 millisecondes) de la voyelle /u/ (les deux premiers), /ø/ (les deux au milieu), et /ø/ (les deux derniers) prononcée par l'apprenante 3 dans une tâche de lecture.

**3) Les voyelles moyennes** Comme attendu, les différences entre les mi-fermées et les mi-ouvertes étaient moins nettement marquées, notamment chez l'apprenante 3, que chez les locuteurs natifs, même si les deux voyelles moyennes de la même série (ex. /ø/-/œ/) avaient été placées successivement sur la liste de lecture, ce qui aurait dû faciliter la tâche de les prononcer de manière distincte.

**4) /ø/** Les apprenants 1 et 2 ont produit un /ø/ avec un F2 élevé (1750-1850 Hz), ce qui rapproche cette voyelle au /y/. Cela serait dû à la prononciation typique de l'orthographe « eu » en anglais /ju:/, qui est associé en général à la séquence /ju(R)/ en japonais (rappelons que le /y/ français est interprété comme /ju/ en japonais : Figure 59).

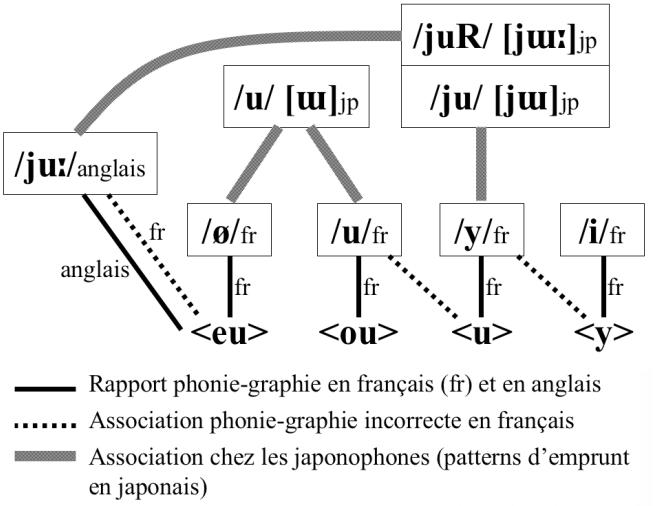



**Figure 59** : Rapport phonie-graphie et des associations susceptibles chez les japonophones apprenant le français L3.

Les données présentées dans la présente section ont été recueillies dans une tâche de lecture. Par conséquent, la production aurait été influencée non seulement par la représentation mentale des apprenants mais aussi par leur interprétation de l'orthographe et la transcription phonémique fournie. Afin d'exclure ces facteurs, nous avons effectué une autre expérience de production avec une tâche de répétition immédiate après un modèle.

### *2.3.2. Répétition*

#### 2.3.2.1. Les matériels

Une des trois répétitions des voyelles isolées prononcées dans la phrase-cadre « *Je dis /V/ comme dans …* » par les 4 locuteurs francophones natifs pour l'Expérience 2 a été extraite. Étant donné que tous les locuteurs ont inséré une brève pause avant et après la voyelle cible, le début et la fin du fichier extrait ont été fixés pendant ces pauses. Les valeurs des 4 premiers formants et la durée de la production utilisée sont représentées dans le Tableau 24. Un bip sonore (440 Hz, 50 millisecondes) a été placé 500 millisecondes avant la voyelle cible. Les fichiers qui contenaient chaque voyelle ont été concaténés dans un ordre semi-aléatoire.

**Tableau 24 :** Valeurs des 4 premiers formants (en Hertz) et la durée (en millisecondes) des 10 voyelles orales du français prononcées par 4 locuteurs natifs (2 hommes, FR1 et FR2, en haut et 2 femmes, FR3 et FR4, en bas) et utilisées dans la tâche de répétition. La durée indiquée est celle de la portion de la voyelle qui a été retenue pour la mesure de formants, pendant laquelle les formants supérieurs étaient clairement visibles.

| FR1 (H) | F1 | F2 | F3 | F4 | durée (ms) |
|---|---|---|---|---|---|
| i | 316 | 2132 | 3143 | 3633 | 141 |
| e | 327 | 2084 | 2656 | 3062 | 168 |
| ɛ | 505 | 1945 | 2432 | 2816 | 141 |
| a | 638 | 1430 | 2429 | 2852 | 141 |
| ɔ | 515 | 1053 | 2526 | 3334 | 206 |
| o | 336 | 743 | 2597 | 3347 | 129 |
| u | 297 | 573 | 2426 | 3495 | 131 |
| y | 289 | 1742 | 2073 | 3113 | 153 |
| ø | 325 | 1393 | 2294 | 3137 | 168 |
| œ | 478 | 1471 | 2412 | 3213 | 192 |
| Moyenne | 403 | 1457 | 2499 | 3200 | 157 |

| FR2 (H) | F1 | F2 | F3 | F4 | durée (ms) |
|---|---|---|---|---|---|
| i | 316 | 2220 | 3218 | 4178 | 163 |
| e | 463 | 1970 | 2618 | 3483 | 202 |
| ɛ | 497 | 2000 | 2627 | 3506 | 175 |
| a | 685 | 1315 | 2477 | 3483 | 170 |
| ɔ | 500 | 950 | 2651 | 3413 | 173 |
| o | 374 | 707 | 2607 | 3513 | 117 |
| u | 321 | 590 | 2711 | 3731 | 121 |
| y | 291 | 1822 | 2152 | 3281 | 173 |
| ø | 402 | 1228 | 2294 | 3202 | 186 |
| œ | 465 | 1290 | 2398 | 3424 | 201 |
| Moyenne | 431 | 1409 | 2575 | 3521 | 168 |

| FR3 (F) | F1 | F2 | F3 | F4 | durée (ms) |
|---|---|---|---|---|---|
| i | 354 | 2530 | 3863 | 4420 | 213 |
| e | 507 | 2491 | 3122 | 4169 | 228 |
| ɛ | 696 | 2181 | 2961 | 3685 | 238 |
| a | 885 | 1268 | 3086 | 4036 | 225 |
| ɔ | 706 | 1014 | 2966 | 4123 | 225 |
| o | 469 | 780 | 2954 | 3980 | 279 |
| u | 302 | 805 | 2472 | 3462 | 274 |
| y | 291 | 2059 | 2525 | 3766 | 286 |
| ø | 497 | 1539 | 2553 | 3991 | 269 |
| œ | 682 | 1721 | 2789 | 4100 | 295 |
| Moyenne | 539 | 1639 | 2929 | 3973 | 253 |

| FR4 (F) | F1 | F2 | F3 | F4 | durée (ms) |
|---|---|---|---|---|---|
| i | 340 | 2427 | 3567 | 4132 | 120 |
| e | 597 | 2238 | 2918 | 3838 | 133 |
| ɛ | 717 | 2108 | 2595 | 3444 | 127 |
| a | 848 | 1421 | 2603 | 3410 | 133 |
| ɔ | 754 | 1134 | 2805 | 3766 | 155 |
| o | 513 | 941 | 2818 | 3926 | 129 |
| u | 385 | 733 | 2210 | 3291 | 103 |
| y | 423 | 2063 | 2640 | 3735 | 133 |
| ø | 554 | 1403 | 2577 | 3988 | 146 |
| œ | 712 | 1565 | 2621 | 3862 | 145 |
| Moyenne | 584 | 1603 | 2735 | 3739 | 132 |

**2.3.2.2. Les procédures**

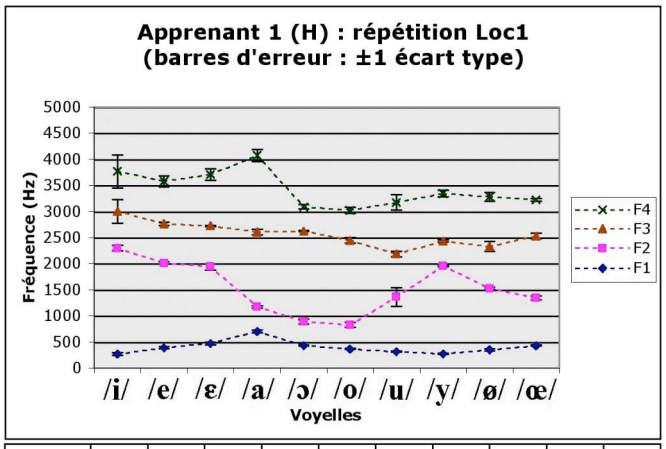
Les mêmes 3 apprenants ont répété les voyelles isolées après le « modèle » (voyelle isolée) prononcé par les 4 locuteurs natifs (FR1-4). Les apprenants ont été invités à écouter le modèle par un casque, et à répéter la voyelle immédiatement. Quand l'expérimentateur a entendu la prononciation de l'apprenant, il a envoyé le stimulus de la voyelle suivante (précédé par le bip sonore) au casque de l'apprenant. L'apprenant 1 a écouté les voyelles prononcées par le locuteur natif FR1, l'apprenant 2, les 2 locuteurs natifs (FR1 et FR2), et l'apprenante 3, les 2 locutrices natives (FR3 et FR4). Chaque apprenant a parcouru deux fois la liste (semi-aléatoire pré-établie) de toutes les voyelles prononcées par un locuteur natif donné avant de passer à un autre locuteur natif. L'enregistrement a été effectué directement sur un ordinateur portable au taux d'échantillonnage de 16 kHz, 16 bits, à travers un microphone attaché au casque.

**L'analyse des données enregistrées** La segmentation et l'analyse ont été effectuées de la même manière que la section précédente (pour les données de la tâche de lecture). Les quatre premiers formants ont été mesurés sur cinq zones différentes de la durée de la voyelle, allant du

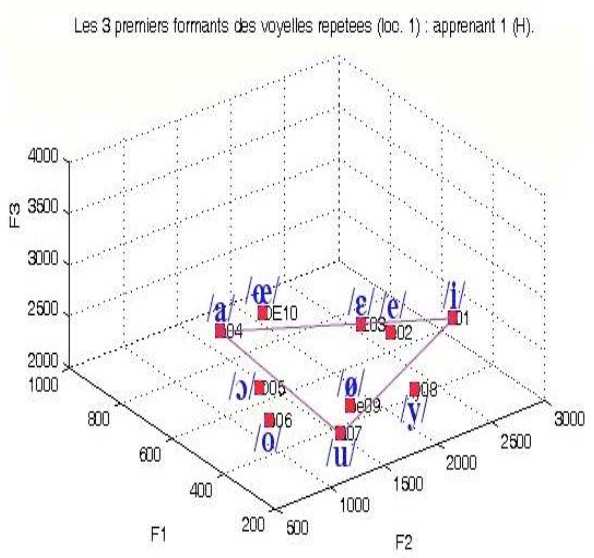
début (0 – 20% de la voyelle) jusqu'à la fin (80 – 100%). La moyenne des cinq valeurs de chacune des deux répétitions a été calculée pour chaque voyelle prononcée par chaque apprenant après chaque locuteur natif modèle.

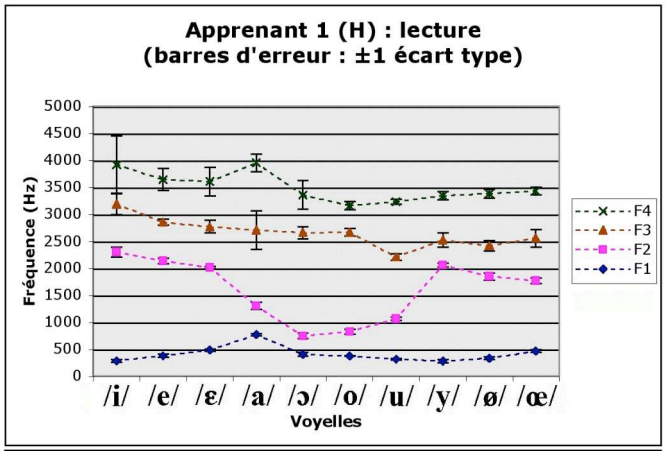
### 2.3.2.3. Les résultats

Les valeurs formantiques sont indiquées dans les Figure 60, Figure 61, et Figure 62. Chaque figure nous permet de comparer les données de la lecture et de la répétition de chaque apprenant. Nous pouvons y observer des tendances similaires, mais aussi des différences, qui seront décrites pour chaque apprenant.




| Appr. 1 Rep1 | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3768 | 3579 | 3707 | 4072 | 3090 | 3022 | 3171 | 3342 | 3280 | 3223 |
| F3 | 3003 | 2762 | 2721 | 2605 | 2625 | 2444 | 2185 | 2434 | 2329 | 2529 |
| F2 | 2294 | 2011 | 1939 | 1171 | 890 | 823 | 1362 | 1954 | 1527 | 1341 |
| F1 | 262 | 385 | 472 | 699 | 427 | 362 | 313 | 268 | 343 | 423 |




| Appr. 1 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3923 | 3643 | 3606 | 3955 | 3360 | 3164 | 3234 | 3346 | 3383 | 3433 |
| F3 | 3188 | 2855 | 2770 | 2709 | 2656 | 2665 | 2212 | 2523 | 2414 | 2555 |
| F2 | 2297 | 2135 | 2015 | 1304 | 749 | 830 | 1065 | 2062 | 1850 | 1770 |
| F1 | 285 | 382 | 486 | 772 | 407 | 376 | 316 | 281 | 334 | 465 |

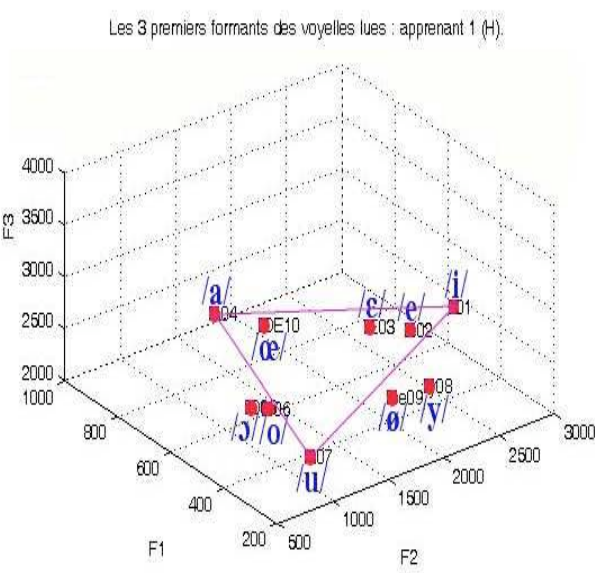


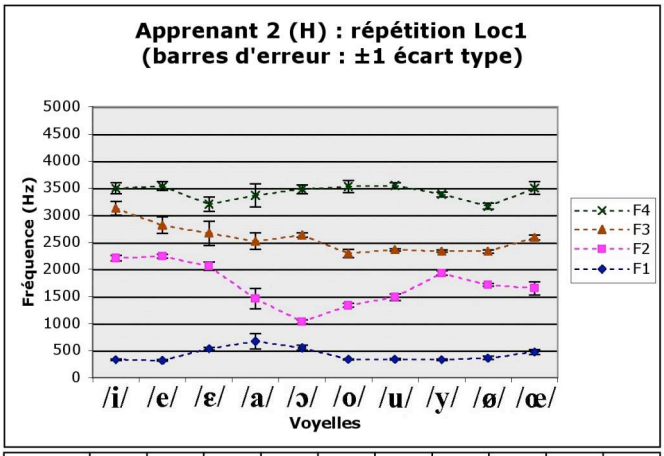

**Figure 60 :** Valeurs formantiques (en Hertz) des voyelles orales du français prononcées par l'apprenant 1 dans la tâche de répétition immédiate (2 répétitions : en haut), et dans la tâche de lecture (2 répétitions : en bas). Les quatre premiers formants (à gauche : les barres d'erreur représentent l'écart-type), et les trois premiers formants dans un triangle vocalique à trois dimensions (à droite).

**Apprenant 1 : /u/** Chez l'apprenant 1 (Figure 60), le F2 du /u/ est plus élevé (1362 Hz) que dans la tâche de lecture (1065 Hz). La tâche de répétition n'a donc pas facilité la production du /u/. Cela suggère que la difficulté liée à la production de cette voyelle ne serait pas seulement due à l'influence de l'orthographe ou de la transcription (le symbole /u/ pourrait évoquer le /u/ [ɯ] japonais), mais également à la difficulté de perception et/ou d'articulation.

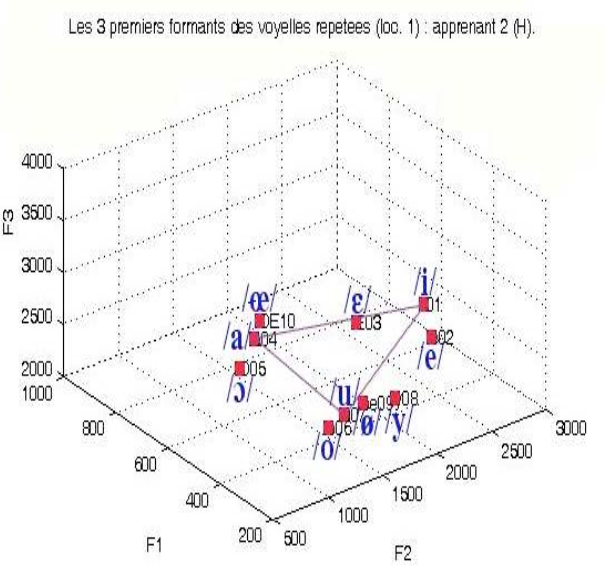
**/y/** La voyelle /y/ a été prononcée avec les F2 et F3 proches, comme dans la tâche de lecture, mais le F2 était légèrement moins élevé (1954 Hz en répétition immédiate contre 2062 Hz en lecture), ce qui rapproche cette production de celle des locuteurs natifs (1806 Hz pour le locuteur natif FR1, et 1749 Hz pour FR2).

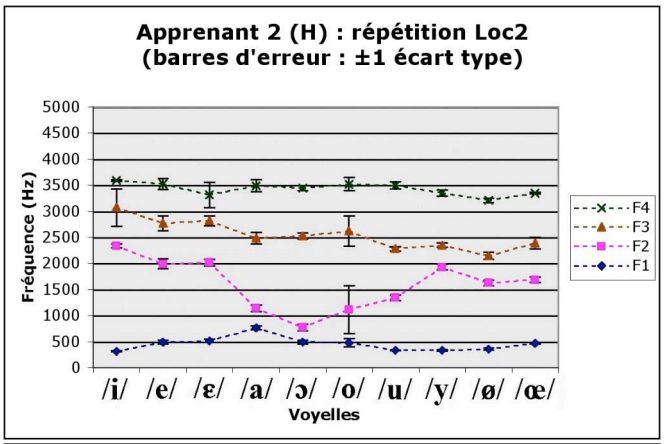
**/ø/ et /œ/** Quant aux /ø/ et /œ/, cet apprenant a mieux réussi à les prononcer en répétition qu'en lecture : le F2 se trouve autour de 1500 Hz. Cette observation nous permettra de supposer que les /ø/ et /œ/ n'ont pas été prononcés correctement en lecture à cause de l'influence de l'orthographe <eu>, ou de la représentation mentale de ces voyelles que l'apprenant avait établie, mais qu'il n'avait pas de difficulté à les percevoir et les reproduire, à la différence du /u/.

**/o/ et /ɔ/** Cet apprenant a mieux différencié les deux voyelles moyennes /o/ et /ɔ/ en répétition qu'en lecture : le F2 du /o/ est moins élevé que celui du /ɔ/, ce qui n'était pas le cas dans la tâche de lecture.

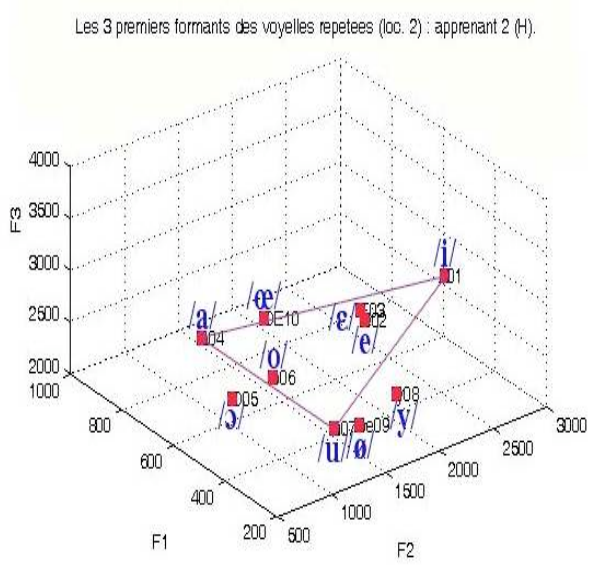


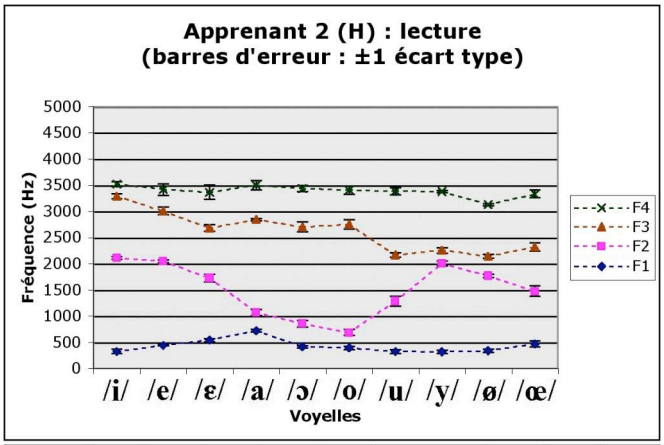

| Appr. 2 Rep1 | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3496 | 3533 | 3201 | 3362 | 3479 | 3525 | 3545 | 3383 | 3163 | 3501 |
| F3 | 3127 | 2810 | 2663 | 2520 | 2633 | 2295 | 2359 | 2335 | 2332 | 2588 |
| F2 | 2204 | 2237 | 2059 | 1455 | 1030 | 1332 | 1485 | 1929 | 1710 | 1645 |
| F1 | 334 | 322 | 533 | 673 | 551 | 337 | 340 | 332 | 362 | 476 |




| Appr. 2 Rep2 | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3588 | 3524 | 3313 | 3491 | 3445 | 3517 | 3493 | 3350 | 3208 | 3343 |
| F3 | 3073 | 2771 | 2814 | 2481 | 2524 | 2615 | 2277 | 2350 | 2140 | 2390 |
| F2 | 2333 | 1989 | 2014 | 1142 | 770 | 1112 | 1338 | 1927 | 1625 | 1686 |
| F1 | 314 | 486 | 513 | 763 | 488 | 475 | 332 | 334 | 352 | 463 |




| Appr. 2 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 3523 | 3420 | 3361 | 3502 | 3434 | 3403 | 3384 | 3375 | 3126 | 3332 |
| F3 | 3290 | 3006 | 2680 | 2843 | 2703 | 2751 | 2168 | 2256 | 2144 | 2318 |
| F2 | 2114 | 2045 | 1729 | 1073 | 858 | 685 | 1283 | 2005 | 1767 | 1474 |
| F1 | 324 | 444 | 543 | 721 | 414 | 391 | 326 | 317 | 336 | 468 |

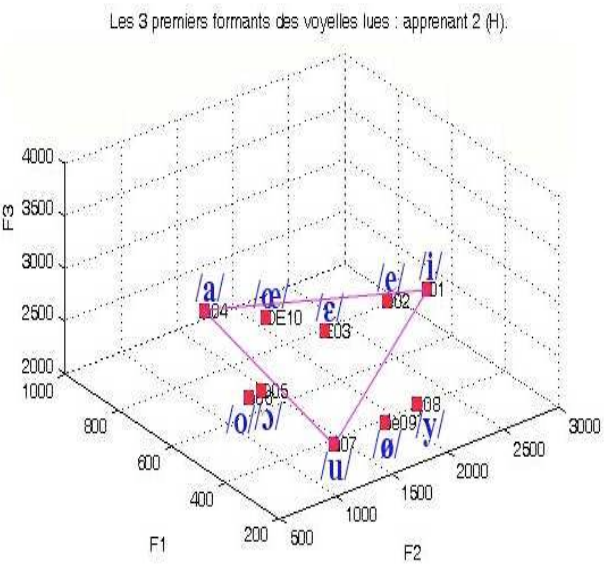


**Figure 61 :** Valeurs formantiques (en Hertz) des voyelles orales du français prononcées par l'apprenant 2 dans la tâche de répétition immédiate (2 répétitions après chacun des deux locuteurs natifs : locuteur FR1 en haut, locuteur FR2 au milieu), et dans la tâche de lecture (2 répétitions : en bas). Les quatre premiers formants (à gauche : les barres d'erreur représentent l'écart-type), et les trois premiers formants dans un triangle vocalique à trois dimensions (à droite).

**Apprenant 2 : /u/** L'apprenant 2 (Figure 61) a montré une tendance similaire à celle de l'apprenant 1 concernant le /u/. Le F2 était plus élevé en répétition (1485 Hz après le locuteur natif FR1, 1338 Hz après FR2, contre 1283 Hz en lecture).

**/y/** Le /y/ a été prononcé, comme chez l'apprenant 1, avec un regroupement des F2/F3, et un F2 inférieur (1929 Hz après le locuteur natif 1, 1927 Hz après le locuteur 2) à celui de la lecture (2005 Hz), même si la différence est plus petite que chez l'apprenant 1.

**/ø/** La voyelle /ø/ a été prononcée avec un F2 plus bas (notamment après le locuteur natif 2 : 1625 Hz) qu'en lecture (1767 Hz). Rappelons que cette tendance a été observée également chez l'apprenant 1, même si la valeur est plus élevée chez l'apprenant 2.

**Les voyelles moyennes** Concernant les voyelles moyennes, cet apprenant a différencié /e/ et /ɛ/ mieux quand il a répété après le locuteur natif FR1 que FR2. La paire /o/-/ɔ/ est bien distincte, mais le F2 du /o/ est supérieur à celui du /ɔ/ (1332 Hz pour /o/ contre 1030 Hz pour /ɔ/ dans les répétitions après le locuteur natif FR1, 1112 Hz contre 770 Hz dans les répétitions après le locuteur natif FR2), contrairement à ce qui est attendu et observé (notre Expérience 2) chez les locuteurs francophone natifs, et à ce que le même apprenant a produit dans la tâche de lecture.

**/i/ F3/F4** Apprenant 2 a prononcé /i/ avec F3 et F4 proches dans toutes les conditions (lecture et répétition immédiate après les deux locuteurs natifs). Rappelons que le /i/ chez les locuteurs francophones natifs est caractérisé par le rapprochement des F3 et F4 (Vaissière 2007)[388], tendance observée dans les données de CALLIOPE (1989 : les valeurs représentées à la Figure 24, p. 81)[389] et celles de notre étude (Expérience 2 : Figure 41, p. 123). En revanche, cela n'était pas le cas pour l'apprenant 1 dans ni l'une ni l'autre des conditions (les F2, F3 et F4 sont quasiment équidistants). Nous retrouvons la même tendance chez l'apprenante 3, qui a produit un /i/ avec un F3 plus proche au F2 qu'au F4 dans toutes les conditions. Cela semble refléter la caractéristique observée concernant le /i/ japonais dans Mokhtari et Tanaka (2000[390] : figure reproduite à la Figure 8, p. 41) et nos données (Figure 39, p. 120) : certains locuteurs présentent un rapprochement F3/F4, et d'autres un F2 élevé.

**Différents types de /i/** Les données acoustiques de la répétition montre que l'apprenant 2 a produit /e/ et /i/ avec des valeurs de F1 et F2 similaires (les deux premiers formants sont de 334

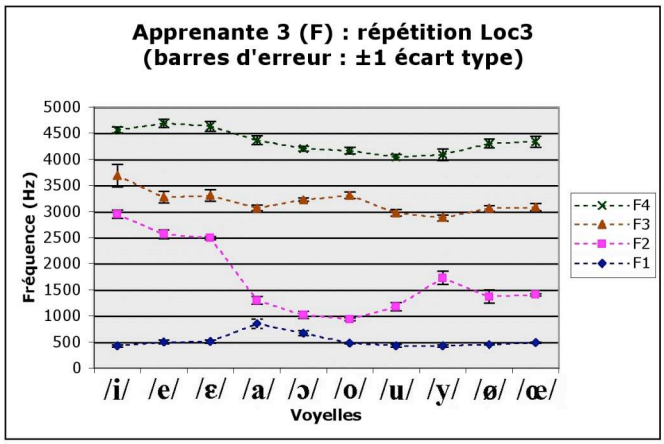
[388] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, and M. Ohala, *Experimental Approaches to Phonology*. Oxford University Press, Oxford, 54-71, 2007.

[389] CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

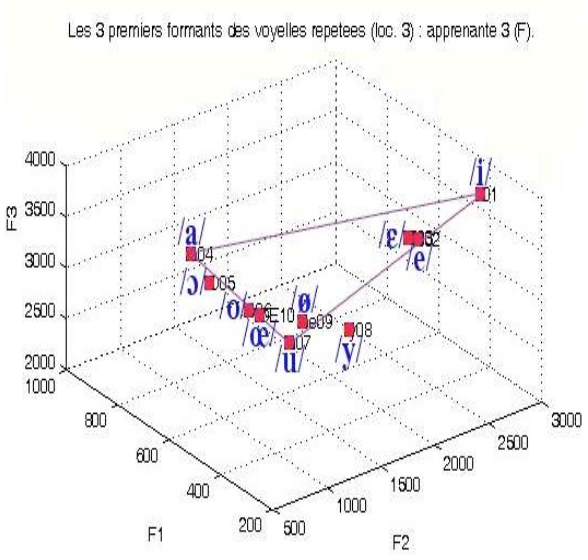
[390] P. Mokhtari, K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66, 2000.

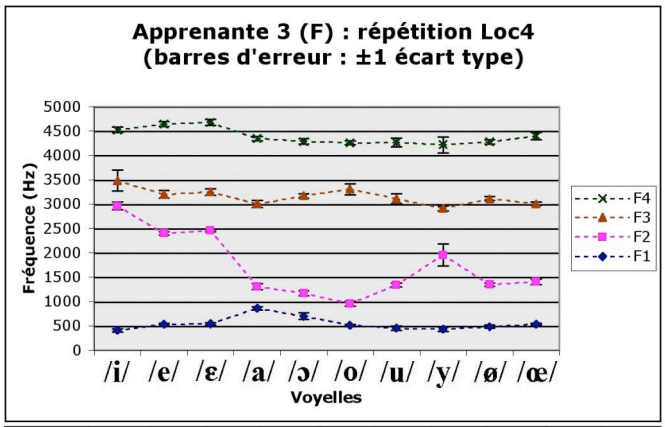
et 2204 Hz pour /i/, 322 et 2237 Hz pour /e/, même si le regroupement F3/F4 est observé uniquement pour /i/) quand il a répété après le locuteur natif FR1, tandis que ce n'était pas le cas pour la répétition après le locuteur natif FR2 et pour la lecture. Il est à noter que le locuteur natif FR1 avait produit des valeurs similaires de F1 et F2 pour /i/ (316 Hz et 2132 Hz) et /e/ (327 Hz et 2084 Hz), et que c'était le F3 qui distinguait les deux voyelles (Tableau 24, p. 143). Il en résulte que l'apprenant 2 a reproduit une structure formantique similaire (F1 et F2 similaires pour /i/ et /e/, F3 élevé uniquement pour /i/), alors que l'apprenant 1 a utilisé une autre stratégie pour différencier les deux voyelles (F1 plus bas et F2 plus élevé pour /i/ : 262 Hz et 2294 Hz ; F1 plus élevé et F2 plus bas pour /e/ : 385 Hz et 2011 Hz).

Nous avons vu dans l'Expérience 3 que /e/ a été perçu occasionnellement comme /i/ (5 réponses sur 40 : Tableau 20, p. 127), mais considérant le produit acoustique de ces deux apprenants (apprenants 1 et 2), nous pouvons supposer qu'ils ont bien identifié perceptivement /i/ et /e/ lors de la tâche de répétition.

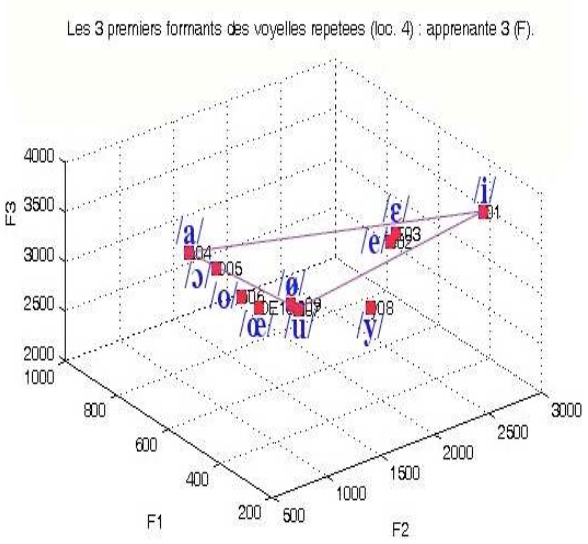


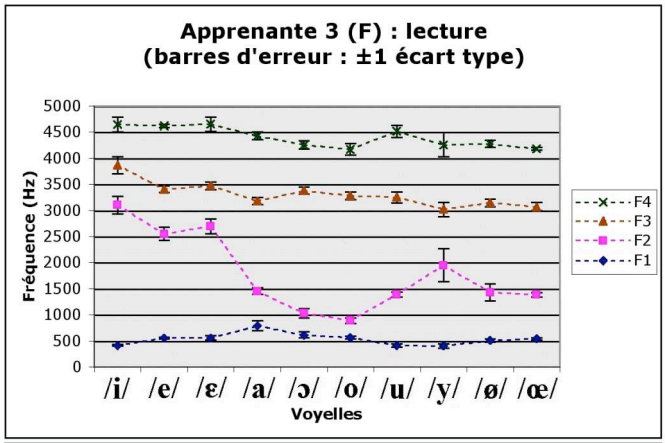

| Appr. 3 Rep1 | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 4564 | 4693 | 4631 | 4365 | 4203 | 4165 | 4041 | 4092 | 4304 | 4338 |
| F3 | 3686 | 3276 | 3304 | 3066 | 3225 | 3308 | 2973 | 2876 | 3061 | 3082 |
| F2 | 2948 | 2560 | 2493 | 1292 | 1015 | 934 | 1176 | 1730 | 1366 | 1403 |
| F1 | 420 | 496 | 507 | 851 | 660 | 474 | 421 | 419 | 447 | 483 |




| Appr. 3 Rep2 | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 4518 | 4636 | 4675 | 4342 | 4284 | 4253 | 4260 | 4213 | 4275 | 4392 |
| F3 | 3479 | 3197 | 3247 | 2999 | 3167 | 3304 | 3106 | 2909 | 3110 | 3003 |
| F2 | 2959 | 2398 | 2453 | 1303 | 1169 | 953 | 1339 | 1952 | 1343 | 1404 |
| F1 | 409 | 535 | 539 | 858 | 697 | 510 | 451 | 428 | 482 | 533 |




| Appr. 3 Lec | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| F4 | 4647 | 4621 | 4653 | 4426 | 4255 | 4172 | 4514 | 4257 | 4276 | 4181 |
| F3 | 3865 | 3405 | 3471 | 3175 | 3380 | 3282 | 3250 | 3018 | 3146 | 3068 |
| F2 | 3101 | 2547 | 2693 | 1452 | 1027 | 888 | 1388 | 1945 | 1429 | 1382 |
| F1 | 411 | 551 | 553 | 785 | 610 | 558 | 406 | 399 | 504 | 538 |

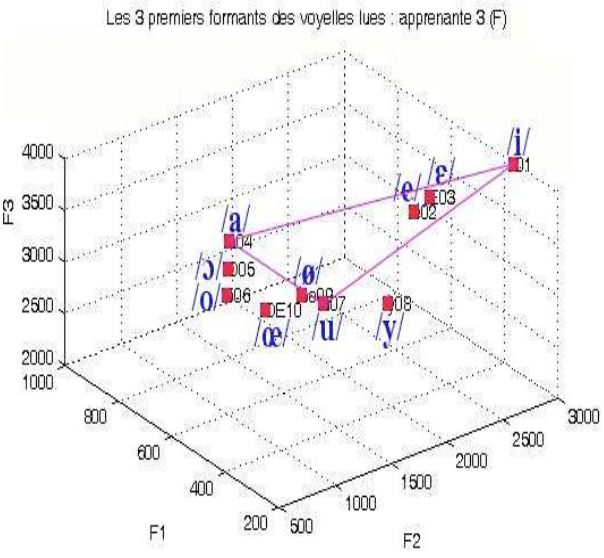



**Figure 62 :** Valeurs formantiques (en Hertz) des voyelles orales du français prononcées par l'apprenante 3 dans la tâche de répétition immédiate (2 répétitions après chacune des deux

locutrices natives : locutrice FR3 en haut, locutrice FR4 au milieu), et dans la tâche de lecture (2 répétitions : en bas). Les quatre premiers formants (à gauche : les barres d'erreur représentent l'écart-type), et les trois premiers formants dans un triangle vocalique à trois dimensions (à droite).

**Apprenante 3 : /u/** Chez l'apprenante 3, le F2 du /u/ n'était pas plus élevé dans la tâche de répétition immédiate qu'en lecture, contrairement aux deux premiers apprenants déjà décrits. Il est moins élevé dans la tâche de répétition après la locutrice native FR3 (1176 Hz) que dans celle de lecture (1388 Hz), mais le même effet ne se retrouve pas pour la répétition après la locutrice native FR4 (1339 Hz). Cependant, les F2 de ces locutrices natives ne sont pas très différents (celui de FR3 est même plus élevé : 805 Hz pour FR3, 733 Hz pour FR4. Voir le Tableau 24, p. 143). Ce qui pourrait expliquer la différence de F2 chez cette apprenante en fonction de la locutrice modèle est celle de la durée de la prononciation modèle qu'elle a entendue. Le /u/ de FR3 était de 274 millisecondes, tandis que celui de FR4 était de 103 millisecondes. La durée plus importante du /u/ de FR3 aurait permis à cette apprenante de mieux percevoir le modèle et aussi de reproduire la voyelle cible de façon plus correcte, même sans être arrivée à la valeur cible de F2.

**/y/** L'apprenante 3, qui a diphtongué le /y/ en lecture (comme le /ju/ du japonais) a fait de même après une voyelle relativement brève (133 ms : FR4. Voir le Tableau 24, p. 143), mais non pas après une voyelle longue (286 ms : FR3. Voir le Tableau 24) dans une des deux répétitions, même si cette dernière condition ne lui a pas permis de reproduire un « beau » /y/ français : l'écart important entre F2 et F3 (1688 Hz et 2863 Hz pour cette répétition non-diphtonguée) fait de cette voyelle une voyelle plutôt acoustiquement centrale. Il est possible que la longue durée ait facilité à l'apprenante de réaliser que le timbre de la voyelle était nettement différent de celui de la deuxième partie de /ju/ japonais, lui permettant de percevoir une monophtongue, et ainsi d'éviter de produire la séquence /ju/ à la japonaise.

**/ø/ /œ/ et /u/** Concernant les voyelles moyennes /ø/ et /œ/, ils ont été prononcés, comme lors de la lecture, de façon similaire au /u/ prononcée par la même apprenante, même si le F1 de /u/ était plus bas que celui de /ø/, et celui de /œ/ était encore plus élevé. Les trois premiers formants des /u/ /ø/ /œ/ étaient de 421, 1176, et 2973 Hz pour /u/, 447, 1366, et 3061 Hz pour /ø/, 483, 1403, et 3082 Hz pour /œ/ lors de la répétition après FR 3, 451, 1339, et 3106 Hz pour /u/, 482, 1343, et 3110 Hz pour /ø/, 533, 1404, et 3003 Hz pour /œ/ après FR4.

**Les autres voyelles moyennes** Quant aux autres voyelles moyennes, /o/-/ɔ/ ont été bien différenciés, mais ce n'était pas le cas pour la paire /e/ et /ɛ/.

Les caractéristiques observées dans ce chapitre se résument comme suit (Tableau 25).

**Tableau 25** : Caractéristiques observées dans les deux tâches de production (lecture et répétition) chez les trois apprenants japonophones (apprenant 1, apprenant 2, apprenante 3).

| Voyelle | Apprenant | | lecture | répétition |
|---|---|---|---|---|
| /u/ | Appr.1 (H) | F2 élevé (> 1000 Hz) | F2 : 1065 Hz < | 1362 Hz |
| | Appr.2 (H) | | F2 : 1283 Hz < | 1485 Hz, 1338 Hz |
| | Appr.3 (F) | | F2 : 1388 Hz > | 1176 Hz (FR3 : 274 ms), 1339 Hz (FR4 : 103 ms) |
| /y/ | Appr.1 (H) | F2/F3 proches | F2 : 2062 Hz > | 1954 Hz (plus proche des natifs : 1806 Hz, 1749 Hz) |
| | Appr.2 (H) | (F2 élevé) | F2 : 2005 Hz > | 1929 Hz, 1927 Hz |
| | Appr.3 (F) | diphtongué [ɯɪ] | | non diphtongué (F2 1688 Hz, F3 2863 Hz) 1 fois après FR4 long |
| /œ ø/ | Appr.1 (H) | | F2 élevé (1850 Hz, 1770 Hz) > | F2 vers 1500 Hz |
| | Appr.2 (H) | | /ø/ : F2 1767 Hz > | 1710 Hz, 1625 Hz |
| | Appr.3 (F) | similaires à /u/ (F1 : /u/ < /ø/ < /œ/) | | |
| /i/ /e/ | Appr.1 (H) | pas de rapprochement F3/F4 pour /i/ ; /i/ et /e/ différenciés par F1 et F2 | | |
| | Appr.2 (H) | non F3/F4 ; /i/ /e/ différnciés par F1 et F2 | | F1 et F2 similaires pour /i/ et /e/, F3 élevé pour /i/ après FR1 |
| | Appr.3 (F) | pas de rapprochement F3/F4 pour /i/ ; /i/ et /e/ différenciés par F1 et F2 | | |
| /e/ /ɛ/ | Appr.1 (H) | différenciés | | |
| | Appr.2 (H) | | différencié | mieux différenciés après FR1 qu'après FR2 |
| | Appr.3 (F) | non-différenciés | | |
| /o/ /ɔ/ | Appr.1 (H) | | F2 : /o/ 830 Hz > /ɔ/ 749 Hz | mieux différenciés : F2 : /o/ 823 Hz < /ɔ/ 890 Hz |
| | Appr.2 (H) | | mieux différenciés | F2 : /o/ 1332 Hz, 1112 Hz > /ɔ/ 1030 Hz, 770 Hz |
| | Appr.3 (F) | différenciés | | |

Nous avons présenté dans ce chapitre les voyelles orales du français prononcées par des apprenants japonophones, en comparant des données recueillies dans deux tâches différentes ; lecture et répétition. En ce faisant, nous avons observé des écarts par rapports à la production des natifs, qui seraient dus à l'influence de l'orthographe et/ou de la transcription (ex. l'association de l'orthographe <eu> avec /ju:/ en anglais, et avec le /ju/ japonais, et ainsi avec le /y/ français : Figure 59, p. 142), d'un côté, et aux difficultés de perception et d'articulation (ex. /u/ prononcé avec un F2 élevé), de l'autre. Comment, alors, les auditeurs francophones natifs percevront-ils ces voyelles produites par des apprenants japonophones ?

## 2.4. IDENTIFICATION : LES VOYELLES ORALES ISOLÉES PRODUITES PAR DES APPRENANTS JAPONOPHONES ET PERÇUES PAR DES AUDITEURS FRANCOPHONES NATIFS : TEST D'IDENTIFICATION ET D'ÉVALUATION (EXPÉRIENCE 5)[391]

Nous nous intéressons maintenant à la question de l'identification par des auditeurs natifs des voyelles isolées prononcées par les apprenants japonophones. L'un des objectifs de l'apprentissage de la prononciation étant d'acquérir les compétences de produire des sons (et des énoncés) compréhensibles auprès des interlocuteurs natifs et non-natifs du français, il est important d'examiner l'aspect perceptif des voyelles produites par les apprenants. Dans le présent chapitre, nous présentons une expérience d'identification dans laquelle des auditeurs francophones natifs ont identifié et évalué les voyelles isolées prononcées par des apprenants japonophones lors de l'expérience précédente (Expérience 4).

### *2.4.1. Les stimuli*

Nous avons utilisé les 10 voyelles orales prononcées dans les tâches de lecture (2.3.1.) et de répétition (2.3.2.) par les 3 apprenants japonophones (2 hommes et 1 femme) pour l'Expérience 4.

[391] Une partie des résultats de cette expérience a été présentée dans la communication suivante :

T. Kamiyama. Acquisition of French vowels by Japanese-speaking learners: close and close-mid rounded vowels. Phonetics and Phonology in Third Language Acquisition (réunion satellite du Congrès International des Sciences Phonétiques), Université de Fribourg (en Brisgau), Allemagne, 3-4 août 2007.

Nous avons extrait uniquement les voyelles isolées des phrases-cadre « *je dis /V/ comme dans …* ». Nous avons ainsi obtenu 160 stimuli :

- 10 voyelles orales x 2 répétitions x
  (3 apprenants x 1 tâche de lecture : 60 stimuli de la lecture
  + 2 apprenants x 2 modèles de répétition
  + 1 apprenant x 1 modèle de répétition). : 100 stimuli de la répétition

Suite à des problèmes techniques, 2 stimuli de la lecture (/ɔ/ et /ø/), ainsi que 5 stimuli de répétition (/u/, 2 x /y/, 2 x /i/) ont été exclus de la liste, ce qui nous a laissé 153 stimuli de voyelles orales (58 de lecture, 95 de répétition). Ces stimuli ont été placés dans un ordre semi-aléatoire pré-établi de sorte que la même voyelle cible ne se succède pas, et que des stimuli prononcés par un apprenant donné ne se suivent pas trois fois ou plus.

### *2.4.2. Les auditeurs*

Les auditeurs étaient des étudiants qui suivaient le cours de japonais (DUFL / UE5) niveau débutant à l'Université Charles de Gaulle Lille III pendant l'année 2006-2007. L'expérience a été effectuée au début de l'année universitaire dans le cadre de la présentation de la langue japonaise, et par conséquent, les participants de ce cours n'avaient pas de connaissances préalables sur le système phonologique du japonais. Nous avons analysé les réponses de 26 francophones natifs.

### *2.4.3. Les procédures*

Les tâches des 26 auditeurs francophones consistaient à écouter les stimuli diffusés par une paire de haut-parleurs, d'identifier une voyelle française, et de l'évaluer sur une échelle de 1 à 5. Pour ce faire, ils disposaient d'une feuille de réponse imprimée, sur laquelle étaient indiquées les 10 voyelles orales avec un numéro (1-10), la transcription phonémique (/i/, /e/, …), et un mot exemple ("lit", "les", … : voir la Figure 63). Les sujets-auditeurs devaient donner le numéro qui correspondait à la voyelle perçue, ainsi que celui qui correspondait à l'évaluation.

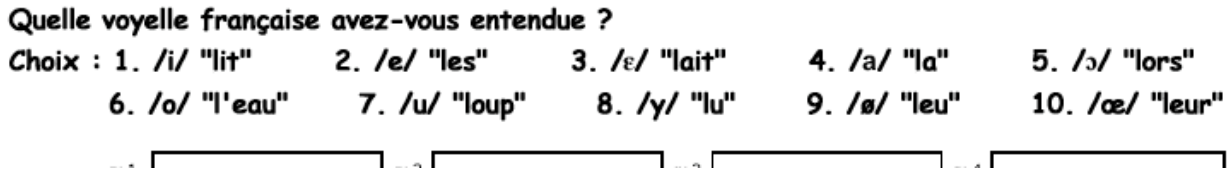



**Figure 63 :** Extrait de la feuille de réponse utilisée dans l'Expérience 5.

L'expérience comportait les trois phases suivantes : l'habituation, l'entraînement et le test.

**L'habituation** Pendant l'habituation, les auditeurs ont écouté des mots composés du /l/ suivi des 10 voyelles orales (/ɔ/ et /œ/ étant suivis par un /ʁ/) prononcés par la locutrice native 4 dans une des 2 répétitions pour l'Expérience 2 (Figure 41), afin que les auditeurs puissent établir une correspondance entre le son vocalique et la notation (numéro, transcription phonémique, exemple de mot) sur la feuille de réponse.

**L'entraînement** Dans la phase d'entraînement, ils ont écouté les 10 voyelles isolées prononcées par la même locutrice native 4. Les stimuli, précédés chacun par un bip sonore (377 Hz, 50 millisecondes) placé 500 millisecondes avant le stimulus, ont été présentés dans un ordre

semi-aléatoire pré-établi. Les tâches des auditeurs étaient d'écouter le stimulus et de l'identifier comme une voyelle orale du français, ainsi que de donner une note (très bien : 5, bien, assez bien, passable, nul : 1). Les auditeurs disposaient de 6 secondes entre deux stimuli. Ils ont écrit la réponse sur la feuille de réponse.

**Le test** Ils sont passés par la suite au test. Ils ont écouté les voyelles isolées prononcées par les 3 apprenants pour l'Expérience 4 (Figure 49, Figure 60, Figure 61, Figure 62), et mises sur la liste décrite ci-dessus (section 2.4.1. les stimuli). Les tâches étaient identiques à celle de l'entraînement. Une pause a été proposée après les 80 premiers stimuli. Le temps de passation était de 20 minutes environ (18 minutes 40 secondes sans compter la pause).

**Le calcul des résultats** Les deux types de réponses (la voyelle identifiée et la note de l'évaluation) ont été traités de manière suivante :

**1)** Le taux d'identification correcte a été d'abord calculé pour chaque stimulus. Ensuite, la valeur moyenne de tous les stimuli d'une voyelle donnée a été calculée. Nous avons ainsi obtenu le taux moyen d'identification de chacune des 10 voyelles orales prononcées par les 3 apprenants japonophones.

**2)** Les notes de l'évaluation (1 : nul – 5 : très bien) ont été intégrées au taux d'identification pour obtenir une note « fréquence x évaluation » (voir Guion *et al*. 2000)[392], calculée de façon suivante : si un stimulus donné a été perçu comme /u/ dans 80% des cas avec une note moyenne de 4,2, et comme /ø/ dans 20% des cas avec une note moyenne de 3,2, la note « fréquence x évaluation » s'élèvera à 3,36 (= 4,2 x 0,8) pour /u/ et 0,64 (= 3,2 x 0,2) pour /ø/. Si un autre stimulus a obtenu 3,8 pour /u/ et 0,3 pour /ø/, nous pourrons considérer que ce dernier (/u/ : 3,8, /ø/ : 0,3) est un meilleur exemplaire du /u/ que n'en est le premier (/u/ : 3,36, /ø/ : 0,64). Par la suite, la note moyenne de chacune des 10 voyelles orales prononcées par les 3 apprenants a été calculée.

### *2.4.4. Les résultats*

Nous présenterons d'abord les résultats globaux des stimuli prononcés dans la tâche de lecture (2.4.4.1.), ceux de la tâche de répétition immédiate (2.4.4.2.), et par la suite, les résultats classés stimulus par stimulus pour chaque voyelle.

#### 2.4.4.1. Les stimuli obtenus dans la tâche de lecture

Le Tableau 26 représente le résultat de l'identification des voyelles produites dans la tâche de lecture. Le taux moyen de l'identification incorrecte (c'est-à-dire, non-correspondance entre le phonème vocalique visé par les apprenants et la perception des auditeurs natifs) était de 43%.

[392] S. G. Guion, J. E. Flege, R. Akahane-Yamada, J. C. Pruitt. 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, 2000.

**Tableau 26** : Résultat de l'identification (matrice de confusion) des voyelles orales produites par les 3 apprenants japonophones dans la tâche de lecture (le taux d'identification en pourcentage). 130 réponses pour /ɔ/ et /ø/ (26 auditeurs x 5 stimuli), 156 réponses pour toutes les autres voyelles (26 auditeurs x 6 stimuli).

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 100 | | | | | | | | | | | 100 |
| e | | 70 | 30 | | | | | | | | | 100 |
| ɛ | 1 | 37 | 55 | 1 | | | | | 3 | 2 | 1 | 100 |
| a | | 1 | 1 | 87 | 1 | | | | 3 | 4 | 4 | 100 |
| ɔ | | 1 | | 5 | 12 | 59 | 8 | 1 | 8 | 5 | 1 | 100 |
| o | | | | | 3 | 94 | 3 | | | | | 100 |
| u | 1 | | | | | | 42 | 13 | 39 | 6 | | 100 |
| y | 4 | 3 | | 1 | | | 14 | 69 | 7 | 1 | 1 | 100 |
| ø | | 1 | 1 | | | | 11 | 62 | 20 | 5 | 1 | 100 |
| œ | | 8 | 13 | | 1 | | 3 | 1 | 56 | 18 | | 100 |
| Total_réponses | 106 | 121 | 100 | 93 | 17 | 153 | 80 | 144 | 137 | 41 | 8 | |

**Les voyelles moyennes** Nous y observons des fréquences élevées de confusion entre les voyelles moyennes. Cette tendance n'est pas surprenante, car les apprenants n'ont pas toujours bien différencié les paires de voyelles moyennes en production (Expérience 4 : Figure 49, Figure 60, Figure 61, Figure 62). Nous y trouvons également une asymétrie entre les voyelles mi-fermées et mi-ouvertes. Ce sont plus souvent les voyelles mi-ouvertes qui ont été perçues comme voyelles mi-fermées que dans le sens inverse. Cette tendance n'est pas très marquée pour la paire /e/-/ɛ/ (37% pour /ɛ/->/e/ contre 30% pour /e/->/ɛ/), mais considérable pour /o/-/ɔ/ (59% pour /ɔ/->/o/ contre 3% pour /o/->/ɔ/) et /ø/-/œ/ (56% pour /œ/->/ø/ contre 4% pour /ø/->/œ/). Nous avons déjà vu que les voyelles mi-ouvertes /ɔ/ et /œ/ ne se prononcent jamais en position finale absolue (rappel : Wioland 2005 : 108[393], entre autres). Les cas de confusions observés ici (notamment /ɔ/->/o/, et /œ/->/ø/) ne poseraient donc aucun problème aux apprenants dans des situations réelles de communication en dehors du laboratoire. Par ailleurs, nous avons constaté une tendance similaire quand des auditeurs natifs du français ont identifié les voyelles du français prononcées par des locuteurs natifs (Expérience 3 : Tableau 20).

**Les voyelles moyennes regroupées** En regroupant les stimuli et les réponses des voyelles moyennes, nous obtenons le Tableau 27. Le taux moyen d'identification incorrecte ainsi calculé baisse à 25% (versus 43%).

[393] F. Wioland. *La vie sociale des sons du français*. Paris, l'Harmattan. 2005.

**Tableau 27** : Résultat de l'identification des voyelles orales produites par les 3 apprenants japonophones dans la tâche de lecture (le taux d'identification en pourcentage). Les voyelles moyennes ont été regroupées. 312 réponses pour /e ɛ/ (26 auditeurs x 12 stimuli), 286 réponses pour /ɔ o/ et /ø œ/ respectivement (26 auditeurs x 11 stimuli), 156 réponses pour toutes les autres voyelles (26 auditeurs x 6 stimuli).

| stimuli\réponses | i | e/ɛ | a | ɔ/o | u | y | ø/œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|
| **i** | **100** | | | | | | | | 100 |
| **e/ɛ** | 0,3 | **96** | 0,3 | | | | 3 | 1 | 100 |
| **a** | | 1 | **87** | 1 | | | 7 | 4 | 100 |
| **ɔ/o** | | 0,3 | 2 | **86** | 5 | 0,3 | 6 | 0,3 | 100 |
| **u** | 1 | | | | **42** | 13 | 45 | | 100 |
| **y** | 4 | 3 | 1 | | 14 | **69** | 8 | 1 | 100 |
| **ø/œ** | | 13 | | 0,3 | 7 | 28 | **52** | 0,3 | 100 |
| Total_réponses | 105 | 114 | 90 | 87 | 67 | 110 | 121 | 6 | |

**/u y ø œ/** Nous y voyons clairement que les cas de confusion sont concentrés sur les voyelles arrondies fermées /u/ /y/ et moyennes /ø œ/. Le seul cas de confusion où le taux dépasse les 10% en dehors de ce groupe est celui de /ø œ/ perçus comme /e ɛ/.

**/y/ relativement bien identifié** Parmi le groupe /u/ /y/ /ø œ/, le /y/ a été relativement bien identifié (69%) par rapport aux autres (42% pour /u/ et 47% pour /ø œ/). La bonne identification du /y/ serait due au fait que les apprenants 1 et 2 ont prononcé cette voyelle avec F2 et F3 proches ; l'apprenante 3 a produit une séquence diphtonguée comparable au /ju/ japonais (Figure 49 à la page 135, Figure 55 à la page 139 : les différences de taux d'identification entre les stimuli seront traitées dans la section 2.4.4.3.). Les apprenants 1 et 2 ont prononcé également le /ø/ avec F2/F3 proches (même si ces deux formants sont plus écartés que pour /y/ : Figure 49), ce qui aurait baissé le taux d'identification correcte de cette voyelle (47% pour /ø/ contre 26 % pour /y/).

**/u/** Quant au /u/, la valeur trop élevée de F2 observée auprès de tous les 3 apprenants (Figure 49) serait responsable du taux faible d'identification correcte (42% pour /u/), et d'un taux de réponse même plus élevé (45%) pour /ø œ/. Ces résultats semblent illustrer l'intérêt d'examiner ces voyelles /u/, /y/, et /ø/ (+ /œ/) en détail.

**/i/ F3/F4 ?** Il est intéressant de noter que le /i/ prononcé par les 3 apprenants a été identifié correctement à 100%. Il est fréquent que deux phonèmes correspondants dans deux langues différentes, transcrits communément avec le même symbole API soient phonétiquement différents, comme nous l'avons vu pour le /u/ français et le /u/ [ɯ] japonais. Delattre (1965)[394] a montré des différences articulatoires entre le /i/ français et le /i/ de l'anglais américain, transcrits

[394] P. Delattre. Comparing the vocalic features of English, German, Spanish and French. *International Review of Applied Linguistics* 2 : 71-97, 1965.

tous les deux avec le même symbole. Vaissière (2007)[395] a modélisé ces différences articulatoires en utilisant la synthèse articulatoire de Maeda. Le fait que les auditeurs francophones aient bien perçu /i/ suggère que les apprenants japonophones n'auraient pas de difficulté à produire le /i/ à la française, avec un rapprochement de F3 et F4. Les données de notre Expérience 3 (Figure 49) montrent effectivement que ces deux formants sont très proches chez l'apprenant 2, du moins, et relativement proches (distance F3/F4 < distance F2/F3) chez l'apprenant 1. Les données sur la production du /i/ japonais (Mokhtari et Tanaka 2000 : figure reproduite à la Figure 8 à la page 41, nos données : Figure 39 à la page 120) montrent souvent un rapprochement F3/F4, malgré des variabilités intra-locuteur. Nous verrons ci-dessous si les [i] prononcés en tant que /i/ français par les 3 apprenants japonais ont obtenu une meilleure évaluation que les autres [i].

**Les notes « fréquence x évaluation »** Le Tableau 28 représente les notes « fréquence x évaluation », qui portent sur le taux d'identification et l'évaluation.

**Tableau 28** : Notes « fréquence x évaluation » moyennes des 10 voyelles orales prononcées par les 3 apprenants japonophones en lecture (max. : 5 – min. : 0). 130 réponses pour /ɔ/ et /ø/ (26 auditeurs x 5 stimuli), 156 réponses pour toutes les autres voyelles (26 auditeurs x 6 stimuli).

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| i | 4,3 | | | | | | | | | |
| e | | 2,7 | 1,0 | | | | | | | |
| ɛ | 0,0 | 1,4 | 1,9 | 0,0 | | | | | 0,1 | 0,0 |
| a | | 0,0 | 0,0 | 2,9 | 0,0 | | | | 0,1 | 0,1 |
| ɔ | | 0,0 | | 0,1 | 0,3 | 2,1 | 0,2 | 0,0 | 0,3 | 0,1 |
| o | | | | | 0,1 | 3,7 | 0,1 | | | |
| u | 0,0 | | | | | | 1,2 | 0,3 | 1,2 | 0,2 |
| y | 0,1 | 0,1 | | 0,0 | | | 0,4 | 2,4 | 0,1 | 0,0 |
| ø | | 0,0 | 0,0 | | | | 0,4 | 2,2 | 0,5 | 0,2 |
| œ | | 0,3 | 0,5 | | 0,0 | | 0,1 | 0,0 | 1,9 | 0,6 |

**Des tendances similaires** Nous pouvons y retrouver des tendances similaires à celles des taux d'identification : des non correspondance (entre l'intention de l'apprenant et la perception native) entre les voyelles moyennes, l'asymétrie entre les voyelles mi-fermées et mi-ouvertes, des confusions entre /u/, /y/, et /ø/.

**/i/ et /o/ : bons exemplaires** Concernant les voyelles /i/ et /o/, elles ont été identifiées non seulement avec des taux élevés (100% et 94%, respectivement : Tableau 26), mais aussi avec une note relativement élevée (4,3 et 3,7), tandis que la note du /a/ n'est pas très élevée (2,9) par

[395] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

rapport à un taux relativement élevé d'identification (87%). Cela suggère que les apprenants ont produit de meilleurs exemplaires de /i/ et de /o/ que de /a/.

**/i/ F3/F4 : phonétiquement et phonémiquement similaire ?** Les bonnes notes que les stimuli du /i/ ont obtenues seraient dues, comme nous l'avons vu ci-dessus, au regroupement de F3 et F4 produit par certains apprenants. Le /i/ représente donc un exemple de voyelle similaire dans la langue cible et la langue source sur le plan à la fois phonétique et phonémique, qui ne pose pas de difficulté de production aux apprenants, du moins pour ceux qui produisent /i/ avec un regroupement de F3/F4 en japonais.

**/o/ /e/ aussi ?** Il en est presque de même pour /o/, et un peu moins pour /a/, qui est suivi de /e/.

#### 2.4.4.2. Les stimuli obtenus dans la tâche de répétition

Le Tableau 29 représente le résultat de l'identification des voyelles produites dans la tâche de répétition. Le taux moyen de l'identification incorrecte (c'est-à-dire la non-correspondance entre le phonème vocalique que les apprenants voulaient prononcer et la perception des auditeurs natifs) était de 50% (contre 43% pour les stimuli de la tâche de lecture).

**Tableau 29** : Résultat de l'identification des voyelles orales produites par les trois apprenants japonophones dans la tâche de répétition immédiate (le taux d'identification en pourcentage). 208 réponses pour /i/ et /y/ (26 auditeurs x 8 stimuli), 234 réponses pour /u/ (26 auditeurs x 9 stimuli), 260 réponses pour toutes les autres voyelles (26 auditeurs x 10 stimuli).

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | 92 | 2 | 3 | 0,5 | | | | 0,5 | | | 2 | 100 |
| e | 30 | 43 | 27 | | | | | | | | | 100 |
| ɛ | | 49 | 51 | | | | | | | 0,4 | | 100 |
| a | 0,4 | | | 86 | 3 | 5 | 0,4 | | 0,4 | 1 | 3 | 100 |
| ɔ | | 1 | | 32 | 17 | 38 | 3 | 0,4 | 5 | 3 | 0,4 | 100 |
| o | | 0,4 | 0,4 | 0,4 | 9 | 65 | 6 | 3 | 13 | 3 | 0,4 | 100 |
| u | | 1 | 0,4 | | 0,4 | 0,4 | 17 | 16 | 52 | 12 | 0,4 | 100 |
| y | | | | | | | 6 | 67 | 18 | 8 | 0,5 | 100 |
| ø | | 2 | | | | 0,4 | 7 | 32 | 47 | 12 | 0,4 | 100 |
| œ | | 3 | 1 | | 0,4 | 1 | 1 | 10 | 60 | 23 | 0,4 | 100 |
| Total_réponses | 122 | 100 | 83 | 119 | 30 | 110 | 40 | 128 | 196 | 64 | 7 | |

**Des tendances déjà observées et des tendances propres** Nous y trouvons non seulement des tendances déjà observées pour les stimuli de la lecture, mais aussi celles qui sont propres aux stimuli de la répétition immédiate.

**Les voyelles moyennes : vers les mi-fermées** Les voyelles moyennes présentent les mêmes patterns de confusion : les voyelles mi-ouvertes ont été perçues comme les voyelles mi-fermées plus fréquemment que dans le sens inverse, sauf que cette tendance est plus marquée pour la paire /e/-/ɛ/ (49% pour /ɛ/->/e/ contre 27% pour /e/->/ɛ/. La différence était moins importante pour les stimuli de la lecture : 37% pour /ɛ/->/e/ contre 30% pour /e/->/ɛ/).

**Les voyelles moyennes regroupées** En regroupant les stimuli et les réponses des voyelles mi-fermées et mi-ouvertes, nous pouvons établir le Tableau 30. Le taux moyen d'identification incorrecte ainsi calculé baisse à 32%.

**Tableau 30** : Résultat de l'identification des voyelles orales produites par les 3 apprenants japonophones dans la tâche de répétition immédiate (le taux d'identification en pourcentage). Les voyelles moyennes ont été regroupées. 208 réponses pour /i/ et /y/ (26 auditeurs x 8 stimuli), 234 réponses pour /u/ (26 auditeurs x 9 stimuli), 520 réponses pour les voyelles moyennes /e ɛ/, /o ɔ/ et /ø œ/ respectivement (26 auditeurs x 20 stimuli), 260 réponses pour toutes les autres voyelles (26 auditeurs x 10 stimuli).

| stimuli\réponses | i | e/ɛ | a | ɔ/o | u | y | ø/œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|
| i | 92 | 5 | 0,5 | | | 0,5 | | 2 | 100 |
| e/ɛ | 15 | 85 | | | | | 0,2 | | 100 |
| a | 0,4 | | 86 | 9 | 0,4 | | 1 | 3 | 100 |
| ɔ/o | | 1 | 16 | 65 | 4 | 2 | 13 | 0,4 | 100 |
| u | | 1 | | 1 | 17 | 16 | 64 | 0,4 | 100 |
| y | | | | | 6 | 67 | 26 | 0,5 | 100 |
| ø/œ | | 3 | | 1 | 4 | 21 | 71 | 0,4 | 100 |
| Total_réponses | 108 | 95 | 103 | 75 | 32 | 106 | 176 | 7 | |

**/u/** Le taux d'identification correcte pour le /u/ est moins élevé (17%) que pour les stimuli de la lecture (47%). Nous avons observé ci-dessus que les 5 apprenants qui ont passé le test d'identification des voyelles françaises (Expérience 3) ont eu de la difficulté à identifier correctement le /u/ français (perçu comme /ø/ dans 12 cas sur 40 : Tableau 22). Nous avons également vu que cette difficulté de perception serait responsable du fait que les apprenants 1 et 2 ont prononcé le /u/ avec le F2 plus élevé dans la tâche de répétition immédiate (1362 Hz chez l'apprenant 1, 1485 Hz et 1338 Hz chez l'apprenant 2 : Figure 60 et Figure 61) que dans la tâche de lecture (1065 Hz chez l'apprenant 1, 1283 Hz chez l'apprenant 2) dans l'Expérience 4. Ces valeurs élevées de F2, à leur tour, auraient favorisé la perception des /ø œ/ (64% : 52% pour /ø/ et 12% pour /œ/) auprès des auditeurs francophones.

**/y/** Les stimuli de la répétition du /y/ ont été presque aussi bien perçus (67%) que ceux de la lecture (69%). Il y a plus de réponses pour /ø/ (18% ; contre 7% pour les stimuli de la lecture), ce qui serait dû au nombre plus petit de stimuli diphtongués (comme le /ju/ japonais), au profit de ceux qui ont un rapprochement insuffisant de F2/F3 chez l'apprenante 3 (Figure 62). Les réponses triées par stimulus seront examinées dans la section suivante.

**/ø/** Par contraste, le /ø/ a été plus correctement perçu (47% ; 71% après regroupement des /ø/ et /œ/) que les stimuli de la lecture (20% ; 52% après regroupement des /ø/ et /œ/ : Tableau 26 et Tableau 27). Cette amélioration par rapport à la production de la lecture pourrait s'expliquer par le fait que les apprenants 1 et 2 n'ont pas produit cette voyelle avec un rapprochement de F2/F3 dans la tâche de répétition immédiate, contrairement à la tâche de lecture (Figure 60 et Figure 61). Nous avons vu que la lecture à l'anglaise de l'orthographe <eu> serait responsable des

productions telles que [y] ou [jɯ] (à la japonaise). La tâche de répétition immédiate aurait ainsi supprimé une source de difficulté, ce qui suggère que la prise de conscience de la valeur phonétique de cette voyelle faciliterait sa production correcte. Cependant, il reste une autre source de difficulté sur le plan perceptif. Les cinq apprenants qui ont participé à l'Expérience 3 (identification des voyelles françaises) ont perçu /y/ comme /ø/ dans 6 cas sur 40, et /ø/ comme /y/ dans 5 cas sur 40 (Tableau 22). Le fait que les stimuli de la répétition du /ø/ utilisés dans la présente expérience aient été perçus comme /y/ dans 32% des cas (21% pour les /ø œ/ regroupés) par les auditeurs natifs (malgré une diminution par rapport à la lecture : 62%) semble refléter cette difficulté de perception.

Notons également que les auditeurs ont choisi cette voyelle presque deux fois plus souvent que la moyenne (le pourcentage total de ce choix s'élève à 196%), ce qui suggère que les apprenants japonophones parviennent à la produire, consciemment ou inconsciemment, assez facilement en répétition, par rapport au /u/ (40% de la moyenne), entre autres.

**/e/** Les stimuli du /e/ répété ont été perçus comme /e/ (43%) et /ɛ/ (27%), et aussi comme /i/ (30%), à la différence de ceux de la lecture (0% perçu comme /i/). L'Expérience 3 nous a montré que le /e/ prononcé par des locuteurs natifs a été perçu comme /i/ par les 5 apprenants japonais dans 5 cas sur 40 (Tableau 22). Certaines des répétitions utilisées dans le présent test de perception semblent refléter cette tendance de perception (les résultats seront analysés stimulus par stimulus dans la section suivante).

**/o/** La voyelle /o/ prononcée par les 3 apprenants dans la tâche de répétition a été perçue correctement dans la majorité des cas (65%), mais aussi comme /ø/ dans 13% des cas (*cf*. aucune réponse de /ø/ pour les stimuli de la lecture). Dans l'Expérience 3, les apprenants ont perçu /o ɔ/ des locuteurs natifs comme /ø œ/ dans 10% des cas (Tableau 23), ce qui serait responsable du résultat de la présente expérience. Notons également que le /o/ prononcé par des natifs a été perçu comme /u/ dans 8 cas sur 40 par les 5 apprenants (Expérience 3 : Figure 17), mais que les 26 auditeurs natifs du présent test ont perçu le /u/ dans 6% des cas seulement. Cette asymétrie serait probablement due à la difficulté de la production du /u/.

**/ɔ/** Cette voyelle a été perçue comme /ɔ/ (17%) et /o/ (38%), et aussi comme /a/ (32%), à la différence des stimuli de la lecture (5% seulement perçu comme /a/). Si l'on tient compte du fait que les cinq auditeurs natifs ont perçu le /ɔ/ prononcé par des natifs comme /a/ dans 5 cas sur 40 (Expérience 3 : Tableau 20), il est raisonnable de supposer que les apprenants aussi ont entendu le /a/ et qu'ils ont produit une voyelle plus ouverte, ce qui aurait dicté le résultat de la présente expérience.

**Les notes « fréquence x évaluation »** Le Tableau 31 représente les notes calculées sur l'identification et l'évaluation des voyelles.

**Tableau 31** : Notes « fréquence x évaluation » moyennes des 10 voyelles orales prononcées par les 3 apprenants japonophones en répétition (max. : 5 – min. : 0). 208 réponses pour /i/ et /y/ (26 auditeurs x 8 stimuli), 234 réponses pour /u/ (26 auditeurs x 9 stimuli), 260 réponses pour toutes les autres voyelles (26 auditeurs x 10 stimuli).

| stimuli\réponses | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|
| i | 3,6 | 0,0 | 0,1 | 0,0 | | | | 0,0 | | |
| e | 1,3 | 1,7 | 1,0 | | | | | | | |
| ɛ | | 1,8 | 1,9 | | | | | | | 0,0 |
| a | 0,0 | | | 3,4 | 0,1 | 0,2 | 0,0 | | 0,0 | 0,0 |
| ɔ | | 0,0 | | 0,8 | 0,6 | 1,3 | 0,1 | 0,0 | 0,2 | 0,1 |
| o | | 0,0 | 0,0 | 0,0 | 0,3 | 2,5 | 0,2 | 0,1 | 0,4 | 0,1 |
| u | | 0,0 | 0,0 | | 0,0 | 0,0 | 0,4 | 0,4 | 1,7 | 0,3 |
| y | | | | | | | 0,2 | 2,2 | 0,5 | 0,2 |
| ø | | 0,0 | | | | 0,0 | 0,2 | 1,0 | 1,5 | 0,4 |
| œ | | 0,1 | 0,0 | | 0,0 | 0,0 | 0,0 | 0,4 | 2,1 | 0,7 |

**Des notes basses par rapport à la lecture** Comme nous l'avons vu pour les stimuli de la lecture, nous y observons une tendance similaire à celle des taux d'identification. Par rapport aux notes des stimuli de la lecture (Tableau 28), nous pouvons remarquer une baisse générale de notes pour les voyelles correctement identifiées, sauf pour /a/ (3,4 ; contre 2,9 pour les stimuli de la lecture) et /ø/ (1,5 ; contre 0,5 pour la lecture). Une des raisons pour cette baisse de notes serait la difficulté relative de la tâche de répétition, qui demande une perception correcte et des gestes articulatoires appropriés.

**/ø/** Concernant la différence de notes entre les stimuli des deux tâches (lecture et répétition immédiate), le /ø/ présente une amélioration nette pour la répétition. Comme nous l'avons déjà vu ci-dessus sur les taux d'identification correcte, cette différence serait due à la non-diphtongaison.

#### 2.4.4.3. Les résultats classés stimulus par stimulus

Nous allons maintenant étudier les résultats classés par stimuli, afin d'examiner mieux le lien entre la perception et la production des apprenants, et l'identification et l'évaluation des auditeurs natifs.

**1) /i/** Les stimuli de la lecture de cette voyelle (Tableau 32) ont été tous perçus correctement par les auditeurs francophones, avec des notes très élevées (≧ 4,0).

**Tableau 32** : Perception des 6 stimuli du /i/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | jp1 | lec1 | 26 | | | | | | | | | | | 26 |
| i | jp1 | lec2 | 26 | | | | | | | | | | | 26 |
| i | jp2 | lec1 | 26 | | | | | | | | | | | 26 |
| i | jp2 | lec2 | 26 | | | | | | | | | | | 26 |
| i | jp3 | lec1 | 26 | | | | | | | | | | | 26 |
| i | jp3 | lec2 | 26 | | | | | | | | | | | 26 |
| Total_réponses | | | 156 | | | | | | | | | | | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | jp1 | lec1 | 4,6 | | | | | | | | | |
| i | jp1 | lec2 | 4,4 | | | | | | | | | |
| i | jp2 | lec1 | 4,4 | | | | | | | | | |
| i | jp2 | lec2 | 4,0 | | | | | | | | | |
| i | jp3 | lec1 | 4,3 | | | | | | | | | |
| i | jp3 | lec2 | 4,4 | | | | | | | | | |
| Moyenne | | | 4,3 | | | | | | | | | |

La moitié (= 4) des stimuli de la répétition (Tableau 33) ont été également perçus à la perfection, avec des notes aussi élevées (> 4,0). L'autre moitié a été évaluée moins bien, ce qui contribue à la baisse de notes moyennes (3,6 ; contre 4,3 pour les stimuli de la lecture).

**Tableau 33** : Perception des 8 stimuli du /i/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | jp1 | rep1 | 26 | | | | | | | | | | | 26 |
| i | jp1 | rep1 | 26 | | | | | | | | | | | 26 |
| i | jp2 | rep1 | 26 | | | | | | | | | | | 26 |
| i | jp2 | rep2 | 21 | | | | | | | 1 | | | 4 | 26 |
| i | jp2 | rep2 | 26 | | | | | | | | | | | 26 |
| i | jp3 | rep1 | 20 | 1 | 4 | 1 | | | | | | | | 26 |
| i | jp3 | rep2 | 23 | 2 | 1 | | | | | | | | | 26 |
| i | jp3 | rep2 | 24 | 1 | 1 | | | | | | | | | 26 |
| Total_réponses | | | 192 | 4 | 6 | 1 | | | | 1 | | | 4 | 208 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | jp1 | rep1 | 4,2 | | | | | | | | | |
| i | jp1 | rep1 | 4,1 | | | | | | | | | |
| i | jp2 | rep1 | 4,2 | | | | | | | | | |
| i | jp2 | rep2 | 3,1 | | | | | | | 0,2 | | |
| i | jp2 | rep2 | 4,3 | | | | | | | | | |
| i | jp3 | rep1 | 2,4 | 0,1 | 0,5 | 0,2 | | | | | | |
| i | jp3 | rep2 | 3,0 | 0,2 | 0,2 | | | | | | | |
| i | jp3 | rep2 | 3,7 | 0,0 | 0,2 | | | | | | | |
| Moyenne | | | 3,6 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |

**2) /e/** Le /e/ de la lecture (Tableau 34) a été entièrement perçu soit comme /e/ soit comme /ɛ/ par les auditeurs natifs. En outre, tous les stimuli ont été identifiés plus fréquemment comme /e/

que comme /ɛ/ (entre 16 pour /e/ contre 10 pour /ɛ/, et 21 pour /e/ et 5 pour /ɛ/), et les notes sont relativement élevées pour /e/ (entre 2,2 et 3,2).

**Tableau 34** : Perception des 6 stimuli du /e/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| e | jp1 | lec1 | | 19 | 7 | | | | | | | | | 26 |
| e | jp1 | lec2 | | 21 | 5 | | | | | | | | | 26 |
| e | jp2 | lec1 | | 17 | 9 | | | | | | | | | 26 |
| e | jp2 | lec2 | | 16 | 10 | | | | | | | | | 26 |
| e | jp3 | lec1 | | 16 | 10 | | | | | | | | | 26 |
| e | jp3 | lec2 | | 20 | 6 | | | | | | | | | 26 |
| Total_réponses | | | | 109 | 47 | | | | | | | | | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| e | jp1 | lec1 | | 2,6 | 1,0 | | | | | | | |
| e | jp1 | lec2 | | 3,2 | 0,7 | | | | | | | |
| e | jp2 | lec1 | | 2,6 | 1,2 | | | | | | | |
| e | jp2 | lec2 | | 2,5 | 1,3 | | | | | | | |
| e | jp3 | lec1 | | 2,2 | 1,3 | | | | | | | |
| e | jp3 | lec2 | | 3,1 | 0,8 | | | | | | | |
| Moyenne | | | | 2,7 | 1,0 | | | | | | | |

Quant à la répétition (Tableau 35), seuls les stimuli que les apprenants 1 et 2 ont produits après le locuteur natif 1 (Tableau 24) ont été identifiés comme /i/ (à l'exception de la deuxième répétition de l'apprenant 1). Nous pouvons supposer que la valeur basse de F1 (327 Hz) et la valeur élevée de F2 (2084 Hz) du /e/ prononcé par le locuteur 1 et utilisé comme modèle de la tâche de répétition (Tableau 24) auraient été responsables des valeurs similaires reproduites par les 2 apprenants (les deux premiers formants sont de 385 Hz et 2011 Hz pour l'apprenant 1, et de 322 Hz et 2237 Hz pour l'apprenant 2 – la moyenne des 2 répétitions : Figure 60 et Figure 61). Par ailleurs, les auditeurs francophones natifs ont identifié le /e/ prononcé par des natifs comme /i/ dans 4 cas sur 40 (Tableau 20) et les apprenants ont fait de même dans 5 cas sur 40 (Tableau 22). Cependant, nous ne pourrons considérer le /e/ prononcé par le locuteur natif FR1 comme un bon exemplaire du /i/ non plus : F3 et F4 seraient trop bas pour /i/ cardinal. De toute manière, l'identification du /e/ de FR1 comme /i/ serait responsable du fait que les 3 stimuli en question, produits par les apprenants, ont été identifiés presque unanimement comme /i/ par les 26 auditeurs natifs de la présente expérience, et cela, avec des notes très élevées (4,6, 4,5, et 4,2).

Les autres stimuli de la répétition présentent la même tendance que celle des stimuli de la lecture, sauf que deux stimuli (la deuxième répétition de l'apprenant 2 d'après le locuteur natif FR2, et la deuxième répétition de l'apprenante 3 d'après la locutrice native FR3) ont été catégorisés plus fréquemment comme /ɛ/ que comme /e/ (10 pour /e/, contre 16 pour /ɛ/).

**Tableau 35** : Perception des 10 stimuli du /e/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| e | jp1 | rep1 | 26 | | | | | | | | | | | 26 |
| e | jp1 | rep1 | | 20 | 6 | | | | | | | | | 26 |
| e | jp2 | rep1 | 26 | | | | | | | | | | | 26 |
| e | jp2 | rep1 | 25 | 1 | | | | | | | | | | 26 |
| e | jp2 | rep2 | | 18 | 8 | | | | | | | | | 26 |
| e | jp2 | rep2 | | 10 | 16 | | | | | | | | | 26 |
| e | jp3 | rep1 | | 17 | 9 | | | | | | | | | 26 |
| e | jp3 | rep1 | | 10 | 16 | | | | | | | | | 26 |
| e | jp3 | rep2 | | 18 | 8 | | | | | | | | | 26 |
| e | jp3 | rep2 | | 18 | 8 | | | | | | | | | 26 |
| Total_réponses | | | 77 | 112 | 71 | | | | | | | | | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| e | jp1 | rep1 | 4,6 | | | | | | | | | |
| e | jp1 | rep1 | | 3,0 | 0,8 | | | | | | | |
| e | jp2 | rep1 | 4,5 | | | | | | | | | |
| e | jp2 | rep1 | 4,2 | 0,2 | | | | | | | | |
| e | jp2 | rep2 | | 2,9 | 1,1 | | | | | | | |
| e | jp2 | rep2 | | 1,5 | 2,2 | | | | | | | |
| e | jp3 | rep1 | | 2,4 | 1,3 | | | | | | | |
| e | jp3 | rep1 | | 1,3 | 2,1 | | | | | | | |
| e | jp3 | rep2 | | 2,7 | 1,1 | | | | | | | |
| e | jp3 | rep2 | | 2,8 | 1,0 | | | | | | | |
| Moyenne | | | 1,3 | 1,7 | 1,0 | | | | | | | |

Pour conclure, nous pouvons considérer que les 3 apprenants ont plutôt bien réussi à prononcer le /e/ dans le sens où les auditeurs natifs ont perçu le /e/ plus fréquemment que le /ɛ/ dans l'ensemble, à l'exception des 3 répétitions qui ont été catégorisées unanimement comme /i/.

**3) /ɛ/** La quasi-totalité des stimuli de la voyelle /ɛ/ a été identifiée comme /e/ ou /ɛ/ (Tableau 36 et Tableau 37). Parmi les stimuli de la lecture, tous ceux qui ont été prononcés par les apprenants 1 et 2 ont été perçus correctement comme /ɛ/ plus souvent que comme /e/, tandis que ceux de l'apprenante 3 ne sont pas parvenus à ce cas de figure : les réponses pour /e/ sont plus nombreuses que celles pour /ɛ/.

**Tableau 36** : Perception des 6 stimuli du /ɛ/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɛ | jp1 | lec1 | | 11 | 15 | | | | | | | | | 26 |
| ɛ | jp1 | lec2 | | 5 | 20 | | | | | | | | 1 | 26 |
| ɛ | jp2 | lec1 | | 4 | 14 | | | | | | 4 | 3 | 1 | 26 |
| ɛ | jp2 | lec2 | | 6 | 19 | | | | | | 1 | | | 26 |
| ɛ | jp3 | lec1 | | 15 | 11 | | | | | | | | | 26 |
| ɛ | jp3 | lec2 | 1 | 17 | 7 | 1 | | | | | | | | 26 |
| Total_réponses | | | 1 | 58 | 86 | 1 | | | | | 5 | 3 | 2 | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɛ | jp1 | lec1 | | 1,7 | 2,2 | | | | | | | |
| ɛ | jp1 | lec2 | | 0,7 | 3,0 | | | | | | | |
| ɛ | jp2 | lec1 | | 0,5 | 1,9 | | | | | | 0,3 | 0,3 |
| ɛ | jp2 | lec2 | | 0,6 | 2,2 | | | | | | 0,1 | |
| ɛ | jp3 | lec1 | | 2,4 | 1,4 | | | | | | | |
| ɛ | jp3 | lec2 | 0,1 | 2,4 | 1,0 | 0,0 | | | | | | |
| **Moyenne** | | | 0,0 | 1,4 | 1,9 | 0,0 | | | | | 0,1 | 0,0 |

Concernant la répétition, le résultat général est moins bon au sens où les stimuli qui ont été perçus correctement comme /ɛ/ plus souvent que comme /e/ sont moins nombreux (5 sur 10) que les stimuli /ɛ/ de la tâche de lecture (4 sur 6).

**Tableau 37** : Perception des 10 stimuli du /ɛ/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɛ | jp1 | rep1 | | 11 | 15 | | | | | | | | | 26 |
| ɛ | jp1 | rep1 | | 20 | 5 | | | | | | | 1 | | 26 |
| ɛ | jp2 | rep1 | | 15 | 11 | | | | | | | | | 26 |
| ɛ | jp2 | rep1 | | 7 | 19 | | | | | | | | | 26 |
| ɛ | jp2 | rep2 | | 5 | 21 | | | | | | | | | 26 |
| ɛ | jp2 | rep2 | | 15 | 11 | | | | | | | | | 26 |
| ɛ | jp3 | rep1 | | 7 | 19 | | | | | | | | | 26 |
| ɛ | jp3 | rep1 | | 17 | 9 | | | | | | | | | 26 |
| ɛ | jp3 | rep2 | | 12 | 14 | | | | | | | | | 26 |
| ɛ | jp3 | rep2 | | 18 | 8 | | | | | | | | | 26 |
| **Total_réponses** | | | | 127 | 132 | | | | | | | 1 | | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɛ | jp1 | rep1 | | 1,5 | 2,2 | | | | | | | |
| ɛ | jp1 | rep1 | | 2,9 | 0,9 | | | | | | | 0,2 |
| ɛ | jp2 | rep1 | | 2,5 | 1,6 | | | | | | | |
| ɛ | jp2 | rep1 | | 1,0 | 2,7 | | | | | | | |
| ɛ | jp2 | rep2 | | 0,7 | 3,0 | | | | | | | |
| ɛ | jp2 | rep2 | | 2,2 | 1,5 | | | | | | | |
| ɛ | jp3 | rep1 | | 0,9 | 2,3 | | | | | | | |
| ɛ | jp3 | rep1 | | 2,5 | 1,2 | | | | | | | |
| ɛ | jp3 | rep2 | | 1,5 | 2,0 | | | | | | | |
| ɛ | jp3 | rep2 | | 2,6 | 1,3 | | | | | | | |
| **Moyenne** | | | | 1,8 | 1,9 | | | | | | | 0,0 |

**4) /a/** Cette voyelle a été relativement bien identifiée avec les deux types de stimuli. La deuxième répétition de la lecture de l'apprenante 3 a été identifiée correctement comme /a/ par 14 auditeurs seulement, 11 autres ayant répondu /ø/ et /œ/ (5 et 6 respectivement : Tableau 38). Ce résultat pourrait être expliqué par le fait que ce stimulus avait un F1 relativement bas (747 Hz, contre 823 Hz pour l'autre répétition de la même apprenante, 885 Hz et 848 Hz pour la production des deux locutrices natives qui a servi comme modèle de la tâche de répétition), et un F2 relativement élevé (1478 Hz, contre 1429 Hz pour l'autre répétition, 1268 Hz et 1421 Hz pour les locutrices natives), ce qui aurait rapproché cette voyelle du /œ/.

**Tableau 38** : Perception des 6 stimuli du /a/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | jp1 | lec1 | | 1 | 1 | 19 | | | | | | | 5 | 26 |
| a | jp1 | lec2 | | | | 25 | | | | | | | 1 | 26 |
| a | jp2 | lec1 | | | | 25 | 1 | | | | | | | 26 |
| a | jp2 | lec2 | | | | 26 | | | | | | | | 26 |
| a | jp3 | lec1 | | | | 26 | | | | | | | | 26 |
| a | jp3 | lec2 | | | | 14 | | | | | 5 | 6 | 1 | 26 |
| Total_réponses | | | | 1 | 1 | 135 | 1 | | | | 5 | 6 | 7 | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | jp1 | lec1 | | 0,0 | 0,0 | 1,2 | | | | | | |
| a | jp1 | lec2 | | | | 3,0 | | | | | | |
| a | jp2 | lec1 | | | | 3,2 | 0,1 | | | | | |
| a | jp2 | lec2 | | | | 3,2 | | | | | | |
| a | jp3 | lec1 | | | | 4,7 | | | | | | |
| a | jp3 | lec2 | | | | 1,8 | | | | | 0,5 | 0,8 |
| Moyenne | | | | 0,0 | 0,0 | 2,9 | 0,0 | | | | 0,1 | 0,1 |

Les stimuli de la répétition ont été perçus presque parfaitement bien avec de très bonnes notes (variant entre 3,7 et 4,5), à l'exception des deux répétitions de l'apprenant 2 après le locuteur natif 1. Étant donné que l'apprenant 1 a prononcé le /a/ de façon tout à fait acceptable après le même modèle (identifié correctement par tous les auditeurs, avec des notes 3,8 et 4,3), nous ne pourrons attribuer le résultat du test aux caractéristiques acoustiques de la production du locuteur natif.

**Tableau 39** : Perception des 10 stimuli du /a/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | jp1 | rep1 | | | | 26 | | | | | | | | 26 |
| a | jp1 | rep1 | | | | 26 | | | | | | | | 26 |
| a | jp2 | rep1 | | | | 2 | 9 | 14 | | | | 1 | | 26 |
| a | jp2 | rep1 | | | | 15 | | | 1 | | 1 | 1 | 8 | 26 |
| a | jp2 | rep2 | | | | 26 | | | | | | | | 26 |
| a | jp2 | rep2 | | | | 26 | | | | | | | | 26 |
| a | jp3 | rep1 | 1 | | | 25 | | | | | | | | 26 |
| a | jp3 | rep1 | | | | 26 | | | | | | | | 26 |
| a | jp3 | rep2 | | | | 26 | | | | | | | | 26 |
| a | jp3 | rep2 | | | | 26 | | | | | | | | 26 |
| Total_réponses | | | 1 | | | 224 | 9 | 14 | 1 | 0 | 1 | 2 | 8 | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | jp1 | rep1 | | | | 3,8 | | | | | | |
| a | jp1 | rep1 | | | | 4,3 | | | | | | |
| a | jp2 | rep1 | | | | 0,2 | 1,2 | 1,8 | | | | 0,1 |
| a | jp2 | rep1 | | | | 0,9 | | | 0,0 | | 0,0 | 0,1 |
| a | jp2 | rep2 | | | | 4,2 | | | | | | |
| a | jp2 | rep2 | | | | 3,7 | | | | | | |
| a | jp3 | rep1 | 0,2 | | | 4,1 | | | | | | |
| a | jp3 | rep1 | | | | 4,0 | | | | | | |
| a | jp3 | rep2 | | | | 4,5 | | | | | | |
| a | jp3 | rep2 | | | | 4,3 | | | | | | |
| Moyenne | | | 0,0 | | | 3,4 | 0,1 | 0,2 | 0,0 | | 0,0 | 0,0 |

**5) /ɔ/** Dans l'ensemble, le /ɔ/ prononcé par les 3 apprenants dans la tâche de lecture a été identifié majoritairement comme /o/ (entre 13 et 24 auditeurs sur 26), à l'exception de la deuxième production de l'apprenant 1, qui a été perçue majoritairement comme /ø œ/ (17 auditeurs sur 26 : Tableau 40).

**Tableau 40** : Perception des 5 stimuli du /ɔ/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɔ | jp1 | lec1 | | | | | 2 | 24 | | | | | | 26 |
| ɔ | jp1 | lec2 | | | | | 5 | 18 | 3 | | | | | 26 |
| ɔ | jp2 | lec1 | | | | 1 | 3 | 22 | | | | | | 26 |
| ɔ | jp2 | lec2 | | 1 | | | | | 7 | 1 | 11 | 6 | | 26 |
| ɔ | jp3 | lec1 | | | | 6 | 6 | 13 | | | | | 1 | 26 |
| Total_réponses | | | | 1 | | 7 | 16 | 77 | 10 | 1 | 11 | 6 | 1 | 130 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɔ | jp1 | lec1 | | | | | 0,2 | 3,2 | | | | |
| ɔ | jp1 | lec2 | | | | | 0,6 | 2,8 | 0,3 | | | |
| ɔ | jp2 | lec1 | | | | 0,0 | 0,3 | 3,3 | | | | |
| ɔ | jp2 | lec2 | | 0,2 | | | | | 0,7 | 0,0 | 1,4 | 0,7 |
| ɔ | jp3 | lec1 | | | | 0,7 | 0,6 | 1,3 | | | | |
| Moyenne | | | | 0,0 | | 0,1 | 0,3 | 2,1 | 0,2 | 0,0 | 0,3 | 0,1 |

Quant à la répétition, nous avons déjà remarqué ci-dessus que les auditeurs ont entendu du /a/ plus souvent que pour la lecture. En effet, 4 stimuli sur 10 ont été identifiés majoritairement comme /a/ (entre 15 et 22 auditeurs sur 26 : Tableau 41). Quatre autres stimuli ont été perçus majoritairement comme /o/ (entre 16 et 25 auditeurs sur 26), 1 comme /ø œ/ (16 auditeurs sur 26). Un seul stimulus (la deuxième répétition de l'apprenante 3 après la locutrice native 3) a été identifié correctement comme /ɔ/ par presque la moitié des auditeurs (12 sur 26, contre 11 sur 26 pour /o/).

**Tableau 41** : Perception des 10 stimuli du /ɔ/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɔ | jp1 | rep1 | | 2 | | | | 1 | 6 | 1 | 13 | 3 | | 26 |
| ɔ | jp1 | rep1 | | | | | 6 | 20 | | | | | | 26 |
| ɔ | jp2 | rep1 | | | | 1 | 5 | 16 | 1 | | | 3 | | 26 |
| ɔ | jp2 | rep1 | | | | 15 | 7 | 1 | | | 1 | 2 | | 26 |
| ɔ | jp2 | rep2 | | | | 1 | 4 | 21 | | | | | | 26 |
| ɔ | jp2 | rep2 | | | | | 1 | 25 | | | | | | 26 |
| ɔ | jp3 | rep1 | | | | 19 | 5 | 2 | | | | | | 26 |
| ɔ | jp3 | rep1 | | | | 2 | 12 | 11 | | | | | 1 | 26 |
| ɔ | jp3 | rep2 | | | | 22 | 2 | 1 | | | | 1 | | 26 |
| ɔ | jp3 | rep2 | | | | 22 | 3 | 1 | | | | | | 26 |
| Total_réponses | | | | 2 | | 82 | 45 | 99 | 7 | 1 | 14 | 9 | 1 | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ɔ | jp1 | rep1 | | 0,3 | | | | 0,1 | 0,7 | 0,1 | 2,0 | 0,4 |
| ɔ | jp1 | rep1 | | | | | 0,7 | 2,3 | | | | |
| ɔ | jp2 | rep1 | | | | 0,1 | 0,5 | 1,8 | 0,1 | | | 0,3 |
| ɔ | jp2 | rep1 | | | | 1,2 | 0,8 | 0,0 | | | 0,2 | 0,1 |
| ɔ | jp2 | rep2 | | | | 0,1 | 0,6 | 2,8 | | | | |
| ɔ | jp2 | rep2 | | | | | 0,2 | 3,8 | | | | |
| ɔ | jp3 | rep1 | | | | 1,7 | 0,6 | 0,2 | | | | |
| ɔ | jp3 | rep1 | | | | 0,2 | 1,7 | 1,3 | | | | |
| ɔ | jp3 | rep2 | | | | 2,4 | 0,3 | 0,2 | | | | 0,1 |
| ɔ | jp3 | rep2 | | | | 2,5 | 0,5 | 0,2 | | | | |
| Moyenne | | | | 0,0 | | 0,8 | 0,6 | 1,3 | 0,1 | 0,0 | 0,2 | 0,1 |

**6) /o/** Le /o/ produit dans la tâche de lecture a été presque parfaitement identifié (par 24-25 auditeurs sur 26) avec des notes relativement élevées (entre 3,2 et 4,0 : Tableau 42).

**Tableau 42** : Perception des 6 stimuli du /o/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| o | jp1 | lec1 | | | | | 1 | 25 | | | | | | 26 |
| o | jp1 | lec2 | | | | | | 24 | 2 | | | | | 26 |
| o | jp2 | lec1 | | | | | 1 | 25 | | | | | | 26 |
| o | jp2 | lec2 | | | | | | 24 | 2 | | | | | 26 |
| o | jp3 | lec1 | | | | | 2 | 24 | | | | | | 26 |
| o | jp3 | lec2 | | | | | 1 | 25 | | | | | | 26 |
| Total_réponses | | | | | | | 5 | 147 | 4 | | | | | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| o | jp1 | lec1 | | | | | 0,1 | 4,0 | | | | |
| o | jp1 | lec2 | | | | | | 3,2 | 0,3 | | | |
| o | jp2 | lec1 | | | | | 0,2 | 3,5 | | | | |
| o | jp2 | lec2 | | | | | | 3,8 | 0,2 | | | |
| o | jp3 | lec1 | | | | | 0,3 | 3,5 | | | | |
| o | jp3 | lec2 | | | | | 0,2 | 4,0 | | | | |
| Moyenne | | | | | | | 0,1 | 3,7 | 0,1 | | | |

Les stimuli de la répétition ont été correctement perçus comme /o/ moins souvent, mais 7 stimuli sur 10 ont néanmoins été identifiés majoritairement comme /o/ (entre 21 et 25 auditeurs sur 26 : Tableau 43). En revanche, les 2 répétitions de l'apprenant 2 après le locuteur natif 1 ont été catégorisées majoritairement comme /ø œ/ (19 et 15 auditeurs sur 26). Comme nous l'avons remarqué ci-dessus, nous pourrons attribuer ce résultat au fait que les apprenants ont perçu /o ɔ/ des locuteurs natifs comme /ø œ/ dans 10% des cas (Expérience 3 : Tableau 23).

**Tableau 43** : Perception des 10 stimuli du /o/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **o** | jp1 | rep1 | | | | | 3 | 21 | 1 | | | | 1 | 26 |
| **o** | jp1 | rep1 | | | | | 4 | 21 | 1 | | | | | 26 |
| **o** | jp2 | rep1 | | | | | | | 5 | 2 | 15 | 4 | | 26 |
| **o** | jp2 | rep1 | | 1 | | | | | 5 | 5 | 14 | 1 | | 26 |
| **o** | jp2 | rep2 | | | | | 1 | 22 | 3 | | | | | 26 |
| **o** | jp2 | rep2 | | | 1 | | 7 | 11 | | | 3 | 4 | | 26 |
| **o** | jp3 | rep1 | | | | 1 | 3 | 22 | | | | | | 26 |
| **o** | jp3 | rep1 | | | | | 4 | 22 | | | | | | 26 |
| **o** | jp3 | rep2 | | | | | | 25 | | | 1 | | | 26 |
| **o** | jp3 | rep2 | | | | | 1 | 25 | | | | | | 26 |
| **Total_réponses** | | | | 1 | 1 | 1 | 23 | 169 | 15 | 7 | 33 | 9 | 1 | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **o** | jp1 | rep1 | | | | | 0,4 | 3,1 | 0,2 | | | |
| **o** | jp1 | rep1 | | | | | 0,5 | 3,1 | 0,2 | | | |
| **o** | jp2 | rep1 | | | | | | | 0,5 | 0,2 | 2,1 | 0,5 |
| **o** | jp2 | rep1 | | 0,1 | | | | | 0,5 | 0,5 | 2,0 | 0,2 |
| **o** | jp2 | rep2 | | | | | 0,2 | 3,3 | 0,4 | | | |
| **o** | jp2 | rep2 | | | 0,2 | | 0,7 | 1,0 | | | 0,2 | 0,4 |
| **o** | jp3 | rep1 | | | | 0,1 | 0,3 | 2,9 | | | | |
| **o** | jp3 | rep1 | | | | | 0,5 | 3,4 | | | | |
| **o** | jp3 | rep2 | | | | | | 3,8 | | | 0,1 | |
| **o** | jp3 | rep2 | | | | | 0,2 | 4,3 | | | | |
| **Moyenne** | | | | 0,0 | 0,0 | 0,0 | 0,3 | 2,5 | 0,2 | 0,1 | 0,4 | 0,1 |

**7) /u/** Parmi les 6 stimuli du /u/ produits dans la tâche de lecture, seuls deux ont été identifiés correctement comme /u/ par plus de la moitié des auditeurs (20 et 15 auditeurs sur 26 : Tableau 44). En revanche, les autres stimuli ont été catégorisés comme /ø œ/ par la moitié des auditeurs ou plus (entre 13 et 19 sur 26). Notons également que les notes sont basses, et c'est bien le cas (2,3) même pour le stimulus qui a été perçu comme /u/ par 20 auditeurs sur 26 (*cf.* les notes des deux stimuli du /o/ en répétition qui ont été identifiés correctement par 21 auditeurs sur 26 étaient de 3,1 pour toutes les deux : Tableau 43).

Concernant la répétition (Tableau 45), 8 stimuli sur 9 ont été identifiés majoritairement comme /ø œ/ (entre 15 et 20 auditeurs sur 26), l'autre majoritairement comme /y/ (16 auditeurs sur 26). Les basses notes pour le /u/ reflètent la faible fréquence d'identification comme /u/ : certains stimuli ont obtenu des notes moyennes (2,5 pour un, 2,6 pour un autre) pour le /ø/, mais non pour le /u/ (< 1,0).

**Tableau 44** : Perception des 6 stimuli du /u/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| u | jp1 | lec1 | | | | | | | 20 | 1 | 5 | | | 26 |
| u | jp1 | lec2 | 1 | | | | | | 15 | 6 | 4 | | | 26 |
| u | jp2 | lec1 | | | | | | | 4 | 3 | 14 | 5 | | 26 |
| u | jp2 | lec2 | | | | | | | 12 | 1 | 11 | 2 | | 26 |
| u | jp3 | lec1 | | | | | | | 7 | 6 | 12 | 1 | | 26 |
| u | jp3 | lec2 | | | | | | | 7 | 3 | 15 | 1 | | 26 |
| Total_réponses | | | 1 | | | | | | 65 | 20 | 61 | 9 | | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| u | jp1 | lec1 | | | | | | | 2,3 | 0,1 | 0,3 | |
| u | jp1 | lec2 | 0,1 | | | | | | 1,5 | 0,4 | 0,3 | |
| u | jp2 | lec1 | | | | | | | 0,4 | 0,3 | 1,8 | 0,5 |
| u | jp2 | lec2 | | | | | | | 1,3 | 0,1 | 1,0 | 0,2 |
| u | jp3 | lec1 | | | | | | | 0,7 | 0,5 | 1,7 | 0,2 |
| u | jp3 | lec2 | | | | | | | 0,8 | 0,2 | 2,3 | 0,1 |
| Moyenne | | | 0,0 | | | | | | 1,2 | 0,3 | 1,2 | 0,2 |

**Tableau 45** : Perception des 9 stimuli du /u/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| u | jp1 | rep1 | | | | | | | 7 | 4 | 12 | 3 | | 26 |
| u | jp1 | rep1 | | | | | | | 2 | 16 | 7 | 1 | | 26 |
| u | jp2 | rep1 | | 2 | | | | | 2 | 5 | 16 | 1 | | 26 |
| u | jp2 | rep1 | | | | | | | 3 | 4 | 15 | 4 | | 26 |
| u | jp2 | rep2 | | | | | | | 8 | 3 | 14 | 1 | | 26 |
| u | jp2 | rep2 | | | 1 | | | | 5 | 2 | 17 | 1 | | 26 |
| u | jp3 | rep1 | | | | | 1 | 1 | 5 | | 11 | 8 | | 26 |
| u | jp3 | rep2 | | | | | | | 4 | 2 | 18 | 2 | | 26 |
| u | jp3 | rep2 | | | | | | | 4 | 2 | 12 | 7 | 1 | 26 |
| Total_réponses | | | | 2 | 1 | | 1 | 1 | 40 | 38 | 122 | 28 | 1 | 234 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| u | jp1 | rep1 | | | | | | | 0,8 | 0,4 | 1,5 | 0,3 |
| u | jp1 | rep1 | | | | | | | 0,2 | 1,8 | 0,5 | 0,1 |
| u | jp2 | rep1 | | 0,2 | | | | | 0,2 | 0,5 | 2,0 | 0,1 |
| u | jp2 | rep1 | | | | | | | 0,2 | 0,5 | 2,2 | 0,5 |
| u | jp2 | rep2 | | | | | | | 0,9 | 0,3 | 1,8 | 0,2 |
| u | jp2 | rep2 | | | 0,1 | | | | 0,4 | 0,2 | 2,6 | 0,1 |
| u | jp3 | rep1 | | | | | 0,0 | 0,0 | 0,5 | | 0,8 | 0,6 |
| u | jp3 | rep2 | | | | | | | 0,4 | 0,2 | 2,5 | 0,2 |
| u | jp3 | rep2 | | | | | | | 0,5 | 0,2 | 1,8 | 0,6 |
| Moyenne | | | | 0,0 | 0,0 | | 0,0 | 0,0 | 0,4 | 0,4 | 1,7 | 0,3 |

**8) /y/** Deux stimuli de la lecture (la deuxième répétition des apprenants 1 et 2) ont été catégorisés correctement comme /y/ quasi unanimement (24 et 25 auditeurs sur 26) avec des notes très élevées (4,2 et 4,1 : Tableau 46). Les deux stimuli diphtongués (comme le /ju/ japonais) de l'apprenante 3 (cf. Figure 54 et Figure 55) ont été catégorisés comme /y/ par presque la moitié des

auditeurs (13 et 12 sur 26), mais les notes sont inférieures à 1,0 (même les notes d'évaluation moyennes brutes, sans considérer la fréquence, sont de 1,4 et 1,7).

**Tableau 46** : Perception des 6 stimuli du /y/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | jp1 | lec1 | 7 | 3 | | | | | 2 | 14 | | | | 26 |
| y | jp1 | lec2 | | | | | | | 2 | 24 | | | | 26 |
| y | jp2 | lec1 | | | | 1 | | | 6 | 19 | | | | 26 |
| y | jp2 | lec2 | | | | | | | 1 | 25 | | | | 26 |
| y | jp3 | lec1 | | 1 | | | | | 4 | 13 | 7 | | 1 | 26 |
| y | jp3 | lec2 | | 1 | | | | | 7 | 12 | 4 | 2 | | 26 |
| Total_réponses | | | 7 | 5 | | 1 | | | 22 | 107 | 11 | 2 | 1 | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | jp1 | lec1 | 0,5 | 0,3 | | | | | 0,3 | 1,6 | | |
| y | jp1 | lec2 | | | | | | | 0,3 | 4,2 | | |
| y | jp2 | lec1 | | | | 0,1 | | | 0,8 | 3,0 | | |
| y | jp2 | lec2 | | | | | | | 0,2 | 4,1 | | |
| y | jp3 | lec1 | | 0,0 | | | | | 0,3 | 0,7 | 0,3 | |
| y | jp3 | lec2 | | 0,1 | | | | | 0,5 | 0,8 | 0,3 | 0,1 |
| Moyenne | | | 0,1 | 0,1 | | 0,0 | | | 0,4 | 2,4 | 0,1 | 0,0 |

Quant aux stimuli de la répétition (Tableau 47), ceux des apprenants 1 et 2 ont été perçus presque parfaitement comme /y/ (24-25 auditeurs sur 26), mais les notes sont inférieures à celles de la lecture (entre 3,6 et 3,8). Deux stimuli (les deux répétitions de l'apprenante 3 après la locutrice native 3) ont été perçus majoritairement comme /ø œ/ (22 et 16 auditeurs sur 26). Ces deux productions n'étaient pas diphtonguées, mais les F2 et F3 n'étaient pas proches non plus (Figure 62). Les deux autres stimuli de l'apprenante 3, qui étaient diphtongués comme le /ju/ japonais, ont été perçus comme /y/ par la moitié des auditeurs ou plus (19 et 13 sur 26), mais les notes étaient basses (1,0 dans les deux cas : même les notes d'évaluation brutes sans considération à la fréquence étaient de 1,4 et 2).

**Tableau 47** : Perception des 8 stimuli du /y/ produits par les 3 apprenants japonophones dans la tâche de répétition : le taux d'identification correcte en pourcentage (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | jp1 | rep1 | | | | | | | 1 | 25 | | | | 26 |
| y | jp1 | rep1 | | | | | | | 2 | 24 | | | | 26 |
| y | jp2 | rep1 | | | | | | | 2 | 24 | | | | 26 |
| y | jp2 | rep2 | | | | | | | 1 | 25 | | | | 26 |
| y | jp3 | rep1 | | | | | | | 2 | 2 | 14 | 8 | | 26 |
| y | jp3 | rep1 | | | | | | | 2 | 8 | 10 | 6 | | 26 |
| y | jp3 | rep2 | | | | | | | 2 | 19 | 5 | | | 26 |
| y | jp3 | rep2 | | | | | | | | 13 | 9 | 3 | 1 | 26 |
| Total_réponses | | | | | | | | | 12 | 140 | 38 | 17 | 1 | 208 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | jp1 | rep1 | | | | | | | 0,2 | 3,8 | | |
| y | jp1 | rep1 | | | | | | | 0,3 | 3,6 | | |
| y | jp2 | rep1 | | | | | | | 0,3 | 3,7 | | |
| y | jp2 | rep2 | | | | | | | 0,2 | 3,6 | | |
| y | jp2 | rep2 | | | | | | | 0,2 | 0,1 | 1,6 | 0,8 |
| y | jp3 | rep1 | | | | | | | 0,2 | 0,7 | 1,0 | 0,7 |
| y | jp3 | rep2 | | | | | | | 0,2 | 1,0 | 0,3 | |
| y | jp3 | rep2 | | | | | | | | 1,0 | 1,0 | 0,2 |
| Moyenne | | | | | | | | | 0,2 | 2,2 | 0,5 | 0,2 |

**9) /ø/** Trois stimuli de la lecture du /ø/ ont été catégorisés majoritairement comme /y/ (entre 21 et 24 auditeurs sur 26) avec des notes relativement bonnes pour /y/ (entre 2,9 et 3,6). Rappelons que les apprenants 1 et 2 ont prononcé le /ø/ avec un rapprochement F2/F3 (Expérience 4 : Figure 49). Les 2 autres ont été perçus correctement comme /ø/ par un peu moins de la moitié des auditeurs (12 et 11 sur 26).

**Tableau 48** : Perception des 5 stimuli du /ø/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ø | jp1 | lec1 | | | 1 | | | | 3 | 21 | 1 | | | 26 |
| ø | jp1 | lec2 | | 1 | | | | | | 11 | 12 | 2 | | 26 |
| ø | jp2 | lec1 | | | | | | | 2 | 24 | | | | 26 |
| ø | jp2 | lec2 | | | | | | | 2 | 21 | 2 | 1 | | 26 |
| ø | jp3 | lec1 | | | | | | | 7 | 3 | 11 | 4 | 1 | 26 |
| Total_réponses | | | | 1 | 1 | | | 0 | 14 | 80 | 26 | 7 | 1 | 130 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ø | jp1 | lec1 | | | 0,1 | | | | 0,5 | 2,9 | 0,0 | |
| ø | jp1 | lec2 | | 0,1 | | | | | | 1,3 | 1,2 | 0,2 |
| ø | jp2 | lec1 | | | | | | | 0,3 | 3,6 | | |
| ø | jp2 | lec2 | | | | | | | 0,2 | 3,0 | 0,1 | 0,1 |
| ø | jp3 | lec1 | | | | | | | 0,7 | 0,3 | 1,3 | 0,6 |
| Moyenne | | | | 0,0 | 0,0 | | | 0,0 | 0,4 | 2,2 | 0,5 | 0,2 |

Quant aux stimuli de la répétition, 6 stimuli sur 10 ont été correctement identifiés comme /ø/ par la moitié des auditeurs ou plus (entre 13 et 22). Les 4 autres stimuli ont été catégorisés majoritairement comme /y/ (entre 15 et 18 auditeurs).

**Tableau 49** : Perception des 10 stimuli du /ø/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ø | jp1 | rep1 | | 1 | | | | | | 5 | 19 | 1 | | 26 |
| ø | jp1 | rep1 | | | | | | | 3 | 3 | 13 | 7 | | 26 |
| ø | jp2 | rep1 | | | | | | | | 15 | 10 | 1 | | 26 |
| ø | jp2 | rep1 | | | | | | | | 18 | 5 | 3 | | 26 |
| ø | jp2 | rep2 | | 1 | | | | | | 17 | 6 | 2 | | 26 |
| ø | jp2 | rep2 | | | | | | | 3 | 17 | 3 | 3 | | 26 |
| ø | jp3 | rep1 | | | | | | | 3 | 3 | 14 | 6 | | 26 |
| ø | jp3 | rep1 | | | | | | | 5 | 4 | 13 | 4 | | 26 |
| ø | jp3 | rep2 | | 1 | | | | 1 | 3 | | 17 | 4 | | 26 |
| ø | jp3 | rep2 | | 1 | | | | | 1 | | 22 | 1 | 1 | 26 |
| Total_réponses | | | | 4 | | | | 1 | 18 | 82 | 122 | 32 | 1 | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ø | jp1 | rep1 | | 0,1 | | | | | | 0,5 | 2,4 | 0,1 |
| ø | jp1 | rep1 | | | | | | | 0,2 | 0,3 | 1,7 | 0,9 |
| ø | jp2 | rep1 | | | | | | | | 1,9 | 1,1 | 0,1 |
| ø | jp2 | rep1 | | | | | | | | 2,5 | 0,4 | 0,3 |
| ø | jp2 | rep2 | | 0,1 | | | | | | 1,8 | 0,8 | 0,2 |
| ø | jp2 | rep2 | | | | | | | 0,3 | 2,1 | 0,3 | 0,3 |
| ø | jp3 | rep1 | | | | | | | 0,4 | 0,4 | 1,8 | 0,6 |
| ø | jp3 | rep1 | | | | | | | 0,4 | 0,4 | 1,4 | 0,5 |
| ø | jp3 | rep2 | | 0,1 | | | | 0,1 | 0,2 | | 2,1 | 0,3 |
| ø | jp3 | rep2 | | 0,1 | | | | | 0,1 | | 3,4 | 0,2 |
| Moyenne | | | | 0,0 | | | | 0,0 | 0,2 | 1,0 | 1,5 | 0,4 |

**10) /œ/** Les 4 stimuli du /œ/ prononcés par les apprenants 2 et 3 dans la tâche de lecture ont été perçus comme /ø/ ou /œ/ par une grande majorité des auditeurs (entre 21 et 26 sur 26), mais les réponses pour /ø/ sont plus nombreuses que celles pour /œ/ (12 pour /œ/ au maximum contre 14 pour /ø/). Quant aux deux stimuli de l'apprenant 1, l'un a été identifié comme /e/ ou /ɛ/ par une grande majorité (22 auditeurs sur 26), et l'autre par une minorité (10 sur 26).

**Tableau 50** : Perception des 6 stimuli du /œ/ produits par les 3 apprenants japonophones dans la tâche de lecture : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| œ | jp1 | lec1 | | 6 | 4 | | | | | | 12 | 4 | | 26 |
| œ | jp1 | lec2 | | 5 | 17 | | | | | | 2 | 2 | | 26 |
| œ | jp2 | lec1 | | 2 | | | | | | | 18 | 6 | | 26 |
| œ | jp2 | lec2 | | | | | | | | | 14 | 12 | | 26 |
| œ | jp3 | lec1 | | | | | 1 | | 3 | 1 | 18 | 3 | | 26 |
| œ | jp3 | lec2 | | | | | | | 2 | | 23 | 1 | | 26 |
| Total_réponses | | | | 13 | 21 | | 1 | | 5 | 1 | 87 | 28 | | 156 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| œ | jp1 | lec1 | | 0,7 | 0,5 | | | | | | 1,3 | 0,3 |
| œ | jp1 | lec2 | | 0,6 | 2,3 | | | | | | 0,2 | 0,2 |
| œ | jp2 | lec1 | | 0,2 | | | | | | | 2,7 | 0,8 |
| œ | jp2 | lec2 | | | | | | | | | 1,7 | 1,4 |
| œ | jp3 | lec1 | | | | | 0,1 | | 0,2 | 0,0 | 2,7 | 0,3 |
| œ | jp3 | lec2 | | | | | | | 0,2 | | 2,9 | 0,2 |
| Moyenne | | | | 0,3 | 0,5 | | 0,0 | | 0,1 | 0,0 | 1,9 | 0,6 |

Concernant les stimuli de la répétition, 8 stimuli sur 10 ont été identifiés presque unanimement comme /ø/ ou /œ/ (entre 24 et 26 auditeurs sur 26), mais les réponses pour le /ø/ sont plus nombreuses dans tous les cas. Un stimulus (la première répétition de l'apprenant 2 après le locuteur natif 1) a été catégorisé comme /y/ par 25 auditeurs sur 26.

**Tableau 51** : Perception des 10 stimuli du /œ/ produits par les 3 apprenants japonophones dans la tâche de répétition : le nombre d'identification correcte (en haut) et les notes « fréquence x évaluation » moyennes (en bas ; max. : 5 – min. : 0). 26 auditeurs francophones natifs x 1 réponse.

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ | autres/sans réponse | Total_stimuli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| œ | jp1 | rep1 | | | | | | | | | 18 | 8 | | 26 |
| œ | jp1 | rep1 | | 1 | | | | | 1 | | 14 | 10 | | 26 |
| œ | jp2 | rep1 | | | | | | | 1 | 25 | | | | 26 |
| œ | jp2 | rep1 | | | 1 | | | | | | 13 | 12 | | 26 |
| œ | jp2 | rep2 | | 5 | 1 | | | | 1 | | 12 | 7 | | 26 |
| œ | jp2 | rep2 | | | | | | | | | 21 | 4 | 1 | 26 |
| œ | jp3 | rep1 | | 1 | | | | 1 | | | 24 | | | 26 |
| œ | jp3 | rep1 | | | | | | | | | 21 | 5 | | 26 |
| œ | jp3 | rep2 | | | | | 1 | 1 | | | 17 | 7 | | 26 |
| œ | jp3 | rep2 | | 1 | | | | 1 | | | 17 | 7 | | 26 |
| Total_réponses | | | | 8 | 2 | | 1 | 3 | 3 | 25 | 157 | 60 | 1 | 260 |

| stimuli_V | loc | tâche | i | e | ɛ | a | ɔ | o | u | y | ø | œ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| œ | jp1 | rep1 | | | | | | | | | 2,5 | 1,0 |
| œ | jp1 | rep1 | | 0,1 | | | | | 0,0 | | 1,8 | 1,0 |
| œ | jp2 | rep1 | | | | | | | 0,2 | 4,0 | | |
| œ | jp2 | rep1 | | | 0,1 | | | | | | 1,6 | 1,5 |
| œ | jp2 | rep2 | | 0,5 | 0,1 | | | | 0,1 | | 1,4 | 0,9 |
| œ | jp2 | rep2 | | | | | | | | | 2,7 | 0,5 |
| œ | jp3 | rep1 | | 0,2 | | | | 0,1 | | | 3,3 | |
| œ | jp3 | rep1 | | | | | | | | | 2,7 | 0,7 |
| œ | jp3 | rep2 | | | | | 0,1 | 0,2 | | | 2,3 | 0,9 |
| œ | jp3 | rep2 | | 0,1 | | | | 0,1 | | | 2,3 | 1,0 |
| Moyenne | | | | 0,1 | 0,0 | | 0,0 | 0,0 | 0,0 | 0,4 | 2,1 | 0,7 |

Nous avons ainsi observé dans la présente série de tests de perception que les cas de non-correspondance entre ce que les apprenants ont voulu prononcer et ce que les auditeurs francophones natifs ont perçu étaient concentrés sur les deux groupes de voyelles suivants :

1) entre les voyelles moyennes : /e ɛ/ /o ɔ/ /ø œ/

2) entre les voyelles arrondies fermées et moyenne(s) : /u y ø (œ)/

## 2.5. CONCLUSION PARTIELLE

Dans la présente partie, trois séries d'expériences ont été présentées :

1) Expérience 3 : La perception des voyelles du français par des apprenants japonophones : identification des voyelles isolées (chapitre 2.2.)

2) Expérience 4 : La production des voyelles orales du français par des apprenants japonophones : lecture et répétition (chapitre 2.3.)

3) Expérience 5 : Les voyelles orales isolées produites par des apprenants japonophones et perçues par des auditeurs francophones natifs : test d'identification et d'évaluation (chapitre 2.4.)

Les résultats ont illustré les trois difficultés majeures suivantes rencontrées par des apprenants japonophones : 1) les voyelles fermées arrondies /u/, /y/ et antérieure mi-fermée arrondie /ø/ (et /œ/) ; 2) les voyelles mi-fermées et mi-ouvertes. 3) les voyelles nasales.

Parmi ces trois groupes de difficulté, nous allons examiner la première dans la troisième partie. Ce choix serait justifié par les raisons suivantes :

1) Premièrement, les deux types de difficulté, à savoir, celles portant sur les voyelles nasales et les voyelles moyennes, sont souvent observées auprès des apprenants d'un nombre important de langues qui ont ou n'ont pas de voyelles nasales, et de celles qui n'ont pas quatre degrés d'aperture (Lauret 2007[396] pour un tableau récapitulatif des difficultés classées par les L1 des apprenants). En revanche, malgré la présence de nombreuses langues qui n'ont pas la série antérieure arrondie, les apprenants confondent en général entre /i/-/y/, ou /y/-/u/ (/e/-/ø/ ou /ø/-/o/ pour les mi-fermées, /ɛ/-/œ/ ou /œ/-/ɔ/ pour les mi-ouvertes : Lauret 2007), et il est plus rare de voir des cas où les voyelles /u y ø/ sont confondues entre elles.

2) Sur le plan de l'importance d'opposition phonémique, tous les francophones natifs ne font pas la distinction entre toutes les trois paires de voyelles moyennes, ou ne produisent pas les trois voyelles nasales comme en français parisien (ex. non pas « nasalisées », comme la réalisation phonétique des voyelles nasales du Midi), alors que la quasi-totalité des locuteurs francophones natifs distinguent les trois voyelles /u y ø/. LeBel (1990)[397] compte les trois (/u y ø/) parmi les sept voyelles (/i E a O u y Œ/, /ø/ faisant partie de l'archiphonème /Œ/ (voir la section 1.5.1. à la page 79 pour la question d'achiphonème). Notons que /œ/ n'apparaît jamais en contexte isolé ou dans les syllabes ouvertes en fin de groupe) qui doivent être distinguées même pour assurer une communication minimale en français.

C'est pour ces raisons que nous examinerons les trois voyelles /u y ø/ en détail. Nous avons vu qu'aucun des 3 apprenants n'a prononcé le /u/ avec un rapprochement de F1/F2 en dessous de 1000 Hz comme les locuteurs natifs l'ont fait, tandis que le /y/ a été prononcé avec un

---

[396] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette. 2007.

[397] J.-G. LeBel. *Traité de correction phonétique ponctuelle*. Québec, Centre international de Recherche en aménagement linguistique. 1990.

rapprochement de F2/F3 par deux apprenants. Il nous reste à vérifier la difficulté relative de ces voyelles sur un plus grand nombre d'apprenants du même niveau, et de différents niveaux afin d'observer une tendance générale.

# Troisième partie : la perception et la production des voyelles /u/, /y/ et /ø/ en contexte isolé

## RÉSUMÉ

L'objectif de cette partie est d'étudier en détail la perception (discrimination) et la production des voyelles /u/, /y/, et /ø/ en contexte isolé.

Nous avons choisi les voyelles /u/, /y/ et /ø/, car ces trois se trouvent parmi les sept voyelles (/i E a O u y Œ/, /ø/ faisant partie de l'archiphonème /Œ/. Notons que le /œ/ n'apparaît jamais en contexte isolé ou dans les syllabes ouvertes en fin de groupe) qui doivent être distinguées même pour assurer une communication minimale en français selon Lebel (1990). C'est pour cette raison qu'il nous paraît plus intéressant de les étudier dans le détail.

1) Afin d'étudier la perception de ces trois voyelles françaises par des apprenants japonais, une expérience AXB de discrimination a été effectuée auprès de 14 auditeurs japonophones apprenant le français. Les paires choisies pour cette étude incluent /u/-/y/, /y/-/ø/, /u/-/ø/, ainsi que /i/-/e/, /u/-/o/, et /ɛ/-/a/. Le résultat révèle que le taux de bonnes réponses pour la paire /u/-/ø/ est significativement inférieur (test-t : 85 %) à celui de /u/-/y/ (95 %), et à /ɛ/-/a/ (99 %, non significativement inférieur à 100 %). Nous pouvons ainsi constater que la paire /u/-/ø/ est effectivement plus difficile à distinguer que la paire /u/-/y/.

2) L'expérience de la production présentée dans la première partie a été complétée par une autre étude qui portait sur des données de 47 apprenants de trois niveaux différents (débutant dans la première année d'apprentissage, et intermédiaire, intermédiaire-avancé), dans une tâche de lecture dans une phrase-cadre (ex. « Je dis /u/ comme dans loup », par exemple), afin de vérifier les tendances observées dans l'expérience précédente. Le résultat montre que seuls 4 apprenants ont produit /u/ avec un F2 inférieur à 1000 Hz, tandis que 20 d'entre eux ont prononcé /y/ au moins une fois avec F2 et F3 proches, et 31 d'entre eux ont produit /ø/ au moins une fois avec les trois premiers formants à peu près équidistants, avec le F2 autour de 1500 Hz. Ce résultat corrobore celui de l'étude de la deuxième partie dans le sens où les apprenants (de tous niveaux) étaient plus nombreux à prononcer /y/ avec F2 et F3 proches, que /u/ avec F1 et F2 proches et bas.

3) Les voyelles /u/, /y/ et /ø/ produites pour les expériences précédentes ont été soumises à un test de perception par des locuteurs francophones. Seize auditeurs francophones natifs ont identifié et évalué 18 productions différentes (5 apprenants) des voyelles /u/, /y/ et /ø/ (4 répétitions de stimuli). La tâche des auditeurs était d'identifier la voyelle et d'indiquer si le stimulus correspondait à un plus ou moins bon exemplaire (1 : mauvais – 5 : bon) de la voyelle choisie. Les productions de /u/ avec un F2 entre 700 Hz et 800 Hz ont obtenu un score élevé pour /u/ (plus de 4 sur 5 en moyenne), tandis que celles de la même voyelle /u/ avec un F2 entre 1000 Hz et 1100 Hz (hommes) ont été perçues aussi souvent comme /u/ que comme /ø/, mais le score était très faible (1 sur 5 environ en moyenne). Les autres stimuli de /u/, avec un F2 plus élevé (entre 1100 Hz et 1600 Hz) ont obtenu des scores qui varient entre 2 et 4,5 en moyenne pour /ø/, tandis que le score moyen pour /u/ était inférieur à 1 pour tous. Ces résultats suggèrent que /u/ produit par une grande majorité des apprenants dans l'expérience précédente, avec un F2 supérieur à 1000 Hz (16 stimuli sur 18), ne serait effectivement pas perçu comme un bon exemplaire de cette voyelle (mais comme /ø/), et confirment la difficulté de production de cette voyelle par des apprenants japonophones. /y/ et /ø/ prononcés par les apprenants ont été identifiés plus correctement (12 stimuli de /y/, 14 stimuli de /ø/ sur 18) par les auditeurs francophones natifs.

4) Afin de mieux étudier la relation entre production, acoustique, et perception des voyelles /u/, /y/, /ø/, deux séries de stimuli synthétisés ont été créées, et utilisées dans une expérience d'identification et d'évaluation.

i) Synthèse articulatoire (VTCalc) avec trois séries de continuum (A. entre /u/ et /ø/ français, et son extension ; B. /u/ français avec la langue mise progressivement vers l'avant ; C. /u/ français avec les lèvres progressivement moins arrondies et moins protruses).

ii) Synthèse par formants (synthétiseur de Klatt) avec deux séries de continuum (A. F1 à 300 Hz avec un F2 qui varie entre 600 Hz et 2200 Hz. B. F1 qui varie entre 300 Hz et 460 Hz, et F2 qui varie entre 600 Hz et 2200 Hz. F3 et F4 ont été fixés à 2350 Hz et 3600 Hz pour les deux séries).

Seize auditeurs francophones natifs ont identifié et évalué ces deux types de stimuli. Les résultats ont montré que : i) la labialisation ET la position de la langue (postériorité) joue un rôle important dans la perception du /u/ français ; ii) des voyelles synthétisées en tant que voyelles fermées postérieures non-arrondies ont été perçues majoritairement comme /œ/ ; iii) les stimuli qui ont un F2 supérieur à 1000 Hz ne sont pas perçus comme bons exemplaires du /u/ français.

En revanche, les 16 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/. Cela suggère que les auditeurs japonophones perçoivent le /u/ [ɯ] japonais dans une zone bien plus large sur l'axe de F2.

Les résultats de ces quatre expériences suggèrent les points suivants :

a) Le F2 élevé du /u/ observé chez les apprenants est responsable du taux élevé de perception de /ø/ (réponse modale pour tous les 16 stimuli avec un F2 supérieur à 1000 Hz) auprès des locuteurs natifs, tandis que les occurrences de /y/ avec les F2/F3 proches ont été plus fréquemment perçues comme /y/ (réponse modale pour 12 stimuli sur 18).

b) Les japonophones perçoivent /u/ dans les zones articulatoires et perceptives dans lesquelles les francophones entendent /u/ et /ø/, ce qui explique le fait que les apprenants japonophones ont des difficultés à distinguer les deux phonèmes.

c) La voyelle phonémiquement « similaire » /u/ du français, qui occupe une place similaire dans l'inventaire phonémique de la langue source (japonais), mais diffère phonétiquement de l'équivalente phonémique dans la langue source (/u/ [ɯ] japonais) serait plus difficile que la voyelle « nouvelle » /y/, qui n'a pas d'équivalent ni phonémique ni phonétique dans la langue source. Ceci corrobore les résultats de Flege (1987). En revanche, la production de la voyelle /ø/, qui est également « nouvelle » sur le plan phonémique, ne semble pas poser beaucoup de difficulté. Ceci pourrait être expliqué par la proximité phonétique entre le /ø/ français et le /u/ japonais. Nous pourrions considérer /ø/ comme une voyelle phonémiquement nouvelle mais phonétiquement similaire.

Les sources de la difficulté du /u/ seraient multiples : 1) la difficulté articulatoire et motrice (le recul de la langue serait plus difficile à maîtriser que l'arrondissement et la protrusion des lèvres pour ceux qui n'en ont pas l'habitude) ; 2) la difficulté de rendre explicite la position de la langue ; 3) la difficulté perceptive des phones qui seraient catégorisés dans la même catégorie phonémique dans le système vocalique de la langue source (Kuhl *et al.* 1992 : « l'Effet Magnétique », Best 1995 : « Modèle d'Assimilation Perceptive ») ; 4) manque de prise de conscience de la difficulté de /u/.

**L'objet d'étude de la présente partie** Nous avons choisi les voyelles /u/, /y/ et /ø/ parmi les trois groupes de difficultés relevés dans la partie précédente. Ces trois voyelles se trouvent parmi les sept (/i E a O u y Œ/, /ø/ faisant partie de l'archi-phonème /Œ/ (qui correspond aux voyelles /ø œ/). Notons que le /œ/ n'apparaît jamais en contexte isolé ou dans les syllabes ouvertes en fin de groupe) qui doivent être distinguées même pour assurer une communication minimale en

français selon LeBel (1990)[398]. C'est pour cette raison qu'il nous paraît important de les étudier en détail.

**Les expériences présentées** Quatre séries d'expériences seront présentées dans cette partie :

1) Expérience 6 : Perception de voyelles par des apprenants japonophones : discrimination AXB sur des voyelles isolées (/u/-/y/, /y/-/ø/, /u/-/ø/, ainsi que /i/-/e/, /u/-/o/, et /ɛ/-/a/).

2) Expérience 7 : Production des voyelles /u y ø/ par des apprenants japonophones de trois niveaux d'apprentissage.

3) Expérience 8 : Perception par des auditeurs francophones natifs des voyelles /u y ø/ produites par des apprenants japonophones.

4) Expérience 9 : Perception de vocoïdes synthétisés (par synthèse articulatoire et synthèse à formants) auprès d'auditeurs natifs francophones et japonophones.

Elles seront suivies d'une discussion.

## 3.1. ANALYSE PERCEPTIVE : PERCEPTION DES VOYELLES /u y ø/ PAR DES APPRENANTS JAPONOPHONES : DISCRIMINATION AXB SUR DES VOYELLES ISOLÉES (EXPÉRIENCE 6)

D'après les résultats de l'Expérience 3 (test d'identification des voyelles orales du français), les voyelles difficiles à identifier pour les apprenants japonophones sont les voyelles arrondies fermées et mi-fermée /u y ø/, entre autres. À part les voyelles nasales et les voyelles moyennes, nous pouvons relever les paires suivantes : /u/-/ø/, /y/-/ø/, /u/-/y/, /o/-/u/, /i/-/e/ (Tableau 22, p. 130). Afin d'examiner la confusion de ces paires plus en détail, nous avons effectué une expérience de discrimination du type AXB.

### *3.1.1. Les stimuli*

**Les paires de voyelles choisies** Nous avons constitué les stimuli en nous fondant sur les paires de voyelles mentionnées ci-dessus (/u/-/y/, /y/-/ø/, /u/-/ø/, /o/-/u/, /i/-/e/) ainsi que la paire /a/-/ɛ/, qui n'a jamais été confondue par les 5 apprenants qui ont participé à l'Expérience 3 (Tableau 22).

**Constitution des triplets** Nous avons choisi les 3 répétitions des voyelles /u y ø i e ɛ a o/ prononcées dans la phrase-cadre « *Je dis* /V/ *comme dans* … » par les locuteurs natifs FR2 (homme), FR3 (femme) et FR4 (femme) pour l'Expérience 2 (Figure 41, p. 123). Nous les avons combinées par la suite afin de constituer des triplets dans 3 conditions différentes de locuteurs : pour les sons A et

[398] J.-G. LeBel. *Traité de correction phonétique ponctuelle*. Québec, Centre international de Recherche en aménagement linguistique, 1990.

B (1er et 3ème dans chaque triplet), nous avons pris les productions de la locutrice FR4, et pour le son X (2ème dans chaque triplet), nous avons les 3 conditions suivantes (Tableau 52, au milieu) :

Condition 1 (sp1) : la même locutrice française (locutrice FR4).

Condition 2 (sp2) : une autre locutrice française du même sexe (locutrice FR3).

Condition 3 (sp3) : un autre locuteur français du sexe opposé (locuteur FR2).

Pour chacune des 6 paires de voyelles et des conditions de locuteurs, nous avons préparé 4 séries de triplets (Tableau 52 : à droite). Si l'on prend la paire /u/-/y/ et la condition 3, par exemple, le premier triplet était composé de : i) la première répétition du /u/ de la locutrice FR4 ; ii) la deuxième répétition du /u/ de FR2 ; iii) la troisième répétition du /y/ de FR4 (la réponse correcte est donc 1) ; le deuxième triplet de : i) la troisième répétition du /u/ de FR4 ; ii) la première répétition du /y/ de FR2 ; iii) la deuxième répétition du /y/ de FR4 (réponse : 3), et ainsi de suite.

Les facteurs décrits ci-dessus sont illustrés dans le Tableau 52. Nous avons ainsi obtenu 72 triplets (6 paires de voyelles x 3 conditions de locuteurs x 4 triplets).

**Tableau 52 :** Les paires de voyelles (à gauche), les conditions de locuteurs (au milieu), et les 4 triplets pour une paire de voyelles donnée (à droite : les indices 1-3, comme dans a1, b3, *etc*., représentent le token), utilisés dans le test de discrimination AXB.

| /u/-/y/ |
|---|
| /y/-/ø/ |
| /u/-/ø/ |
| /o/-/u/ |
| /i/-/e/ |
| /a/-/ɛ/ |

| | A | X | B |
|---|---|---|---|
| Condition 1 | loc. 4 (F) | loc. 4 (F) | loc. 4 (F) |
| Condition 2 | loc. 4 (F) | loc. 3 (F) | loc. 4 (F) |
| Condition 3 | loc. 4 (F) | loc. 2 (H) | loc. 4 (F) |

| | A | X | B |
|---|---|---|---|
| 1 | a1 | a2 | b3 |
| 2 | a3 | b1 | b2 |
| 3 | b1 | b3 | a2 |
| 4 | b2 | a1 | a3 |

**La liste des triplets** Les triplets ont été disposés dans un ordre semi-aléatoire de sorte que la même paire de voyelles ne se succède pas, et que la même réponse attendue (1er ou 3ème stimulus) ne se suive pas quatre fois ou plus.

**L'intervalle entre les stimuli** Les trois stimuli dans le même triplet ont été placés avec un intervalle d'une seconde environ entre eux.

### *3.1.2. Les procédures*

L'expérience a été effectuée auprès d'apprenants japonophones, de locuteurs natifs du français et de locuteurs natifs du japonais n'apprenant pas le français à titre de comparaison.

**Les tâches** Les tâches des auditeurs consistaient à écouter les triplets de voyelles (isolées), et de juger si le deuxième son était identique au premier ou au troisième. Dans le test effectué auprès des apprenants, les auditeurs ont écouté le son diffusé par des haut-parleurs. Ils ont encadré leur choix sur une feuille imprimée (Figure 64). Les auditeurs francophones natifs et

japonophones non-apprenants ont écouté les stimuli à travers un casque et disposaient d'une souris et d'un écran pour répondre sur un ordinateur.

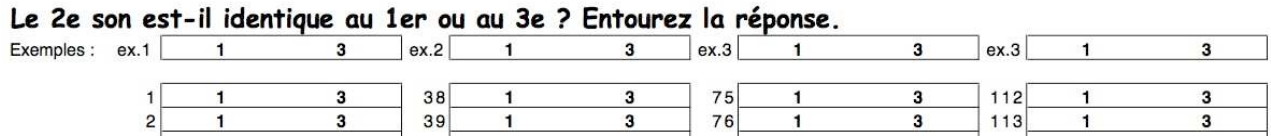



**Figure 64 :** Extrait de la feuille de réponse distribuée aux sujets-auditeurs.

**L'entraînement** L'expérience comportait 2 phases : l'entraînement et le test. Pendant l'entraînement, les sujets-auditeurs ont écouté les quatre combinaisons de triplets possibles de la paire /a/-/i/ pour qu'ils puissent s'habituer aux procédures d'une expérience AXB.

**Le test** Après avoir écouté tous les 72 triplets du test, ils ont eu une pause, et ensuite, nous leur avons présenté les mêmes triplets dans un autre ordre. Le temps de passation était fixé à 19 minutes (pour les 2 répétitions de ces 72 triplets) pour les apprenants ; les auditeurs disposaient de 5 secondes par stimulus pour donner leur réponse. Le nombre des répétitions a été restreint à 2, après avoir considéré la charge cognitive des apprenants de niveaux relativement élémentaires (et des non-apprenants). Le temps de passation pour les deux autres groupes (auditeurs natifs français et japonais) variait selon le rythme de chaque auditeur, mais ne dépassait pas une demi-heure.

### *3.1.3. La perception par des apprenants japonophones*

Nous présenterons dans cette section les résultats des apprenants japonophones.

#### 3.1.3.1. Les auditeurs

Les participants étaient tous des élèves du lycée Musashi (Tokyo, Japon). Ils constituaient les deux groupes suivants :

1) 7 élèves en première année ("niveau I" : 15 mois d'apprentissage à raison de 100 minutes de leçons par semaine)[399] ;

2) 7 élèves en années supérieures ("niveau II" : 5 en deuxième année, et 2 en troisième année).

#### 3.1.3.2. L'analyse des résultats

**Le taux de réponses correctes** Les résultats ont été calculés de façon suivante : nous avons compté le nombre de réponses correctes pour chacune des 6 paires de voyelles, chacune des 3 conditions de locuteurs, et chacun des 14 auditeurs (8 réponses au total = 4 triplets x 2 répétitions).

[399] Dans cet établissement collège (*chuugakkou*) – lycée (*koutou-gakkou*), les élèves commencent l'apprentissage du français en 3ème et dernière année de collège (14-15 ans) avant la première année de lycée. Ils apprennent le français en tant que deuxième langue étrangère facultative après l'anglais (première langue étrangère obligatoire), qu'ils commencent dès la première année de collège. Pendant la troisième année de collège seulement, une deuxième langue étrangère au choix (français, allemand, chinois, coréen) est obligatoire.

Le taux de réponses correctes sera donc représenté en terme du nombre des réponses correctes sur 8 pour une paire, une condition de locuteurs, et un auditeur donné.

**La moyenne par groupe** Nous avons calculé par la suite la moyenne des taux de réponses correctes pour chacun des deux groupes d'apprenants (niveaux I et II), et effectué des comparaisons selon les facteurs suivants : les groupes (niveaux I et II), les conditions de locuteurs (sp1, sp2, sp3), et les paires de voyelles (/u/-/y/, /y/-/ø/, /u/-/ø/, /o/-/u/, /i/-/e/, /a/-/ɛ/).

### 3.1.3.3. Les résultats

**Les conditions de locuteurs et les niveaux d'apprentissage** La Figure 65 montre les résultats en fonction des conditions de locuteurs et de l'expérience d'apprentissage. Nous y observons que la difficulté s'accroît au fur et à mesure que l'hétérogénéité des stimuli augmente (sp3 > sp2 > sp1). Pour la condition sp1 (même locutrice), le taux global de réponses correctes (97,6 % pour le niveau I, et 98,8 % pour le niveau II) s'approche à 100 %, mais reste à peine significativement inférieur à 100 % ($t_{41}$ = 2,08 p = 0,044 pour les deux groupes). Nous y trouvons un effet significatif de la condition de locuteurs (sp1, sp2, sp3 : $F_{(2, 249)}$ = 30,36 p < 0,0001), la différence entre chaque catégorie contribuant à cet effet (sp1 > sp2 > sp3).

**L'hétérogénéité des stimuli et la difficulté** Cet effet de condition de locuteurs serait dû à la difficulté de normaliser perceptivement les phones vocaliques en tant que phonèmes de la L2 en considérant les variations individuelles. Nous pouvons également noter une différence de taux de bonnes réponses selon le niveau d'apprentissage. Cette tendance s'intensifie au fur et à mesure que la tâche devient difficile (sp1 < sp2 < sp3), mais la différence n'est pas significative pour aucune des 3 conditions de locuteurs. Ce résultat suggère que les apprenants plus avancés seraient capables de mieux extraire les informations acoustiques pertinentes à la distinction phonémique de la L2.

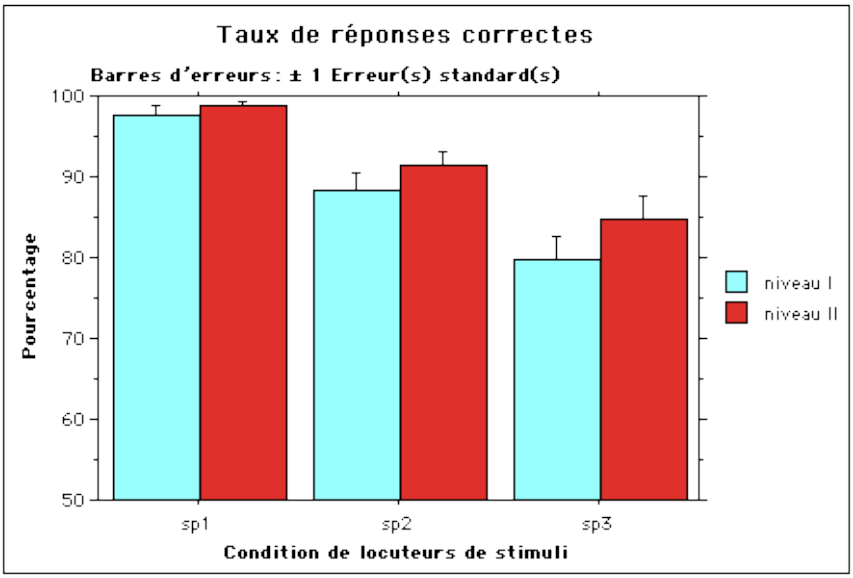



**Figure 65 :** Taux de bonnes réponses (niveau du hasard : 50%. 7 auditeurs pour chacun des groupes de sujets, 2 réponses, 24 triplets pour chaque condition de locuteurs, sp1-3, toutes les paires de voyelles confondues : 336 réponses pour chaque groupe et chaque condition de locuteur) trié par les conditions de locuteurs (sp1 : la même locutrice, sp2 : une autre locutrice du même

sexe, sp3 : un autre locuteur du sexe opposé), et par les groupes de sujets ("niveau I" : les élèves en première année, "niveau II" : ceux en années supérieures). Les barres d'erreur représentent l'erreur-type.

**Les paires de voyelles et les niveaux d'apprentissage** Maintenant, comparons les différentes paires de voyelles que nous avons examinées dans l'expérience (Figure 66).

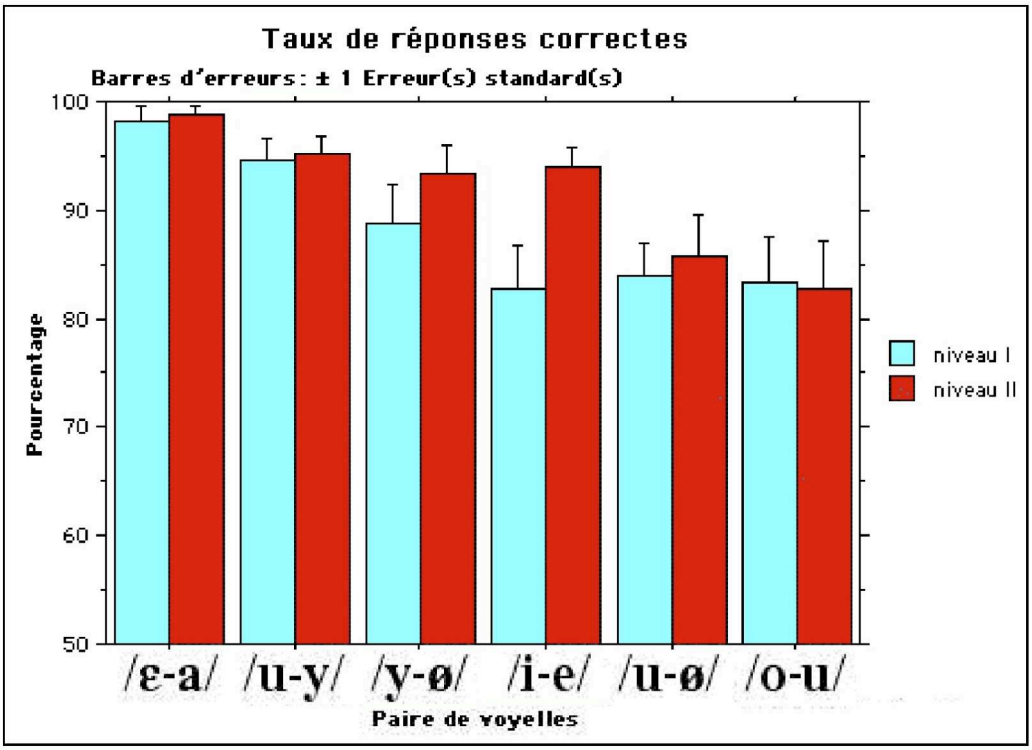



**Figure 66 :** Taux de bonnes réponses (niveau du hasard : 50%. 7 auditeurs pour chacun des groupes de sujets, 2 réponses, 12 triplets pour chaque paire, toutes les 3 conditions de locuteurs confondues : 168 réponses pour chaque groupe et chaque paire de voyelles) trié par les paires de voyelles, et par les groupes de sujets ("niveau I" : les élèves en première année, "niveau II" : ceux qui sont en années supérieures). Les barres d'erreur représentent l'erreur-type.

**La mieux distinguée : /ɛ/-/a/** Concernant la paire /ɛ/-/a/, pour laquelle nous n'avons pas prévu de difficulté d'après le résultat du test d'identification (Expérience 3 : Tableau 22, p. 130), le taux de bonnes réponses (98,3 % pour le niveau I, 98,8 % pour le niveau II) s'approche à 100 % pour les deux groupes, sans être significativement inférieur (test t univarié : $t_{20} = 1{,}37$ $p = 0{,}186$ pour le niveau I, $t_{20} = 1{,}45$ $p = 0{,}162$ pour le niveau II). Cela montre que les consignes de l'expérience ont bien été comprises par les auditeurs, et que cette paire ne pose pas de difficulté à ces apprenants, car ce taux élevé contraste avec toutes les autres paires, qui ont obtenu un score significativement inférieur (/ɛ/-/a/ *vs.* (/i/-/e/ U /o/-/u/ U /u/-/y/ U /u/-/ø/ U /y/-/ø/)[400] : $t_{13} = 7{,}65$ $p < 0{,}0001$).

[400] Comparaison entre la paire /ɛ/-/a/, d'un côté, et l'ensemble des paires /i/-/e/, /o/-/u/, /u/-/y/, /u/-/ø/, /u/-/ø/ combinées, de l'autre.

**Les pires : /o/-/u/ et /u/-/ø/** À l'extrémité opposée se trouvent les paires /o/-/u/ et /u/-/ø/, qui présentent des scores autour de 85 % pour les deux groupes d'apprenants, qui sont significativement inférieurs à ceux de tous les autres contrastes ensemble ( (/o/-/u/ U /u/-/ø/) *vs.* (/i/-/e/ U /ɛ/-/a/ U /u/-/y/ U /u/-/ø/ U /y/-/ø/) : $t_{13} = 6,45$ $p < 0,0001$), mais aucune différence significative n'a été observée entre ces deux paires (/o/-/u/ vs. /u/-/ø/ : $t_{13} = 0,56$ $p = 0,589$).

**/u/-/y/** En ce qui concerne la paire /u/-/y/, le taux de bonnes réponses est relativement élevé, supérieur à 90 % (94,6% pour le niveau I, 95,2% pour le niveau II, mais significativement inférieur à la paire /ɛ/-/a/. /u/-/y/ *vs.* /ɛ/-/a/ : $t_{13} = 2,28$ $p = 0,042$), et significativement supérieur à /o/-/u/ et /u/-/ø/ mis ensemble (/u/-/y/ *vs.* (/o/-/u/ U /u/-/ø/) : $t_{13} = 6,04$ $p < 0,0001$).

**/y/-/ø/ et /i/-/e/** En revanche, les deux autres paires /y/-/ø/ (91,1%) et /i/-/e/ (88,4%) présentent des scores meilleurs que ceux des paires /o/-/u/ et /u/-/ø/, mais cela n'est qu'une tendance, sauf la différence entre /y/-/ø/, d'un côté, et /o/-/u/ (83,0%) et /u/-/ø/ (84,8%) mis ensemble, de l'autre (/y/-/ø/ vs. (/o/-/u/ U /u/-/ø/) : $t_{13} = 3,07$ $p = 0,01$), ainsi que celle qui se trouve entre /y/-/ø/ et /o/-/u/ ($t_{13} = 3,99$ $p = 0,002$).

**/y/-/ø/ et /i/-/e/ : les niveaux d'apprentissage** Pour ces deux contrastes (/y/-/ø/ et /i/-/e/), nous observons également une différence relativement importante de taux de bonnes réponses entre les deux groupes d'apprenants : les apprenants plus avancés ont répondu presque aussi bien que la paire /u/-/y/ (la différence n'étant pas significative), tandis que les moins avancés n'ont pas distingué beaucoup mieux que les paires /o/-/u/ et /u/-/ø/ (la différence n'est pas significative, non plus). Cependant, cette tendance entre les deux groupes d'apprenants n'est significative que pour la paire /i/-/e/ ($t_{13} = 2,63$ $p = 0,02$).

**Comparaison avec le test d'identification : /u y ø/** Une comparaison entre ces résultats et ceux du test d'identification (Expérience 3 : Tableau 22, p. 130) nous permet de retrouver la hiérarchie de difficulté des paires /u/-/y/, /y/-/ø/, et /u/-/ø/. Dans le test d'identification, le nombre des cas de confusion était le plus petit pour la paire /u/-/y/ (2 sur 40 pour /u/->/y/, aucun cas pour /y/->/u/), suivie de /y/-/ø/ (6 sur 40 pour /y/->/ø/, 5 sur 40 pour /ø/->/y/), et de /u/-/ø/ (12 cas sur 40 dans les deux sens). Le taux de réponses correctes décroît dans ce même ordre : /u/-/y/ > /y/-/ø/ > /u/-/ø/.

**/u/-/o/ : identification et AXB** En revanche, nous observons une différence entre les résultats des deux tests concernant la paire /u/-/o/. Dans le test d'identification, /o/ a été identifié comme /u/ dans 8 cas sur 40, sans aucune confusion dans l'autre sens. Le taux de bonnes réponses de cette paire est aussi faible que celui de /u/-/ø/ dans le test AXB, ce qui suggère que la paire /u/-/o/ était plus difficile dans le test AXB. Nous pouvons nous poser la question de savoir si les apprenants, qui ne sont d'ailleurs pas les mêmes dans les deux tests, ont adopté des stratégies différentes, au moins concernant cette paire. D'un côté, une écoute plutôt « phonétique » ou physique dans le test AXB, malgré la présence du silence d'une seconde entre les stimuli, aurait permis aux apprenants de remarquer la similarité acoustique entre les deux voyelles postérieures (mi-)fermées (F1 et F2 proches et bas). De l'autre côté, une écoute plutôt « phonémique » pendant le test d'identification leur aurait fait se fonder sur le /u/ phonémique de leur interlangue (qui se

situerait probablement entre les /u/ et /ø/ français), qui est assez différent du /o/ (notons que les apprenants ont produit de bons exemplaires du /o/ dans les tâches de lecture et de répétition : Tableau 42, p. 167 et Tableau 43, p. 168).

**La difficulté de /i/-/e/ et /u/-/o/** La confusion entre /i/-/e/ et /u/-/o/ pourrait être expliquée par le fait qu'il y a un chevauchement de zones de dispersion acoustique (du moins sur le plan F1 et F2) de chacune de ces paires de voyelles, sauf pour /u/-/o/ prononcés par les femmes (CALLIOPE 1989 : 85. Figure reproduite à la Figure 67[401]). Nous retrouvons également dans nos données un chevauchement de /i/-/e/ et /u/-/o/ sur la plan F1-F2 chez les deux hommes FR1 et FR2 (Figure 42, p. 123 ; Figure 43, p. 124 pour les femmes FR3 et FR4).

Concernant l'opposition /i/-/e/, les auditeurs francophones natifs distingueraient ces deux voyelles sans difficulté grâce aux informations sur les formants supérieurs (regroupement F3/F4 : Figure 41, p. 123 ; F3 différencie /i/ et /e/ ; Figure 42, p. 123), alors que les apprenants ne feraient pas de même. Notons que les japonophones ont de la difficulté à ditinguer perceptivement /l/ et /r/ de l'anglais, qui se différencient essentiellement par F3 sur le plan acoustique (Hattori et Iverson 2009[402], entre autres ; F3 extra-bas, inférieur à 2000 Hz pour le /r/ anglais).

Quant à la paire /u/-/o/, notons également que le /o/ japonais est la seule voyelle dans le système vocalique du japonais dont le F2 est très bas (inférieur à 1000 Hz), ce qui réduirait la distance perceptive entre /u/ et /o/ français, à l'écart du /u/ prototypique dans l'espace perceptif des apprenants japonophones (effet magnétique de Kuhl *et al*.[403]). Ce qui en résulte est la difficulté de distinction de la paire /u/-/o/ française, car les indices acoustiques sont en conflit pour les auditeurs japonophones : le F1 plus bas du /u/ français pointe vers /u/ (japonais), mais le F2 bas du /u/ français pointe vers /o/ (F2 est plus bas que celui du /u/ en japonais).

---

[401] CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

[402] K. Hattori, P. Iverson. English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America* 125(1) : 469–479, 2009.

[403] P. K. Kuhl, P. Iverson. "Linguistic experience and the 'Perceptual Magnet Effect'", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 121-154, 1995.

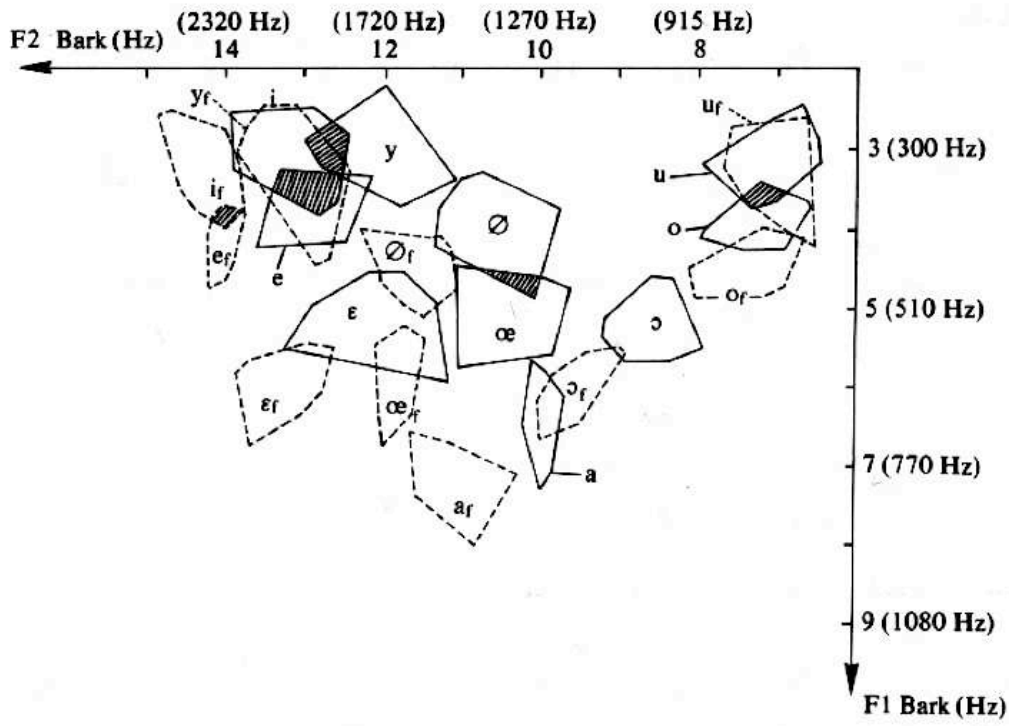



FIG. III.6. — *Zones de dispersion des voyelles non nasales du français sur le plan $F_1/F_2$* (échelle de Bark).
Sujets masculins : ————
Sujets féminins : ----------
Les hachures délimitent les zones de recouvrement pour un même sexe.

**Figure 67 :** Zones de dispersion des voyelles orales du français sur le plan F1/F2. Deux répétitions des 10 voyelles orales du français chez 10 hommes et 9 femmes. Contexte : /p_/ pour /e o u y ø/, /p_ʀ/ pour /i ɛ a ɔ œ/ (CALLIOPE 1989 : 85).

### *3.1.4. La perception par des non-apprenants : locuteurs natifs du français et natifs du japonais n'apprenant pas le français*

La même expérience a été effectuée auprès de deux groupes de locuteurs : les locuteurs natifs du français et les locuteurs natifs du japonais qui n'apprennent pas le français. L'objectif de l'expérience avec le premier groupe était de vérifier la discriminabilité des stimuli auprès des locuteurs natifs du français, et ainsi la validité du test. Quant au second, il s'agissait de savoir si certaines des paires étudiées dans cette série d'expériences seraient plus ou moins faciles à distinguer pour les locuteurs natifs du japonais qui ne sont pas familiers avec les phones et les oppositions cibles du français. Les sujets ont répondu sur un ordinateur avec une souris.

#### 3.1.5.1. Les stimuli et les procédures

Les stimuli utilisés pour l'expérience étaient identiques à ceux de l'étude décrite ci-dessus, mais les auditeurs ont passé le test individuellement sur un écran d'ordinateur, en écoutant le son par un casque. Ils disposaient d'une souris pour choisir la case sur l'écran qui correspondait à leur réponse (1 : premier ou 3 : troisième). Les consignes ont été fournies en français aux auditeurs français, et en japonais aux auditeurs japonais. Ces derniers étaient informés qu'ils entendraient des voyelles, mais le nom de la langue n'a pas été précisé. Ils avaient à cliquer sur le bouton « ok »

qui a apparu en bas de l'écran chaque fois qu'ils avaient fini leur choix, afin de finaliser la réponse et de passer au triplet suivant (Figure 68). Ils pouvaient changer leur choix de réponse avant de cliquer sur le bouton « ok ». Le triplet suivant a été présenté aux auditeurs 1,5 secondes après qu'ils ont finalisé leur réponse.

Tout comme pour le test effectué auprès des apprenants, les auditeurs ont passé d'abord un entraînement, qui leur permettait de s'habituer à la tâche de discrimination AXB. Au milieu du test, ils avaient une pause.

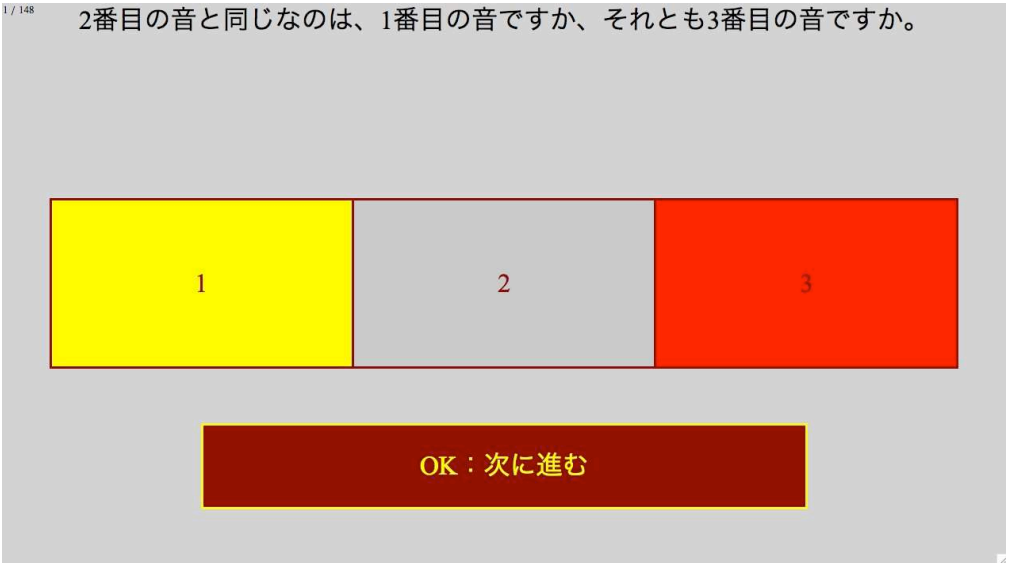



**Figure 68** : Copie de l'écran présenté aux auditeurs japonais pendant l'expérience. La case foncée (3 dans cet exemple) indique la réponse sélectionnée. La question posée aux auditeurs : « le deuxième son est-il identique au premier ou au troisième ? » (en haut). Les cases dans lesquelles figurent les chiffres 1, 2, et 3 correspondent aux premier, deuxième, et troisième sons. Le bouton « ok » (« ok : passer au suivant ») en bas.

#### 3.1.5.2. Les sujets-auditeurs

Les 8 auditeurs natifs du japonais qui ont participé à l'expérience étaient tous originaires de la région de Tokyo et de ses environs. Ils étaient âgés de 26 à 61 ans. Les 5 auditeurs natifs du français résidaient dans la région parisienne. Ils étaient âgés de 23 à 49 ans.

#### 3.1.5.3. Les résultats des auditeurs natifs francophones

Les auditeurs francophones ont distingué toutes les 6 paires de voyelles quasi parfaitement, avec un taux de réponses correctes de 99% ou plus, ce qui montre que les stimuli utilisés dans cette expérience sont distingués sans difficulté par les auditeurs natifs.

#### 3.1.5.4. Les résultats des auditeurs japonais non-apprenants

La Figure 69 montre les résultats des auditeurs japonais non-apprenants.

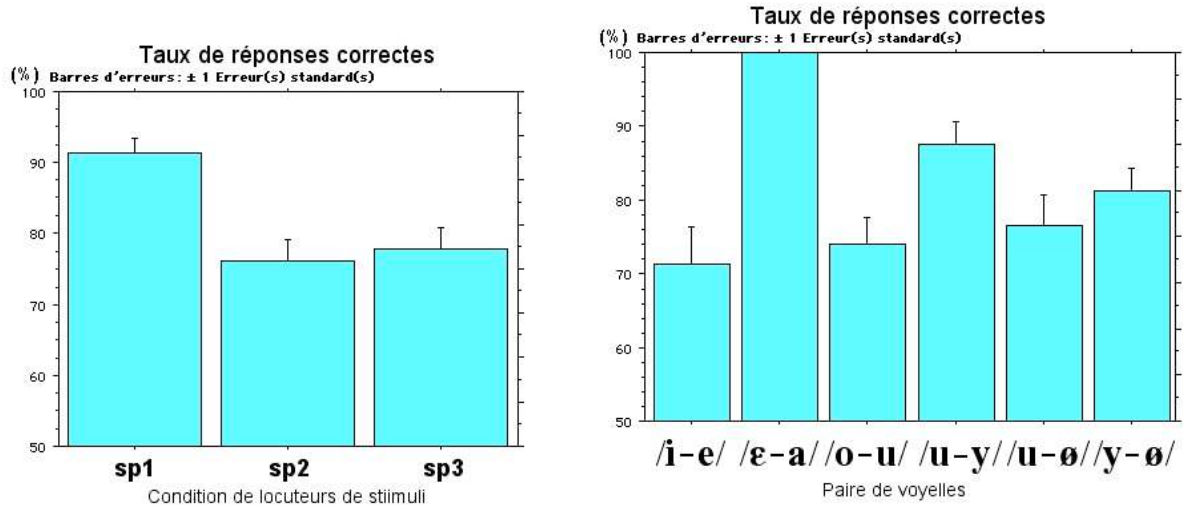



**Figure 69 :** Résultats de l'expérience AXB effectuée auprès des 8 auditeurs natifs du japonais n'apprenant pas le français. Le taux de bonnes réponses en pourcentage (niveaux du hasard : 50%) classé par les conditions de locuteurs des stimuli (à gauche), et par les paires de voyelles (à droite). Les barres d'erreurs représentent une erreur type. 72 triplets x 2 répétitions pour chaque auditeur.

**Locuteurs des stimuli** Nous voyons que le taux de bonnes réponses était plus élevé quand tous les trois stimuli du même triplet ont été prononcés par le/la même locuteur/locutrice (sp1 : 91,4% ; contre 76% pour sp2 et 77,9% pour sp3). Il y a un effet significatif des conditions de locuteurs (ANOVA : $F_{(2, 141)} = 9{,}77$, $p = 0{,}001$), et les différences entre sp1 et sp2 ($p < 0{,}001$), ainsi que sp1 et sp3 ($p = 0{,}005$) sont significatives (PLSD de Fisher). Cette tendance a été observée chez les apprenants (Figure 65). En revanche, à la différence des données des apprenants, la différence entre sp2 (stimuli prononcés par deux femmes différentes) et sp3 (stimuli prononcés par une femme et un homme) n'était pas significative (PLSD de Fisher : $p = 0{,}63$). Cet écart entre les non-apprenants et les apprenants suggère que les non-apprenants dépendraient plus largement de la différence acoustique entre les stimuli, et que les apprenants auraient appris à percevoir les voyelles du français de façon phonémique, ne serait-ce qu'imparfaitement.

**Paires de voyelles /u y ø/** Les tendances que nous observons sont similaires à celles des apprenants (Figure 66) en ce qui concerne les voyelles /u y ø/ : le taux de bonnes réponses était le meilleur pour la paire /u/-/y/ (87,5%), suivie de /y/-/ø/ (81,3%), et de /u/-/ø/ (77,9%). La différence entre ces 3 paires n'est pas significative sauf entre /u/-/y/ et /u/-/ø/ (PLSD de Fisher : $p = 0{,}03$). Cela suggère que les locuteurs natifs du japonais distingueraient le /y/ plus facilement des deux autres même sans avoir appris le français.

**/i/-/e/ et /o/-/u/** Les paires /i/-/e/ (71,4%) et /o/-/u/ (74%) étaient plus difficiles à distinguer que les trois paires composées de /u y ø/. Cette tendance est similaire à celle des apprenants des deux niveaux concernant la paire /u/-/o/, mais uniquement à celle du niveau I concernant la paire /i/-/e/ : les apprenants du niveaux II ont obtenu un meilleur résultat pour cette paire. Ces données suggèrent qu'il existerait un effet important d'apprentissage pour la paire /i/-/e/, mais que ce ne serait pas le cas pour /o/-/u/, concernant au moins les apprenants des niveaux étudiés dans la présente étude. Cette tendance pourrait s'expliquer par le fait que certains locuteurs japonais produisent un [i] avec un rapprochement F3/F4 en japonais, alors qu'un [u] avec un F2 bas (inférieur à 1000 Hz) est très rare dans la langue (Mokhtari et Tanaka 2000 : figures

reproduites aux Figure 8, p. 41, Figure 9, p. 42 ; nos données : Figure 39, p. 120) : les apprenants auraient appris à percevoir un phénomène qui existe dans leur langue native, même s'il n'est pas pertinent du point de vue phonémique.

En résumé, les résultats de la présente expérience suggèrent que la paire /u/-/y/ est relativement facile à distinguer, que /y/-/ø/ et /i/-/e/ sont plus difficiles mais deviennent plus faciles avec l'apprentissage, et que /o/-/u/ et /u/-/ø/ restent relativement difficiles même pour les apprenants plus avancés.

Ces données seront discutées et interprétées plus en détail dans le chapitre 3.5.

## 3.2. ANALYSE ACOUSTIQUE : PRODUCTION DES VOYELLES /u y ø/ PAR DES APPRENANTS JAPONOPHONES DE DIVERS NIVEAUX D'APPRENTISSAGE (EXPÉRIENCE 7)[404]

Dans la deuxième partie, nous avons présenté une série d'études préliminaires sur la production des voyelles orales du français par des apprenants japonophones. Nous examinons maintenant de plus près la production des voyelles arrondies fermées /u y/ et mi-fermée /ø/ auprès d'un nombre plus important d'apprenants japonophones de différents niveaux d'expérience d'apprentissage.

### *3.2.1. Le corpus*

Les 10 voyelles orales et les 3 nasales du français parisien ont été placées dans la phrase-cadre « *Je dis* /V/ *comme dans* … », tout comme dans les expériences précédentes (Expériences 2 et 4). Les phrases ont été disposées dans un ordre semi-aléatoire pré-établi de sorte que les voyelles qui seraient considérées similaires par les apprenants (ex. /u/ et /ø/ ; /e/ et /ɛ/) ne se succèdent pas.

Les phrases ont été écrites en orthographe, à part la voyelle cible transcrite en API, comme le montre la Figure 70.

[404] Une partie des résultats de cette expérience a été présentée dans :

T. Kamiyama. Acquisition of the production of 'new' and 'similar' vowels: the case of /u/ and /y/ in French by Japanese-speaking learners. *Journal of the Acoustical Society of America* 123(5) : 3741, 2008.

Je dis [i] comme dans « lit ».
Je dis [u] comme dans « loup ».
Je dis [ɔ] comme dans « lors ».
Je dis [ɛ̃] comme dans « lin ».
Je dis [œ] comme dans « leur ».
Je dis [õ] comme dans « long ».
Je dis [e] comme dans « les ».
Je dis [o] comme dans « l'eau ».
Je dis [a] comme dans « la ».
Je dis [ø] comme dans « leu ».
Je dis [ɑ̃] comme dans « lent ».
Je dis [y] comme dans « lu ».
Je dis [ɛ] comme dans « lait ».
Je dis [i] comme dans « lit ».

**Figure 70 :** Phrases présentées aux apprenants pour la tâche de lecture.

### *3.2.2. Les locuteurs*

Les 50 apprenants qui ont participé à cette expérience étaient des étudiants qui suivaient tous des cours de français langue étrangère dans 4 classes différentes dans des universités privées de la région de Tokyo (Tableau 53)[405]. Les 4 groupes étaient de niveaux différents d'apprentissage :

1) Groupe 1 : niveau débutant. Les étudiants suivaient en moyenne 2 cours de 90 minutes par semaine depuis un semestre. Il y avait une étudiante (1-13f) qui avait appris le français au collège et au lycée (2 heures par semaine pendant 6 ans).

2) Groupe 2 : niveau intermédiaire. Les apprenants suivaient en moyenne 5-8 cours de 90 minutes par semaine depuis 5 semestres. Quatre d'entre eux avaient appris le français au lycée (21-02f, 21-03f, 21-07f, 21-91m : 1-4 heures par semaine pendant 1 an ou 2). Deux d'entre eux avaient appris le français également dans une école de langue (2 heures par semaine pendant 1 an pour un apprenant, 3 ans pour l'autre).

3) Groupe 3 : niveau intermédiaire. Les apprenants suivaient en moyenne 6 cours de 90 minutes depuis 3 semestres. Une d'entre eux seulement avaient appris le français au lycée (2 heures par semaine pendant 2 ans).

4) Groupe 4 : niveau intermédiaire-avancé. Les apprenants suivaient en moyenne 5 cours de 90 minutes depuis un semestre à l'université, mais ils avaient tous appris le français avant d'y entrer : au lycée (6 personnes : 2-8 heures par semaine pendant 2-3 ans), dans une école de langue (une personne : 20 heures par semaine pendant 6 mois), dans un pays francophone (4 personnes y avaient vécu 10 mois – 4 ans), ou avec une combinaison d'éléments ci-dessus.

---

[405] Nous remercions très vivement Madame Miho Inoue, qui a effectué les enregistrements auprès des étudiants qui suivaient ses cours.

Après avoir rejeté des enregistrements non valables de 3 apprenants (ceux des apprenants qui n'ont pas prononcé les voyelles cibles, ou qui ont été aidés par leurs voisins), nous avons retenu les données de 47 apprenants : 14, 10, 13, et 10 étudiants de chaque groupe.

**Tableau 53** : Apprenants japonophones qui ont participé à la présente étude.

| Groupe 1 : élémentaire | sexe | heures / sem | années | séjour dans un pays francophone | d'autres expériences d'apprentissage du français |
|---|---|---|---|---|---|
| 1-01f | f | 3 | ? | 9 jours en France | |
| 1-02f | f | 3 | 0,5 | | |
| 1-03f | f | 3 | 0,5 | | |
| 1-04f | f | 3 | 0,5 | 1 mois en GB et France, mais français non parlé | |
| 1-05f | f | 3 | 3 mois | | |
| 1-06f | f | 3 | 0,5 | | |
| 1-07f | f | 3 | 0,5 | | |
| 1-008f (données écartées) | f | 3 | 0,5 | | 2h x 0,5 an (lycée) |
| 1-09f | f | 3 | 0,5 | | |
| 1-10f | f | 3 | 0,5 | | |
| 1-11f | f | 3 | 1 | | |
| 1-12f | f | 3 | ? | | |
| 1-13f | f | 3 | 0,5 | | 2h x 6 ans (collège-lycée) |
| 1-14f | f | 3 | 0,5 | | |
| 1-15f | f | 3 | 0,5 | | |

| Groupe 2 : intermédiaire I | sexe | heures / sem | | séjour dans un pays francophone | d'autres expériences d'apprentissage du français |
|---|---|---|---|---|---|
| 21-02f | f | 9 | 4 | 1 mois x plusieurs fois en France | 1h x 1 an (lycée), 1h x 2 ans (école de langue) |
| 21-03f | f | 7,5 | 2 | | 4h x 2 ans (lycée) |
| 21-04f | f | 9 | 2 | | |
| 21-05f | f | 9 | 2 | | |
| 21-06f | f | 9 | 2 | | |
| 21-07f | f | 12 | 2 | | 1h x 1 an (lycée) |
| 21-08m | m | 9 | 2,5 | | |
| 21-09m | m | 12 | 2 | | |
| 21-10f | f | ? | 2 | 2 semaines en France | |
| 21-11m | m | 9 | 3,5 | 1 mois en France | 2h x 1 an (lycée), 3h x 2 ans (école de langue) |

| Groupe 3 : intermédiaire II | sexe | heures / sem | | séjour dans un pays francophone | d'autres expériences d'apprentissage du français |
|---|---|---|---|---|---|
| 22-01f | f | 9 | 2 | | |
| 22-02f | f | 9 | 1 | | |
| 22-03f | f | 9 | 1,5 | | |
| 22-04f | f | 9 | 1 | | |
| 22-05f | f | 9 | 1,5 | | |
| 22-006f (données écartées) | f | 9 | 1 | | |
| 22-07f | f | 9 | 1,5 | 1 semaine en France | |
| 22-08f | f | 9 | 1,5 | 6 jours en France | |
| 22-09m | m | 9 | 1 | 3 mois en France | |
| 22-10f | f | 9 | 1 | | |
| 22-11f | f | 9 | 1 | | |
| 22-12f | f | 9 | 1,5 | | 2h x 2 ans (lycée) |
| 22-13f | f | 9 | 1 | | |
| 22-14f | f | 9 | 1,5 | | |

| Groupe 4 : intermédiaire-avancé | sexe | heures / sem | | séjour dans un pays francophone | d'autres expériences d'apprentissage du français |
|---|---|---|---|---|---|
| 3-01f | f | 7,5 | 0,5 | 3 semaines en France | 2h x 3 ans (lycée) |
| 3-02f | f | 7,5 | 0,5 | | 2h x 2 ans (lycée) |
| 3-003m (données écartées) | m | 7,5 | 1 | 1 an en Belgique | |
| 3-04f | f | 7,5 | 0,5 | moins de 3 semaines en France | 8h x ? (lycée) |
| 3-05f | f | 7,5 | 0,5 | 3 semaines au Canada | 1h x 6 ans (primaire), 2h x 3 ans (collège), 6h x 3 ans (lycée) |
| 3-06m | m | ? | ? | 10 mois à Paris | 20h x 0,5 an (école de langue) |
| 3-07f | f | 7,5 | 0,5 | 3 semaines en France | 8h x 3 ans (lycée) |
| 3-08f | f | 7,5 | 0,5 | 2 semaines en France | 7h x 3 ans (lycée) |
| 3-09f | f | 7,5 | 0,5 | 4 ans en France, 1 semaine en Beligique et Suisse | |
| 3-10m | m | 7,5 | 0,5 | oui (manque de détail) | 5h x 6 ans (primaire), 5h x 3 ans (collège) |
| 3-11m | m | ? | ? | 2 ans en France | ?h x 1,5 ans (primaire), ?h x 0,5 an (collège) |

### *3.2.3. Les procédures*

Les apprenants se trouvaient dans un laboratoire d'enseignement de langue assisté par ordinateur (ELAO) lors de l'enregistrement. Les consignes suivantes, écrites en japonais, ont été présentées aux apprenants, avec les phrases-cadre à lire (Figure 70) sous format électronique :

*1) Le symbole phonétique représente la voyelle contenue dans le mot à la fin de chaque ligne.*

*2) Lisez ligne par ligne à un débit qui vous facilite la lecture.*

*3) Quand vous aurez terminé toutes les phrases, du haut en bas, répétez la même chose 5 fois.*

En réalité, tous les étudiants n'ont suivi pas strictement le nombre de répétitions demandé. Une grande majorité d'entre eux ont répété de 4 à 6 fois, mais 2 apprenants avancés dans la classe intermédiaire-avancée n'ont répété qu'une seule fois.

Les étudiants disposaient tous d'un micro-casque. Ils ont effectué la tâche pendant le cours en présence de leur enseignante. Les données ont été enregistrées sur le serveur (22050 Hz, 16 bits).

### *3.2.4. Les analyses*

**Segmentation** Les fichiers son ont été segmentés à la main à partir de la visualisation simultanée de l'oscillogramme, du spectrogramme (à bande large : fenêtre à 5 millisecondes), et du suivi de formants (sous Praat). La portion des voyelles cibles où les formants supérieurs, notamment le F2, étaient clairement visibles sur le spectrogramme a été retenue pour l'analyse. Les parties où les formants n'étaient pas stables à cause de la coarticulation éventuelle avec les sons environnants ont été exclues, à l'exception des cas du [jɯ] à la japonaise (prononcé pour /y/, /ø/ ou /u/ français), durant lesquels le F2 baisse progressivement.

**Mesures de formants : 3 mesures + moyenne** Les 4 premiers formants ont été mesurés en enregistrant les données sur un « *Log file* »[406] sous Praat. En ce faisant, on peut visualiser d'abord du suivi de formants détectés automatiquement par le logiciel, en modifiant des paramètres (nombre de formants à détecter dans une zone fréquentielle donnée) après vérification si nécessaire, et ensuite d'enregistrer les valeurs formantiques à un moment donné ou les valeurs moyennes de la sélection. Nous pouvons ainsi éviter la mauvaise détection de formants. Les formants ont été mesurés vers le début du segment (à 10,4% du début du segment en moyenne, avec un écart-type de 3,1% sur 681 mesures), au milieu (à 50,9% du début du segment en moyenne, avec un écart-type de 2,1% sur 681 mesures), vers la fin (à 89% du début du segment en moyenne, avec un écart-type de 3% sur 681 mesures), et sur l'ensemble (moyenne de toutes les valeurs détectées toutes les 6 millisecondes).

**Confrontation des deux moyennes calculées** Par la suite, la moyenne arithmétique des 3 mesures de formants (début, milieu, fin) de chaque voyelle a été calculée et comparée avec la valeur obtenue sur l'ensemble de la voyelle. La différence (en valeur absolue) entre les deux types de mesures (la moyenne des 3 mesures et la valeur sur l'ensemble), calculée d'abord pour la moyenne de toutes les répétitions de chacune des 3 voyelles (/u y ø/) de chaque apprenant, et ensuite moyennée pour tous les apprenants, était de 0,7%. Dans 77% des cas (435 cas sur 564 : 4 formants x 3 voyelles x 47 locuteurs), la différence était inférieure à 1%. Nous considérons cette différence négligeable, et par conséquent, nous présenterons les valeurs sur l'ensemble de la voyelle.

**Changement F2** Les valeurs de F2 au début et à la fin de voyelle ont été utilisées afin de juger si la voyelle en question était diphtonguée, car le /y/ (ou /ø/, /u/) prononcé par certains apprenants japonophones comme le /ju/ japonais est caractérisé par une baisse progressive de F2. Nous avons calculé la valeur relative de F2 à la fin de voyelle par rapport au début.

### *3.2.5. Les résultats*

Les valeurs des 4 premiers formants des voyelles /u y ø/ prononcées par les 47 apprenants japonophones sont représentées dans les Figure 73 (Groupe 1), Figure 74 (Groupe 2), Figure 75

[406] Pour les détails, voir le site : http://www.fon.hum.uva.nl/praat/manual/Log_files.html

(Groupe 3), et Figure 76 (Groupe 4), avec la valeur de F2 relatif (fin / début), qui suggèrerait le degré de diphtongaison. Nous avons également caluculé pour chacune des trois voyelles /u/, /y/, et /ø/ la moyene des 40 apprenantes réparties dans chaque groupe (14, 7, 12, et 7 femmes, respectivement). Ces données sont affichées en comparaison avec la moyenne des deux locutrices francophones natives de l'Expérience 2 (FR3 et FR4 : Figure 41, p. 123, Figure 43, p. 124) aux Figure 72, Figure 78 (p. 199), et Figure 80 (p. 202). Les données des 7 apprenants masculins n'ont pas été incluses dans le calcul de la moyenne à cause des différences de valeurs formantiques entre les hommes et les femmes (Figure 67, p. 185 pour les données de CALLIOPE 1989[407] ; Figure 42, p. 123, et Figure 43, p. 124 pour nos données).

**/u/** La tendance générale est représentée par les trois premiers formants de toutes les 201 occurrences des 40 femmes (sur 47 apprenants : Figure 71). Nous y observons que la quasi-totalité des occurrences (sauf les 4 qui se trouvent près des valeurs des 2 locutrices natives) présentent des valeurs de F1 et de F2 entre celles de /y/ et de /ø/ des locutrices natives (mais le F3 est trop élevé par rapport aux valeurs des natives).

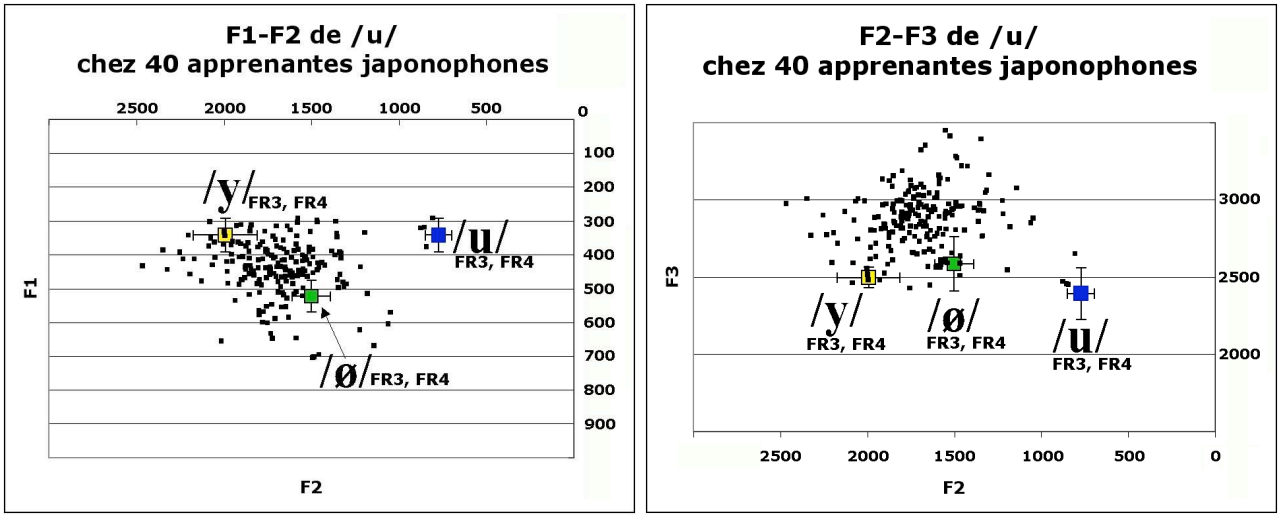



**Figure 71** : Trois premiers formants de /u/ prononcé dans la phrase cadre « je dis /u/ comme dans 'loup' » par 40 apprenantes japonophones (petits carreaux : 201 occurrences), et les moyenne de /u/, /y/ et /ø/ prononcés par 2 locutrices francophones natives (FR3 et FR4 : Figure 43, p. 124 ; 5 mesures par occurrence x 3 répétitions x 2 locutrices ; les barres d'erreur représentent l'écart-type), sur le plan F1-F2 (à gauche) et sur le plan F2-F3 (à droite).

Les valeurs moyennes des quatre premiers formants pour chacun des 4 groupes de femmes (Figure 72) montrent que tous les 4 groupes d'apprenants présentent des valeurs très éloignées de celles des 2 locutrices natives : F2 des apprenantes est supérieur à 1500 Hz (1625 Hz, 1719 Hz, 1719 Hz, et 1706 Hz pour chacun quatre des groupe), alors que celui des 2 locutrices natives est en dessous de 1000 Hz (772 Hz), et proche de F1 (342 Hz), qui est plus bas que celui des apprenantes (455 Hz, 463 Hz, 434 Hz, 381 Hz pour chacun des quatre groupes).

[407] CALLIOPE. *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson, 1989.

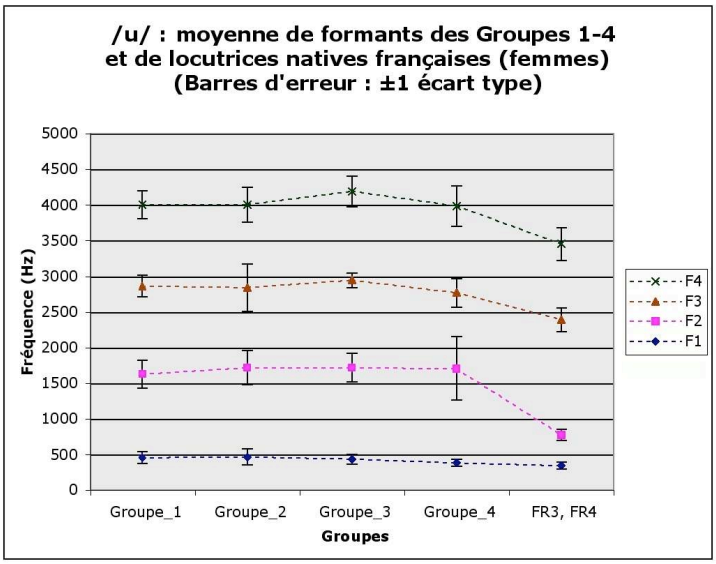



**Figure 72** : La moyenne des quatre premiers formants de la voyelle /u/ prononcé par des femmes pour les groupes suivants : Groupe 1 (14 femmes, 70 répétitions), Groupe 2 (7 femmes, 34 répétitions), Groupe 3 (12 femmes, 63 répétitions), Groupe 4 (7 femmes, 34 répétitions), locutrices natives FR3 et FR4 (6 répétitions au total : Figure 43, p. 124). Les barres d'erreur représentent l'écart-type.

Plus concrètement, seules 4 personnes parmi les 47 apprenants étudiés ont prononcé le /u/ au moins une fois avec un F2 bas, inférieur à 1000 Hz : 22-11m dans le Groupe 2 (Figure 74), 3-09f, 3-10m, et 3-11m dans le Groupe 4 (Figure 76). Ces 4 apprenants avaient tous une expérience relativement riche d'apprentissage du français avant d'entrer à l'université : 1 an au lycée et 2 ans dans une école de langue pour l'apprenant 2-11m, et les 3 autres (3-09f, 3-10m, 3-11m) avaient tous vécu dans un pays francophone (Tableau 53, p. 190). Tous les autres apprenants ont prononcé cette voyelle avec un F2 supérieur à 1000 Hz. Un certain nombre d'apprenants ont produit [y] (ou proche de [y] : 9 apprenants : 1-14f, 21-03f, 21-04f, 21-06f, 22-09m, 3-01f, 3-04f, 3-05f, 3-07f) ou [jɯ] (comme le /ju/ japonais : 10 apprenants : 1-15f, 21-05f, 21-07f, 21-09m, 22-07f, 22-09m[408], 22-10f, 22-11f, 3-02f, 3-08f), probablement à cause de la confusion entre la graphie <u> /y/ et la transcription phonémique /u/ (Figure 59, p. 142). Un apprenant (1-12f) a produit un F2 relativement bas (1047Hz, 1060 Hz), mais avec un F1 élevé (571 Hz, 605 Hz), ce qui fait un [o]. Rappelons que le F2 du /o/ japonais est plus bas que celui du /u/ [ɯ] japonais (Expérience 1 : Figure 39, p. 120).

[408] Notons que cet apprenant a prononcé les deux types, un vocoïde proche de [y] et [jɯ].

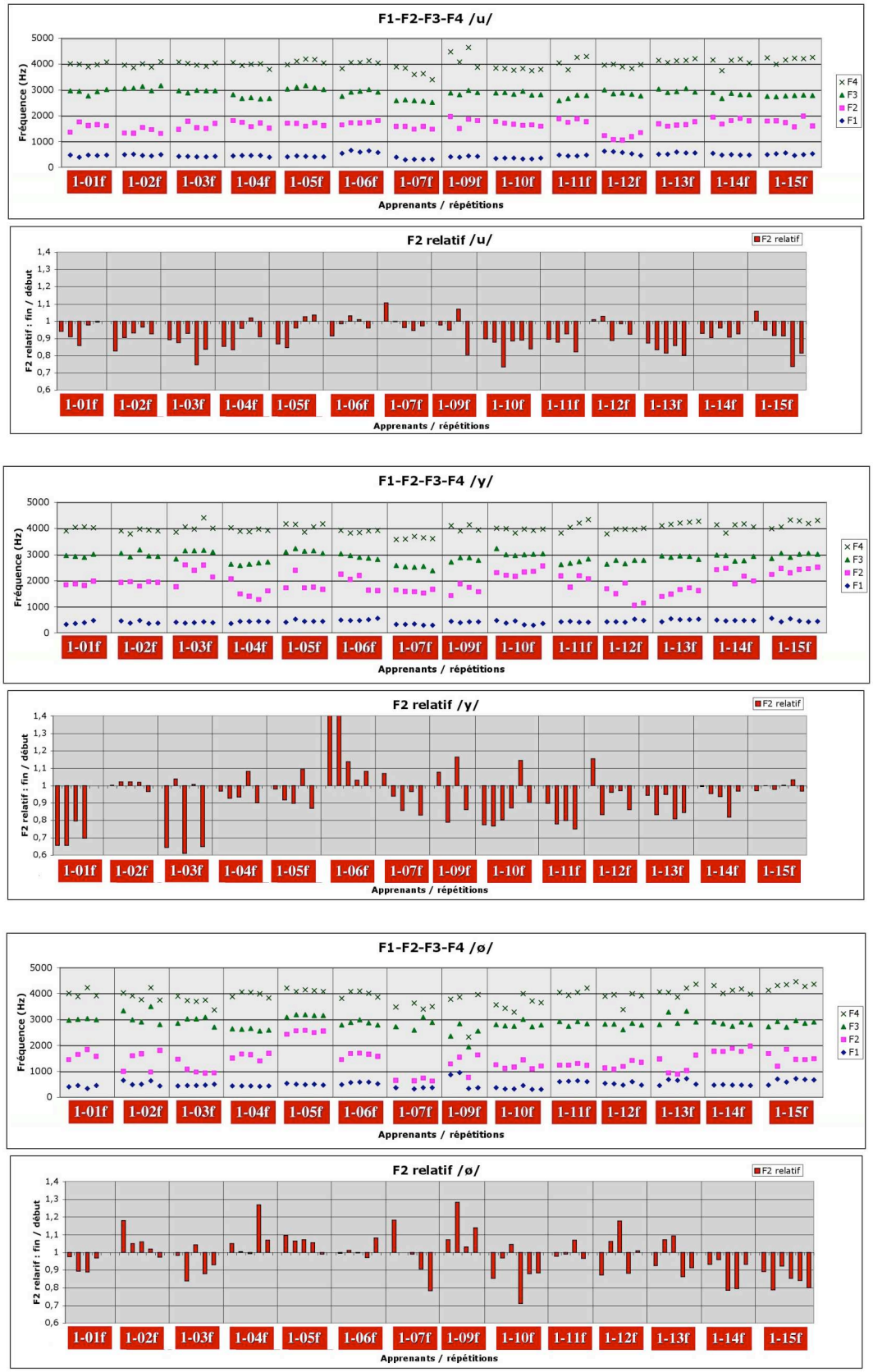



**Figure 73 :** Les quatre premiers formants des voyelles /u y ø/ (en haut de chaque paire de graphiques), et la valeur relative du F2 de la fin par rapport au début (F2 relatif : en bas) des trois voyelles prononcées par les 14 apprenants du Groupe 1 (niveau élémentaire). Chaque colonne correspond à une répétition.

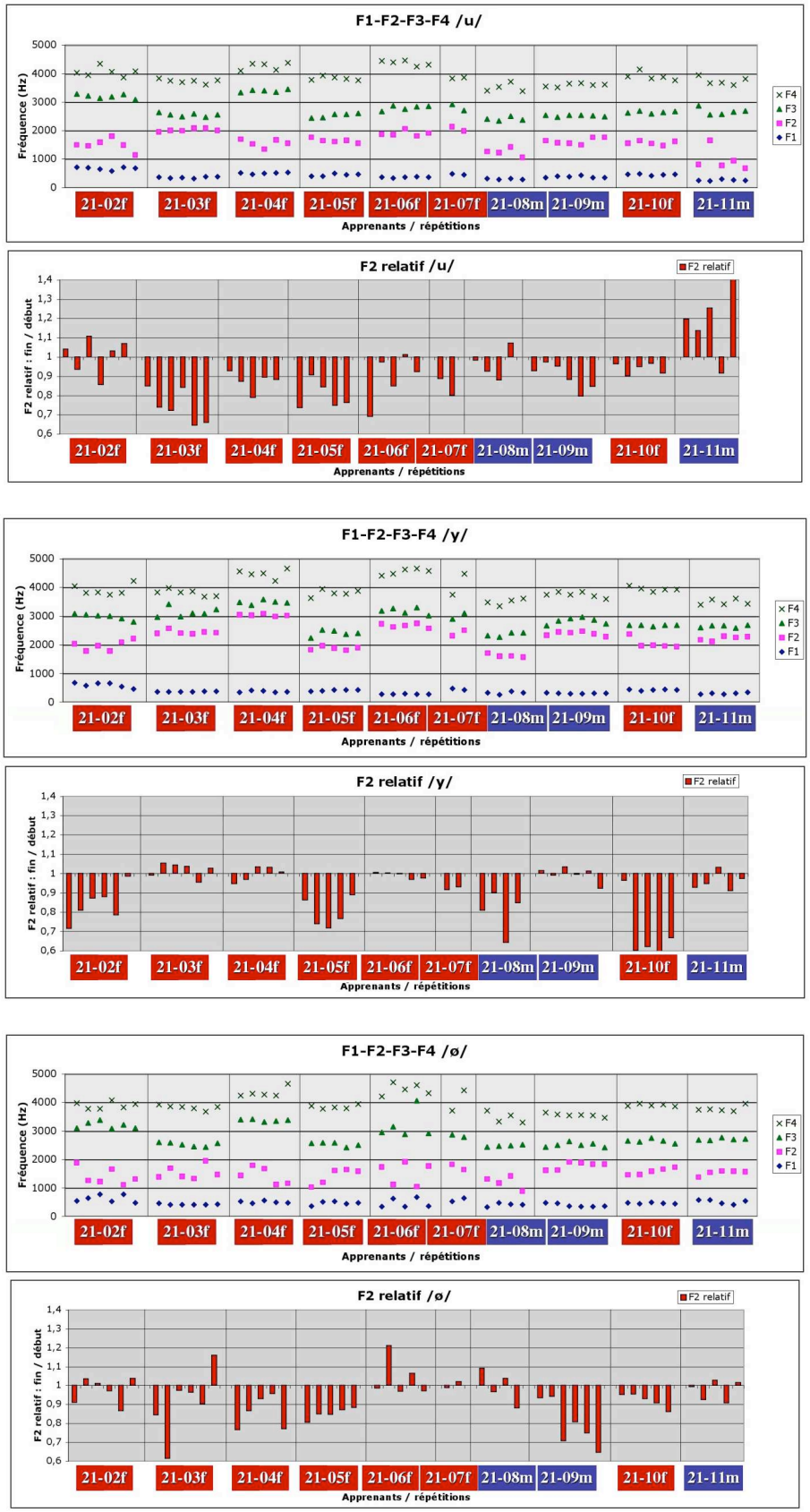



**Figure 74 :** Les quatre premiers formants des voyelles /u y ø/ (en haut de chaque paire de graphiques), et la valeur relative du F2 de la fin par rapport au début (F2 relatif : en bas) des trois voyelles prononcées par les 10 apprenants du Groupe 2 (niveau intermédiaire I). Chaque colonne correspond à une répétition.

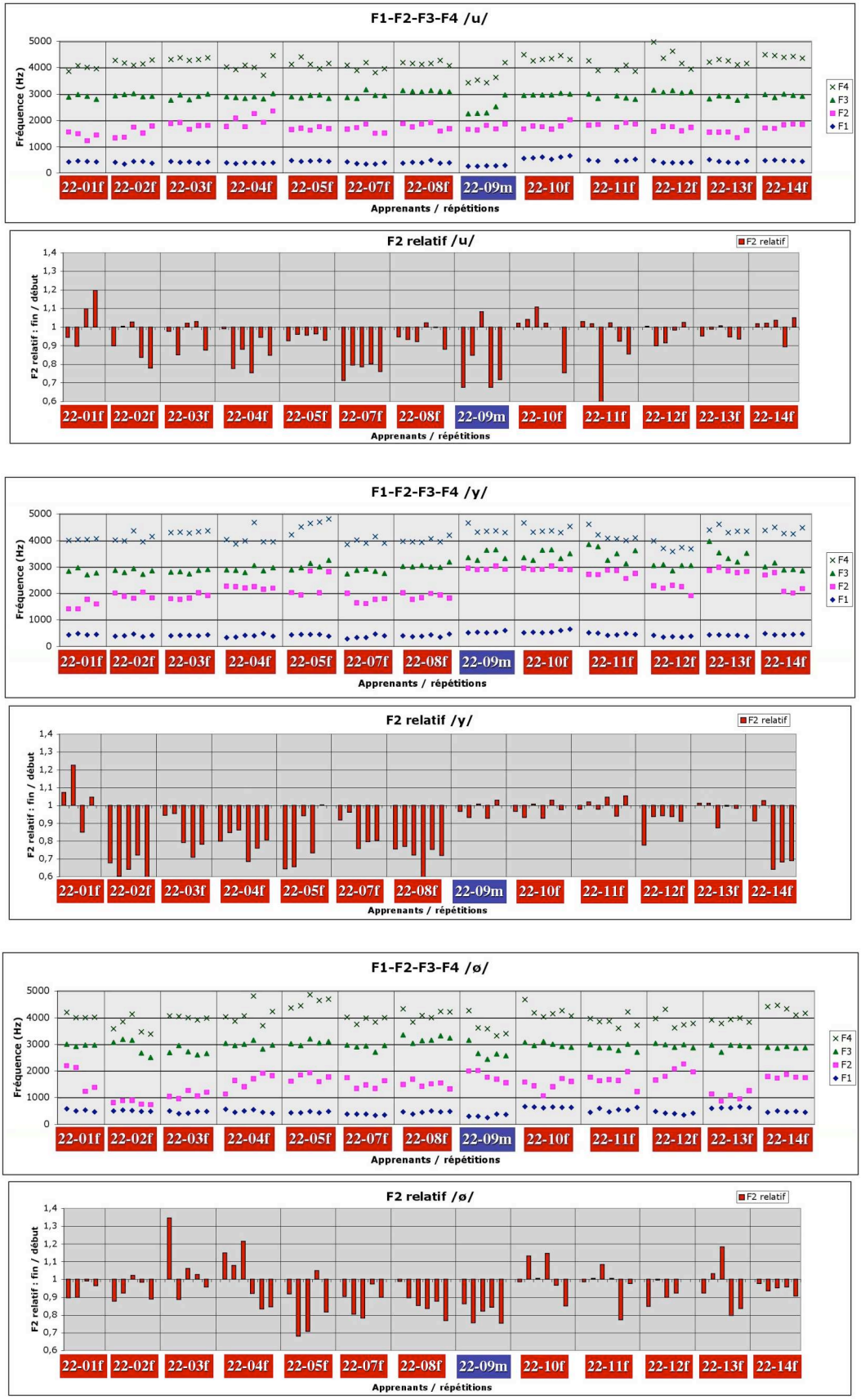



**Figure 75 :** Les quatre premiers formants des voyelles /u y ø/ (en haut de chaque paire de graphiques), et la valeur relative du F2 de la fin par rapport au début (F2 relatif : en bas) des trois voyelles prononcées par les 13 apprenants du Groupe 3 (niveau intermédiaire II). Chaque colonne correspond à une répétition.

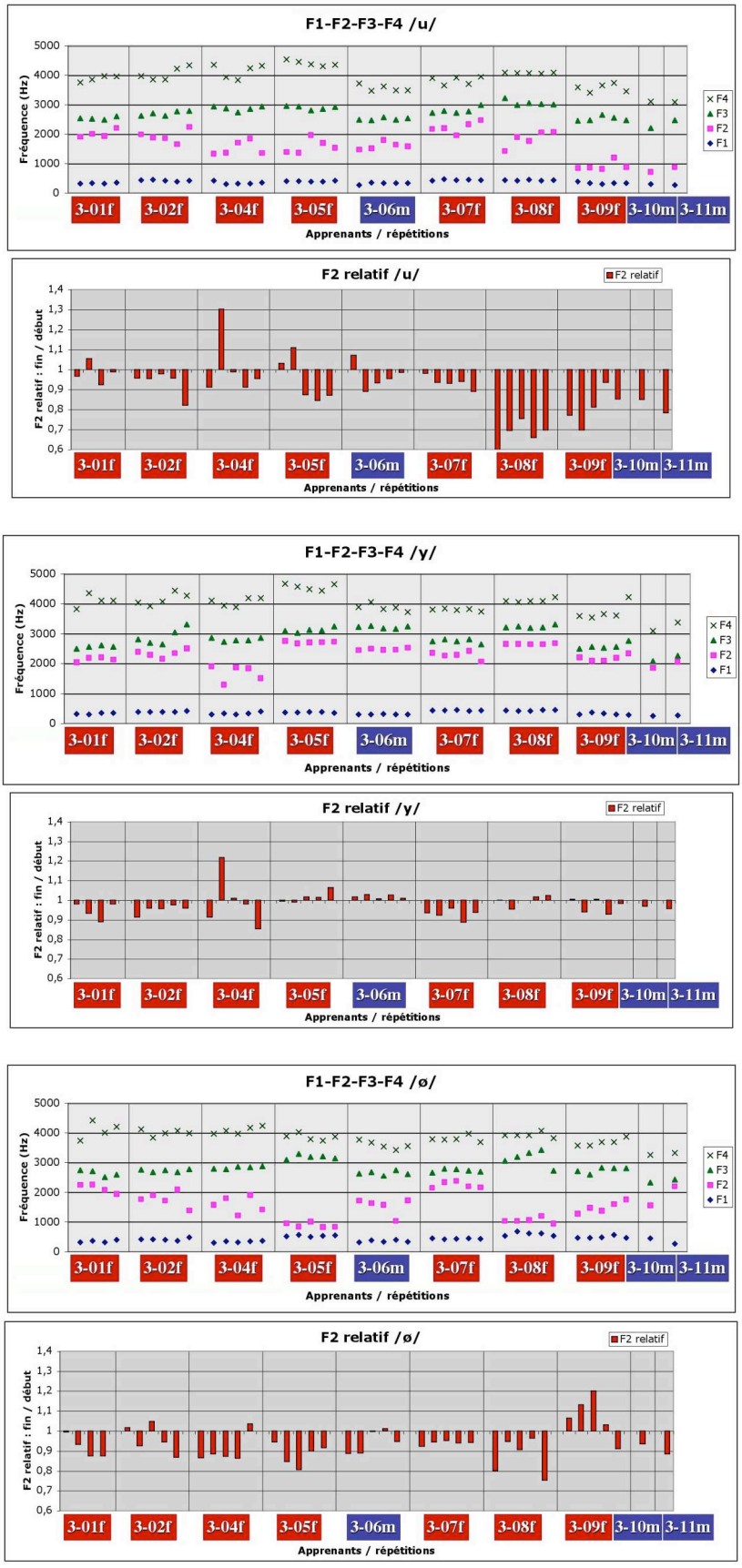



**Figure 76 :** Les quatre premiers formants des voyelles /u y ø/ (en haut de chaque paire de graphiques), et la valeur relative du F2 de la fin par rapport au début (F2 relatif : en bas) des trois voyelles prononcées par les 10 apprenants du Groupe 4 (niveau intermédiaire–avancé). Chaque colonne correspond à une répétition.

**/y/ avec F2/F3 proches** La Figure 77 montre les trois premiers formants de toutes les 200 occurrences de /y/ prononcé par les 40 femmes (sur 47 appreants). Sur le plan F1-F2, il exsite de nombreuses occurrences qui se situent près des valeurs de /y/ des loctruces native FR3 et FR4, mais sur le plan F2-F3 on observe une concentration moins importante d'occurrences autour des valeur de /y/ des natives : F3 est en général plus élevé. Nous trouvons également de nombreuses occurrences montrant F2 et F3 très élevés (47 sur 200 : F2 > 2500 Hz. F3 > 3000 Hz), qui sont relativement proches (41 sur 47 : distance F2-F3 < 700 Hz).

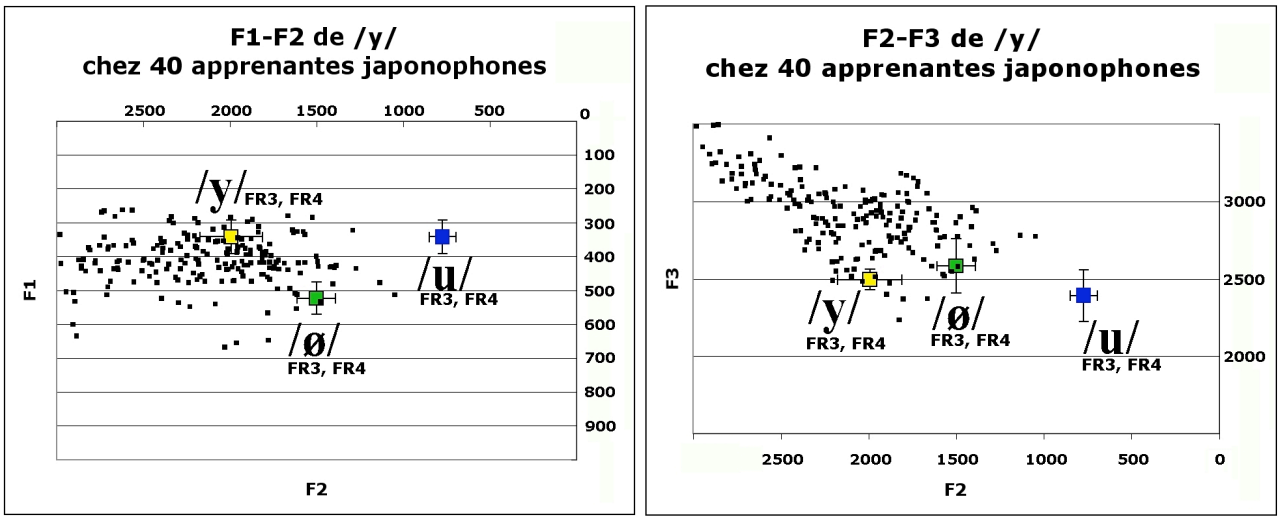



**Figure 77** : Trois premiers formants de /y/ prononcé dans la phrase cadre « je dis /y/ comme dans 'lu' » par 40 apprenantes japonophones (petits carreaux : 200 occurrences), et les moyenne de /u/, /y/ et /ø/ prononcés par 2 locutrices francophones natives (FR3 et FR4 : Figure 43, p. 124 ; 5 mesures par occurrence x 3 répétitions x 2 locutrices ; les barres d'erreur représentent l'écart-type), sur le plan F1-F2 (à gauche) et sur le plan F2-F3 (à droite).

Les valeurs moyennes des productions des femmes (Figure 78 : 40 apprenantes et 2 locutrices natives FR3 et FR4 à titre de comparaison) montrent que l'écart entre les 4 groupes d'apprenants et les deux locutrices natives (rapprochement F2/F3, et la hauteur de ce regroupement) est plus petit (F2 et F3 : 1918 Hz et 2880 Hz pour Groupe 1, 2331 Hz et 2975 Hz pour Groupe 2, 2238 Hz et 3071 Hz pour Groupe 3, 2285 Hz et 2849 Hz pour Groupe 4, 1992 Hz et 2496 Hz pour FR3 et FR4) que pour /u/ (Figure 72, p. 193).

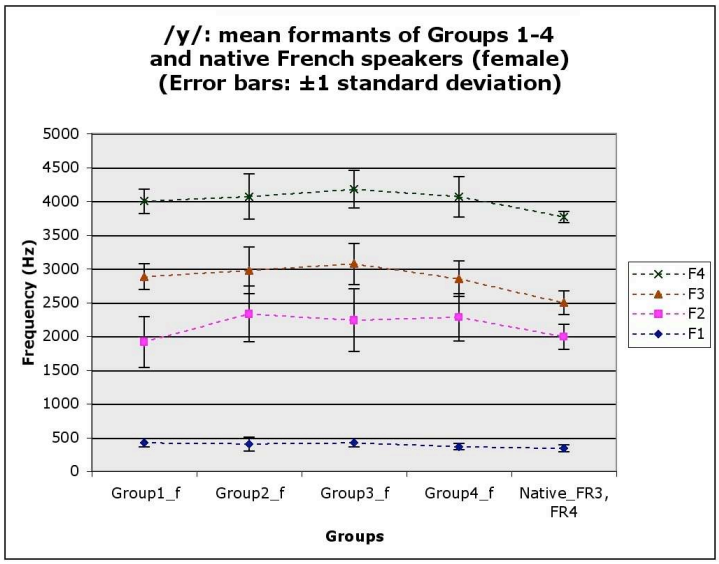



**Figure 78** : La moyenne des quatre premiers formants de la voyelle /y/ prononcé par des femmes pour les groupes suivants : Groupe 1 (14 femmes, 69 répétitions), Groupe 2 (7 femmes, 34 répétitions), Groupe 3 (12 femmes, 63 répétitions), Groupe 4 (7 femmes, 34 répétitions), locutrices natives FR3 et FR4 (6 répétitions au total). Les barres d'erreur représentent l'écart-type.

Les apprenants ont été plus nombreux à produire le /y/ avec des valeurs proches de celles des locuteurs natifs (c'est-à-dire, F2 et F3 proches autour de 1900-2000 Hz pour les hommes) que pour le /u/ (F2 inférieur à 1000 Hz) : 20 personnes ont produit cette voyelle au moins une fois avec F2 et F3 proches (distance F2/F3 < 750 Hz ; *cf*. 504 Hz pour la moyenne des locutrices natives FR3 et FR4 ; écart-type : 257 Hz), et F2 pas trop élevé (< 2200 Hz pour les hommes, < 2400 Hz pour les femmes ; *cf*. moyenne des locutrices natives FR3 et FR4 + 2 écarts-type = 2356 Hz).

Groupe 1 (élémentaire, Figure 73) : 3 apprenantes sur 14.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 1-04f | 1 | |
| 1-12f | 1 | |
| 1-14f | 3, 4 | |

Groupe 2 (intermédiaire I, Figure 74) : 6 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 21-02f | 6 | |
| 21-05f | 1 | 2-6 diphtonguées |
| 21-07f | 1 | |
| 21-08m | 2 | 1, 3, 4 diphtonguées |
| 21-10f | 1 | |
| 21-11m | 1, 2 | 3, 4, 5 : F2 trop élevé ? |

Groupe 3 (intermédiaire II, Figure 75) : 4 apprenantes sur 13.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 22-01f | 3 | |
| 22-04f | 1, 2, 3, 4, 5 | (diphtonguées ?) |
| 22-07f | 1 | |
| 22-12f | 3 | |

Groupe 4 (intermédiaire-avancée, Figure 76) : 7 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 3-01f | 1, 2, 3, 4 | |
| 3-02f | 1, 2, 3 | |
| 3-04f | 1, 3, 4 | |
| 3-07f | 1, 2, 3 | |
| 3-09f | 1, 2, 3, 4, 5 | |
| 3-10m | 1 | |
| 3-11m | 1 | |

**/y/ : diphtongaison** Quand les apprenants n'ont pas prononcé le /y/ comme les locuteurs natifs, deux tendances majeures ont été observées. La première est le phénomène de diphtongaison, tel que nous l'observons dans la séquence /ju/ [jɯ] du japonais. Parmi les 47 apprenants, 16 personnes ont produit le /y/ au moins une fois avec cette caractéristique, suggérée par une valeur basse du « F2 relatif ».

Groupe 1 (élémentaire, Figure 73) : 5 apprenantes sur 14.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 1-01f | 1, 2, 3, 4 | |
| 1-03f | 1, 3, 5 | |
| 1-09f | 2 | |
| 1-10f | 1, 2, 3 | |
| 1-11f | 2, 3, 4 | |

Groupe 2 (intermédiaire I, Figure 74) : 3 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 21-05f | 2, 3, 4, 5, 6 | |
| 21-08m | 1, 3, 4 | |
| 21-10f | 2, 3, 4, 5 | |

Groupe 3 (intermédiaire II, Figure 75) : 8 apprenantes sur 13.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 22-02f | 1, 2, 3, 4, 5 | |
| 22-03f | 3, 4, 5 | |
| 22-04f | 1, 2, 3, 4, 5, 6 | |
| 22-05f | 1, 2, 4 | |
| 22-07f | 2, 3, 4, 5 | |
| 22-08f | 1, 2, 3, 4, 5, 6 | |
| 22-12f | 1 | |
| 22-14f | 3, 4, 5 | |

Groupe 4 (intermédiaire-avancée, Figure 76) : aucun apprenant (sur 10).

**« [i] F2 »** L'autre type fréquemment observé de prononciation différente de la cible était une sorte de [i] : il s'agissait souvent d'un [i] avec un F2 élevé, avec F2/F3 relativement proches. Il a été observé dans les productions suivantes :

Groupe 1 (élémentaire, Figure 73) : 4 apprenantes sur 14.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 1-03f | 2, 4 | |
| 1-10f | 4, 5, 6 | |
| 1-14f | 1, 2 | |
| 1-15f | 4, 5, 6 | |

Groupe 2 (intermédiaire I, Figure 74) : 4 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 21-03f | 1, 2, 3, 4, 5, 6 | |
| 21-04f | 1, 2, 3, 4, 5 | |
| 21-06f | 1, 2, 3, 4, 5 | |
| 21-09m | 1, 2, 3, 4, 5, 6 | |

Groupe 3 (intermédiaire II, Figure 75) : 5 apprenants sur 13.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 22-09m | 1, 2, 3, 4, 5 | |
| 22-10f | 1, 2, 3, 4, 5, 6 | |
| 22-11f | 1, 2, 3, 4, 5, 6 | |
| 22-13f | 1, 2, 3, 4, 5 | |
| 22-14f | 1, 2 | |

Groupe 4 (intermédiaire-avancée, Figure 76) : 4 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 3-02f | 5 | |
| 3-05f | 1, 2, 3, 4, 5 | |
| 3-06m | 1, 2, 3, 4, 5 | |
| 3-08f | 1, 2, 3, 4, 5 | |

Nous ne pouvons déterminer si c'est une déviation phonétique par rapport à la cible [y] (F2 est trop élevé pour un beau [y]), ou que c'est dû à la confusion entre la transcription phonémique /y/ et la graphie <y> /i/ (Figure 59, p. 142). Dans l'Expérience 4, aucun des 3 apprenants n'a prononcé une voyelle ressemblant à [i] dans la tâche de répétition (Figure 60, p. 144, Figure 61, p. 147, Figure 62, p. 149).

***/ø/*** La Figure 79 montre 199 occurrences de /ø/ prononcés par 40 femmes (sur 47 apprenants). Les deux premiers formants, notamment le F2, sont distribués de part et d'autre des valeurs des locutrices natives, mais le F3 est globalement trop élevé par rapport aux valeurs des locutrices natives.

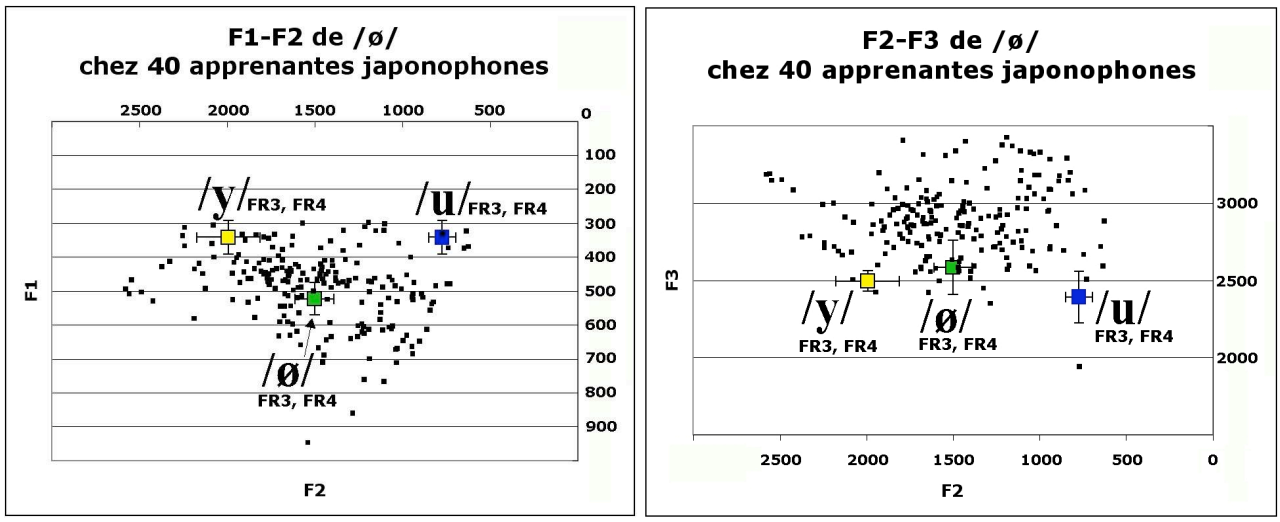



**Figure 79** : Trois premiers formants de /ø/ prononcé dans la phrase cadre « je dis /ø/ comme dans 'leu' » par 40 apprenantes japonophones (petits carreaux : 199 occurrences), et les moyenne de /u/, /y/ et /ø/ prononcés par 2 locutrices francophones natives (FR3 et FR4 : Figure 43,

p. 124 ; 5 mesures par occurrence x 3 répétitions x 2 locutrices ; les barres d'erreur représentent l'écart-type), sur le plan F1-F2 (à gauche) et sur le plan F2-F3 (à droite).

Les valeurs moyennes des productions des femmes (40 apprenantes et 2 locutrices natives à titre de comparaison : Figure 80) montrent que la structure formantique des quatre groupes d'apprenants est ressemblante à celle des deux locutrices natives, plus que pour /u/ (Figure 72, p. 193) et pour /y/ (Figure 78, p. 199). Tous les groupes présentent une voyelle acoustiquement centrale : les formants sont à peu près équidistants, avec F2 autour de 1500 Hz. F3 a tendance cependant à être plus élevé que celui des deux locutrices françaises (2864 Hz, 2897 Hz, 2943 Hz, 2861 Hz, *versus* 2583 Hz pour FR3 et FR4).

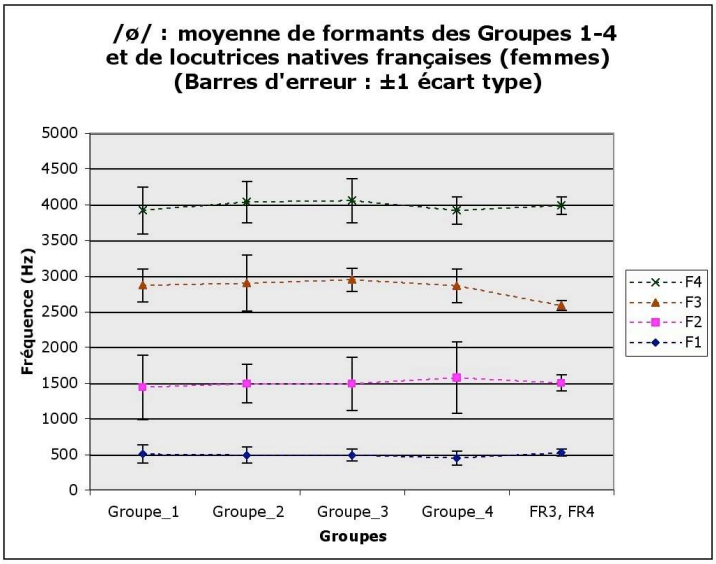



**Figure 80** : La moyenne des quatre premiers formants de la voyelle /ø/ prononcé par des femmes pour les groupes suivants : Groupe 1 (14 femmes, 68 répétitions), Groupe 2 (7 femmes, 34 répétitions), Groupe 3 (12 femmes, 63 répétitions), Groupe 4 (7 femmes, 34 répétitions), locutrices natives FR3 et FR4 (6 répétitions au total). Les barres d'erreur représentent l'écart-type.

La majorité des apprenants ont prononcé le /ø/ avec une répartition à peu près équidistante des 4 premiers formants, avec F2 autour de 1500 Hz. Les apprenants suivants ont produit une telle voyelle au moins une fois :

Groupe 1 (élémentaire, Figure 73) : 7 apprenantes sur 14.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 1-01f | 1, 2, 3, 4 | |
| 1-03f | 1 | |
| 1-06f | 1, 2, 3, 4, 5 | |
| 1-10f | 4 | |
| 1-13f | 1, 5 | |
| 1-14f | 1, 2 | |
| 1-15f | 1 | |

Groupe 2 (intermédiaire I, Figure 74) : 10 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 21-02f | 4, 6 | |
| 21-03f | 1, 3, 4, 6 | |

| | | |
|---|---|---|
| 21-04f | 1, 3 (2 ?) | |
| 21-05f | 3, 4, 5 | |
| 21-06f | 1, 6 | |
| 21-07f | 1 ? | |
| 21-08m | 1, 3 | |
| 21-09m | 1, 2 | |
| 21-10f | 1, 2, 3, 4 (5 ?) | |
| 21-11m | 1, 2, 3, 4, 5 | |

Groupe 3 (intermédiaire II, Figure 75) : 9 apprenantes sur 13.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 22-01f | 3, 4 | |
| 22-04f | 2, 4 | |
| 22-05f | 1, 5 | |
| 22-07f | 1, 2, 3, 4, 5 | |
| 22-08f | 1, 2, 3, 4, 5 | |
| 22-10f | 1, 2, 4, 5, 6 | |
| 22-11f | 1, 2, 3, 4 | |
| 22-12f | 1 (2 ?) | |
| 22-14f | 1, 2, 3, 4, 5 | |

Groupe 4 (intermédiaire-avancée, Figure 76) : 5 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 3-02f | 5 (3 ?) | |
| 3-04f | 5 (1 ?) | |
| 3-06m | 2, 3 | |
| 3-09f | 1, 2, 3, 4 | |
| 3-10m | 1 | |

Parmi les types de prononciation non conformes à la cible, le plus fréquemment observé était [o] ou [ɔ]. Notons que les apprenants ont souvent produit une voyelle du type [o] avec un F2 inférieur à 1000 Hz.

Classe 1 (élémentaire, Figure 73) : 6 apprenantes sur 14.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 1-02f | 1, 4 | |
| 1-03f | 2, 3, 4, 5 | |
| 1-11f | 1, 2, 3, 4 | |
| 1-12f | 1, 2, 3 | |
| 1-13f | 2, 3, 4 | |
| 1-15f | 2 | |

Classe 2 (intermédiaire I, Figure 74) : 3 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 21-05f | 1, 2 | |
| 21-06f | 2, 4 | |
| 21-08m | 4 | |

Classe 3 (intermédiaire II, Figure 75) : 5 apprenantes sur 13.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 22-02f | 1, 2, 3, 4, 5 | |
| 22-03f | 1, 2, 3, 4, 5 | |
| 22-04f | 1 | |
| 22-11f | 6 | |
| 22-13f | 2, 3, 4 | |

Classe 4 (intermédiaire-avancée, Figure 76) : 3 apprenants sur 10.

| Apprenant(e) | Répétition(s) | |
|---|---|---|
| 3-05f | 1, 2 ,3 4, 5 | |
| 3-06m | 4 | |
| 3-08f | 1 | |

Nous avons également observé un nombre moins important de cas de [y] bien prononcé (4 apprenants : 21-03f, répétition –rép.– 5 ; 3-02f, rép. 4 ; 3-07f, rép. 1 ; 3-93m, rép. 1), et de [jɯ] à la japonaise (4 apprenants : 1-14f, rép. 3, 4 ; 21-03f, rép. 2 ; 22-05f, rép. 2, 3 ; 22-09m, rép. 2, 5). Cela serait à cause de l'association entre la graphie <eu> et la séquence /juː/ de l'anglais (interprétée comme /ju/ japonais : Figure 59, p. 142).

Les données de production des voyelles /u y ø/ présentées dans ce chapitre montrent que les apprenants ont mieux réussi à produire le /y/ (produit par 20 apprenants au moins une fois) que le /u/ (produit par 4 apprenants), et encore mieux le /ø/ (produit par 31 apprenants au moins une fois). Les autres types principaux de prononciation des apprenants sont résumés au Tableau 54 ci-dessous.

**Tableau 54 :** Principaux types de prononciation des voyelles /u/, /y/, et /ø/ observés chez les 47 apprenants étudiés dans cette expérience. N.B. La somme des nombres d'apprenants peut dépasser 47, car certains apprenants ont produits plusieurs types pour une voyelle cible donnée.

| **/u/** | **[u] : F2 < 1000 Hz** | **4 apprenants** |
|---|---|---|
| | [y] ou proche de [y] | 9 apprenants |
| | [jɯ] à la japonaise (/ju/ jp) | 10 apprenants |
| | [o] | 1 apprenante |
| | | |
| **/u/** | **[y] : F2/F3 proche** (distance < 750 Hz)**, F2 pas trop élevé** (2400 Hz) | **20 apprenants** |
| | [y] diphthongué et [jɯ] à la japonaise (/ju/ jp) | 16 apprenants |
| | [i] avec F2 élevé | 17 apprenants |
| | | |
| **/ø/** | **[ø] : F2 autour de 1500 Hz, sans rapprochement F2/F3** | **31 apprenants** |
| | [o ~ ɔ] | 17 apprenants |
| | [y] ou proche de [y] | 4 apprenants |
| | [jɯ] à la japonaise (/ju/ jp) | 4 apprenants |

Nous pouvons maintenant nous poser la question de savoir comment les valeurs formantiques des voyelles prononcées par les apprenants sont liées plus précisément avec la perception par des locuteurs natifs francophones. Comment le /u/, par exemple, prononcé majoritairement avec un F2 supérieur à 1000 Hz par les apprenants, sera-t-il identifié et évalué par les locuteurs natifs en fonction de la hauteur de F2 ?

## 3.3. IDENTIFICATION : PERCEPTION DES VOYELLES /u y ø/ PRODUITES PAR DES APPRENANTS JAPONOPHONES AUPRÈS D'AUDITEURS FRANCOPHONES NATIFS (EXPÉRIENCE 8)

Nous allons présenter dans ce chapitre une expérience d'identification et d'évaluation de la qualité des voyelles /u y ø/ prononcées par des apprenants japonais auprès d'auditeurs natifs du français. L'objectif de cette étude est d'examiner de plus près la relation entre les valeurs formantiques des stimuli prononcés par les apprenants et la perception des auditeurs natifs, afin de déterminer les indices acoustiques pertinents de ces trois voyelles sur le plan perceptif, que les apprenants doivent à produire.

### *3.3.1. Les stimuli*

Les stimuli utilisés dans cette expérience sont représentés dans le Tableau 55. Ils ont été prononcés par 3 apprenants dans deux tâches, de lecture et de répétition (Expérience 4) : sp1 correspond à l'apprenant 1, sp2 à l'apprenant 2, et sp3 à l'apprenante 3. Rappelons que ces trois apprenants ont montré les tendances suivantes (Figure 60, p. 144 ; Figure 61, p. 147 ; Figure 62, p. 149) :

/u/ : tous les 3 apprenants ont produit le /u/ avec un F2 supérieur à 1000 Hz.

/y/ : les apprenants 1 et 2 (sp1 et sp2 ici) ont prononcé le /y/ avec F2 et F3 proches, tandis que l'apprenante 3 (sp3) a produit une séquence diphtonguée à la japonaise (/ju/ [jɯ]) dans la plupart des temps (5 cas sur 6 : y-11, y-12, y-14, y15, y-16 au Tableau 55 et à la Figure 84).

/ø/ : les apprenants 1 et 2 (sp1 et sp2) ont prononcé le /ø/ comme [y] dans la tâche de lecture (ø-01, ø-02, ø-05, ø-06 : Tableau 55, Figure 85).

À ces stimuli prononcés par les 3 apprenants a été rajoutée une répétition d'un apprenant et d'une apprenante qui ont produit des valeurs assez proches de celles des locuteurs natifs dans l'Expérience 7 (sp4 : l'apprenant 21-11m du Groupe 2 ; sp5 : l'apprenante 3-09f du Groupe 4 ; Tableau 53, p. 190), afin d'équilibrer la liste des stimuli. Le nombre des stimuli s'élève ainsi à 18 pour chaque voyelle, soit 54 au total. La Figure 82 représente ces 54 stimuli sur les plan F1-F2 et F2-F3, en comparaison avec les valeurs moyennes des locuteurs natifs (Figure 42, p. 123 ; Figure 43, p. 124).

**Tableau 55 :** Liste des stimuli des voyelles /u/ (en haut à gauche), /y/ (en haut à droite), et /ø/ (en bas) prononcés par 5 apprenants japonophones et utilisés dans le test d'identification et d'évaluation.

| ID stimulus | Locuteur | Sexe | Tâche | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|---|---|---|---|---|---|---|---|
| u-01 | sp1 | M | lecture | 313 | 1051 | 2237 | 3228 |
| u-02 | sp1 | M | lecture | 319 | 1080 | 2186 | 3240 |
| u-03 | sp1 | M | répétition | 304 | 1215 | 2157 | 3053 |
| u-04 | sp1 | M | répétition | 322 | 1509 | 2212 | 3288 |
| u-05 | sp2 | M | lecture | 326 | 1363 | 2168 | 3361 |
| u-06 | sp2 | M | lecture | 325 | 1204 | 2168 | 3406 |
| u-07 | sp2 | M | répétition | 331 | 1541 | 2371 | 3551 |
| u-08 | sp2 | M | répétition | 350 | 1429 | 2348 | 3540 |
| u-09 | sp2 | M | répétition | 331 | 1292 | 2318 | 3456 |
| u-10 | sp2 | M | répétition | 332 | 1384 | 2235 | 3529 |
| u-11 | sp3 | F | lecture | 376 | 1415 | 3164 | 4533 |
| u-12 | sp3 | F | lecture | 436 | 1360 | 3336 | 4495 |
| u-13 | sp3 | F | répétition | 409 | 1123 | 3004 | 4013 |
| u-14 | sp3 | F | répétition | 433 | 1230 | 2941 | 4070 |
| u-15 | sp3 | F | répétition | 468 | 1349 | 3077 | 4225 |
| u-16 | sp3 | F | répétition | 433 | 1329 | 3134 | 4294 |
| u-17 | sp4 | M | lecture | 290 | 775 | 2574 | 3680 |
| u-18 | sp5 | F | lecture | 292 | 806 | 2651 | 3643 |

| ID stimulus | Locuteur | Sexe | Tâche | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|---|---|---|---|---|---|---|---|
| y-01 | sp1 | M | lecture | 268 | 2088 | 2587 | 3288 |
| y-02 | sp1 | M | lecture | 294 | 2035 | 2458 | 3404 |
| y-03 | sp1 | M | répétition | 272 | 1946 | 2438 | 3297 |
| y-04 | sp1 | M | répétition | 263 | 1961 | 2430 | 3387 |
| y-05 | sp2 | M | lecture | 319 | 2042 | 2302 | 3369 |
| y-06 | sp2 | M | lecture | 315 | 1968 | 2209 | 3381 |
| y-07 | sp2 | M | répétition | 336 | 1907 | 2325 | 3365 |
| y-08 | sp2 | M | répétition | 327 | 1951 | 2344 | 3402 |
| y-09 | sp2 | M | répétition | 338 | 1962 | 2312 | 3344 |
| y-10 | sp2 | M | répétition | 330 | 1891 | 2388 | 3356 |
| y-11 | sp3 | F | lecture | 394 | 1928 | 3019 | 4305 |
| y-12 | sp3 | F | lecture | 404 | 1961 | 3017 | 4208 |
| y-13 | sp3 | F | répétition | 405 | 1688 | 2863 | 4091 |
| y-14 | sp3 | F | répétition | 433 | 1772 | 2889 | 4092 |
| y-15 | sp3 | F | répétition | 438 | 1970 | 2922 | 4233 |
| y-16 | sp3 | F | répétition | 418 | 1935 | 2897 | 4193 |
| y-17 | sp4 | M | lecture | 273 | 2290 | 2664 | 3416 |
| y-18 | sp5 | F | lecture | 357 | 2081 | 2562 | 3541 |

| ID stimulus | Locuteur | Sexe | Tâche | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|---|---|---|---|---|---|---|---|
| ø-01 | sp1 | M | lecture | 311 | 1904 | 2411 | 3373 |
| ø-02 | sp1 | M | lecture | 356 | 1796 | 2418 | 3393 |
| ø-03 | sp1 | M | répétition | 359 | 1535 | 2401 | 3308 |
| ø-04 | sp1 | M | répétition | 327 | 1520 | 2257 | 3253 |
| ø-05 | sp2 | M | lecture | 332 | 1783 | 2144 | 3134 |
| ø-06 | sp2 | M | lecture | 340 | 1752 | 2144 | 3118 |
| ø-07 | sp2 | M | répétition | 382 | 1704 | 2342 | 3116 |
| ø-08 | sp2 | M | répétition | 341 | 1715 | 2321 | 3211 |
| ø-09 | sp2 | M | répétition | 369 | 1642 | 2188 | 3227 |
| ø-10 | sp2 | M | répétition | 335 | 1608 | 2093 | 3189 |
| ø-11 | sp3 | F | lecture | 511 | 1543 | 3211 | 4278 |
| ø-12 | sp3 | F | lecture | 497 | 1314 | 3081 | 4274 |
| ø-13 | sp3 | F | répétition | 451 | 1295 | 3103 | 4293 |
| ø-14 | sp3 | F | répétition | 444 | 1437 | 3019 | 4315 |
| ø-15 | sp3 | F | répétition | 479 | 1345 | 3139 | 4264 |
| ø-16 | sp3 | F | répétition | 486 | 1341 | 3080 | 4286 |
| ø-17 | sp4 | M | lecture | 449 | 1586 | 2754 | 3717 |
| ø-18 | sp5 | F | lecture | 456 | 1468 | 2584 | 3574 |

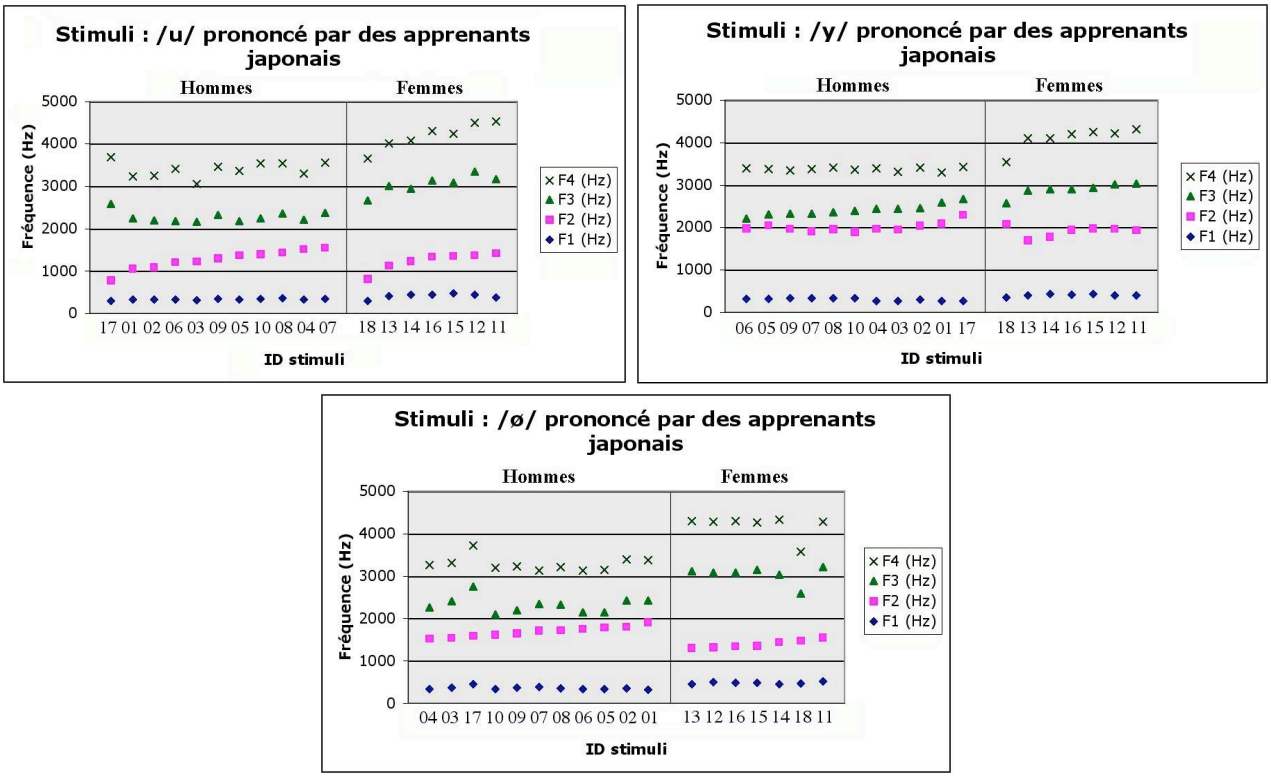



**Figure 81 :** Les quatre premiers formants des stimuli des trois voyelles /u/, /y/, et /ø/ (représentation graphique du Tableau 55), ordonnés en fonction de la valeur croissante de F2 pour ceux du /u/ (en haut à gauche), de F3 pour ceux du /y/ (en haut à droite), et de F2 pour ceux du /ø/ (en bas).

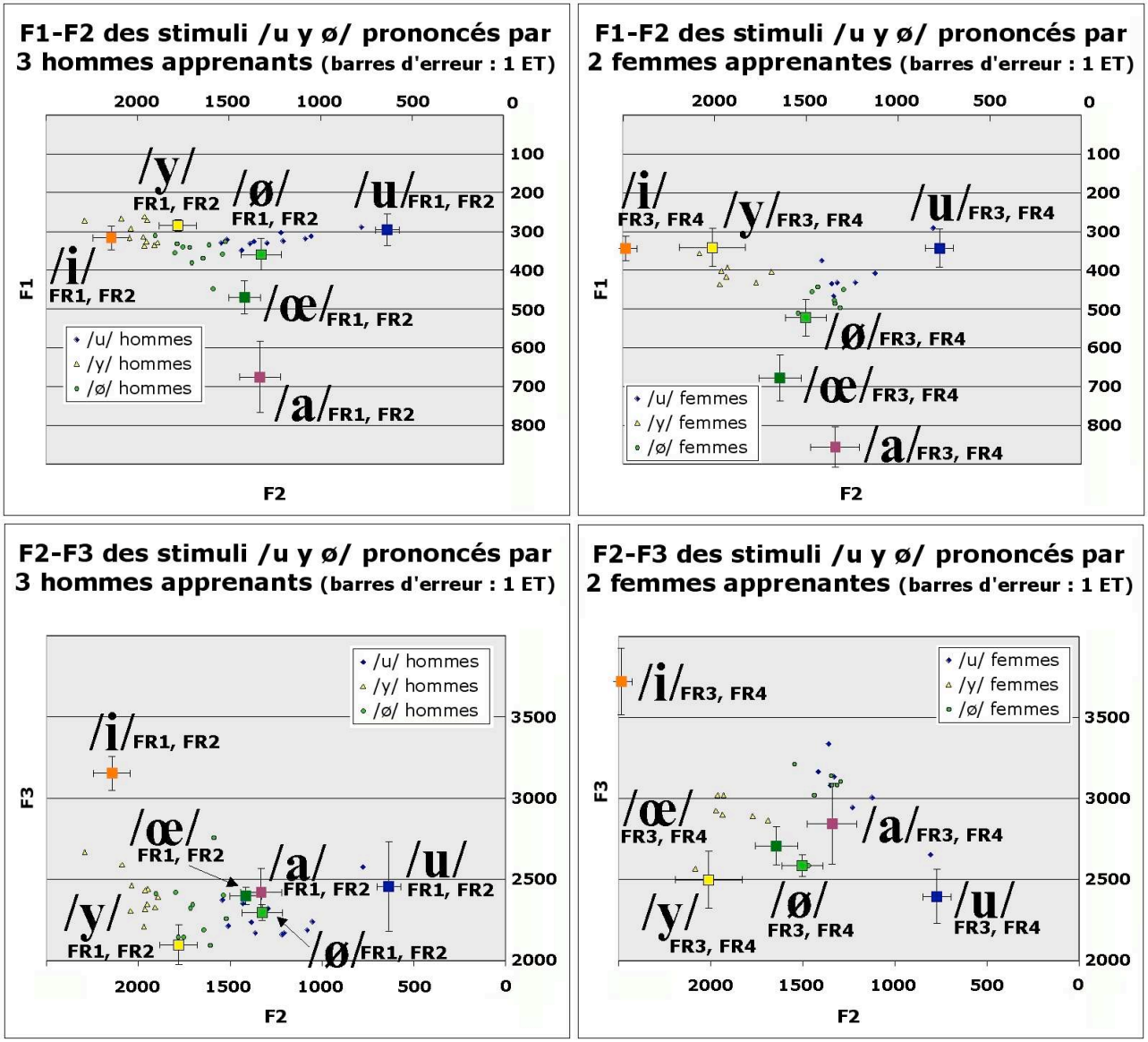



**Figure 82 :** Les trois premiers formants des stimuli de /u/ (petits carreaux), de /y/ (petits triangles) et de /ø/ (petits cercles) prononcés par 3 hommes apprenants japonophones (à droite : 11 stimuli pour chaque voyelle) et 2 femmes apprenantes (à gauche : 7 stimuli pour chaque voyelle), et les valeurs moyennes de /i/, /a/, /u/, /y/, /ø/, /œ/ prononcés par deux locuteurs francophones natifs FR1 et FR2 (à gauche ; Figure 42, p. 123) et par deux locutrices francophones natives (à droite : Figure 43, p. 124). Les barres d'erreur représentent l'écart-type.

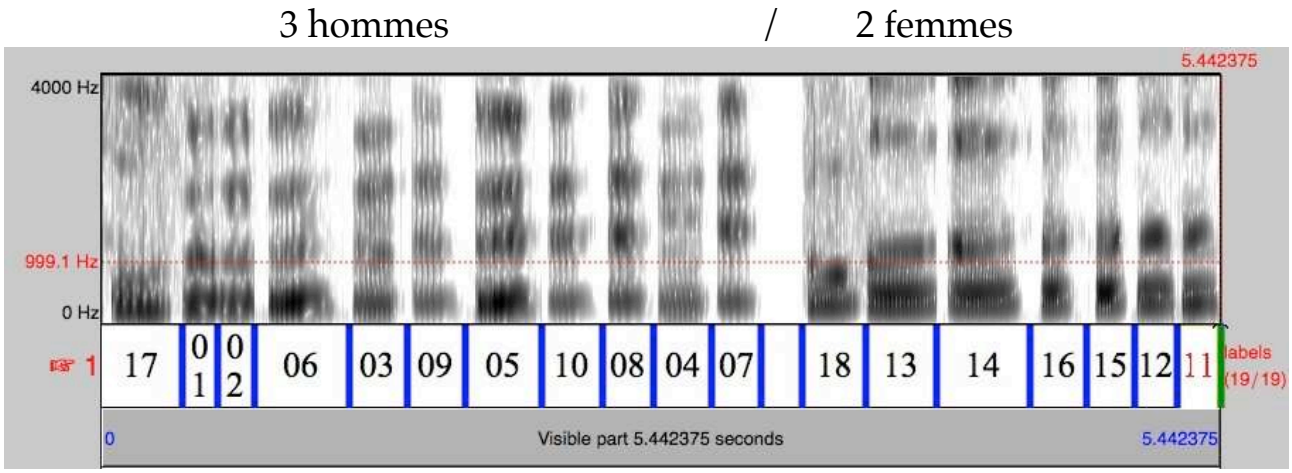



**Figure 83 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 18 stimuli de /u/ prononcés par 5 apprenants japonophones (ceux de 3 hommes à gauche, ceux de 2 femmes à droite).

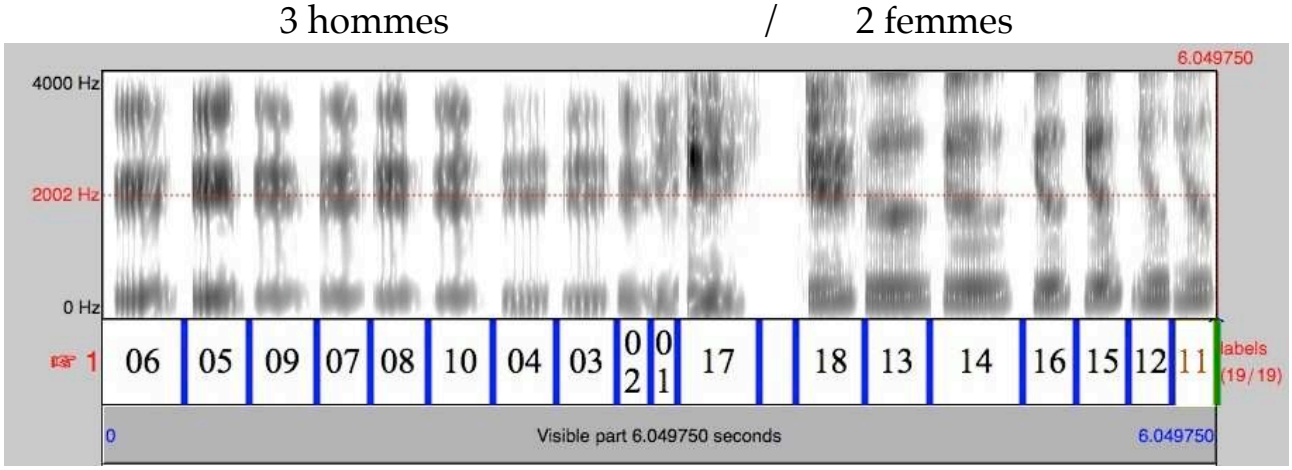



**Figure 84 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 18 stimuli de /y/ prononcés par 5 apprenants japonophones (ceux de 3 hommes à gauche, ceux de 2 femmes à droite).

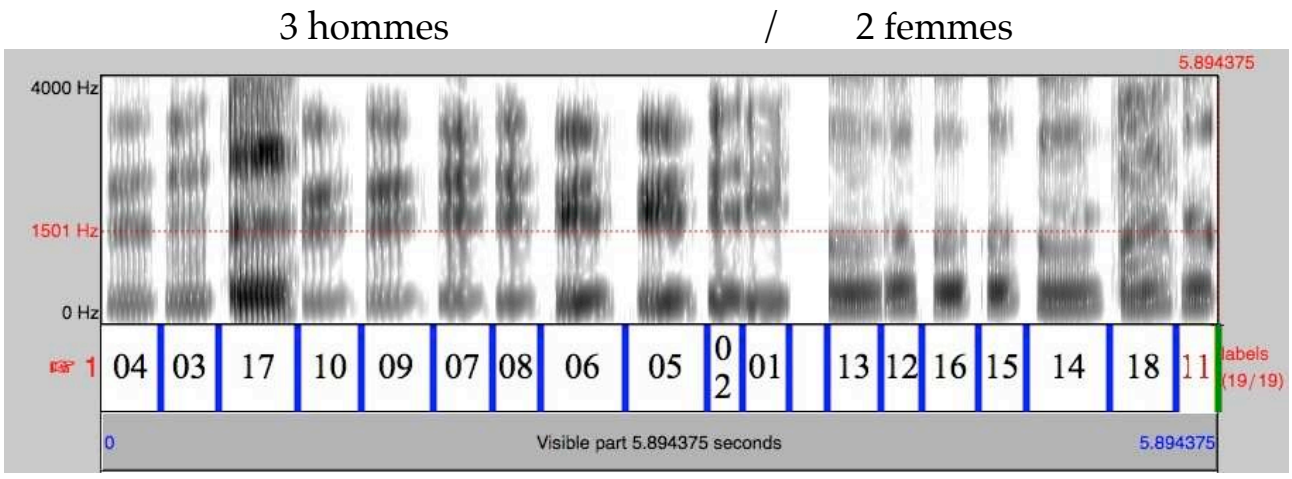



**Figure 85 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 18 stimuli du /ø/ prononcés par 5 apprenants japonophones (ceux de 3 hommes à gauche, ceux de 2 femmes à droite).

Les stimuli ont été disposés dans 4 listes différentes dans un ordre semi-aléatoire de sorte que la même voyelle cible ne se succède pas, et qu'il n'y ait pas de séquence de 4 stimuli ou plus prononcés par des apprenants du même sexe.

### *3.3.2. Les auditeurs*

Les 16 auditeurs qui ont participé à l'expérience étaient tous des locuteurs natifs du français résidant dans la région parisienne. Ils étaient âgés de 23 à 34 ans.

### *3.3.3. Les procédures*

Les tâches des auditeurs consistaient à écouter les stimuli, précédés par un bip sonore (onde sinusoïdal : 377 Hz, 50 millisecondes de longueur, placé environ 1 seconde avant chaque stimulus), et à indiquer quelle voyelle du français ils ont entendue, et à indiquer si le vocoïde était un plus ou moins bon exemplaire de la voyelle qu'ils ont choisie sur une échelle de 1 (mauvais) à 5 (bon). Ils étaient informés qu'ils allaient entendre des voyelles du français prononcées par des étudiants étrangers, et qu'ils n'entendraient pas nécessairement toutes les voyelles. Pour effectuer les tâches, ils disposaient d'un casque à travers lequel les stimuli ont été diffusés, un écran d'ordinateur et d'une souris. Ils ont donné la réponse en cliquant sur les cases correspondant à leur choix (la voyelle et la note d'évaluation). Ils devaient cliquer sur le bouton « ok » pour finaliser leur réponse afin de passer au stimulus suivant. Ils ne pouvaient pas réécouter le stimulus, mais ils pouvaient changer leur réponse, en cas de mauvaise manipulation par exemple, avant de la finaliser avec le bouton « ok ». L'écran présenté aux auditeurs a été reproduit à la Figure 86 : la présentation des stimuli et des cases de réponse, ainsi que l'enregistrement des réponses a été effectuée en utilisant Praat (« *ExperimentMFC 4* »).

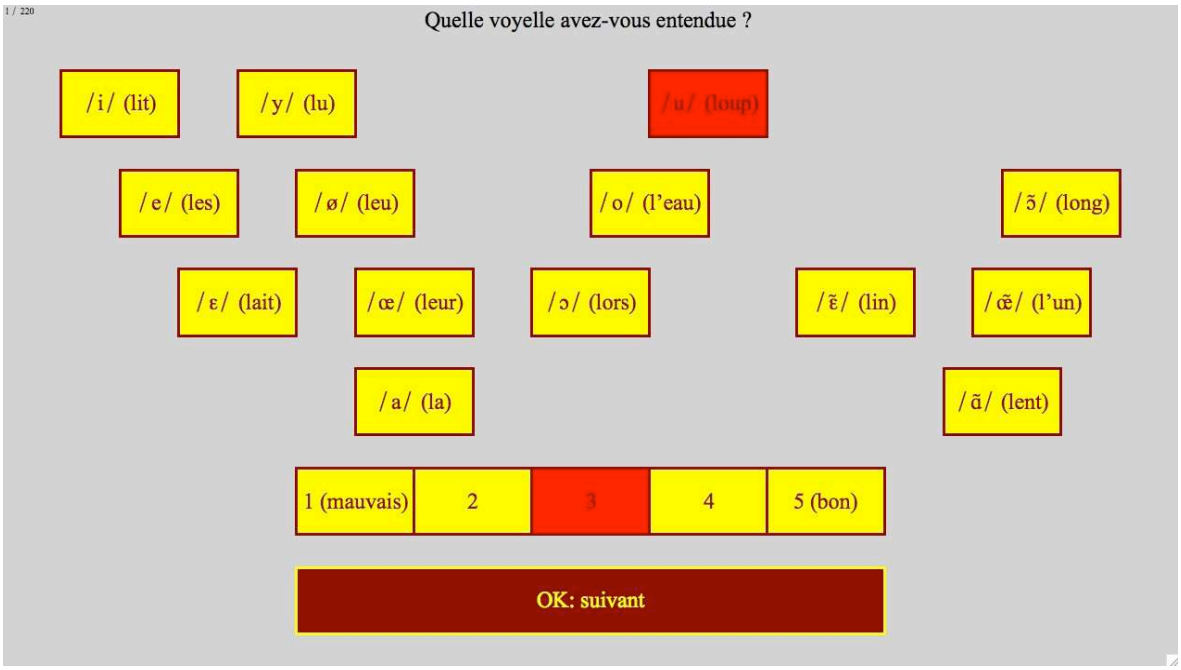



**Figure 86 :** Copie de l'écran présenté aux sujets-auditeurs pour la présente expérience. Les cases foncées indiquent les réponses sélectionnées (la voyelle /u/ « loup » et la note de 3 sur 5).

Les sujets-auditeurs ont commencé par une session d'entraînement, pendant laquelle ils ont entendu 6 stimuli : les trois voyelles cibles /u y ø/ prononcées par un locuteur natif et par une

apprenante japonophone. Les stimuli de ces deux locuteurs n'allaient pas réapparaître dans le test qui suivait.

Ils sont passés par la suite au test, qui comportait les 4 listes de stimuli décrites ci-dessus (3.3.1. Les stimuli). Ils ont ainsi entendu les 54 stimuli 4 fois, ce qui a donné lieu à 216 réponses par sujet. Après la deuxième liste, une pause a été proposée aux auditeurs. Ils ont entendu des stimuli *dummy* (stimuli non pris en compte pour les analyses) au début du test, avant et après la pause, et à la fin du test. Les stimuli ont été présentés dans l'ordre inverse à la moitié des sujets (8 personnes).

### *3.3.4. Les analyses des réponses*

Les réponses enregistrées ont été traitées de manière suivante (*cf.* Guion *et al.* 2000[409]). Premièrement, le taux d'identification correcte (c'est-à-dire, la correspondance entre la voyelle que les apprenants devaient prononcer, et celle que les auditeurs natifs ont perçue) a été calculé pour chaque stimulus des 3 voyelles cibles. Ensuite, le score, qui tient compte de l'identification et de l'évaluation, a été calculé pour chaque auditeur et chaque stimulus en multipliant la note d'évaluation par la proportion de réponses. Ainsi, un stimulus qui a été identifié 3 fois comme /ø/ et 1 fois comme /u/ par un auditeur donné, avec des notes de 4, 3, 4, et 2, respectivement, a obtenu des notes de 2,75 (= (4 + 3 + 4) / 4) pour le /ø/, et de 0,5 (= 2 / 4) pour le /u/ concernant cet auditeur. Ensuite, la moyenne des notes de tous les auditeurs a été calculée pour chaque stimulus.

### *3.3.5. Les résultats*

#### 3.3.5.1. Les stimuli /u/

La Figure 87 représente la proportion d'identification des stimuli /u/.

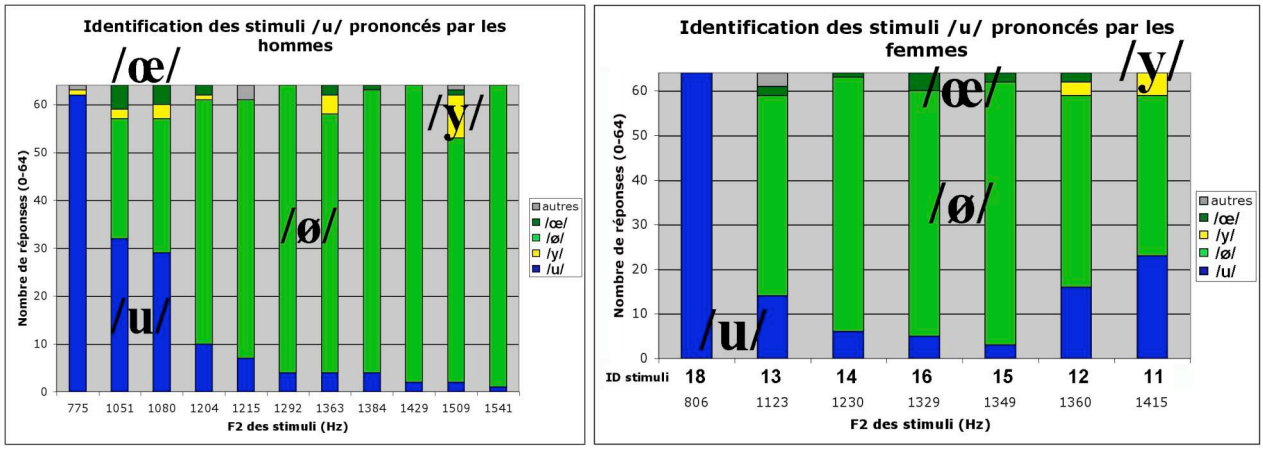



**Figure 87 :** Nombre de réponses d'identification (0-64) des stimuli /u/ prononcés par les 5 apprenants japonais (3 hommes à gauche, et 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de

[409] S. G. Guion, J. E. Flege, R. Akahane-Yamada, J. C. Pruitt. 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, 2000.

F2. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Les deux stimuli, qui ont un F2 inférieur à 1000 Hz (u-17 à 775 Hz et u-18 à 806 Hz), ont été reconnus parfaitement bien (ou presque, pour le premier : 62 sur 64) comme /u/. En revanche, les deux avec un F2 entre 1000 et 1100 Hz (u-01 à 1051 Hz et u-02 à 1080 Hz) ont été identifiés presque aussi souvent comme /ø/ (25 et 28 réponses sur 64, respectivement) que comme /u/ (32 et 29 réponses sur 64, respectivement). Tous les autres ont été perçus majoritairement comme /ø/ (36 – 63 réponses sur 64).

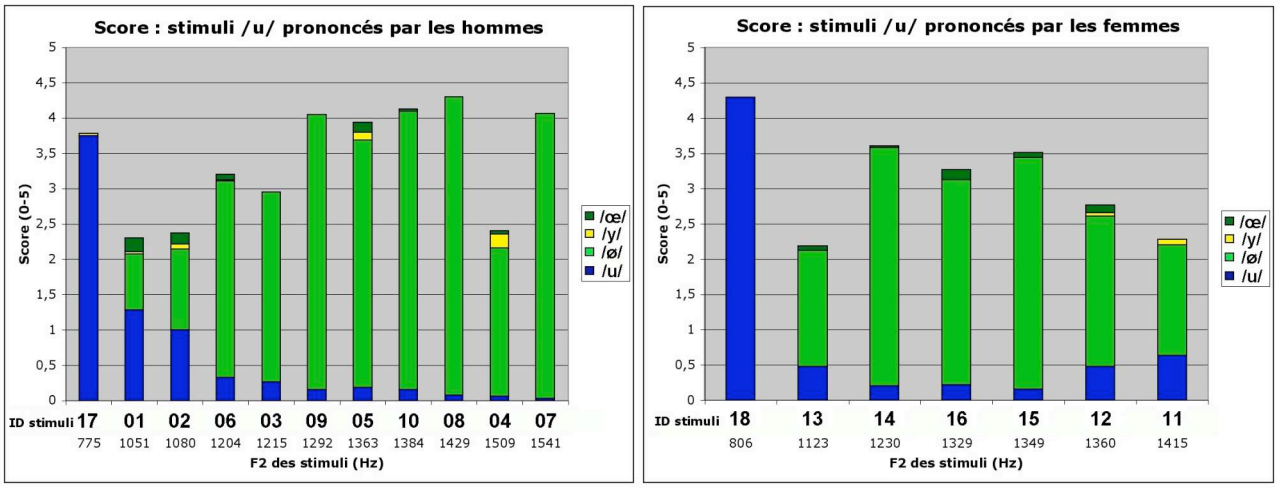



**Figure 88 :** Score calculé sur le nombre de réponses d'identification et la note d'évaluation. Le /u/ prononcé par les 5 apprenants (3 hommes à gauche, 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de F2. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Le score calculé sur le nombre de réponses d'identification et la note d'évaluation montre une tendance similaire (Figure 88). Le score pour /ø/ est globalement plus élevé pour les stimuli qui ont un F2 élevé (notons que le F2 de ces stimuli ne dépasse pas 1600 Hz). Le stimulus u-11 (F2 à 1415 Hz) a obtenu un score bas pour /ø/ (1,6), probablement à cause d'un F1 relativement bas (376 Hz) par rapport aux autres stimuli prononcés par la même apprenante (409 – 468 Hz : Tableau 55).

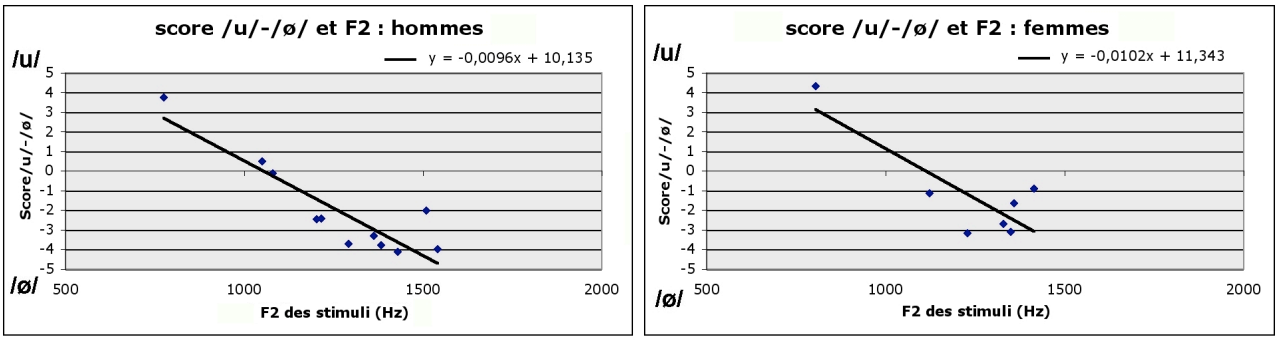



**Figure 89 :** Valeur « score pour la réponse /u/ - score pour la réponse /ø/ » (en ordonnée : Figure 88) et F2 des stimuli (ceux qui ont été prononcés par les 3 hommes à gauche, par les 2 femmes à droite : en abscisse). 64 réponses par stimulus (16 auditeurs x 4 répétitions).

La Figure 89 illustre le rapport entre le F2 des stimuli et les scores obtenus pour /u/ et /ø/. La valeur « score pour /u/ - score pour /ø/ », qui permet de situer les stimuli sur un continuum qui varie entre un très bon exemplaire du /u/ (c'est-à-dire, 5) et un très bon exemplaire du /ø/ (c'est-à-dire, -5), et celle de F2 montrent une bonne corrélation : 0,89 pour les stimuli des hommes, et 0,83 pour ceux des femmes (coefficient de corrélation de Pearson).

### 3.3.5.2. Les stimuli /y/

Le nombre de réponses d'identification (0-64) aux stimuli /y/ est représenté à la Figure 90.

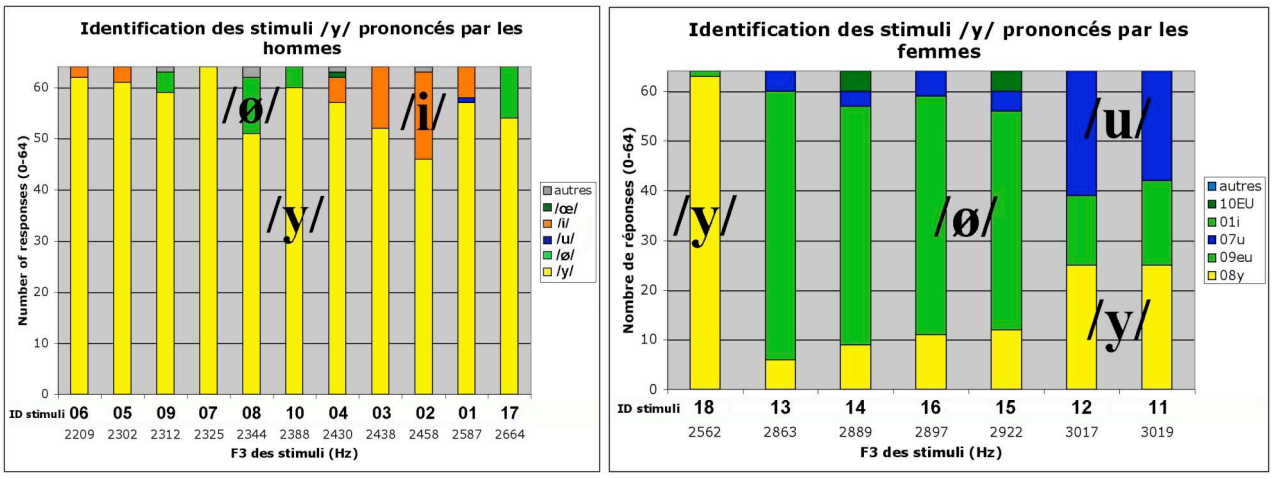



**Figure 90 :** Nombre de réponses d'identification (0-64) des stimuli /y/ prononcés par les 5 apprenants japonais (3 hommes à gauche, et 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de F3. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Tous les stimuli des hommes, qui avaient des F2/F3 proches, ont été identifiés majoritairement comme /y/ (de 46 réponses sur 64 pour le stimulus y-02, à 64 sur 64 pour y-07). Ceux qui ont été prononcés par l'apprenant 1 (sp1 : de y-01 à y-04) ont été perçus comme /i/ dans 5 (y-04) à 17 (y-02) fois sur 64. Ces stimuli ont un F2 relativement élevé (de 1946 Hz pour y-03 à 2088 Hz pour y-01), ce qui suggère que les auditeurs francophones entendent /i/ plutôt que /y/ si le F2 est trop élevé, même si les F2 et F3 sont relativement proches.

Concernant les stimuli prononcés par les femmes, ceux de l'apprenante 3 (sp 3) n'ont pas été reconnus majoritairement comme /y/, à la différence de y-18, qui a été prononcé par l'autre apprenante (sp 5) avec un rapprochement F2/F3. Ils ont été identifiés soit majoritairement comme /ø/ (les stimuli y-13, 14, 15, 16), soit comme /u/, /y/ et /ø/ de façon équilibrée (y-11 et y-12).

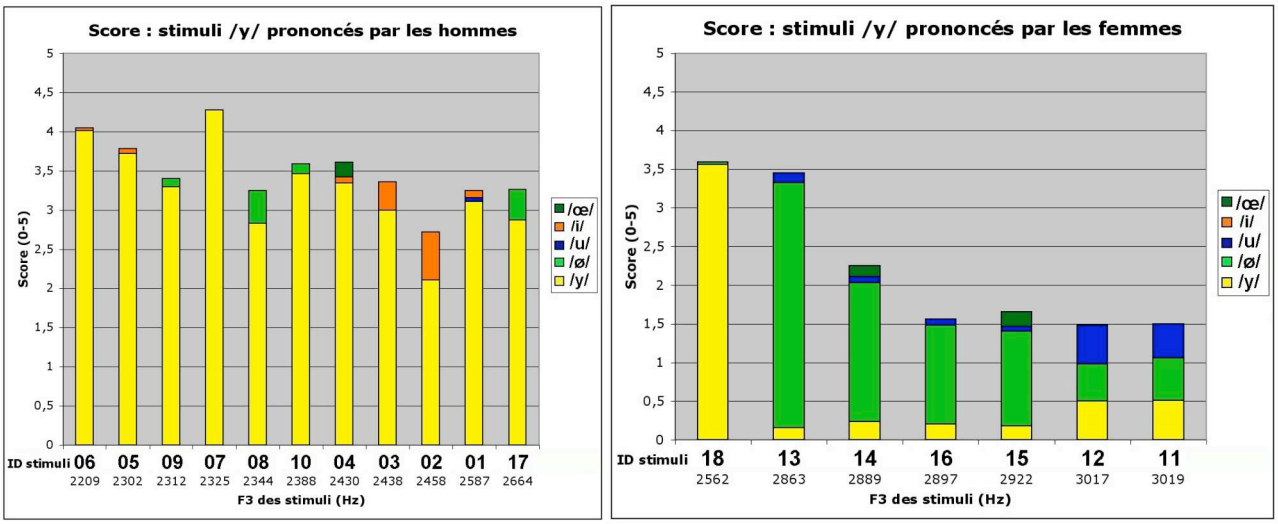



**Figure 91 :** Score calculé sur le nombre de réponses d'identification et la note d'évaluation. Le /y/ prononcé par les 5 apprenants (3 hommes à gauche, 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de F3. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Les scores présentent grosso modo des tendances similaires (Figure 91). Mais les stimuli y-13, 14, 16, et 15, qui ont été perçus tous majoritairement comme /ø/ (54, 48, 48, et 44 sur 64 réponses, respectivement), ont obtenu des scores différents pour /ø/ (3,2, 1,8, 1,3, et 1,2). Les y-11 et 12, qui étaient clairement diphtongués (comme /ju/ [jɯ] japonais), ont obtenu des scores également bas pour /u/, /y/, et /ø/ (y-11 : 0,5 pour /u/, /y/, et /ø/. y-12 : 0,4 pour /u/, 0,5 pour /y/ et /ø/). Il est possible que les auditeurs aient entendu plusieurs timbres vocaliques durant la séquence, et aient hésité entre plusieurs voyelles (comme certains d'entre eux l'ont dit après l'expérience), mais aient jugé que le stimulus ne correspondait pas à un bon exemplaire de la voyelle choisie. Notons que y-13, qui n'était pas diphtongué, a obtenu un score pour /ø/ le plus élevé (3,2, sur 5).

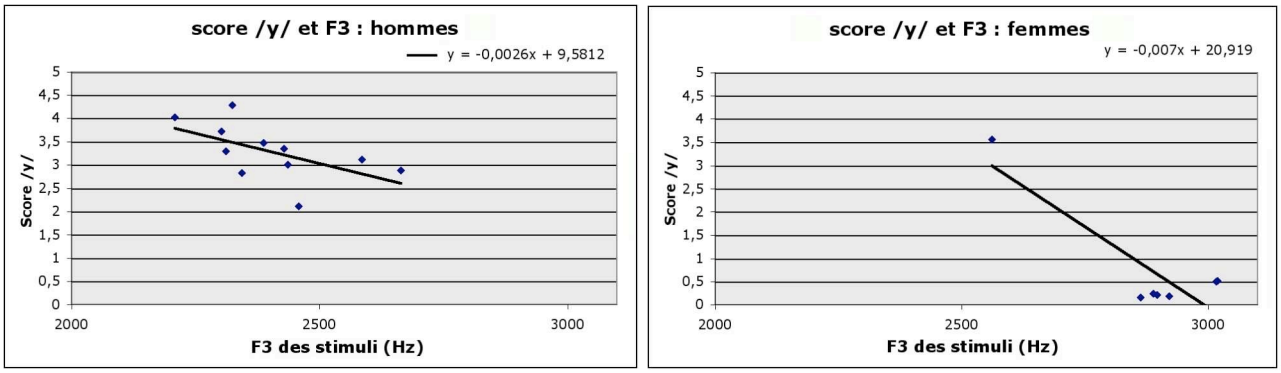



**Figure 92 :** Score pour la réponse /y/ (en ordonnée : Figure 91) et F3 des stimuli (ceux qui ont été prononcés par les 3 hommes à gauche, par les 2 femmes à droite : en abscisse). 64 réponses par stimulus (16 auditeurs x 4 répétitions).

La Figure 92 représente le score pour /y/ et le F3 des stimuli. Nous y observons la tendance suivante : plus le F3 est bas, plus le score pour le /y/ est élevé (notons que le F3 le plus bas observé ici est de 2209 Hz). Ces deux variables présentent une bonne corrélation, notamment pour les stimuli des femmes : 0,57 pour les stimuli des hommes, 0,86 pour ceux des femmes (coefficient de corrélation de Pearson).

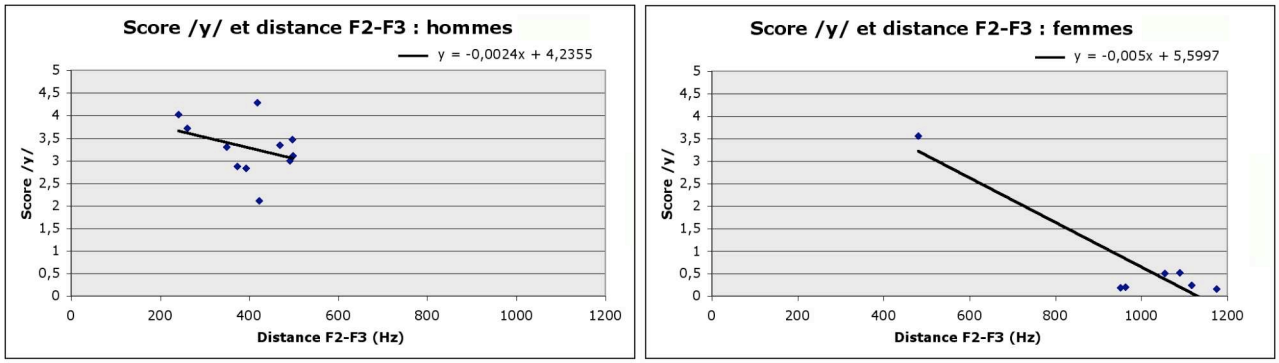



**Figure 93** : Score pour la réponse /y/ (en ordonnée : Figure 91) et distance entre F2 et F3 des stimuli (ceux qui ont été prononcés par les 3 hommes à gauche, par les 2 femmes à droite : en abscisse). 64 réponses par stimulus (16 auditeurs x 4 répétitions).

La Figure 93 montre le score pour /y/ et la distance entre F2 et F3. Le score est en général plus élevé pour les stimuli avec F2 et F3 proches. Ces deux variables présentent une corrélation de 0,36 pour les stimuli des hommes, et de 0,93 pour ceux des femmes (coefficient de corrélation de Pearson).

#### 3.3.5.3. Les stimuli /ø/

Le nombre de réponses d'identification des stimuli /ø/ est représenté à la Figure 94.

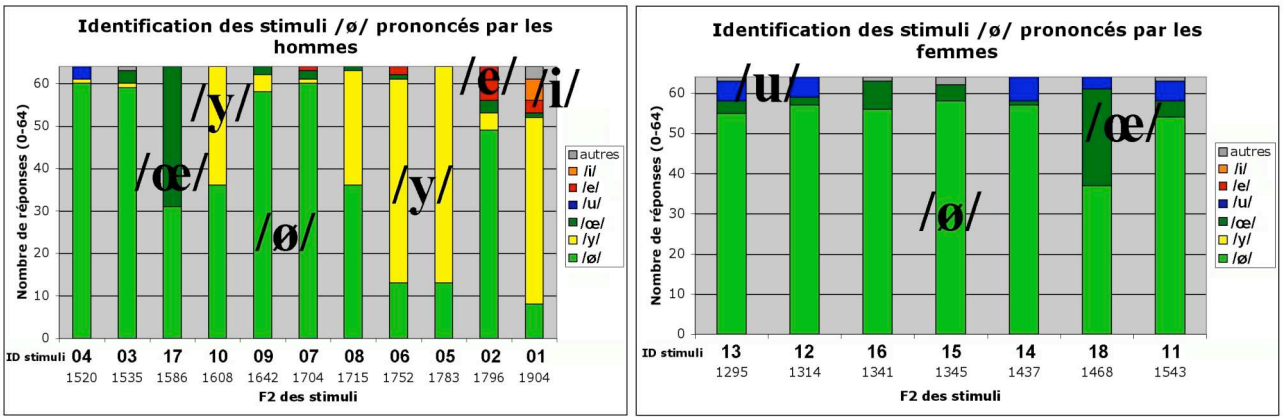



**Figure 94 :** Nombre de réponses d'identification (0-64) des stimuli /ø/ prononcés par les 5 apprenants japonais (3 hommes à gauche, et 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de F2. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Parmi les stimuli prononcés par les 3 hommes, 3 ont été reconnus majoritairement comme /y/ (44 réponses sur 64 pour ø-01, 51 sur 64 pour ø-05, et 48 sur 64 pour ø-06), et 2 autres dans presque la moitié des cas (27 sur 64 pour ø-08, et 28 sur 64 pour ø-10). Ces 5 stimuli ont été prononcés avec F2/F3 proches : 1904 et 2411 Hz pour ø-01, 1783 et 2144 Hz pour ø-05, 1752 et 2144 Hz pour ø-06, 1715 et 2321 Hz pour ø-08, et 1608 et 2093 Hz pour ø-10. Tous les autres stimuli des hommes ont été identifiés majoritairement comme /ø/ (entre 49 et 60 réponses sur 64), à l'exception de ø-17, qui a été classifié 31 fois comme /ø/ et 33 fois comme /œ/. Le F1 de ce dernier (449 Hz) est plus élevé que celui des autres stimuli des hommes (entre 311 et 382 Hz), ce qui aurait favorisé la perception du /œ/ plutôt que du /ø/. Le F2 des stimuli des hommes qui ont été identifiés majoritairement comme /ø/ variait de 1520 à 1704 Hz, à l'exception de y-02 (1796 Hz), qui a été classifié comme /e/, ne serait-ce que minoritairement (8 réponses sur 64).

Quant aux stimuli prononcés par les femmes, ils ont été reconnus comme /ø/ dans la majorité des cas (entre 54 et 58 réponses sur 64, sauf le ø-18 : 37 réponses pour /ø/ et 24 réponses pour /œ/). Le F2 de ces stimuli varie de 1295 à 1543 Hz.

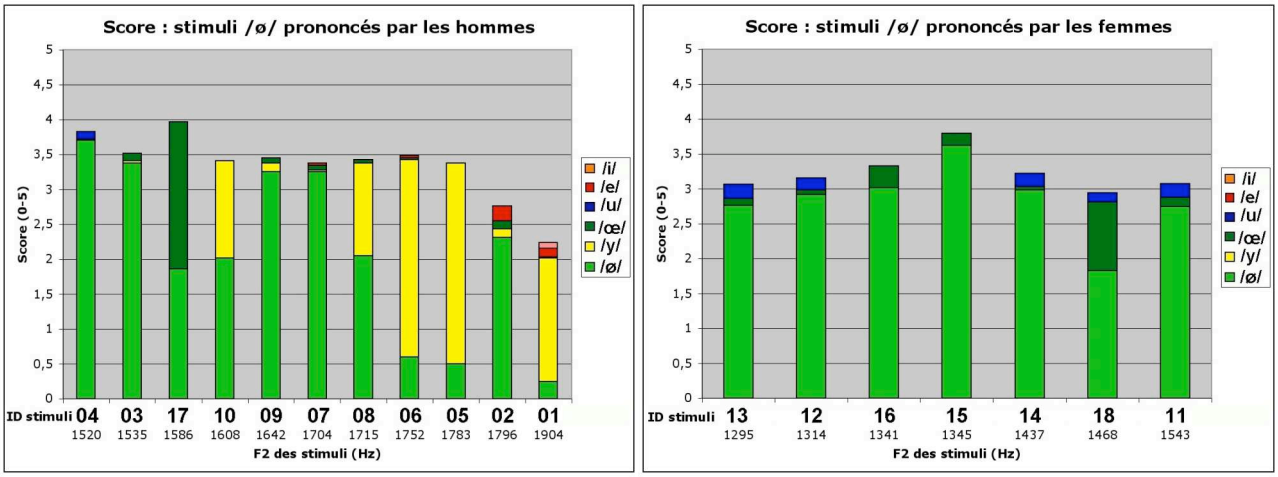



**Figure 95 :** Score calculé sur le nombre de réponses d'identification et la note d'évaluation. Le /ø/ prononcé par les 5 apprenants (3 hommes à gauche, 2 femmes à droite). 64 réponses par stimulus (16 auditeurs x 4 répétitions). Les stimuli sont ordonnés horizontalement par la valeur croissante de F2. Les identifiants des stimuli correspondent à ceux du Tableau 55 (p. 205) et de la Figure 81 (p. 206).

Le score présenté à la Figure 95 illustre une tendance similaire, mais les stimuli qui avaient un F2 pas très éloigné de 1500 Hz environ ont obtenu de bonnes notes par rapport aux autres. Le stimulus ø-04 (F2 à 1520 Hz), par exemple, a été identifié comme /ø/ dans 60 réponses sur 64, et a obtenu une note de 3,7. En revanche, les stimuli ø-09 et ø-07 (F2 à 1642 et 1704 Hz) ont été catégorisés également comme /ø/ dans 58 et 60 cas sur 64 respectivement, mais leur score pour le /ø/ ne s'élève qu'à 3,25.

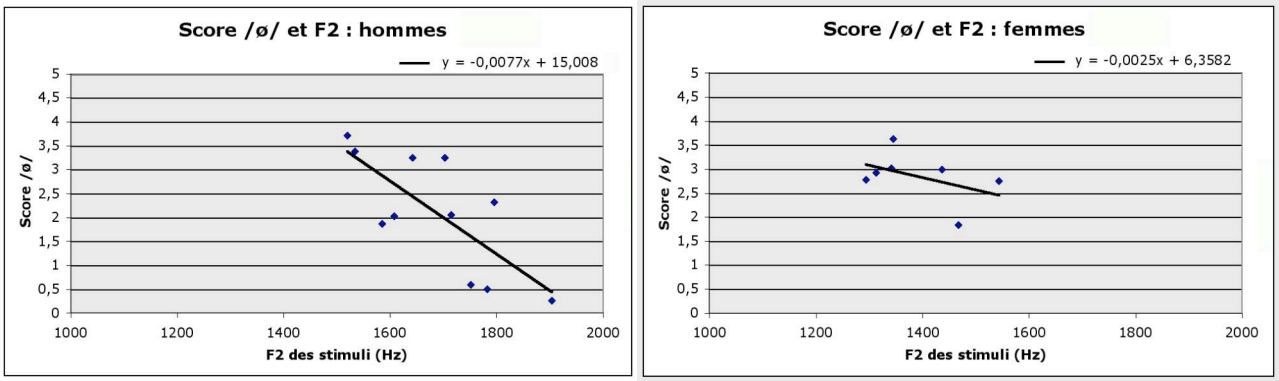



**Figure 96 :** Score pour la réponse /ø/ (en ordonnée : Figure 95) et F2 des stimuli (ceux qui ont été prononcés par les 3 hommes à gauche, par les 2 femmes à droite : en abscisse). 64 réponses par stimulus (16 auditeurs x 4 répétitions).

La Figure 96 illustre le rapport entre le score et le F2 des stimuli. En ce qui concerne les stimuli des hommes, les deux variables montrent une bonne corrélation (0,74), le F2 étant supérieur à 1500 Hz pour tous les stimuli. En revanche, ce n'est pas le cas pour les stimuli des femmes (coefficient de corrélation de Pearson : 0,44).

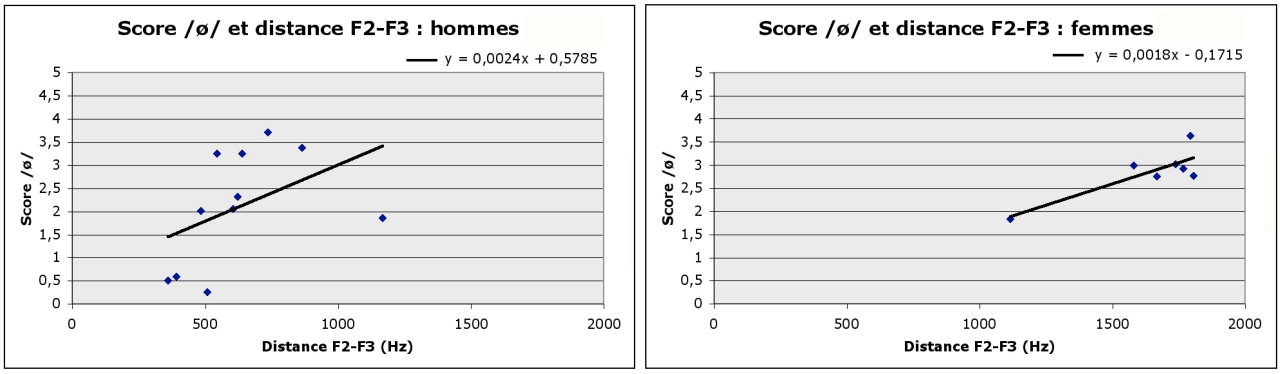



**Figure 97** : Score pour la réponse /ø/ (en ordonnée : Figure 95) et F2 des stimuli (ceux qui ont été prononcés par les 3 hommes à gauche, par les 2 femmes à droite : en abscisse). 64 réponses par stimulus (16 auditeurs x 4 répétitions).

La Figure 97 montre le score pour /ø/ et la distance entre F2 et F3. Le score est globalement plus élevé pour les stimuli avec F2 et F3 écartés (sans rapprochement F2/F3). Les deux variables présentent une corrélation de 0,45 pour les stimuli des hommes, de 0,84 pour ceux des femmes (coefficient de corrélation de Pearson).

Les résultats de cette expérience ont montré les suivants :

1) /u/ : les stimuli avec un F2 inférieur à 1000 Hz ont été identifiés presque unanimement comme /u/, avec une bonne note d'évaluation. Ceux qui avaient un F2 entre 1000 et 1100 Hz ont été classifiés presque aussi souvent comme /ø/ que comme /u/. Avec un F2 supérieur à 1100 Hz, les stimuli ont été perçus majoritairement comme /ø/.

2) /y/ : les stimuli avec F2 et F3 proches, avec un F2 pas trop élevé, ont été reconnus majoritairement comme /y/. Les réalisations diphtonguées, comme /ju/ [jɯ] japonais, ont été identifiées majoritairement comme /ø/, ou également comme /u/, /y/, et /ø/, mais avec une mauvaise note dans tous les cas.

3) /ø/ : les stimuli avec un F2 autour de 1500 Hz environ ont été classifiés majoritairement comme /ø/ avec une bonne note.

Nous retrouvons ainsi les caractéristiques acoustiques des cibles des locuteurs natifs, comme nous les avons constatées et observées au début de la deuxième partie de cette étude (Expérience 2 : Figure 41, p. 123).

Il nous convient maintenant d'étudier le rapport entre l'acoustique (c'est-à-dire, les valeurs formantiques des trois voyelles /u/, /y/, et /ø/) et la perception auprès des auditeurs natifs d'une manière plus détaillée. C'est ainsi qu'une autre expérience de perception a été effectuée en utilisant des stimuli synthétisés.

## 3.4. PERCEPTION DE VOYELLES SYNTHÉTISÉES (SYNTHÈSE ARTICULATOIRE ET SYNTHÈSE À FORMANTS) AUPRÈS D'AUDITEURS NATIFS FRANCOPHONES ET JAPONOPHONES (EXPÉRIENCE 9)

Nous présenterons dans ce chapitre une autre expérience de perception effectuée en utilisant des stimuli synthétisés, afin d'éclaircir mieux le rapport entre l'articulation (synthèse articulatoire) et l'acoustique (synthèse à formants), d'une part, et la perception des auditeurs francophones et japonophones natifs, de l'autre, concernant les voyelles arrondies fermées /u y/ et mi-fermée /ø/.

### *3.4.1. Les stimuli*

Deux types de stimuli synthétisés ont été utilisés dans la présente expérience, créés par synthèse articulatoire (VTCalc) et par synthèse à formants (synthétiseur de Klatt).

#### 3.4.1.1. Synthèse articulatoire (VTCalc)

Le logiciel de synthèse articulatoire VTCalc calcule les caractéristiques acoustiques de l'ensemble d'un conduit vocal, de la glotte jusqu'aux lèvres, spécifié par les 7 paramètres suivants (voir Maeda 1982[410] pour les détails) :

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-*

Les fichiers son de sortie représentent une voyelle de 340 millisecondes environ, avec les caractéristiques spectrales calculées par le logiciel à partir des 7 paramètres articulatoires. La fréquence fondamentale (F0) du début est de 120 Hz environ, et atteint 130 Hz environ à la fin du premier tiers, et descend à 100 Hz environ à la fin.

En utilisant ce synthétiseur, nous avons généré trois séries de continua de stimuli, qui partent tous des paramètres articulatoires d'un [u] français.

**1) Série 1 (/u/ -> /ø/ -> )**

Nous avons choisi, comme point de départ, les paramètres articulatoires d'un [u] français : 0,5 cm pour « *jaw* », 2,0 cm pour « *tongue* », 1,5 cm pour « *shape* », -2,0 cm pour « *apex* », -1,0 cm pour « *lip_ht* », 1,5 cm pour « *lip_pr* », et 0 cm pour « *larynx* ». Ces paramètres induisent les valeurs de formant suivantes : 284 Hz pour F1, 882 Hz pour F2, 2301 Hz pour F3, et 3593 Hz pour F4 (Tableau 56 : première ligne). Nous avons nommé ce stimulus « ST0 ».

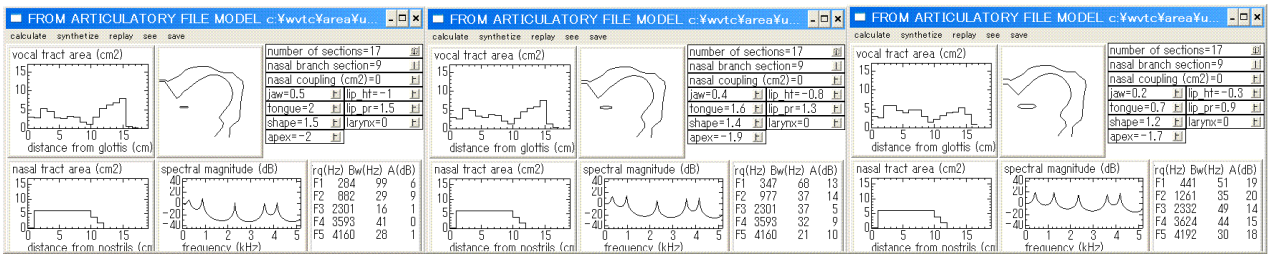
[410] S. Maeda. A digital simulation method of the vocal tract system. *Speech Communication* 1(3-4) : 199-229, 1982.

Ensuite, nous avons pris les paramètres correspondant à un [ø] français : 0 cm pour « *jaw* », -0,2 cm pour « *tongue* », 1 cm pour « *shape* », -1,5 cm pour « *apex* », 0,25 cm pour « *lip_ht* », 0,5 cm pour « *lip_pr* », et 0 cm pour « *larynx* ». Nous l'avons légèrement modifié (0,25 cm -> 0,3 cm pour « *lip_ht* » afin d'obtenir une valeur que l'on pouvait entrer manuellement) pour créer « ST5 ». Nous avons par la suite pris 4 valeurs intermédiaires de chacun des 7 paramètres articulatoires pour créer « ST1 » - « ST4 ». Pour terminer, les stimuli « ST6 » - « ST8 » ont été créés sur le même continuum, en variant les paramètres de même façon. Ce continuum permet de simuler les mouvements de délabialisation (ouverture et absence progressive de protrusion des lèvres) et l'avancement de la langue en même temps, qui caractériseraient le passage d'un [u] à un [ø], puis plus loin vers un [e]. Ces stimuli sont représentés dans le Tableau 56, la Figure 98 et la Figure 100.

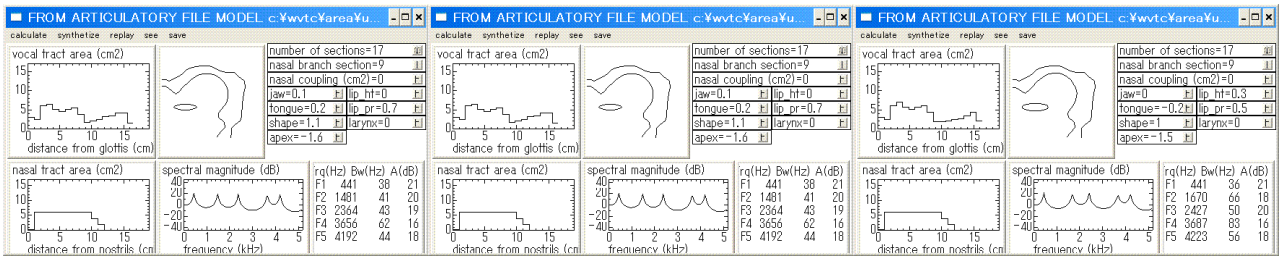
**Tableau 56** : Sept paramètres articulatoires d'entrée (à gauche) et les formants (en Hertz), la largeur de bande et l'amplitude des formants des stimuli de la série 1 (de [u] à [ø] et plus loin) calculés par VTCalc.

Série 1 [u] à [ø], puis plus loin

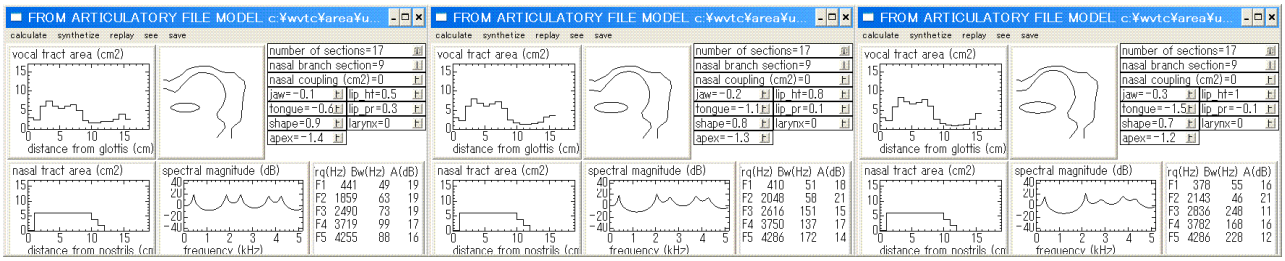
| | jaw | tongue | shape | apex | lip_ht | lip_pr | larynx | F1 | F2 | F3 | F4 | F5 | Largeur de la bande (Hz) | | | | | Amplitude (dB) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| st0 | 0,5 | 2,0 | 1,5 | -2,0 | -1,0 | 1,5 | 0,0 | 284 | 882 | 2301 | 3593 | 4160 | 99 | 29 | 16 | 41 | 28 | 6 | 9 | 1 | 0 | 1 |
| st1 | 0,4 | 1,6 | 1,4 | -1,9 | -0,8 | 1,3 | 0,0 | 340 | 960 | 2321 | 3590 | 4148 | 68 | 37 | 37 | 32 | 21 | 13 | 14 | 5 | 9 | 10 |
| st2 | 0,3 | 1,1 | 1,3 | -1,8 | -0,5 | 1,1 | 0,0 | 402 | 1114 | 2321 | 3590 | 4179 | 46 | 35 | 17 | 32 | 45 | 18 | 18 | 17 | 14 | 12 |
| st3 | 0,2 | 0,7 | 1,2 | -1,7 | -0,3 | 0,9 | 0,0 | 433 | 1269 | 2352 | 3621 | 4179 | 51 | 35 | 49 | 44 | 30 | 19 | 20 | 14 | 15 | 18 |
| st4 | 0,1 | 0,2 | 1,1 | -1,6 | 0,0 | 0,7 | 0,0 | 433 | 1486 | 2383 | 3652 | 4210 | 38 | 41 | 43 | 62 | 44 | 21 | 20 | 19 | 16 | 18 |
| st5 | 0,0 | -0,2 | 1,0 | -1,5 | 0,3 | 0,5 | 0,0 | 433 | 1671 | 2414 | 3683 | 4210 | 36 | 66 | 50 | 83 | 56 | 21 | 18 | 20 | 16 | 18 |
| st6 | -0,1 | -0,6 | 0,9 | -1,4 | 0,5 | 0,3 | 0,0 | 433 | 1857 | 2476 | 3714 | 4240 | 49 | 63 | 73 | 99 | 88 | 19 | 19 | 19 | 17 | 16 |
| st7 | -0,2 | -1,1 | 0,8 | -1,3 | 0,8 | 0,1 | 0,0 | 402 | 2043 | 2631 | 3745 | 4271 | 51 | 58 | 151 | 137 | 172 | 18 | 21 | 15 | 17 | 14 |
| st8 | -0,3 | -1,5 | 0,7 | -1,2 | 1,0 | -0,1 | 0,0 | 371 | 2167 | 2817 | 3776 | 4302 | 55 | 46 | 248 | 168 | 228 | 16 | 21 | 11 | 16 | 12 |



ST0 ST1 ST2



ST3 ST4 ST5



ST6 ST7 ST8

**Figure 98 :** Neuf stimuli (ST0 – ST8) de la série 1 ([u] français -> [ø] français -> plus loin), synthétisés en utilisant VTCalc, obtenus en modifiant la labialisation et l'avancement de la langue (Tableau 56).

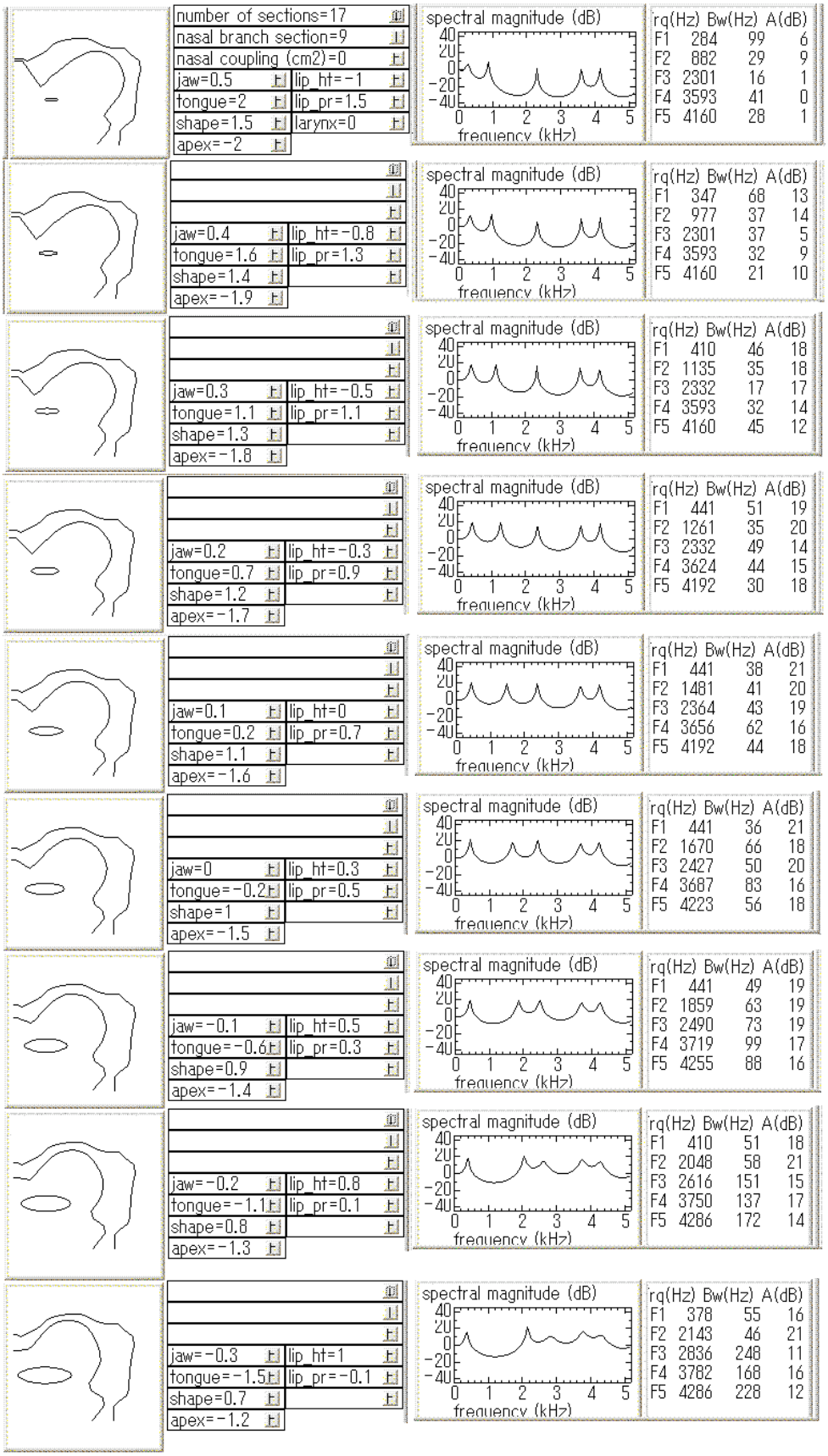



**Figure 99 :** Neuf stimuli (ST0 : en haut – ST8 : en bas) de la série 1 ([u] français -> [ø] français -> plus loin), synthétisés en utilisant VTCalc, obtenus en modifiant la labialisation et l'avancement de la langue (Figure 98).

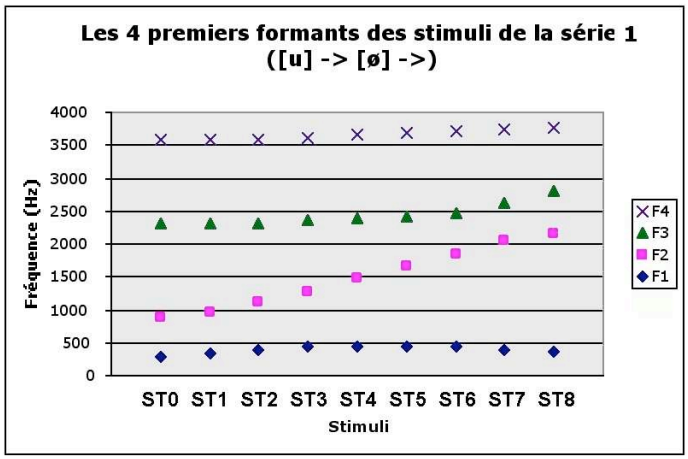



**Figure 100 :** Les quatre premiers formants des 9 stimuli de la série 1 ([u] -> [ø] ->) synthétisés en utilisant VTCalc, obtenus en modifiant la labialisation et l'avancement de la langue (Figure 98 et Figure 99).

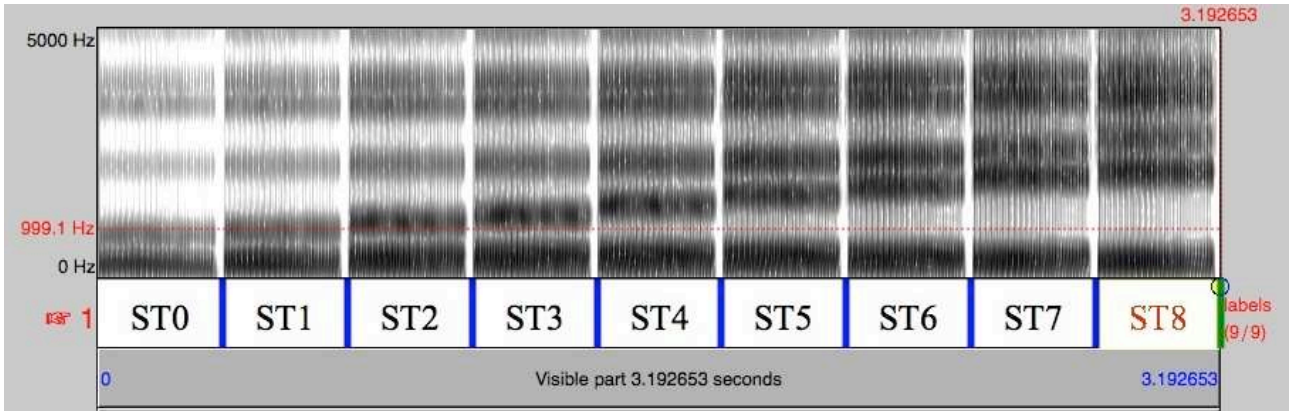


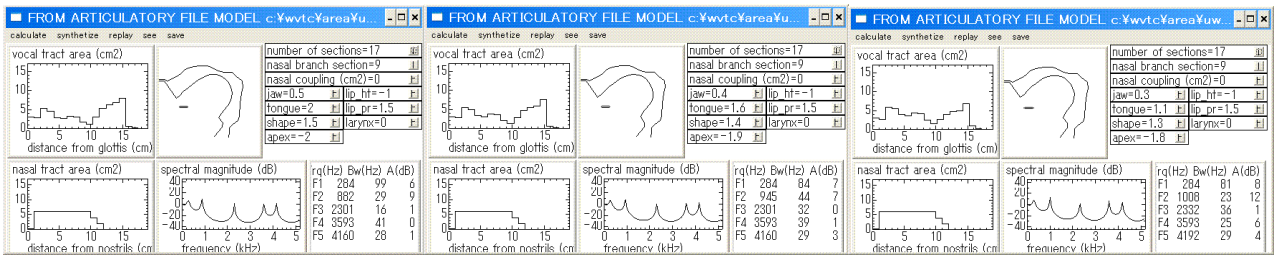

**Figure 101 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 9 stimuli de la série 1 ([u] français -> [ø] ->) synthétisés en utilisant VTCalc, correspondant aux Figure 88, Figure 89, Figure 90.

### 2) Série 2 (avancement de la langue)

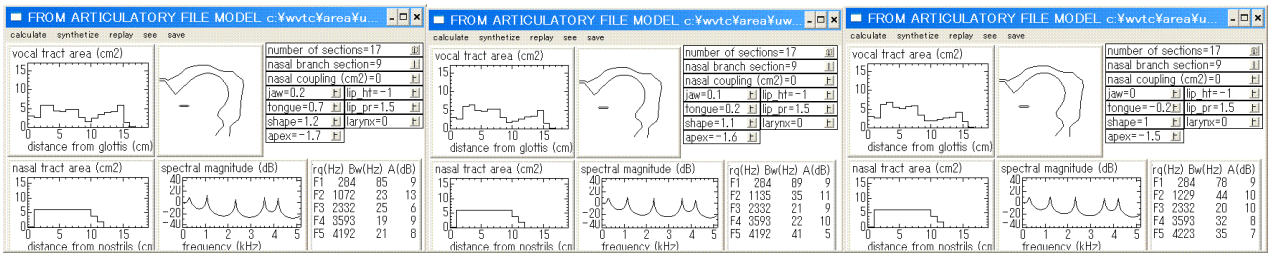
Pour la série 2, nous avons fixé les 2 paramètres des lèvres (la hauteur « *lip_ht* » et la protrusion « *lip_pr* » des lèvres) aux valeurs du stimulus ST0 (-1,0 cm et 1,5 cm respectivement), qui simule un [u] français, et fait varier les 4 paramètres de la mâchoire (« *jaw* ») et de la langue (« *tongue* », « *shape* », et « *apex* ») de la même façon que la série 1 (le larynx « *larynx* » reste inchangé à 0 cm). Ce continuum permettrait de simuler l'effet de l'avancement de la langue (et la baisse de la mâchoire), tout en conservant l'arrondissement et la protrusion des lèvres d'un [u] français. Les stimuli de cette série (ST0-ST08) sont représentés dans le Tableau 57, la Figure 102 et la Figure 104. Notons que l'avancement de la langue contribue à augmenter le F2 par rapport à une constriction vers le milieu du conduit vocal (région vélaire) quand les lèvres sont arrondies et protruses (voir le nomogramme de Fant 1960, reproduit à la Figure 12, p. 46).

**Tableau 57** : Sept paramètres articulatoires d'entrée (à gauche) et les formants (en Hertz), la largeur de bande et l'amplitude des formants des stimuli de la série 2 (avancement de la langue) calculés par VTCalc.

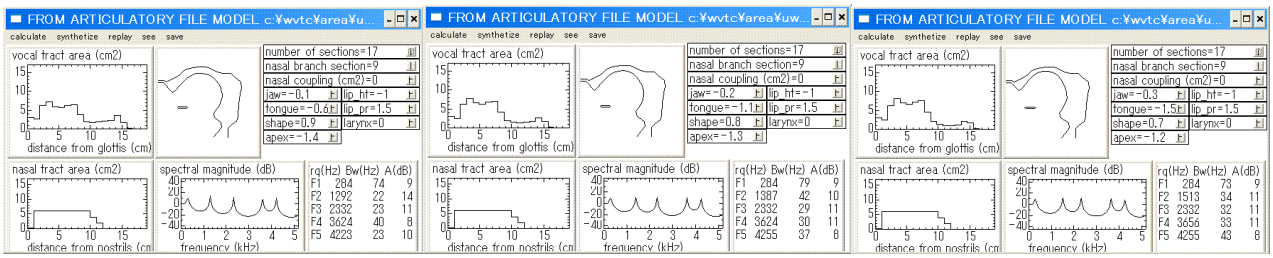
| Série 2 | avancement de langue | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | jaw | tongue | shape | apex | lip_ht | lip_pr | larynx | F1 | F2 | F3 | F4 | F5 | Largeur de la bande (Hz) | | | | | Amplitude (dB) | | | | | | | | | |
| st0 | 0,5 | 2,0 | 1,5 | -2,0 | -1,0 | 1,5 | 0,0 | 284 | 882 | 2301 | 3593 | 4160 | 99 | 29 | 16 | 41 | 28 | 6 | 9 | 1 | 0 | 1 | | | | | |
| st01 | 0,4 | 1,6 | 1,4 | -1,9 | -1,0 | 1,5 | 0,0 | 284 | 945 | 2301 | 3593 | 4160 | 84 | 44 | 32 | 39 | 29 | 7 | 7 | 0 | 1 | 3 | | | | | |
| st02 | 0,3 | 1,1 | 1,3 | -1,8 | -1,0 | 1,5 | 0,0 | 284 | 1008 | 2332 | 3593 | 4192 | 81 | 23 | 36 | 25 | 29 | 8 | 12 | 1 | 6 | 4 | | | | | |
| st03 | 0,2 | 0,7 | 1,2 | -1,7 | -1,0 | 1,5 | 0,0 | 284 | 1072 | 2332 | 3593 | 4192 | 85 | 23 | 25 | 19 | 21 | 9 | 13 | 6 | 9 | 8 | | | | | |
| st04 | 0,1 | 0,2 | 1,1 | -1,6 | -1,0 | 1,5 | 0,0 | 284 | 1135 | 2332 | 3593 | 4192 | 89 | 35 | 21 | 22 | 41 | 9 | 11 | 9 | 10 | 5 | | | | | |
| st05 | 0,0 | -0,2 | 1,0 | -1,5 | -1,0 | 1,5 | 0,0 | 284 | 1229 | 2332 | 3593 | 4223 | 78 | 44 | 20 | 32 | 35 | 9 | 10 | 10 | 8 | 7 | | | | | |
| st06 | -0,1 | -0,6 | 0,9 | -1,4 | -1,0 | 1,5 | 0,0 | 284 | 1292 | 2332 | 3624 | 4223 | 74 | 22 | 23 | 40 | 23 | 9 | 14 | 11 | 8 | 10 | | | | | |
| st07 | -0,2 | -1,1 | 0,8 | -1,3 | -1,0 | 1,5 | 0,0 | 284 | 1387 | 2332 | 3624 | 4255 | 79 | 42 | 29 | 30 | 37 | 9 | 10 | 11 | 11 | 8 | | | | | |
| st08 | -0,3 | -1,5 | 0,7 | -1,2 | -1,0 | 1,5 | 0,0 | 284 | 1513 | 2332 | 3656 | 4255 | 73 | 34 | 32 | 33 | 43 | 9 | 11 | 11 | 11 | 8 | | | | | |



ST0 ST01 ST02



ST03 ST04 ST05



ST06 ST07 ST08

**Figure 102 :** Neuf stimuli (ST0 – ST08) de la série 2 (avancement de la langue), synthétisés en utilisant VTCalc, obtenus en simulant l'avancement progressif de la langue (Tableau 57).

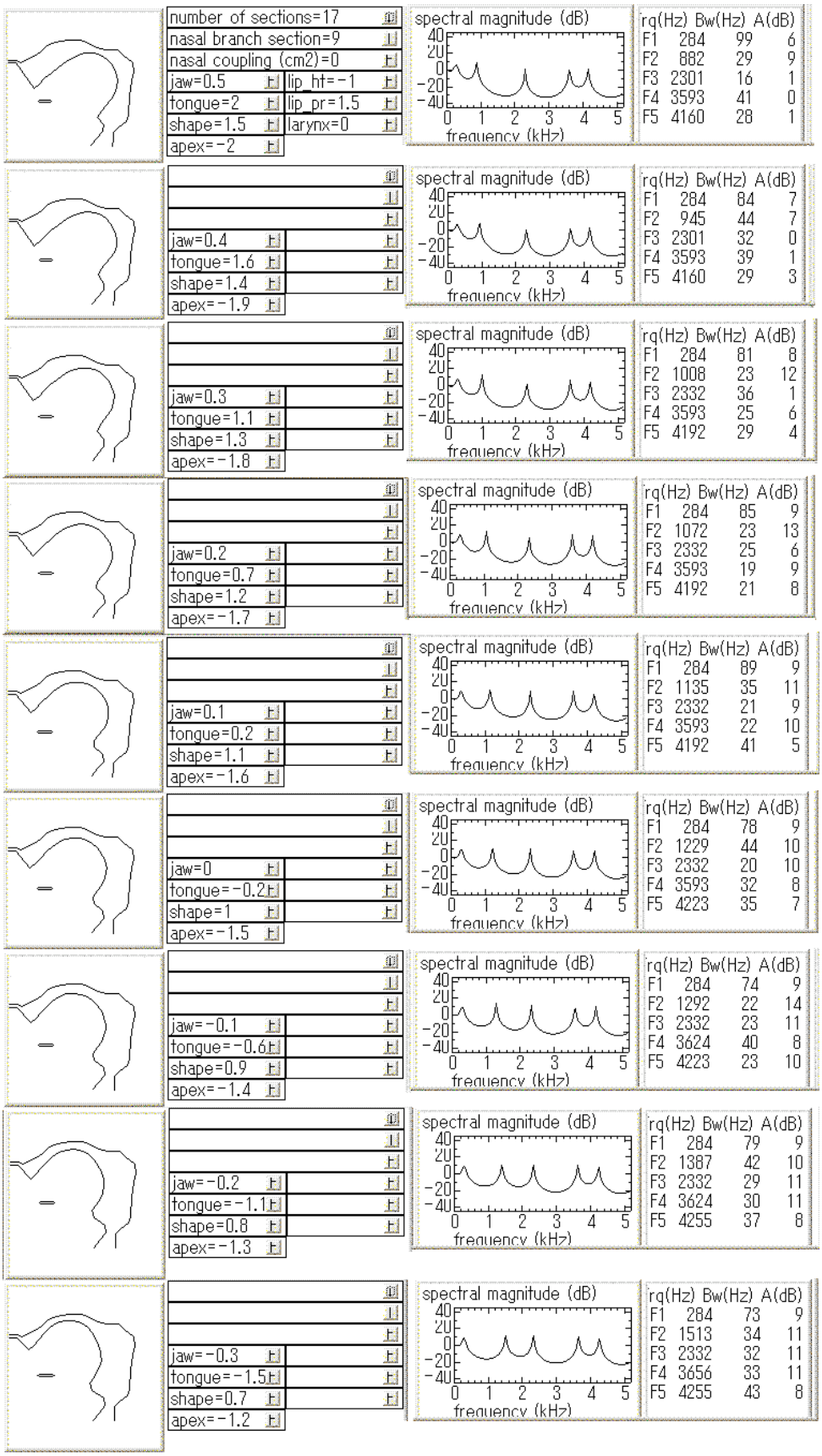



**Figure 103 :** Neuf stimuli (ST0 : en haut – ST08 : en bas) de la série 2 (avancement de la langue), synthétisés en utilisant VTCalc, obtenus en simulant l'avancement progressif de la langue (Figure 102).

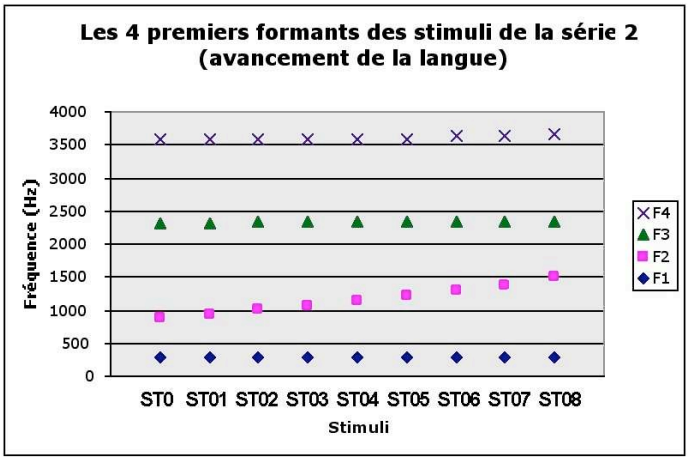



**Figure 104 :** Les quatre premiers formants des 9 stimuli de la série 2 (avancement de la langue) synthétisés en utilisant VTCalc, obtenus en simulant l'avancement progressif de la langue (Figure 102 et Figure 103).

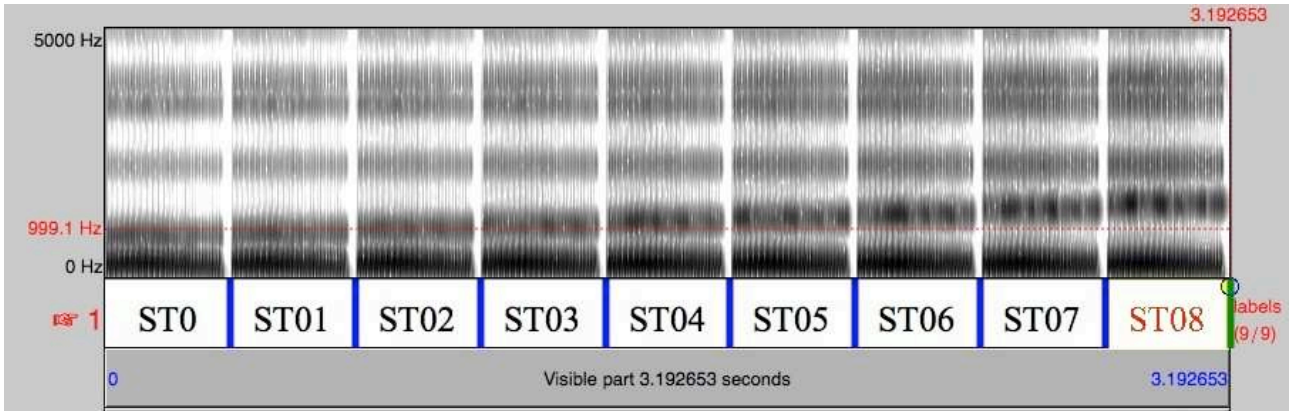


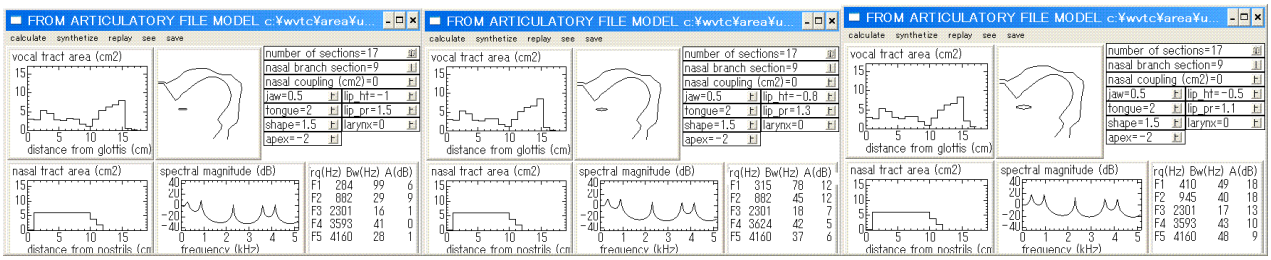

**Figure 105 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 9 stimuli de la série 2 (avancement de la langue) synthétisés en utilisant VTCalc, obtenus en simulant l'avancement progressif de la langue (Figure 102, Figure 103, et Figure 104).

### 3) Série 3 (délabialisation)

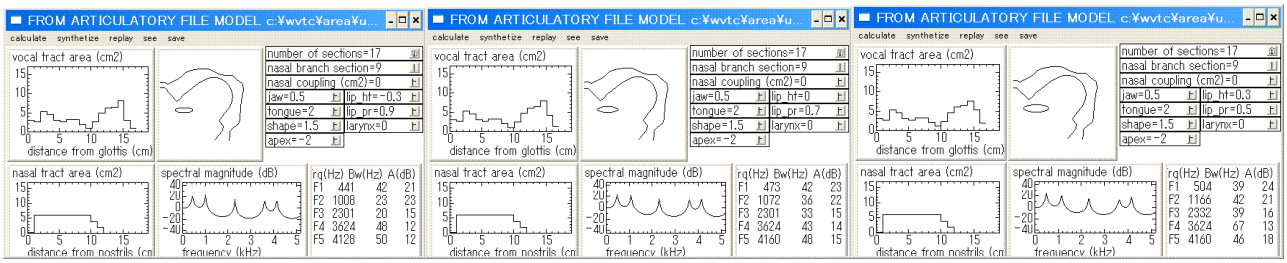
Pour la troisième série, à la différence de la série 2, nous avons fixé les 4 paramètres de la mâchoire (« *jaw* ») et de la langue (« *tongue* », « *shape* », et « *apex* ») aux valeurs du stimulus ST0, qui simule un [u] français, et avons fait varier les 2 paramètres des lèvres (la hauteur « *lip_ht* » et la protrusion « *lip_pr* » des lèvres) de la même façon que le continuum de la série 1 (le larynx « *larynx* » reste inchangé). Ce continuum permettrait de simuler l'effet de délabialisation (aperture et absence progressive de la protrusion des lèvres), tout en conservant la position de la langue d'un [u] français, ce qui correspondrait à des vocoïdes postérieurs plus ou moins arrondis. Les stimuli (ST0-ST98) sont représentés dans le Tableau 58, la Figure 106 et la Figure 108. Rappelons que la délabialisation contribue à augmenter les deux premiers formants, notamment le F2, quand la constriction se trouve vers le milieu du conduit vocal (région vélaire ; nomogramme de Fant 1960 reproduit à la Figure 12, p. 46).

**Tableau 58 :** Sept paramètres articulatoires d'entrée (à gauche) et les formants (en Hertz), la largeur de bande et l'amplitude des formants des stimuli de la série 2 (avancement de la langue) calculés par VTCalc.

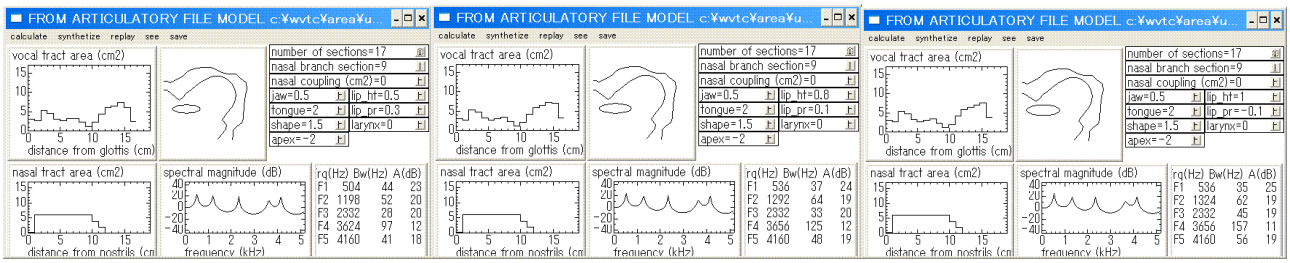
| Série 3 | délabialisation | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | jaw | tongue | shape | apex | lip_ht | lip_pr | larynx | F1 | F2 | F3 | F4 | F5 | Largeur de la bande (Hz) | | | | | Amplitude (dB) | | | | |
| st0 | 0,5 | 2,0 | 1,5 | -2,0 | -1,0 | 1,5 | 0,0 | 284 | 882 | 2301 | 3593 | 4160 | 99 | 29 | 16 | 41 | 28 | 6 | 9 | 1 | 0 | 1 |
| st91 | 0,5 | 2,0 | 1,5 | -2,0 | -0,8 | 1,3 | 0,0 | 315 | 882 | 2301 | 3624 | 4160 | 78 | 45 | 18 | 42 | 37 | 12 | 12 | 7 | 5 | 6 |
| st92 | 0,5 | 2,0 | 1,5 | -2,0 | -0,5 | 1,1 | 0,0 | 410 | 945 | 2301 | 3593 | 4160 | 49 | 40 | 17 | 43 | 48 | 18 | 18 | 13 | 10 | 9 |
| st93 | 0,5 | 2,0 | 1,5 | -2,0 | -0,3 | 0,9 | 0,0 | 441 | 1008 | 2301 | 3624 | 4128 | 42 | 23 | 20 | 48 | 50 | 21 | 23 | 15 | 12 | 12 |
| st94 | 0,5 | 2,0 | 1,5 | -2,0 | 0,0 | 0,7 | 0,0 | 473 | 1072 | 2301 | 3624 | 4160 | 42 | 36 | 33 | 43 | 48 | 23 | 22 | 15 | 14 | 15 |
| st95 | 0,5 | 2,0 | 1,5 | -2,0 | 0,3 | 0,5 | 0,0 | 504 | 1166 | 2332 | 3624 | 4160 | 39 | 42 | 39 | 67 | 46 | 24 | 21 | 16 | 13 | 18 |
| st96 | 0,5 | 2,0 | 1,5 | -2,0 | 0,5 | 0,3 | 0,0 | 504 | 1198 | 2332 | 3624 | 4160 | 44 | 52 | 28 | 97 | 41 | 23 | 20 | 20 | 12 | 18 |
| st97 | 0,5 | 2,0 | 1,5 | -2,0 | 0,8 | 0,1 | 0,0 | 536 | 1292 | 2332 | 3656 | 4160 | 37 | 64 | 33 | 125 | 48 | 24 | 19 | 20 | 12 | 19 |
| st98 | 0,5 | 2,0 | 1,5 | -2,0 | 1,0 | -0,1 | 0,0 | 536 | 1324 | 2332 | 3656 | 4160 | 35 | 62 | 45 | 157 | 56 | 25 | 19 | 19 | 11 | 19 |



ST0 ST91 ST92



ST93 ST94 ST95



ST96 ST97 ST98

**Figure 106 :** Neuf stimuli (ST0 – ST98) de la série 3 (délabialisation), synthétisés en utilisant VTCalc, obtenus en simulant la délabialisation progressive (Tableau 58).

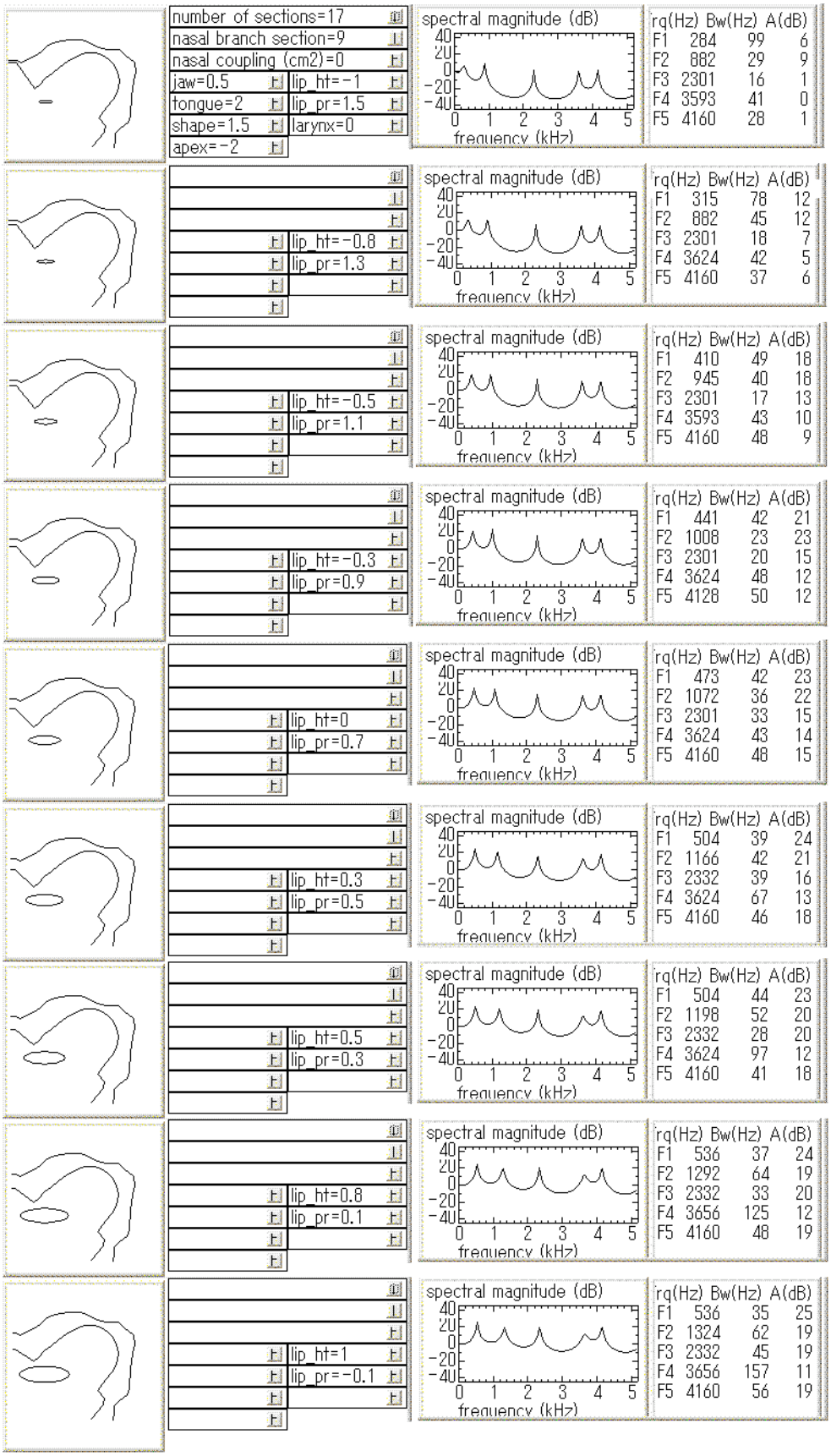



**Figure 107 :** Neuf stimuli (ST0 : en haut – ST98 : en bas) de la série 3 (délabialisation), synthétisés en utilisant VTCalc, obtenus en simulant la délabialisation progressive (Figure 106).

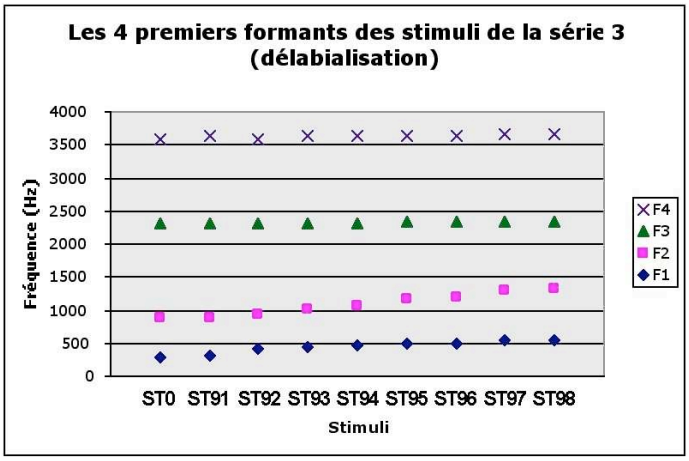



**Figure 108 :** Les 4 premiers formants des 9 stimuli de la série 3 (délabialisation) synthétisés en utilisant VTCalc, obtenus en simulant la délabialisation progresssive (Figure 106 et Figure 107).

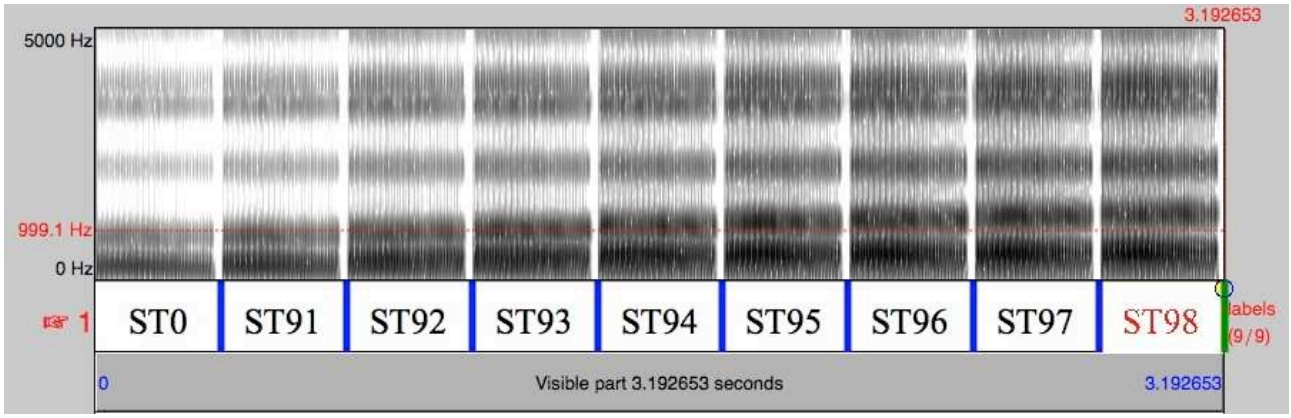



**Figure 109 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 9 stimuli de la série 3 (délabialisation) synthétisés en utilisant VTCalc, obtenus en simulant la délabialisation progressive (Figure 106, Figure 107, Figure 108).

Nous avons ainsi obtenu 25 stimuli (= 9 x 3 – 2, le stimulus ST0 étant commun aux trois séries : Figure 110). Ces stimuli ont été disposés dans 4 listes différentes dans un ordre semi-aléatoire de sorte que les stimuli avec des F2 proches ne se succèdent pas. Plus concrètement, tous les stimuli ont été séparés en 4 groupes selon leur F2 : inférieur à 1000 Hz, 1000 – 1200 Hz, 1200 – 1500 Hz, supérieur à 1500 Hz. Ensuite, nous avons évité de proposer des séquences de deux stimuli appartenant au même groupe. L'amplitude de tous les stimuli a été multipliée par 5 afin d'ajuster le niveau sonore pour l'expérience.

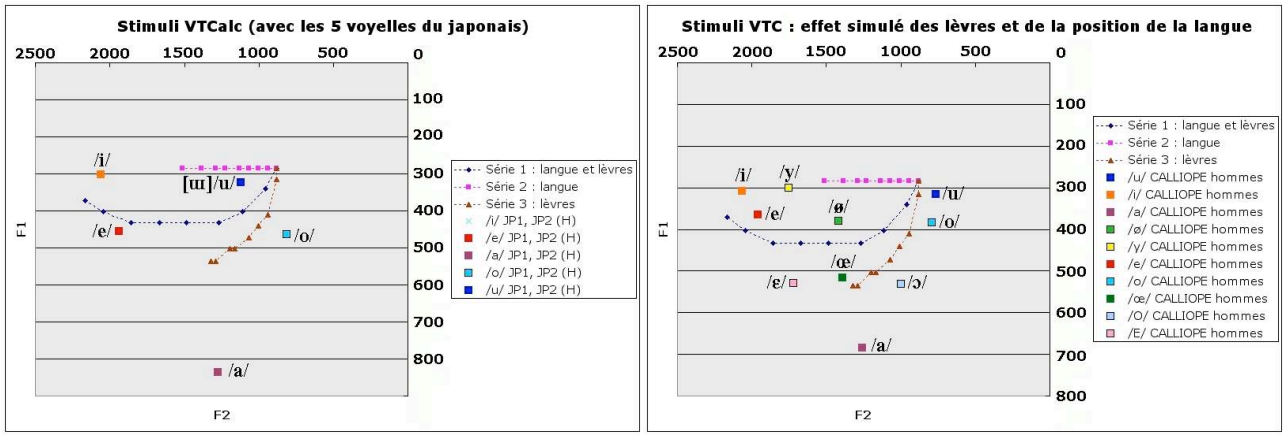



**Figure 110** : Les deux premiers formants des trois séries de stimuli synthétisées en utilisant VTCalc : Série 1 (modification de la position de la langue et délabialisation : Figure 98-Figure 101) ; Série 2 (simulation de l'avancement progressif de la langue : Figure 102-Figure 105) ; Série 3 (simulation de la délabialisation progressive : Figure 106-Figure 109). Les grands carreaux représentent les valeurs des 5 voyelles du japonais de Tokyo (à gauche : moyenne des locuteurs JP1 et JP2, Figure 39, p. 120) et celles des 10 voyelles orales du français (à droite : données des hommes dans CALLIOPE 1989, représentées à la Figure 24, p. 81).

### 3.4.1.2. Synthèse à formants (synthétiseur Klatt)

L'autre type de stimuli a été créé en utilisant le synthétiseur de Klatt[411]. Nous avons choisi la synthèse en cascade, comme cela est généralement fait pour la synthèse de voyelles. Les paramètres que nous avons fait varier sont les suivants : F1, F2, F3, F4, F0. Les valeurs par défaut ont été utilisées pour les autres paramètres. La durée a été fixée à 400 millisecondes, la F0 du début à 110 Hz, et celui de la fin à 120 Hz. Nous avons généré deux séries de stimuli.

#### 1) Série 1 (variation de F2)

Pour la première série de stimuli, nous avons fixé le F1 à 300 Hz, F3 à 2350 Hz, F4 à 3600 Hz, et fait varier le F2 entre 600 et 2200 Hz, à des intervalles de 100 Hz. Nous avons ainsi généré 17 stimuli dont F2 varie de 600 à 2200 Hz (Figure 111).

[411] D. H. Klatt. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67(3) : 971-95, 1980.

L'interface utilisée pour la création de nos stimuli a été développée par H. Timothy Bunnell (http://www.asel.udel.edu/speech/tutorials/synthesis/Klatt.html).

**Série 1 : variation de F2**

| ID stimuli | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|---|---|---|---|---|
| kl-00 | 300 | 600 | 2350 | 3600 |
| kl-01 | 300 | 700 | 2350 | 3600 |
| kl-02 | 300 | 800 | 2350 | 3600 |
| kl-03 | 300 | 900 | 2350 | 3600 |
| kl-04 | 300 | 1000 | 2350 | 3600 |
| kl-05 | 300 | 1100 | 2350 | 3600 |
| kl-06 | 300 | 1200 | 2350 | 3600 |
| kl-07 | 300 | 1300 | 2350 | 3600 |
| kl-08 | 300 | 1400 | 2350 | 3600 |
| kl-09 | 300 | 1500 | 2350 | 3600 |
| kl-10 | 300 | 1600 | 2350 | 3600 |
| kl-11 | 300 | 1700 | 2350 | 3600 |
| kl-12 | 300 | 1800 | 2350 | 3600 |
| kl-13 | 300 | 1900 | 2350 | 3600 |
| kl-14 | 300 | 2000 | 2350 | 3600 |
| kl-15 | 300 | 2100 | 2350 | 3600 |
| kl-16 | 300 | 2200 | 2350 | 3600 |

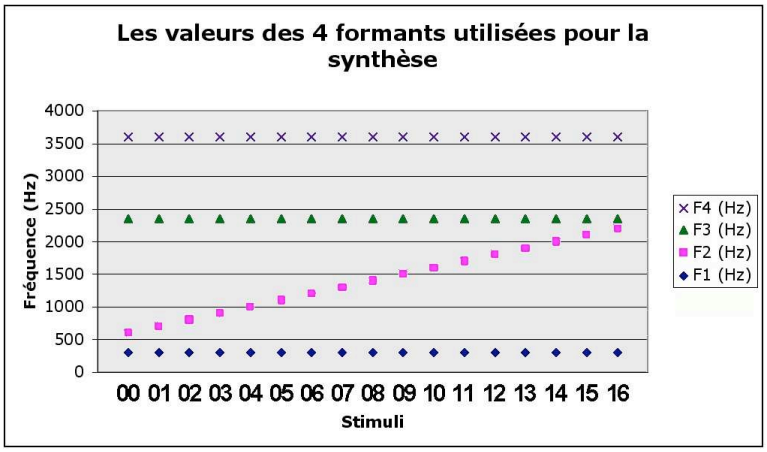



**Figure 111 :** Valeurs des 4 formants utilisées pour la synthèse à formants de Klatt (stimuli de la série 1 : modification de F2).

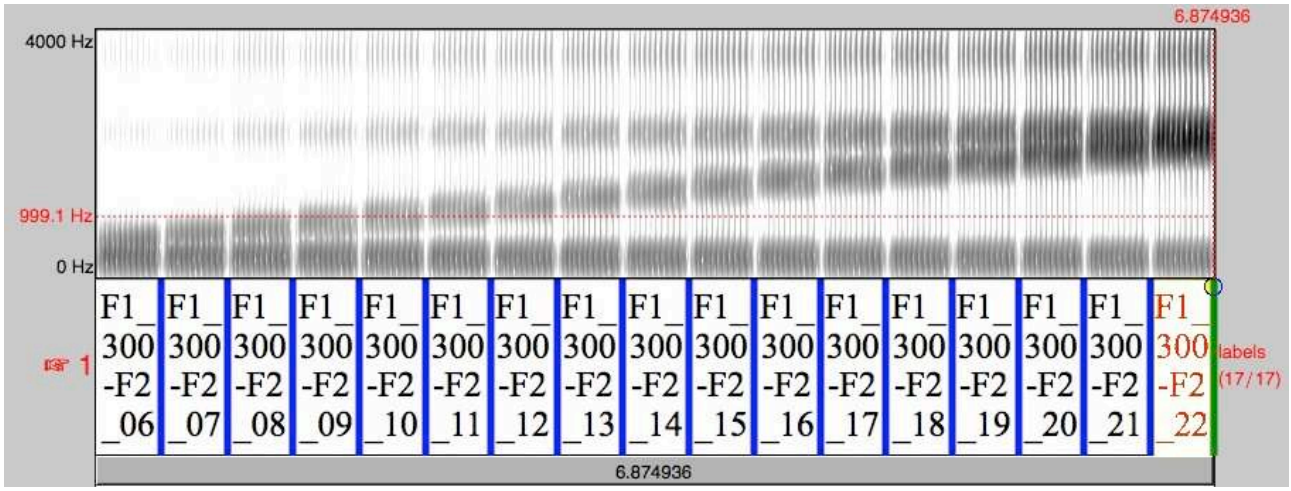



**Figure 112 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 17 stimuli de la série 1 (variation de F2) synthétisés en utilisant le synthétiseur de Klatt (Figure 111).

**2) Série 2 (variation de F1 et de F2)**

Pour la seconde, nous avons fait varier non seulement le F2 mais aussi le F1 en même temps, à des intervalles de 40 Hz. Nous avons ainsi généré 16 stimuli (mis à part le stimulus kl-00 : F1 à 300 Hz et F2 à 600 Hz) dont F1 varie de 310 à 460 Hz, et le F2, de 700 à 2200 Hz (Figure 113).

**Série 2 : variation de F1 et F2**

| ID stimuli | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|---|---|---|---|---|
| kl-00 | 300 | 600 | 2350 | 3600 |
| kl-21 | 310 | 700 | 2350 | 3600 |
| kl-22 | 320 | 800 | 2350 | 3600 |
| kl-23 | 330 | 900 | 2350 | 3600 |
| kl-24 | 340 | 1000 | 2350 | 3600 |
| kl-25 | 350 | 1100 | 2350 | 3600 |
| kl-26 | 360 | 1200 | 2350 | 3600 |
| kl-27 | 370 | 1300 | 2350 | 3600 |
| kl-28 | 380 | 1400 | 2350 | 3600 |
| kl-29 | 390 | 1500 | 2350 | 3600 |
| kl-30 | 400 | 1600 | 2350 | 3600 |
| kl-31 | 410 | 1700 | 2350 | 3600 |
| kl-32 | 420 | 1800 | 2350 | 3600 |
| kl-33 | 430 | 1900 | 2350 | 3600 |
| kl-34 | 440 | 2000 | 2350 | 3600 |
| kl-35 | 450 | 2100 | 2350 | 3600 |
| kl-36 | 460 | 2200 | 2350 | 3600 |

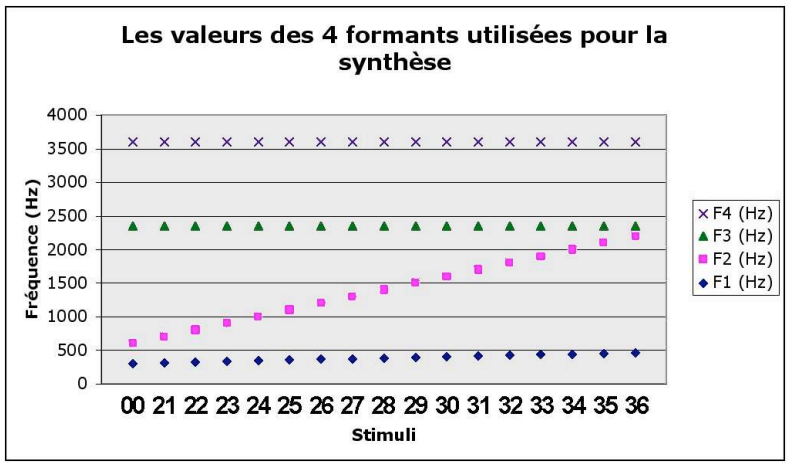



**Figure 113 :** Valeurs des 4 formants utilisées pour la synthèse à formants de Klatt (stimuli de la série 2 : modification de F1 et F2).

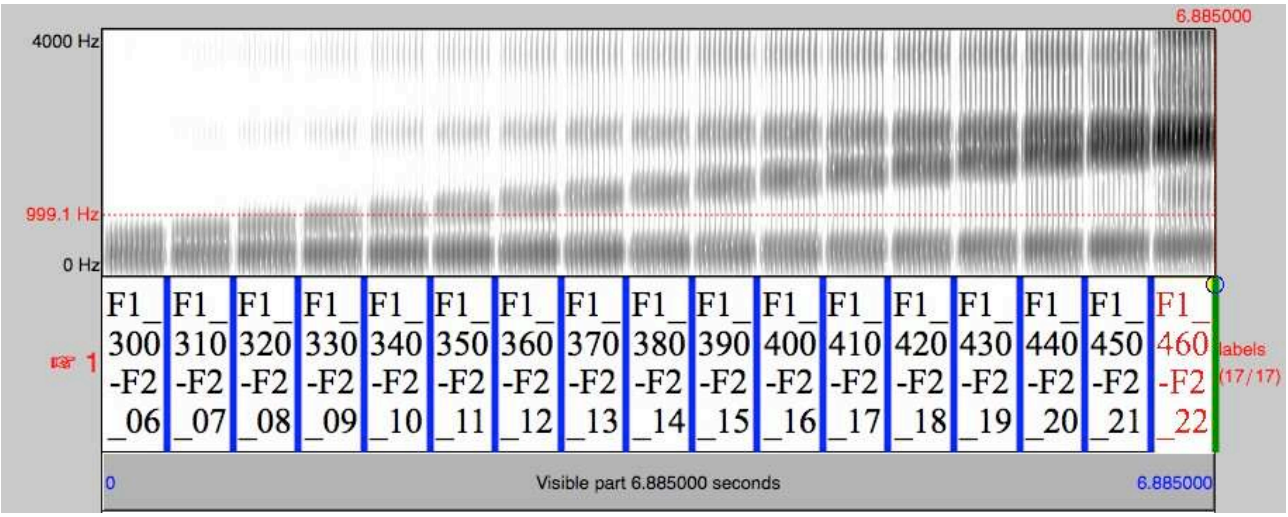



**Figure 114 :** Spectrogramme à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) des 17 stimuli de la série 2 (variation de F1 et F2) synthétisés en utilisant le synthétiseur de Klatt (Figure 113).

L'amplitude de tous les stimuli a été multipliée par 0,6 afin d'ajuster le niveau sonore pour l'expérience. L'effet fondu (« *fade* ») a été appliqué sur les 50 ms au début et à la fin de chaque stimulus en utilisant Sound Forge 7.0[412] pour éviter de créer des clics.

Nous avons ainsi obtenu 33 stimuli (= 17 x 2 - 1, le stimulus kl-00 étant commun aux deux séries : Figure 115), qui ont été disposés dans 4 listes différentes dans un ordre semi-aléatoire de sorte que les stimuli qui ont des structures formantiques proches ne se succèdent pas. Plus concrètement, les stimuli des deux séries ont été séparés en 4 groupes : 1) F2 inférieur à 1000 Hz inclus, 2) F2 entre 1100 Hz et 1500 Hz, 3) F1 à 300 Hz et F2 supérieur à 1600 Hz inclus (série 1), 4)

[412] Site web du logiciel Sound Forge : http://www.sonycreativesoftware.com/soundforge

F1 supérieur à 400 Hz et F2 supérieur à 1600 Hz inclus (série 2). Nous avons évité des séquences de deux stimuli appartenant au même groupe.

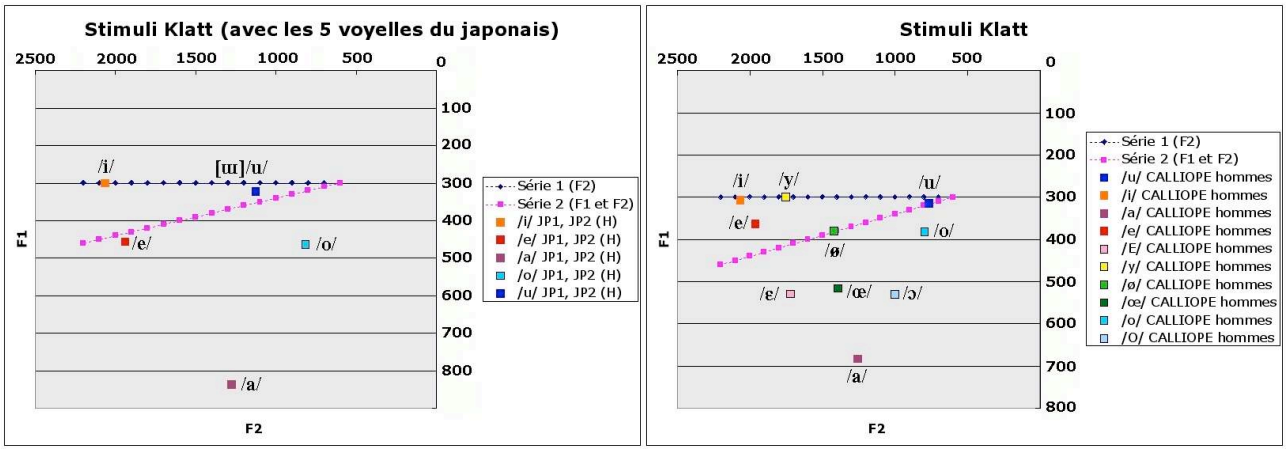



**Figure 115** : Les deux premiers formants des deux séries de stimuli synthétisées en utilisant le synthétiseur de Klatt : Série 1 (modification de F2 : Figure 111, Figure 112) ; Série 2 (modification de F1 et F2 : Figure 113, Figure 114). Avec les valeurs des 5 voyelles du japonais de Tokyo (à droite : moyenne des locuteurs JP1 et JP2, Figure 39, p. 120) et celles des 10 voyelles orales du français (à gauche : données des hommes dans CALLIOPE 1989, représentées à la Figure 24, p. 81).

### *3.4.2. Les auditeurs*

Deux groupes d'auditeurs ont participé à cette expérience de perception. Le premier est composé de 16 auditeurs natifs du japonais. Ils étaient âgés de 24 à 64 ans. Ils étaient tous originaires de Tokyo et de ses environs. Aucun d'entre eux n'avait vécu plus d'un an à l'étranger. Le second est composé de 16 auditeurs natifs du français. Ils étaient âgés de 18 à 38 ans, et résidaient tous dans la région parisienne lors de l'expérience.

### *3.4.3. Les procédures*

L'expérience a été effectuée sur un ordinateur, et les auditeurs disposaient d'une souris et d'un casque. Les tâches des auditeurs étaient d'écouter les stimuli, de les identifier comme une des voyelles de leur langue maternelle (français ou japonais), et de les évaluer sur une échelle de 1 (mauvais) à 5 (bon). Ils étaient informés qu'ils allaient entendre des voyelles prononcées par des étrangers.

Les auditeurs japonais disposaient comme choix de réponse non seulement des 5 voyelles du japonais (/i e a o u/) mais aussi de 3 syllabes avec le /j/ (/ja jo ju/[413]). Ces dernières ont été rajoutées afin de permettre la réponse /ju/, qui est utilisé dans des mots d'emprunt pour représenter la voyelle /y/ du français (ex. *Hugo* /ygo/ > /'juRgoR/ [jɯːgoː], ou /'jugoR/ [jɯgoː] ; Tableau 11). Ces voyelles et syllabes ont été transcrites en *katakana* dans l'ordre alphabétique japonais (アイウエオ, ヤユヨ : /a i u e o/, /ja ju jo/ : Figure 116 à gauche).

[413] Notons qu'il n'existe pas de syllabaires non-composés qui représentent /ji/ et /je/, que /ji/ n'est pas autorisé dans la langue, et que /je/ est utilisé uniquement dans des mots d'emprunt dans le système phonémique de nos jours.

Les auditeurs devaient choisir la voyelle, puis la note d'évaluation en cliquant sur les cases correspondant à leur choix. Ils devaient par la suite cliquer sur le bouton « ok » qui se trouvait en bas de l'écran (Figure 116) afin de finaliser leurs réponses et de passer au stimulus suivant. Ils pouvaient changer leurs réponses avant de cliquer sur le bouton « ok ».

Quand ils ont cliqué sur le bouton « ok », ils ont entendu le stimulus suivant 800 millisecondes plus tard.

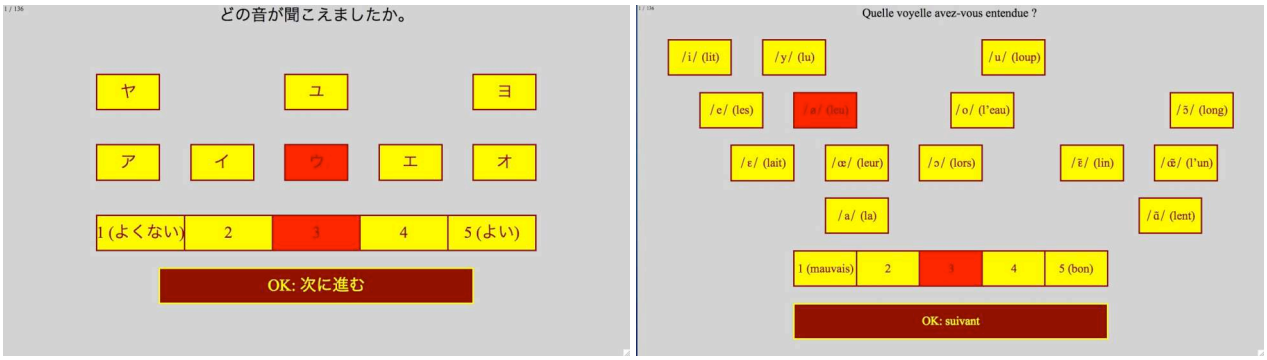



**Figure 116** : Copie de l'écran présenté aux sujets-auditeurs japonais (à gauche) et français (à droite) pour la présente expérience. Les cases foncées indiquent que ces réponses (la voyelle /ø/ « leu » et la note de 3 sur 5, sur l'écran présenté aux francophones : à droite) sont sélectionnées.

L'expérience était composée de l'entraînement et de deux parties de test.

Après avoir lu des consignes, les auditeurs ont passé un entraînement. Ils ont effectué les mêmes tâches que le test en écoutant 5 stimuli synthétisés par VTCalc afin de s'habituer aux tâches. Les 5 stimuli présentés ont été synthétisés à partir des données articulatoires d'un [i], [e], [a], [o] et [ə] français.

Dans la première partie du test, les 4 listes de stimuli de VTCalc décrites ci-dessus (3.4.1. Les stimuli) ont été présentées aux auditeurs. Les auditeurs ont eu une pause après la deuxième liste. Un stimulus *dummy* (stimulus non pris en compte pour les analyses) a été inséré au début et à la fin du test, et immédiatement avant et après la pause. Les sujets-auditeurs ont ainsi entendu 104 stimuli (25 x 4 répétitions + 4 stimuli *dummy*).

La deuxième partie était composée des stimuli du synthétiseur de Klatt. Comme dans la première partie, les auditeurs ont entendu 2 listes avant et après une pause. Un stimulus *dummy* a été également mis en place au début et à la fin du test ainsi qu'au milieu, avant et après la pause. Les sujets ont entendu 136 stimuli (33 x 4 répétitions + 4 stimuli *dummy*). Comme pour la première partie, la moitié des participants (8 sur 16) ont passé le test avec l'ordre inverse de stimuli.

Le temps de passation de l'ensemble de l'expérience était de 20 à 25 minutes environ.

### *3.4.4. Les analyses des données*

Les analyses des données ont été effectuées comme dans l'expérience précédente (Expérience 8). Sur les 64 réponses obtenues pour chaque stimulus (16 auditeurs x 4 réponses), la proportion (le nombre de fois) dans laquelle une voyelle donnée a été sélectionnée comme réponse a été calculée.

Par la suite, un score qui tient compte de la proportion de réponses et de l'évaluation a été calculé, en multipliant les deux valeurs pour chaque réponse de chaque auditeur, puis en prenant la moyenne de ces valeurs pour chaque stimulus. La proportion (le nombre) de réponses varie donc de 0 à 64, le score de 0 (mauvais) à 5 (bon).

### *3.4.5. Les résultats*

#### 3.4.5.1. Les résultats des auditeurs japonais non apprenants

**1) Les stimuli de VTCalc**

Les résultats des stimuli de la série 1 ([u] français -> [ø] -> : Tableau 56, p. 219 ; Figure 110, p. 229) sont représentés à la Figure 117. Les deux premiers stimuli (ST0 et ST1) ont été identifiés presque unanimement comme /u/ japonais (61 et 64 réponses sur 64), avec des notes élevées (3,5, et 3,6). Les deux suivants (ST2 et ST3) ont été perçus majoritairement comme /u/ (54 et 44 réponses sur 64), mais les notes pour le /u/ baissent considérablement (2,6 et 1,8). Avec le suivant (ST4), le nombre de réponses pour /e/ dépasse celui de /u/ (30 contre 25), mais les notes sont égales pour les deux voyelles (0,9). Notons que, comme nous le verrons plus tard, ces 3 stimuli (ST2, ST3, et ST4) ont obtenu de bonnes notes pour /ø/ auprès des auditeurs francophones. Les 4 autres stimuli ont été classifiés presque unanimement (63 réponses sur 64 pour ST5, 64 réponses pour ST6) ou majoritairement (57 réponses sur 64 pour ST7, 37 réponses pour ST8) comme /e/, avec la note plus élevée pour ST6 (3,8, contre 2,8 pour ST5, 3,1 pour ST7, et 1,9 pour ST8). Les deux derniers (ST7 et ST8) ont été perçus minoritairement comme /i/ japonais, mais comme l'on verra ci-dessous, les auditeurs francophones ne les ont presque jamais identifiés comme /i/ français (mais comme /e/).

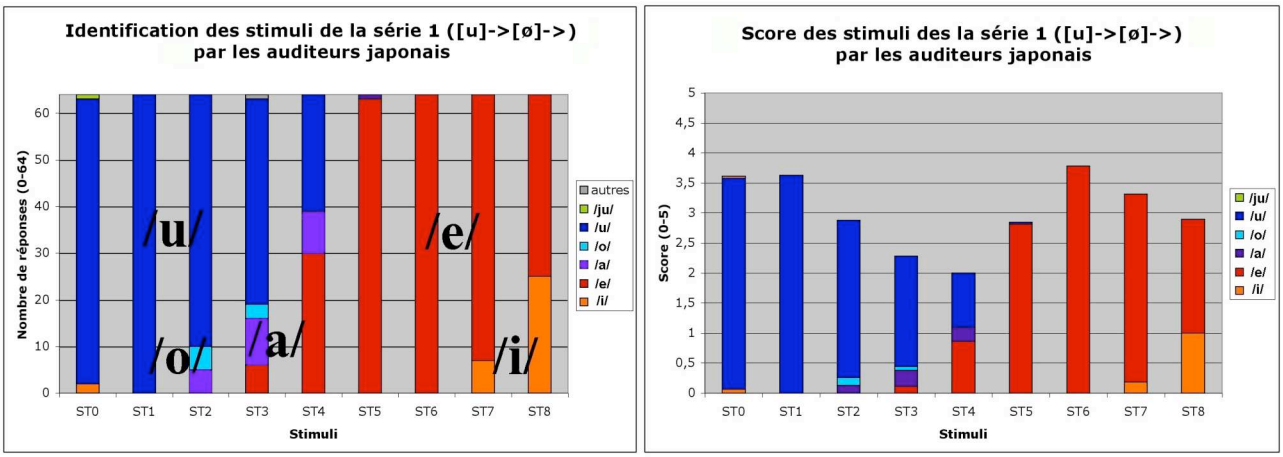



**Figure 117 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 1 ([u] français -> [ø] -> : Tableau 56, p. 219 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs japonais x 4 répétitions) par stimulus. Résultats des 16 auditeurs francophones sont représentés à la Figure 122 (p. 239).

La Figure 118 illustre les résultats des stimuli de la série 2 (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229). Nous y observons que tous les stimuli ont été identifiés comme /u/ japonais presque unanimement, ou majoritairement, du moins (49 – 63 réponses sur 64). Les notes restent entre 3,3 et 3,6 pour les 6 premiers stimuli (ST0 – ST05), puis baissent

légèrement pour les deux suivants (3,0 pour ST06, 2,9 pour ST07), encore plus pour ST8 (2,2). Même si l'on trouve un effet significatif des stimuli sur les scores ($F_{(8, 135)} = 2{,}36$, $p = 0{,}02$), il n'y a pas de différence significative entre les 8 premiers stimuli (ST0 – ST7 : PDSL de Fisher). Ce cas de figure contraste, comme nous le verrons ci-dessous, avec celui de la perception des francophones, qui ont montré un passage du /u/ au /ø/. Notons également que le stimulus ST08 a été identifié comme /ju/ 12 fois sur 64, tandis que le même stimulus n'a été perçu que marginalement comme /y/ par les auditeurs francophones.

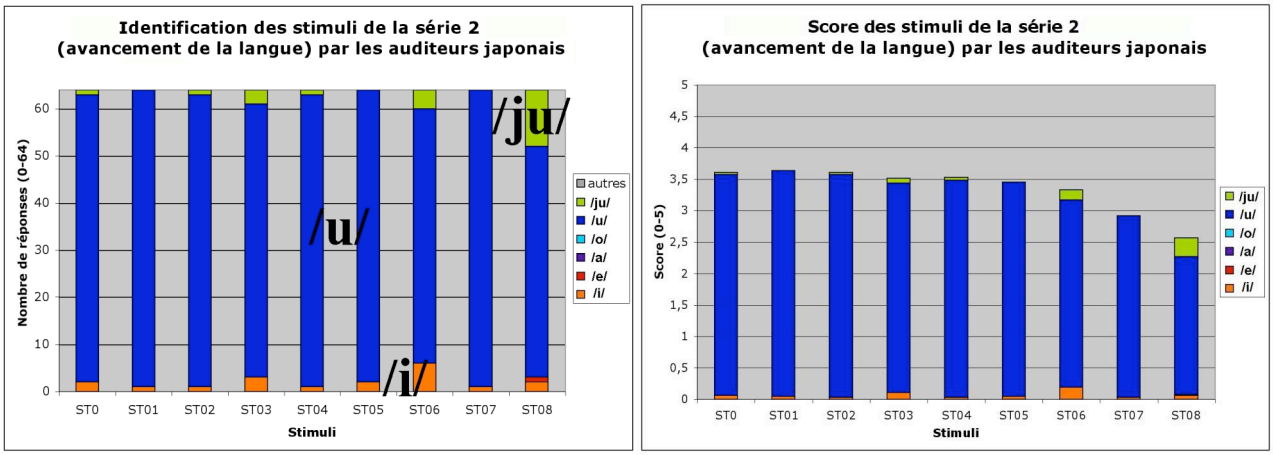



**Figure 118 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 2 (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs japonais x 4 répétitions) par stimulus. Résultats des 16 auditeurs francophones sont représentés à la Figure 123 (p. 240).

Les résultats de la série 3 (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229) sont représentés à la Figure 119. Les deux premiers stimuli (ST0 et ST91) ont été identifiés presque toujours comme /u/ japonais (61 et 64 réponses sur 64) avec des notes élevées (3,5 pour les deux). Le suivant (ST92) a été perçu majoritairement comme /u/ (42 réponses sur 64), mais aussi comme /o/ un tiers des temps environ (21 sur 64). La note pour /o/ est basse en conséquence. Concernant le stimulus ST93, les réponses pour /o/ étaient les plus nombreuses (41 sur 64), mais la note n'était pas élevée (1,3). Le suivant (ST94) a été identifié presque également comme /a/ (29 sur 64) et comme /o/ (26 sur 64), avec des notes semblables (0,8 pour /a/, 0,7 pour /o/). Les autres stimuli (ST95 – ST98) ont été classifiés majoritairement comme /a/ (48 – 56 réponses sur 64), mais avec des notes relativement basses (entre 1,6 et 1,9). Comme l'on verra ci-dessous, ces stimuli ont été identifiés en général comme /œ/ (ou /ø/) par les auditeurs francophones.

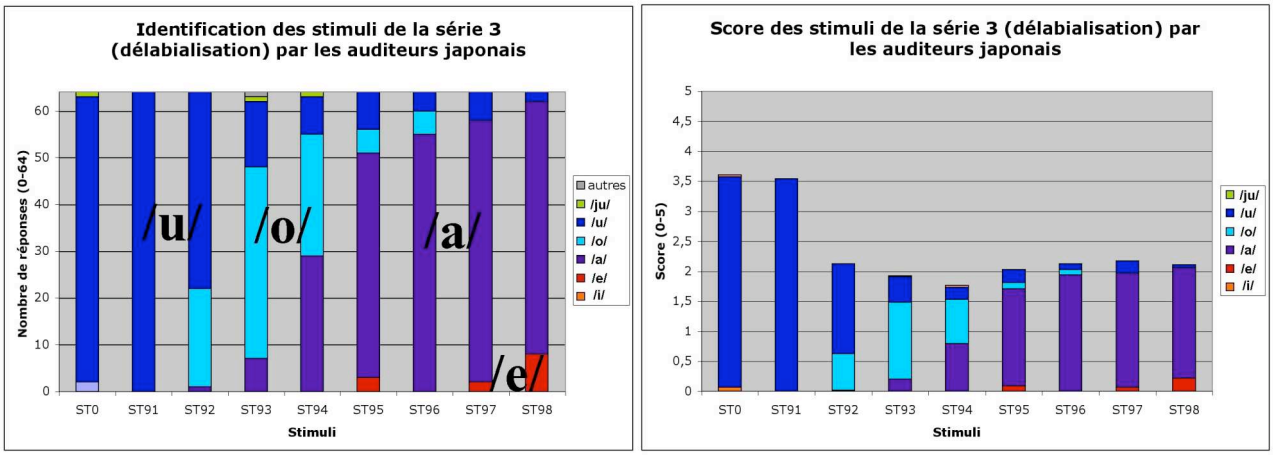



**Figure 119 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 3 (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs japonais x 4 répétitions) par stimulus. Résultats des 16 auditeurs francophones sont représentés à la Figure 124 (p. 241).

## 2) Les stimuli du synthétiseur de Klatt

Les résultats des stimuli de la série 1 (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232) sont représentés à la Figure 120. Les 4 stimuli qui ont un F2 entre 800 et 1300 Hz ont été identifiés presque unanimement comme /u/ japonais (62 – 64 réponses sur 64) avec de bonnes notes (entre 3,4 et 3,9, la plus élevée étant celle du stimulus avec F2 à 1000 Hz). Les deux stimuli qui les entourent (F2 à 700 et 1400 Hz) ont été également perçus presque parfaitement comme /u/ (59 et 62 réponses sur 64), mais les notes sont moins élevées (3,1 et 3,0). Le nombre de réponses d'identification de /u/ et le score pour /u/ baisse progressivement sur les deux côtés. Pour les valeurs basses de F2, la réponse /o/ devient importante, ne serait-ce que minoritairement (23 réponses sur 64 avec une note de 0,9 pour le stimulus avec F2 à 600 Hz). Avec le F2 plus élevé, le nombre de réponses et la note de /i/ deviennent de plus en plus dominants au fur et à mesure que ceux de /u/ baissent (à 2200 Hz, aucune réponse pour /u/, contre 55 réponses sur 64 pour /i/ avec une note de 2,6), mais nous observons également des réponses pour /e/ et /ju/, mais cela reste dans la minorité (16 réponses sur 64 pour /e/ pour le stimulus avec F2 à 2200 Hz, 8 réponses sur 64 pour /ju/ pour ceux avec F2 à 1800 et 1900 Hz).

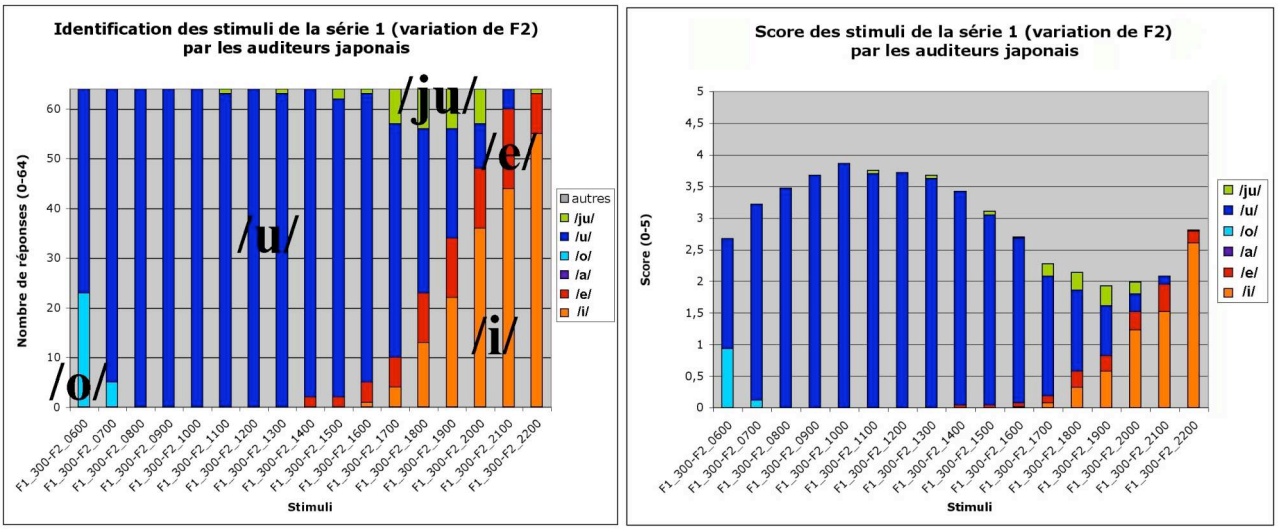



**Figure 120 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 17 stimuli de la série 1 (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232) du synthétiseur de Klatt. 64 réponses (16 auditeurs japonais x 4 répétitions) par stimulus. Résultats des 16 auditeurs francophones sont représentés à la Figure 125 (p. 242).

La Figure 121 illustre les résultats des stimuli de la série 2 (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232). Le choix du /u/ japonais est majoritaires (entre 41 et 64 réponses sur 64) pour les stimuli entre F1 à 300 Hz et F2 à 600 Hz, d'un côté, et F1 à 390 Hz et F2 à 1500 Hz, de l'autre. Le nombre de réponses et la note pour /u/ augmentent au fur et à mesure que les deux premiers formants augmentent jusqu'à 370 Hz pour F1 et 1300 Hz pour F2 (64 réponses sur 64, avec une note de 3,7 : celui d'avant atteint la fréquence de 64 sur 64, mais la note ne s'élève qu'à 3,5), au détriment de la diminution de ceux du /o/ (maximum de 23 réponses sur 64, avec une note de 0,9 pour le premier stimulus avec F1 à 300 Hz et F2 à 600 Hz). De là, c'est le /e/ qui devient de plus en plus dominant. La proportion parvient à la majorité avec F1 à 400 Hz et F2 à 1600 Hz, et atteint la totalité avec F1 à 430 Hz, et F2 à 1900 Hz. C'est le stimulus suivant (F1 à 440 Hz et F2 à 2000 Hz) qui a obtenu le meilleur score (4,3), et le score baisse légèrement pour les stimuli qui ont les deux premiers formants plus élevés, malgré l'unanimité de choix pour cette voyelle. Nous y observons ainsi un passage du /u/ (vers le meilleur exemplaire du /u/) au /e/, dans les zones où nous nous attendrions à un passage du /u/ au /e/ en passant par le /ø/ chez les auditeurs francophones.

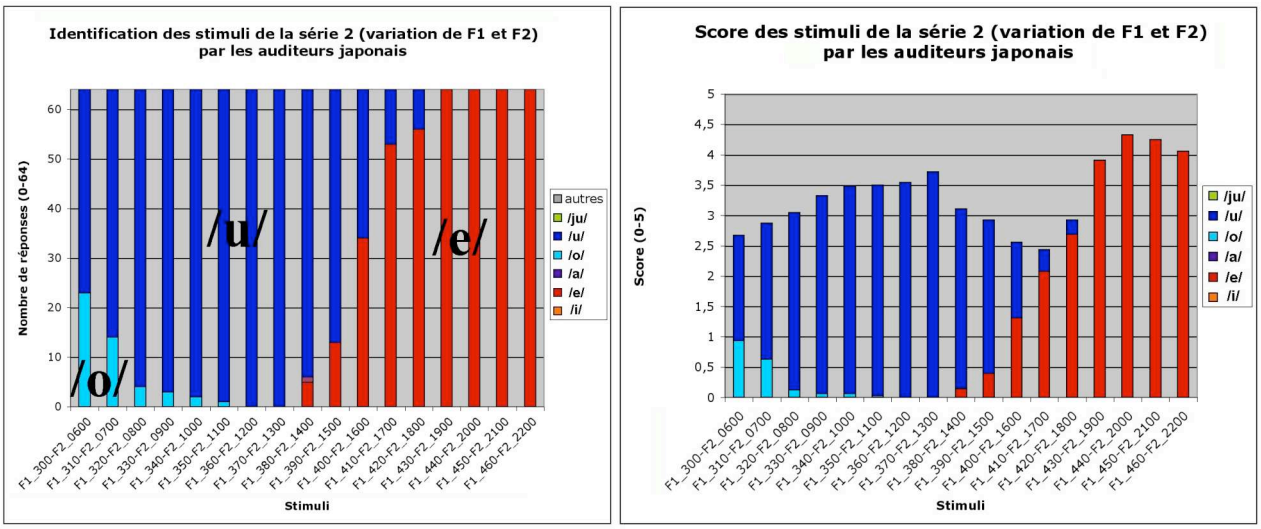



**Figure 121 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 17 stimuli de la série 2 (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232) du synthétiseur de Klatt. 64 réponses (16 auditeurs japonais x 4 répétitions) par stimulus. Résultats des 16 auditeurs francophones sont représentés à la Figure 126 (p. 244).

Nous pouvons résumer les comportements des auditeurs japonophones de la manière suivante :

1) Avec la série 1 des stimuli VTCalc ([u]->[ø]- > : Tableau 56, p. 219 ; Figure 110, p. 229), nous avons observé un passage du /u/ au /e/ autour des stimuli ST4 et ST5 (Figure 117, p. 234).

2) Avec la série 2 des stimuli VTCalc (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229), les auditeurs ont majoritairement perçu le /u/ pour tous les stimuli, avec une baisse de note pour les 3 derniers (ST6 – ST8 : Figure 118, p. 235).

3) Avec la série 3 des stimuli VTCalc (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229), nous avons trouvé un passage du /u/ au /a/ en passant par le /o/ (Figure 119, p. 236).

4) Pour la série 1 des stimuli Klatt (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232), nous avons observé un passage du /u/ au /e/, les notes pour le /u/ étant au niveau maximal pour les stimuli avec F2 à 800-1300 Hz (Figure 120, p. 237).

5) Pour la série 2 des stimuli Klatt (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232), nous avons également observé un passage du /u/ au /e/, avec la note pour le /u/ baissant à partir du stimulus avec F1 à 380 Hz et F2 à 1400 Hz (Figure 121, p. 238).

### 3.4.5.2. Les résultats des auditeurs francophones natifs

**1) Les stimuli de VTCalc**

Les résultats des stimuli de la série 1 ([u]->[ø]-> : Tableau 56, p. 219 ; Figure 110, p. 229) sont représentés à la Figure 122. Le stimulus ST0 (un [u] français) a en effet été identifié presque unanimement (63 réponses sur 64) comme /u/ par les auditeurs francophones, avec un score élevé

(4,0 sur 5). En revanche, dès le stimulus suivant (ST1 : F1 à 340 Hz, F2 à 960 Hz), le nombre de réponses baisse quasiment à la moitié (33 réponses sur 64), avec une note moyenne très basse (1,3 sur 5). Les autres réponses minoritaires incluent /o/ (15 sur 64) et /ø/ (11 sur 64). Les trois stimuli suivants (ST2 – ST4) ont été perçus comme /ø/ (entre 34 et 49 sur 64) ou /œ/ (entre 12 et 30 sur 64), la somme pour les deux voyelles étant de 61 à 64 sur 64. Les notes moyennes ne sont pas très élevées (entre 2 et 2,4 sur 5 pour /ø/) à cause de la dispersion des réponses entre ces deux voyelles. Les notes moyennes sans considération de la proportion de réponses s'élèveraient à 3,8 pour /ø/ et 3,7 pour /œ/ pour ST3, 3,7 pour /ø/ et 3,9 pour /œ/ pour ST4. Le stimulus ST5 a également été classifié majoritairement comme /ø/ ou /œ/ (23 et 25 sur 64, soit 48 en tout sur 64), mais minoritairement comme /ɛ/ ou /e/ (12 et 4 sur 64). Pour ST6, les réponses pour ces deux voyelles deviennent majoritaires (27 pour /e/ et 26 pour /ɛ/, soit 53 sur 64), avec une minorité de réponses pour /ø/ et /œ/ (10 et 1, respectivement). Les deux autres stimuli ST7 et ST8 ont été identifiés quasi unanimement comme /e/ (62 et 64 réponses sur 64), avec des scores très élevés (4,3 et 4,7 sur 5).

Nous pouvons ainsi observer un passage de la zone de perception de /u/ à celle de /ø/ et /œ/, puis à celle de /e/ dans cette série de stimuli.

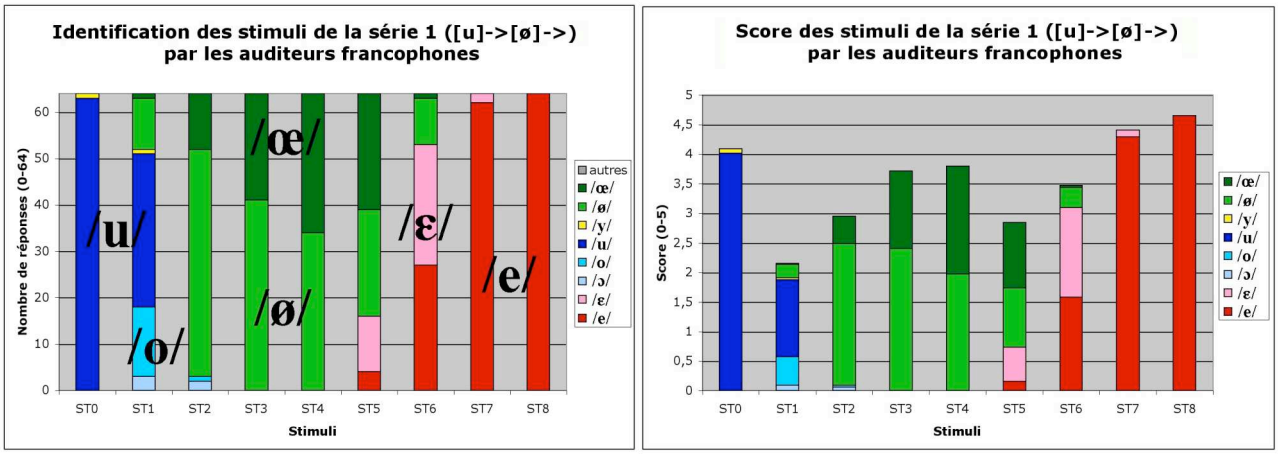



**Figure 122 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 1 ([u] français -> [ø] ->: Tableau 56, p. 219 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs francophones x 4 répétitions) par stimulus. Résultats des 16 auditeurs japonophones sont représentés à la Figure 117 (p. 234).

La Figure 123 montre les résultats obtenus pour les stimuli de la série 2 (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229). Dans cette série, le nombre de réponses pour /u/ diminue progressivement, et le score obtenu suit le même mouvement. Avec le quatrième (ST03 : F1 à 284 Hz et F2 à 1072 Hz), le nombre de réponses pour /u/ baisse environ à la moitié (37 réponses sur 64) avec une note de 1,3, alors que les réponses pour le /ø/ comptent 26 sur 64. Tous les autres stimuli ont été identifiés majoritairement comme /ø/ (entre 43 et 51 réponses sur 64), le nombre de réponses et la note augmentant progressivement jusqu'au ST07 (F1 à 284 Hz, F2 à 1387 Hz), où ils atteignent les maxima (54 sur 64, et 3,1 sur 5, respectivement). Pour le ST08, ils baissent au profit des réponses pour /y/, qui restent toujours minoritaires (19 réponses sur 64, contre 43 sur 64 pour /ø/).

Ces résultats suggèrent donc que cette série de stimuli correspondrait aux zones de perception de /u/ et de /ø/ chez les auditeurs francophones, même si les notes pour /ø/ ne sont pas très élevées (maximum : 2,5 pour ST3).

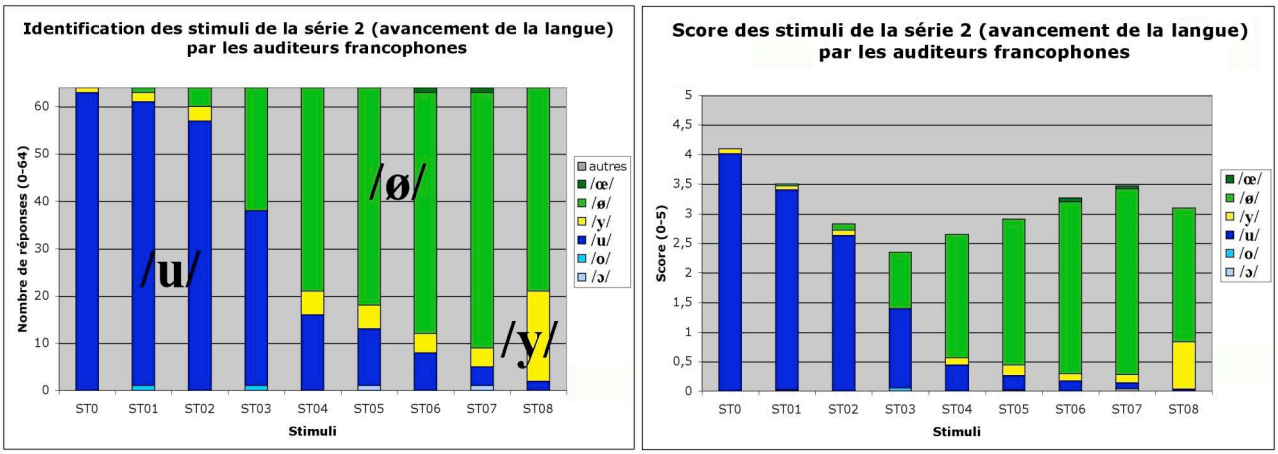



**Figure 123 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 2 (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs francophones x 4 répétitions) par stimulus. Résultats des 16 auditeurs japonophones sont représentés à la Figure 118 (p. 235).

Les résultats des stimuli de la série 3 (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229) sont illustrés à la Figure 124. Tout comme le stimulus ST0, le ST91 a été identifié majoritairement comme /u/ (49 réponses sur 64), mais également comme /o/ dans une minorité de cas (12 sur 64), à la différence du ST0. Les deux stimuli suivants, ST92 et ST93, ont été classifiés majoritairement comme /o/ ou /ɔ/ (46 et 11 sur 64 pour ST92, 22 et 37 sur 64 pour ST93). Avec le ST94, le nombre des réponses pour /o/ et /ɔ/ dépasse à peine la moitié (4 et 32, soit 36 sur 64), l'autre moitié étant occupée par /ø/ et /œ/ (7 et 21, soit 28 sur 64). Tous les 4 autres stimuli ont été perçus majoritairement comme /œ/ (entre 48 et 54 sur 64), avec une note qui augmente progressivement et atteint au maximum pour le dernier stimulus ST98 (3,6 sur 5). On trouve également des réponses minoritaires pour /ø/ pour ces stimuli (entre 7 et 12 sur 64).

Nous observons ainsi un passage de /u/ à /œ/ en passant par /o/ et /ɔ/ pour cette série de stimuli, qui simule la délabialisation progressive avec la position de la langue d'un [u] français.

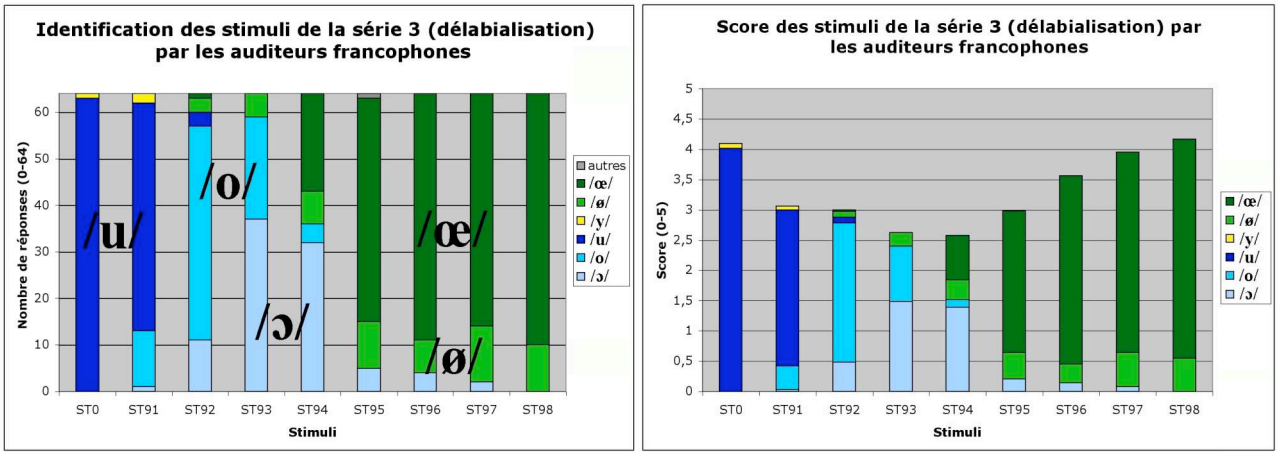



**Figure 124 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 9 stimuli de la série 3 (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229) de VTCalc. 64 réponses (16 auditeurs francophones x 4 répétitions) par stimulus. Résultats des 16 auditeurs japonophones sont représentés à la Figure 119 (p. 236).

Les auditeurs francophones ont donc identifié les vocoïdes postérieurs non-arrondis de la série 3 (délabialisation), avec la position de la langue d'un [u] français, majoritairement comme /œ/. Le F2 du dernier stimulus ST98, qui a obtenu la meilleure note pour /œ/ (3,6 sur 5) est de 1324 Hz (Tableau 58, p. 226), et cette valeur correspond à une voyelle « acoustiquement centrale » typique (F2 vers 1500 Hz ; Vaissière 2006 : 73[414]). Le résultat de la présente expérience suggère que les auditeurs francophones se comportent effectivement de la même manière pour les voyelles antérieures arrondies et les voyelles postérieures non-arrondies.

## 2) Les stimuli du synthétiseur de Klatt

Les résultats des stimuli de la série 1 (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232) sont représentés à la Figure 125. Les 5 premiers stimuli (F2 à 600-1000 Hz) ont été identifiés majoritairement comme /u/ (50-61 sur 64 réponses), avec le nombre maximum de réponses (61 sur 64) pour le F2 à 900 Hz. Les notes sont assez élevées pour les 4 premiers stimuli (F2 à 600-900 Hz, entre 3,4 et 4,0 sur 5), la meilleure (4,0 sur 5) étant celle du stimulus qui a son F2 à 800 Hz. La note pour /u/ baisse à 2,9 pour le 5ème stimulus (F2 à 1000 Hz), même si le nombre de réponses pour /u/ reste grand (52 sur 64) pour ce stimulus. Les 5 premiers stimuli ont été perçus également comme /o/ dans une minorité de cas (3 à 14 sur 64), mais de moins en moins au fur et à mesure que le F2 augmente, le nombre maximum étant 14 (sur 64) pour le F2 à 600 Hz. Avec le 6ème stimulus (F2 à 1100 Hz), les réponses pour /u/ dépassent la moitié (38 sur 64), mais celles du /ø/ comptent presque un tiers (23 sur 64). Les réponses pour /ø/ sont majoritaires (36-49 réponses sur 64) pour les stimuli qui ont leur F2 de 1200 Hz à 1600 Hz, le maximum atteint à 1400 Hz (49 sur 64). Le meilleur score a été également obtenu par ce stimulus, mais il est moins élevé (2,7) que celui du /u/ (4,0 sur 5), ce qui correspond au nombre maximal de réponses moins important (49 sur 64, contre 61 sur 64 pour /u/). Ces stimuli, qui ont été reconnus majoritairement comme /ø/, sont des

[414] J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France, 2006.

voyelles acoustiquement centrales, avec une répartition équilibrée des formants, F2 autour de 1400-1500 Hz. En revanche, le F1 est trop bas pour un [ø] français typique. Cela expliquerait la perception majoritaire du /ø/, mais avec une note pas très élevée (et un petit nombre de réponses pour /y/). Concernant les stimuli avec leur F2 de 1700 à 2000 Hz, les auditeurs ont identifié majoritairement le /y/ (43-56 réponses sur 64), le maximum étant atteint par le F2 à 1900 Hz. Les notes correspondent approximativement à cette tendance, sauf que la meilleure note a été obtenue par le stimulus avec F2 à 1800 Hz (3,2 sur 5, contre 3,1 pour celui qui a son F2 à 1900 Hz). Ces stimuli sont marqués par un regroupement de F2 et F3, qui caractérise effectivement le /y/ français. Pour les stimuli avec un F2 plus élevé (2100-2200 Hz), les réponses pour /y/ diminuent au fur et à mesure que celle du /i/ augmente (38 sur 64 pour F2 à 2200 Hz). En revanche, la note pour /i/ reste basse (1,9 sur 5) pour le stimulus avec le F2 à 2200 Hz, et il y a également des réponses pour /e/ (11 sur 64). Cela suggère que ces stimuli ont un F2 trop élevé pour percevoir un beau [y], mais le F3 (fixé à 2350 Hz) serait trop bas et éloigné du F4 pour un beau [i] à la française (Vaissière 2007[415] ; CALLIOPE 1989, les données représentées à la Figure 24, p. 81 ; nos données représentées à la Figure 41, p. 123).

Nous observons ainsi un passage de /u/ à /ø/ (avec une note pas très élevée pour /ø/, à savoir 2,7 au maximum : F2 à 1200-1600 Hz), puis à /y/ (F2 à 1700-2000 Hz), et ensuite au /i/ avec une note très basse, au fur et à mesure que le F2 du stimulus augmente.

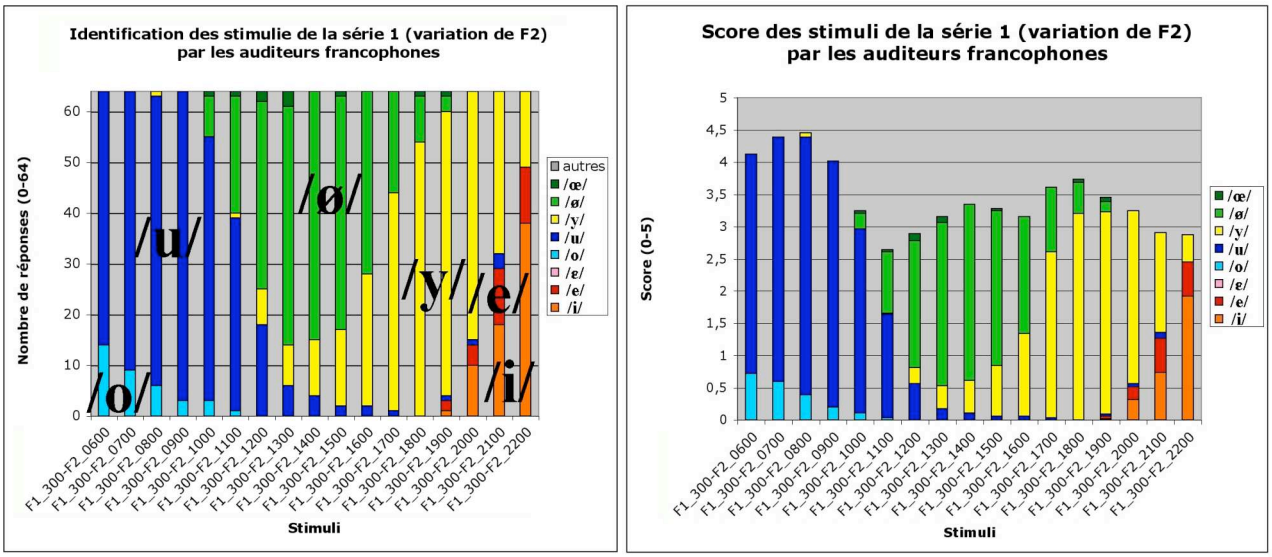


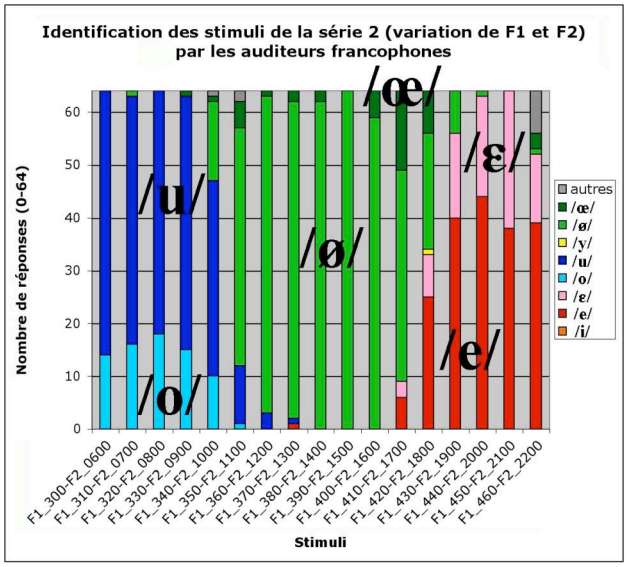

**Figure 125 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 17 stimuli de la série 1 (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232) du synthétiseur de Klatt. 64 réponses (16 auditeurs francophones x 4 répétitions) par stimulus. Résultats des 16 auditeurs japonophones sont représentés à la Figure 120 (p. 237).

La Figure 126 représente les résultats de la série 2 (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232). Les réponses pour /u/ sont majoritaires pour les 5 premiers stimuli (F1 à 300

[415] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

Hz et F2 à 600 Hz – F1 à 340 Hz et F2 à 1000 Hz), le nombre restant relativement constant (entre 46 et 50 sur 64) pour les 4 premiers, mais baissant considérablement (37 sur 64, légèrement supérieur à la moitié) pour le dernier (F1 à 340 Hz et F2 à 1000 Hz). Le score baisse au fur et à mesure que les deux formants montent (de 3,4 à 2,8 pour les 4 premiers, et 1,7 pour le cinquième). Pour ces 5 premiers stimuli, on trouve également des réponses pour /o/, qui sont plus nombreuses que dans la série 1 (notons que F2 augmente dans la série 2 : Figure 113, p. 231 ; Figure 115, p. 232), mais reste minoritaires (entre 10 et 18 sur 64, le maximum étant atteint avec F1 à 320 Hz et F2 à 800 Hz). Les réponses pour le /ø/, enregistrées pour la première fois à plusieurs (15 sur 64) pour le 5ème stimulus (F1 à 340 Hz et F2 à 1000 Hz), deviennent majoritaires pour le suivant (F1 à 350 Hz et F2 à 1100 Hz), au détriment des réponses pour le /u/ (45 pour /ø/ contre 11 sur 64 pour /u/), puis restent presque unanimes (entre 59 et 64 sur 64) dans la zone entre F1 à 360 Hz et F2 à 1200 Hz, d'un côté, et F1 à 400 Hz et F2 à 1600 Hz, de l'autre. Le maximum (réponses unanimes pour le /ø/ : 64 sur 64) est atteint par le stimulus avec F1 à 390 Hz et F2 à 1500 Hz. Le score montre une tendance similaire, parvenant au maximum (4,1 sur 5) pour ce même stimulus, et restant élevé pour cette même zone (F1 à 360 Hz et F2 à 1200 Hz ~ F1 à 400 Hz et F2 à 1600 Hz : entre 3,4 et 4,1 sur 5). Avec le stimulus suivant (F1 à 410 Hz et F2 à 1700 Hz), les réponses pour le /ø/ sont toujours majoritaires (40 sur 64) avec une note de 2,1, mais celles de /œ/ (15 sur 64), et /e/ et /ɛ/ (6 et 3 sur 64) sont également observées. Au-delà, en passant par une zone d'ambiguïté (F1 à 420 Hz et F2 à 1800 Hz) entre /ø/ /œ/ (22 et 8 sur 64) et /e/ /ɛ/ (25 et 8 sur 64), on entre dans une zone où /e/ et /ɛ/ sont majoritaires (F1 à 430 Hz et F2 à 1900 Hz ~ F1 à 460 Hz et F2 à 2200 Hz) : 38-44 réponses sur 64 pour /e/, et 13-26 réponses pour /ɛ/. Le nombre de réponses pour /e/ augmente jusqu'au stimulus avec F1 à 440 Hz et F2 à 2000 Hz, au fur et à mesure que le F2 s'élève. Celui du /ɛ/ augmente également, mais il est toujours plus petit que celui du /e/. Le nombre maximum de réponses pour /ɛ/ est atteint par l'avant-dernier stimulus (F1 à 450 Hz et F2 à 2100 Hz). Les 8 réponses « autres » pour le dernier stimulus (F1 à 460 Hz et F2 à 2200 Hz) pourraient être dues à la dégradation de la qualité de son causée par un rapprochement trop important de paramètres de formants (F2 à 2200 Hz et F3 à 2350 Hz).

Nous y observons ainsi un passage de /u/ (avec des réponses plus nombreuses pour /o/ par rapport à la série 1 : Figure 125) à /ø/ (F1 à 350 Hz et F2 à 1100 Hz – F1 à 410 Hz et F2 à 1700 Hz), puis à /e/ et /ɛ/. Notons également que la meilleure note pour /ø/ est supérieure (4,1) à celle de la série 1 (2,7 : F1 était fixé à 300 Hz), probablement grâce à un F2 plus élevé (390 Hz pour le stimulus qui a obtenu le meilleur score), et donc plus proche des valeurs observées dans la production des locuteurs natifs.



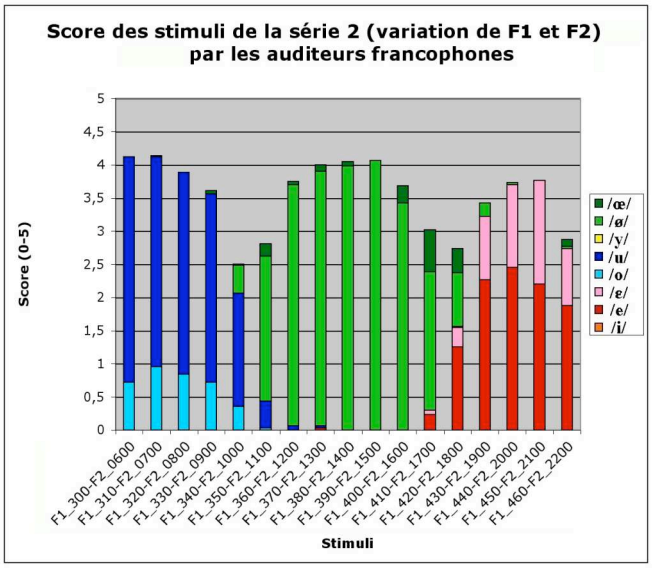




**Figure 126 :** Nombre de réponses d'identification (0-64 : à gauche) et score (proportion x note d'évaluation : 0-5 : à droite) des 17 stimuli de la série 2 (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232) du synthétiseur de Klatt. 64 réponses (16 auditeurs francophones x 4 répétitions) par stimulus. Résultats des 16 auditeurs japonophones sont représentés à la Figure 121 (p. 238).

Nous pouvons résumer les comportements des 16 auditeurs francophones de la manière suivante :

1) Avec la série 1 des stimuli VTCalc ([u]->[ø]- > : Tableau 56, p. 219 ; Figure 110, p. 229), nous avons observé un passage de la zone de perception de /u/ à celle de /ø/ et /œ/ (de ST1 à ST2), puis à celle de /e/ (de ST5 à ST6 : Figure 122, p. 239).

2) La série 2 des stimuli VTCalc (avancement de la langue : Tableau 57, p. 223 ; Figure 110, p. 229), correspond aux zones de perception de /u/ et de /ø/ chez les auditeurs francophones, même si les notes pour /ø/ ne sont pas très élevées (maximum : 2,5 pour ST3 : Figure 123, p. 240).

3) Avec la série 3 des stimuli VTCalc (délabialisation : Tableau 58, p. 226 ; Figure 110, p. 229), nous avons observé un passage de /u/ à /o/ et /ɔ/ (de ST91 à ST92), et ensuite à /œ/ (de ST94 à ST95 : Figure 124, p. 241).

4) Pour la série 1 des stimuli Klatt (variation de F2 : Figure 111, p. 230 ; Figure 115, p. 232), nous avons observé un passage de /u/ à /ø/ (avec une note pas très élevée pour /ø/, à savoir 2,7 au maximum : F2 à 1200-1600 Hz), puis à /y/ (F2 à 1700-2000 Hz), et ensuite au /i/ avec une note très basse, au fur et à mesure que le F2 du stimulus augmente (Figure 125, p. 242).

5) Pour la série 2 des stimuli Klatt (variation de F1 et F2 : Figure 113, p. 231 ; Figure 115, p. 232), nous avons observé un passage de /u/ (avec des réponses plus nombreuses pour /o/ par rapport à la série 1) à /ø/ (F1 à 350 Hz et F2 à 1100 Hz – F1 à 410 Hz et F2 à 1700 Hz), puis à /e/ et /ɛ/. Notons également que la meilleure note pour /ø/ est supérieure (4,1) à celle de la série 1 (2,7).

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 du /u/ japonais, avec une évaluation relativement bonne. Les francophones sont très sensibles à des écarts par rapport au /u/ français, sur le plan articulatoire (avancement de la langue et délabialisation) et acoustique (F2 élevé), tandis que la zone du /u/ japonais est plus large. Ces résultats aux Figure 127 et Figure 128, qui montrent le regroupement des stimuli par leur réponse modale par les auditeurs japonophones et francophones.

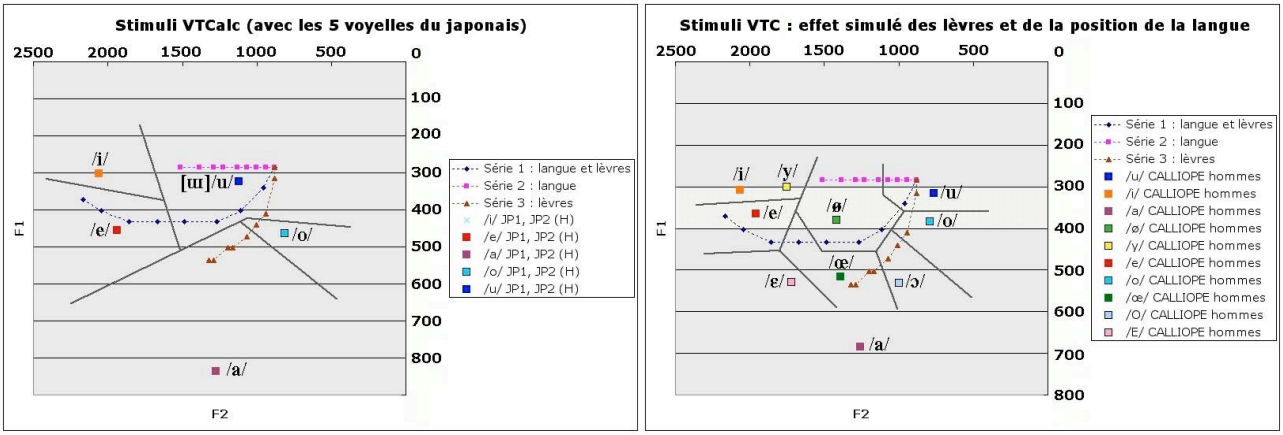



**Figure 127** : Regroupement des trois sérires de stimuli de VTCalc par leur réponse modale (réponse la plus représentative) par les 16 auditeurs japonophones (à gauche) et les 16 auditeurs francophones (à droite). Série 1 (modification de la position de la langue et délabialisation : Figure 98-Figure 101) ; Série 2 (simulation de l'avancement progressif de la langue : Figure 102-Figure 105) ; Série 3 (simulation de la délabialisation progressive : Figure 106-Figure 109). Les grands carreaux représentent les valeurs des 5 voyelles du japonais de Tokyo (à gauche : moyenne des locuteurs JP1 et JP2, Figure 39, p. 120) et celles des 10 voyelles orales du français (à droite : données des hommes dans CALLIOPE 1989, représentées à la Figure 24, p. 81).

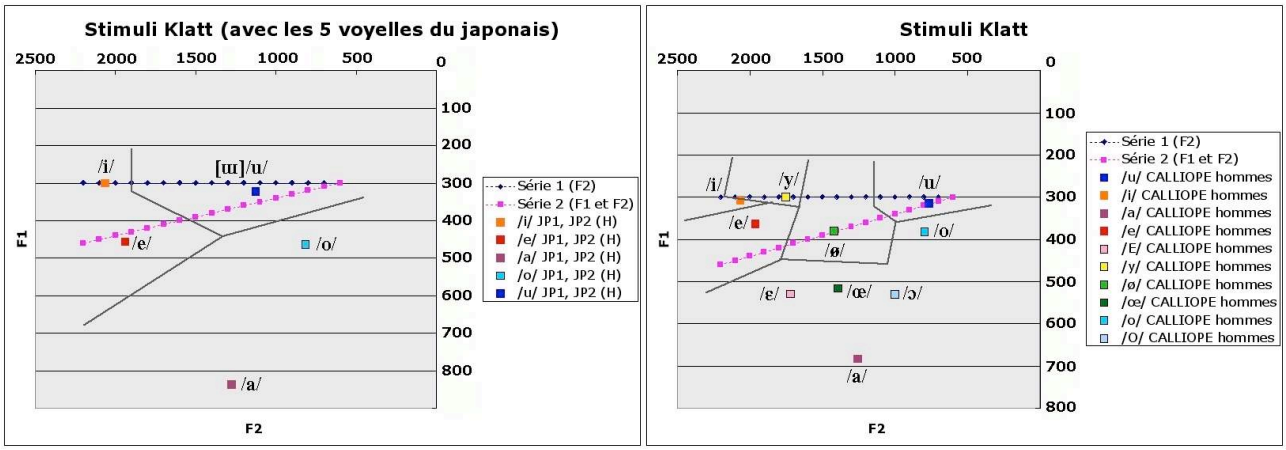



**Figure 128** : Regroupement des deux sérires de stimuli du synthétiseur de Klatt par leur réponse modale (réponse la plus représentative) par les 16 auditeurs japonophones (à gauche) et les 16

auditeurs francophones (à droite). Série 1 (modification de F2 : Figure 111, Figure 112) ; Série 2 (modification de F1 et F2 : Figure 113, Figure 114). Avec les valeurs des 5 voyelles du japonais de Tokyo (à droite : moyenne des locuteurs JP1 et JP2, Figure 39, p. 120) et celles des 10 voyelles orales du français (à gauche : données des hommes dans CALLIOPE 1989, représentées à la Figure 24, p. 81). Notons que F3 n'a pas été modifié.

### 3.5. DISCUSSION

Les 4 expériences présentées dans cette partie ont montré les résultats suivants.

1) Expérience 6 (discrimination AXB) : la paire /u/-/y/ est relativement facile à distinguer, et /y/-/ø/ et /i/-/e/ sont plus difficiles mais deviennent plus faciles avec de l'expérience d'apprentissage, tandis que /o/-/u/ et /u/-/ø/ restent relativement difficiles même pour les apprenants plus avancés.

2) Expérience 7 (production en lecture) : les apprenants ont mieux réalisé le /y/ (F2/F3 proches, sans que le F2 soit trop élevé – inférieur à 2400Hz – : produit par 20 apprenants au moins une fois) que le /u/ (F2 inférieur à 1000 Hz : produit par 4 apprenants), et encore mieux le /ø/ (pas de rapprochement de formants, F2 autour de 1500 Hz : produit par 31 apprenants au moins une fois).

3) Expérience 8 (identification et évaluation par des francophones de /u/ /y/ /ø/ prononcés par des apprenants) :

/u/ : les stimuli avec un F2 inférieur à 1000 Hz ont été identifiés presque unanimement comme /u/, avec de bons socres (3,8 et 4,4 sur 5). Ceux avec un F2 entre 1000 et 1100 Hz ont été identifiés presque aussi souvent comme /ø/ que comme /u/. Avec un F2 supérieur à 1100 Hz, les stimuli ont été perçus majoritairement (de 36 à 62 réponses sur 64) comme /ø/. Ces résultats confirment le rôle primordial de F2 comme indice.

/y/ : les stimuli avec F2 et F3 proches, avec un F2 pas trop élevé (inférieur à 2100 Hz pour les stimuli d'hommes), ont été reconnus majoritairement (de 46 à 64 réponses sur 64) comme /y/. Les réalisations diphtonguées, comme la séquence /ju/ [jɯ] japonaise, ont été identifiées majoritairement (de 44 à 54 réponses sur 64) comme /ø/, ou également comme /u/, /y/, et /ø/, mais avec de mauvais scores dans tous les cas (1,8 sur 5 au maximum).

/ø/ : les stimuli avec un F2 autour de 1500 Hz environ (de 1520 Hz à 1796 Hz pour les stimuli des 3 hommes, de 1295 Hz à 1543 Hz pour les stimuli des 2 femmes) ont été catégorisés majoritairement (de 36 à 60 réponses sur 64) comme /ø/.

4) Expérience 9 (identification et évaluation de stimuli synthétisés –synthése articulatoire et synthèse à formants– par d'auditeurs japonophones non-apprenants et francophones) : quand les auditeurs francophones ont montré un passage du /u/ au /ø/, les auditeurs japonophones ont perçu en général /u/ japonais, avec des notes relativement bonnes. Les francophones seraient très sensibles à des écarts par rapport au /u/ français, sur le plan articulatoire (avancement de la

langue et délabialisation) et acoustique (F2 élevé), tandis que la zone perceptive du /u/ japonais est plus large (F2 jusqu'à 1800-1900 Hz).

**La perception de /u/ /y/ et /ø/** Les résultats de l'Expérience 6 (discrimination AXB) ont montré l'ordre de difficulté suivant : /u/-/ø/ > /y/-/ø/ > /u/-/y/ (Figure 66, p. 182). En nous fondant sur ces résultats, nous pouvons supposer que le /ø/ serait le plus difficile à distinguer des deux autres, ensuite le /u/, et le plus facile serait le /y/. Cette hiérarchie correspond également aux résultats de l'Expérience 3 (test d'identification : Tableau 22, p. 130) : le nombre de réponses correctes était de 17 sur 40 pour /ø/, 22 sur 40 pour /u/, et 33 sur 40 pour /y/.

Une analyse traditionnelle en traits ne semble pas permettre d'expliquer cette hiérarchie de difficulté (Tableau 59). Il existe deux différences de traits entre /u/ et /ø/ (aperture et antériorité), alors qu'il n'en existe qu'une seule pour les paires /u/-/y/ (antériorité) et /y/-/ø/ (aperture), ce qui nous permettrait de prédire que la paire /u/-/ø/ serait la plus facile à distinguer, étant donné que la distance est la plus importante. Or, c'est cette paire qui s'est avérée la plus difficile.

**Tableau 59** : Différence de distinction perceptive aobservée et différences de traits.

| **Difficulté de distinction perceptive** | **Paires de voyelles** | **Trais différents** |
|---|---|---|
| +++ | /u/-/ø/ | Antériorité, aperture |
| ++ | /y/-/ø/ | Aperture |
| + | /u/-/y/ | Antériorité |

La facilité relative de distinction du /y/ serait due à la possibilité de l'assimiler à la séquence /ju/ du japonais. Si nous employons les termes du Modèle d'assimilation perceptive (PAM : *Perceptual Assimilation Model*) de Best et ses collègues (Best *et al.* 1988[416], Best 1995[417]), le /y/ français serait un exemplaire « acceptable mais non idéal », ou « notoirement différent » de la catégorie /ju/ du japonais. En revanche, le /u/ français serait un exemplaire « acceptable mais non idéal » du /u/ [ɯ] japonais, et le /ø/ comme un « bon exemplaire » ou un exemplaire « acceptable mais non idéal » du /u/ [ɯ] japonais (Tableau 60 ci-dessous ; Figure 3, p. 32).

**Tableau 60** : Pattern d'assimilation perceptive de /u y ø/ par les japonophones.

| **Français** | | **Japonais** |
|---|---|---|
| /y/ | « exemplaire acceptable mais non idéal » ou « notoirement différent » de | /ju/ [jɯ] |
| /u/ | « exemplaire acceptable mais non idéal » de | /u/ [ɯ] |
| /ø/ | « bon exemplaire » ou « exemplaire acceptable mais non idéal » de | /u/ [ɯ] |

[416] C. T. Best, G. W. McRoberts, N. N. Sithole. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click perception by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14 : 345-60, 1988.

[417] C. T. Best. "A Direct Realist View of Cross-Language Speech Perception", in Winifred Strange, *Speech perception and linguistic experience: Issues in cross-language research*. York Press, Baltimore, 171-204, 1995.

Dans l'Expérience 9 (perception de stimuli synthétisés) de notre étude, les stimuli synthétisés à partir des profils articulatoires proches de celui d'un [u] français (F2 bas, inférieur à 1000 Hz) n'ont pas obtenu des notes moins élevées (3,5 pour ST0 simulant [u], 3,6 pour ST01) que ceux avec la langue plus avancée (3,5 pour ST04, 3,4 pour ST05, par exemple) par les auditeurs japonophones (Figure 117, p. 234 pour la série 1, Figure 118, p. 235 pour la série 2 des stimuli de VTCalc), tandis qu'avec les stimuli du synthétiseur Klatt, nous avons observé des meilleures notes pour les stimuli avec F2 relativement élevé par rapport aux valeurs des francophones natifs (à partir de 900Hz-1000 Hz) que pour ceux avec un F2 plus bas (600 Hz-700 Hz ; Figure 120 pour la série 1, Figure 121 pour la série 2). Ces résultats dans les basses fréquences seraient dus aux caractéristiques du /o/ japonais, qui a un F2 plus bas que celui du /u/ [ɯ] japonais (Figure 39, p. 120), mais c'est tout de même majoritairement /u/ (japonais) qui est reconnu.

En terme de paires de voyelles en opposition (Tableau 61 ci-dessous ; Figure 4, p. 34), la paire /u/-/y/ serait considérée comme un cas d' « assimilation à deux catégories[418] », si le /y/ est assimilé au /ju/ [jɯ] japonais, alors que le /u/ est assimilé au /u/ [ɯ] japonais.

Le PAM prédit que ce type d'assimilation permet une bonne discrimination, et effectivement, les non-apprenants et apprenants japonophones ont relativement bien réussi à distinguer cette paire. En revanche, la paire /y/-/ø/ serait également considérée comme un exemple d' « assimilation à deux catégories », étant donné que le /ø/, comme le /u/ français, serait assimilé au /u/ [ɯ] japonais, mais le résultat de discrimination s'est avérée moins bonne.

La paire /u/-/ø/ serait considérée comme exemple du type « assimilation à une seule catégorie »[419] ou une version faible de « différence de degré de bon exemplaire »[420], et la difficulté prédite pour ces cas de figure est plus grande que pour le cas d' « assimilation à deux catégories ». Cette prédiction nous permet d'expliquer la difficulté effectivement observée pour cette paire auprès des non apprenants et des apprenants japonophones (Expérience 6 : discrimination AXB ; Figure 69, p. 187 pour les non-apprenants, Figure 66, p. 182 pour les apprenants).

**Tableau 61** : Types d'assimilation perceptive des oppositions /u/-/y/, /y/-/ø/, et /u/-/ø/.

| **Oppositions** | |
|---|---|
| /u/-/y/ | « assimilation à deux catégories » (type TC) |
| /y/-/ø/ | « assimilation à deux catégories » (type TC) |
| /u/-/ø/ | « assimilation à une seule catégorie » (type SC)<br>ou « différence de degré de bon exemplaire » (type CG) |

[418] *Two-category assimilation :* « *Each non-native segment is assimilated to a different native category, and discrimination is expected to be excellent* » (Best *op. cit.* p. 194)

[419] *Single-category assimilation :* « *Both non-native sounds are assimilated to the same native category, but are equally discrepant from the native "ideal", that is, both are equally acceptable or both equally deviant. Discrimination is expected to be poor (although it may be somewhat above chance level).* » (Best *op. cit.* p. 194)

[420] *Category-goodness difference* : « *Both non-native sounds are assimilated to the same native category, but they differ in discrepancy from native "ideal" (e.g., one is acceptable, the other deviant). Discrimination is expected to be moderate to very good, depending on the magnitude of difference in category goodness for each of the non-native sounds.* » (Best *op. cit.* p. 194)

Si l'on considère la notion de l'effet « magnétique » de la perception (*Perceptual Magnet Effect*, *Native Language Magnet model*) de Kuhl et ses collègues (Kuhl *et al.* 1992[421], Kuhl 2000[422]), le /u/ et le /ø/ se trouveraient trop proches du protoype du /u/ [ɯ] japonais pour que les japonophones natifs (les non apprenants, du moins) puissent les distinguer perceptivement, mais le /y/ serait suffisamment éloigné du prototype du /u/ [ɯ] japonais pour être distingué du /u/ et du /ø/ français.

**/u/ et /ø/** Les résultats de l'Expérience 9 (perception de stimuli synthétisés) suggèrent que le /u/ français et le /u/ [ɯ] japonais sont en effet phonétiquement différents sur le plan perceptif, au sens où les auditeurs francophones et japonophones ont identifié comme leur /u/ respectif des zones articulatoires et acoustiques différentes, même s'il y a des zones communes (autour du /u/ français, mais avec un F2 pas trop bas : 800-900 Hz).

Il existe des zones que les auditeurs francophones ont catégorisées comme /ø/ (ou /œ/) mais que les auditeurs japonophones ont considérées majoritairement comme leur /u/ [ɯ]. Cela suggère qu'au moins une partie des zones articulatoires et acoustiques du /ø/ français correspondent à celles du /u/ [ɯ] japonais, et que le /ø/ français et le /u/ [ɯ] japonais sont effectivement phonétiquement similaires sur le plan perceptif.

**Le /u/ [ɯ] japonais** Concernant le /u/ [ɯ] japonais, comme nous l'avons vu dans le profil articulatoire d'Uemura (1990 : reproduit à la Figure 11, p. 45), le lieu d'articulation (dans le sens traditionnel du terme : la position de la partie de la langue plus élevée) n'est pas complètement postérieur, et les lèvres ne sont pas complètement étirées, dans le contexte isolé, du moins. Les résultats de notre Expérience 9 (perception de stimuli synthétisés) suggèrent effectivement que sur le plan perceptif, les auditeurs japonophones n'entendraient pas du /u/ [ɯ] japonais quand la langue se trouve en arrière (comme le /u/ français) avec les lèvres ouvertes et non protruses (configuration qui correspondrait à la définition de la voyelle cardinale numéro 16 [ɯ]) : la réponse dominante pour le dernier stimulus ST98 de cette série était le /a/, mais avec une mauvaise note (1,8 : Figure 119, p. 236). En revanche, les stimuli de la série 2 (avancement de la langue) ont vu un maintien de bonnes notes pour /u/ [ɯ] sur la quasi-totalité de la série. Ce résultat suggère que les auditeurs japonais accepteraient comme /u/ japonais un [ɯ ~ ʉ] prononcé avec la langue assez avancée (jusqu'au niveau du /ø/ français), mais sans trop de délabialisation. Il nous conviendrait donc de considérer le /u/ japonais comme une voyelle centrale-postérieure, avec les lèvres légèrement arrondies, quand il est prononcé dans le contexte isolé.

Nous avons vu des données acoustiques du /u/ [ɯ] japonais (Sugitô 1995[423] à la Figure 6, p. 38 ; Mokhtari et Tanaka 2000[424] à la Figure 8, p. 41 et Figure 9, p. 42) et observé que le F2 était très élevé par rapport au /u/ français (supérieur à 1000 Hz), et que le F1 tend à être plus élevé que

---

[421] P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens, B. Lindblom. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255 : 606 – 608, 1992.

[422] P. K. Kuhl. A new view of language acquisition. *Proceedings of the National Academy of Science USA* 97 : 11850-7, 2000.

[423] M. Sugitô. *Ôsaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [CD-ROM Accent dictionary of Spoken Osaka and Tokyo Japanese]. Tokyo, Maruzen, 1995.

[424] P. Mokhtari, K. Tanaka. A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66, 2000.

celui du /i/ japonais. L'aspect perceptif de cette valeur élevée de F2 pourrait être expliqué par les résultats des stimuli de VTCalc utilisés dans l'Expérience 9 (perception de stimuli synthétisés) : nous pouvons voir que l'avancement de la langue (Tableau 57, p. 223) et la délabialisation (Tableau 58, p. 226), qui caractérisent le /u/ [ɯ] japonais par rapport au /u/ français, contribuent à augmenter le F2. Concernant le F1, les stimuli de VTCalc montrent également que l'effet de délabialisation (ouverture des lèvres et absence de protrusion) contribue à une hausse de F1 (Tableau 58, p. 226). De plus, la constriction moins étroite du /u/ [ɯ] japonais que celle du /u/ français (Figure 11, p. 45) contribuerait à une augmentation de F1.

**Tableau 62** : Principales caractéristiques articulatoires et acoustiques du /u/ [ɯ] japonais, du /u/ français, de la voyelle cardinale numéro 16 [ɯ], du /ø/ français, du /y/ français, et du /uː/ américain.

| | **Position de la langue** | **Lèvres** | **Constriction** | **Acoustique** | |
|---|---|---|---|---|---|
| /u/ [ɯ] jp | centrale-postérieure | légèrement arrondies | moins fermée que [u] fr | F2 > 1000 Hz, non focal | N.B. F1 /u/ jp > F1 /i/ jp |
| /u/ [u] fr | postérieure | arrondies | fermées | F2 < 1000 Hz, F1/F2 focal | N.B. F1 /u/ fr = F1 /i/ fr |
| [ɯ] : cadinale no. 16 | postérieure | étirées | fermées | F2 > 1000 Hz (Jones, Ladefoged, Wells) | |
| /ø/ fr | antérieure | arrondies | mi-fermée | F2 vers 1500 Hz, non focal | |
| /y/ fr | antérieure | arrondies | fermée | F2/F3 vers 1900 Hz, focal | N.B. F1 /y/ = F1 /i/ |
| /uː/ américain | centrale-postérieure | arrondies, mais moins que /u/ fr | fermée | F2 descend vers 1000 Hz | souvent diphtonguée |

**La production de /u/, /y/, et /ø/** Les résultats de l'Expérience 7 (production de 47 apprenants : Figure 73 – Figure 76, pages 194 – 197) suggèrent que les apprenants japonophones apprendraient à prononcer correctement le /ø/ plus facilement que le /y/, qui serait plus facile que le /u/. Ceux de l'Expérience 8 sur l'identification et l'évaluation de ces voyelles prononcées par des apprenants japonais auprès des auditeurs francophones (Figure 87, p. 210 – Figure 96, p. 216) semblent confirmer cette tendance.

***/ø/*** Le /ø/ est caractérisé par une distribution équilibrée des formants, avec le F2 localisé vers 1500 Hz (Figure 41, p. 123. *Cf.* les données de CALLIOPE 1989 : 84[425], représentées à la Figure 24, p. 81, et celles de Gendrot et Adda 2004[426], représentées à la Figure 25, p. 81), tout comme les voyelles postérieures non-arrondies (Vaissière 2006 : 73[427]). Nos Expériences 8 (perception de

[425] CALLIOPE. *La parole et son traitement automatique*. Masson, Paris, Milano, Barcelona, Mexico, 1989.

[426] C. Gendrot, M. Adda-Decker. Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande. *Proceedings of the Workshop MIDL*, Paris, France, 7-12, 2004.

[427] J. Vaissière. *La phonétique*. Paris, Presses Universitaires de France, 2006.

francophones) et 9 (perception de stimuli synthétisés) semblent confirmer cette caractéristique sur le plan perceptif[428]. Les apprenants devraient donc produire pour /ø/, sur le plan articulatoire, une voyelle avec la langue avancée (Expérience 9 : série 2 des stimuli de VTCalc : Figure 123, p. 240), ou avec une délabialisation (série 3 de VTCalc : Figure 124, p. 241), ou avec une combinaison des deux (série 1 de VTCalc : Figure 122, p. 239) par rapport au /u/ français, et il se trouve qu'en le faisant le lieu d'articulation s'approche à celui du /u/ japonais. Sur le plan acoustique, des productions avec le F1 vers 400 Hz et le F2 vers 1500 Hz seraient considérées comme de bons exemplaires de /ø/ par les francophones (série 2 de Klatt : Figure 126, p. 244), mais le F1 pourrait être plus bas pour identifier cette voyelle, même si l'évaluation est moins bonne (vers 300 Hz : série 1 de Klatt : Figure 125, p. 242). Ce sont des conditions qui sont similaires à celles requises pour le /u/ [ɯ] japonais : le F2 pourrait être également plus bas pour cette voyelle japonaise. Nous pourrons donc considérer le /ø/ français et le /u/ [ɯ] japonais comme similaires phonétiquement, ce qui explique la bonne identification du /ø/ produit par les apprenants (Expérience 8 : Figure 94, p. 215) et le mauvais résultat pour le /u/ (perçu plutôt comme /ø/ : Figure 87, p. 210) auprès des auditeurs francophones.

Le lieu d'articulation n'est évidemment pas le même pour le /u/ [ɯ] japonais, qui serait caractérisé comme une voyelle centrale-postérieure fermée avec un arrondissement faible, et pour le /ø/, décrit comme une voyelle antérieure mi-fermée arrondie, mais les résultats de nos expériences suggèrent que le produit acoustique et le résultat perceptif des deux voyelles seraient assez ressemblants. Cela nous semble naturel si l'on considère la remarque de Vaissière (2006, *op. cit.* : 73) sur la distance acoustique très réduite entre les voyelles antérieures arrondies et les voyelles postérieures non arrondies.

D'autres exemples de prononciation incorrecte du /ø/ incluent un [y] ou [jɯ], d'une part, et un [o], de l'autre. Nous avons observé dans l'Expérience 4 (production de 3 apprenants) que les apprenants 1 et 2 ont prononcé cette voyelle avec F2 et F3 proches dans la tâche de lecture, mais ce n'était pas le cas dans la tâche de répétition (F2 était plus bas : Figure 60, p. 144 pour l'apprenant 1 et Figure 61, p. 147 pour l'apprenant 2), et ce sont principalement ces stimuli de lecture, produits avec un rapprochement de F2 et F3, qui ont été perçus majoritairement comme /y/ par les auditeurs francophones dans l'Expérience 8 (perception de francophones : Figure 94, p. 215, Figure 95, p. 216). Le fait que la voyelle ait été prononcée plus correctement dans la tâche de répétition suggère qu'il s'agirait principalement de la difficulté de la phonie-graphie, due à l'association de l'orthographe <eu> avec le /juː/ de l'anglais, qui serait considéré comme un son proche du /y/ français, même si le facteur perceptif n'est pas complètement à exclure étant donné que l'identification et la distinction perceptives de /y/ et /ø/ ne sont pas aussi faciles que la paire /u/ et /y/ (Tableau 21, p. 129 et Figure 47, p. 131 pour l'identification de l'Expérience 3, Figure 66, p. 182 pour la discrimination AXB de l'Expérience 6). Le digraphe <eu> /juː/ de l'anglais est adapté en japonais sous la forme /ju/ ou /juR/ (ex. <euro> /juRro/ : /R/ représentant la deuxième more d'une voyelle longue). Dans les classes de français langue étrangère au Japon,

[428] Cet aspect a été présenté dans Kamiyama (2009).

T. Kamiyama. Acoustic similarities between front rounded and back unrounded vowels as evidenced by French /ø/ and /u/ produced by Japanese-speaking learners. *Journal of the Acoustical Society of America* 125(4) : 2756, 2009.

nous observons souvent des apprenants qui utilisent cette séquence pour l'orthographe <eu> lors d'une tâche de lecture. La production [jɯ], observée pour /ø/ dans l'Expérience 7 (production de 47 apprenants), pourrait être expliquée par la même difficulté de la phonie-graphie (Figure 59, p. 142).

Dans cette même expérience, nous avons également observé des apprenants qui ont produit pour /ø/ un [o] ou [ɔ] avec un F2 bas (vers 1000 Hz ou inférieur : Figure 73 – Figure 76, p. 194 – p. 197). Cette confusion serait due à la mauvaise interprétation du symbole phonémique /ø/. Aucun des trois apprenants étudiés dans l'Expérience 4 (production de 3 apprenants : Figure 60 – Figure 62, p. 144 – p. 149) a produit un F2 bas, proche de, ou inférieur à 1000 Hz pour le /ø/ dans la tâche de répétition (et de lecture également). Sur le plan perceptif, le /ø/ et le /œ/ n'ont jamais été identifiés comme /o/ ou /ɔ/ par les 5 apprenants dans l'Expérience 3 (identification des voyelles orales : Tableau 23, p. 132). Nous pouvons donc supposer que cette confusion ne serait pas liée à la perception. En revanche, nous ne pouvons pas complètement exclure le facteur de difficulté purement articulatoire : il est possible de supposer que les apprenants auraient bien réalisé la labialisation, mais avec la langue en arrière (notons que le /o/ est la voyelle la plus arrondie du japonais, comme nous pouvons le constater dans les profils sagittaux d'Uemura 1990[429]).

**/y/** Nous avons vu que le /y/ est prononcé par les natifs avec un regroupement de F2 et F3 vers 1900 Hz (pour les hommes : Figure 41, p. 123 de notre étude, Vaissière 2007[430]. *Cf.* les données de CALLIOPE 1989 : 84, représentées à la Figure 24, p. 81, et celles de Gendrot et Adda 2004, représentées à la Figure 25, p. 81). Cette voyelle a été produite par les apprenants principalement soit avec un rapprochement de F2 et F3, soit comme la séquence /ju/ [jɯ] du japonais (ou parfois comme une voyelle acoustiquement centrale sans rapprochement F2/F3, ou encore avec un F2 très élevé comme un [i] anglais) par les apprenants japonophones (Figure 73 – Figure 76, p. 194 – p. 197). Les apprenants étaient moins nombreux à prononcer le /y/ avec un rapprochement de F2 et F3 sans diphtongaison qu'à produire le /ø/ correctement (F2 vers 1500 Hz), mais tout de même plus nombreux qu'à produire le /u/ avec un F2 bas, inférieur à 1000 Hz. Les résultats de l'Expérience 8 (perception de francophones : Figure 90 et Figure 91, p. 212 – p. 213) montrent que les stimuli prononcés avec un rapprochement de F2 et F3 (sans que le F2 soit trop élevé) ont été effectivement bien identifiés comme /y/ avec de bonnes notes. L'Expérience 9 (perception de stimuli synthétisés) a également montré que les stimuli avec un F1 bas (300 Hz), et un rapprochement de F2 et F3 (1800 – 2000 Hz, et 2350 Hz respectivement) ont été catégorisés majoritairement comme /y/ avec des notes relativement élevées.

Parmi les cas de prononciation déviante, les similitudes entre le /y/ français et le /ju/ [jɯ] japonais seront discutées plus loin. Concernant les cas d'un F2 trop élevé, une des explications possibles serait la confusion entre le symbole phonémique /y/ et l'orthographe <y> /i/. L'autre

---

[429] 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]. Tokyo, Shûei shuppan, 1990.

[430] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

serait de l'ordre articulatoire : les apprenants ont visé un regroupement de F2 et de F3, mais cet effort a donné lieu à un F2 trop élevé, comme pour le /i/ anglais, pour que la voyelle produite soit perçue comme /y/. En revanche, il est peu probable que cette confusion soit due à la perception. Rappelons qu'aucun cas de confusion n'a été observé entre /i/-/y/ : les 5 apprenants qui ont participé au test d'identification (Expérience 3) n'ont jamais confondu /i/ et /y/ dans ni l'un ni l'autres des sens (Tableau 22, p. 130).

**/u/** La voyelle /u/ française, caractérisée par un regroupement des deux premiers formants en dessous de 1000 Hz (Figure 41, p. 123 de notre étude, Vaissière 2007, *op. cit*. *Cf*. les données de CALLIOPE 1989 : 84, représentées à la Figure 24, p. 81, et celles de Gendrot et Adda 2004 représentées à la Figure 25, p. 81). Aucun des trois apprenants qui ont participé à l'Expérience 4 (production de 3 apprenants) n'a produit une telle voyelle dans ni l'une ni l'autre des deux tâches (lecture et répétition : Figure 60, p. 144 – Figure 62, p. 149). Dans l'Expérience 7 (production de 47 apprenants), seuls quatre apprenants sur 47 ont prononcé le /u/ avec un F2 inférieur à 1000 Hz (Figure 73, p. 194 – Figure 76, p. 197). Les résultats de l'Expérience 8 (perception de francophones) montrent que seuls les stimuli avec un F2 inférieur à 1000 Hz ont effectivement été identifiés majoritairement comme /u/ avec de bonnes notes (Figure 87, p. 210 et Figure 88, p. 211). Ces données corroborent celles de l'Expérience 9 (perception de stimuli synthétisés) au sens où les stimuli synthétisés qui avaient le F2 supérieur à 1000 Hz n'ont pas été reconnus majoritairement comme /u/. Ce cas de figure contrasterait très clairement avec le cas du /uː/ de l'anglais américain appris par les apprenants japonophones, étant donné que le /u/ [ɯ] japonais est catégorisé par des locuteurs américains comme /uː/ de l'anglais américain (Nishi *et al*. 2008[431]).

Les cas de prononciation déviante consistent majoritairement en un vocoïde avec un F2 supérieur à 1000 Hz, qui serait perçu majoritairement comme /ø/ par les auditeurs francophones. Une autre tendance fréquemment observée dans la tâche de lecture de l'Expérience 7 (production de 47 apprenants : Figure 73, p. 194 – Figure 76, p. 197) est la confusion de la transcription /u/ avec l'orthographe <u> /y/. Nous avons ainsi eu des apprenants qui ont prononcé un [y] ou un [jɯ] à la japonaise (Figure 59, p. 142).

La difficulté de production de cette voyelle serait expliquée par des raisons perceptive et articulatoire. Comme nous l'avons vu ci-dessus, les apprenants japonophones ont des difficultés à identifier le /u/ et de distinguer cette voyelle du /ø/. Par ailleurs, le F2 du /u/ était plus bas, donc plus proche de la cible native (tout en restant au dessus de 1000 Hz) dans la tâche de lecture que dans celle de répétition chez les apprenants 1 et 2 de l'Expérience 4 (production de 3 apprenants : Figure 60, p. 144 et Figure 61, p. 147). Il est possible que ces apprenants aient été conscients des caractéristiques articulatoires et acoustiques du /u/ en le prononçant dans la tâche de lecture, mais que la confusion perceptive entre /u/ et /ø/ les ait empêchés de faire de même dans la tâche de répétition immédiate. Sur le plan de la production, comme la cible articulatoire (/u/ français) est assez éloignée du /u/ [ɯ] japonais, les apprenants ne seraient pas habitués aux gestes articulatoires de l'arrondissement des lèvres et du recul de la langue.

---

[431] K. Nishi, W. Strange, R. Akahane-Yamada, R. Kubo, S. A. Trent-Brown. Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America* 124(1) : 576-588, 2008.

**Nouveau ou similaire ?** Comparons maintenant nos résultats avec ceux de Flege (1987)[432], qui a montré que des locuteurs natifs de l'anglais américain ont produit le /y/ avec des valeurs formantiques plus proches de celles des locuteurs francophones natifs que pour le /u/ (F2 était trop élevé). Ce cas de figure ressemble à nos résultats. Selon les hypothèses de son Modèle d'apprentissage de la parole (SLM : *Speech Learning Model* : voir Flege 1995[433] et le chapitre 1.3. de notre étude), une nouvelle catégorie sera établie si les bilingues discernent au moins quelques-unes des différences phonétiques entre les sons de L1 et de L2 (Hypothèse 2)[434], et plus la dissimilarité phonétique perçue entre un son de L2 et le son le plus proche de L1 est grande, avec plus de chance les différences phonétiques entre les sons seront discernées (Hypothèse 3)[435]. En utilisant les termes de Flege, une nouvelle catégorie a été établie pour le son « nouveau » /y/ au cours de l'apprentissage, qui est, plus que le /u/ français, écarté du /uː/ américain, alors que la similitude (relative : notons que le /uː/ anglais a un F2 élevé, et est diphtongué) du /u/ français (un son « similaire ») avec le /uː/ américain a empêché aux locuteurs anglo-américains d'établir une nouvelle catégorie /u/ française.

Suivant le même principe, nous pouvons interpréter les résultats de nos expériences comme suit : la nouvelle catégorie /y/, qui est, plus que le /u/ français, éloignée du /u/ [ɯ] japonais, aurait été établie au cours de l'apprentissage. En revanche, la catégorie du /u/ français, qui est « similaire », du moins sur le plan perceptif des apprenants, comme nous l'avons vu ci-dessus, n'a pas été établie, sauf un petit nombre d'entre eux (4 sur 47, notamment ceux qui avaient vécu en France quand ils étaient plus jeunes).

Dans ce cas-là, comment pourrait-on caractériser le /ø/ ? Sur le plan phonétique (acoustique et perceptif, du moins), cette voyelle serait la plus proche du /u/ [ɯ] japonais, et les apprenants japonophones ont mieux réussi à produire un vocoïde proche de la cible acoustique et perceptive du /ø/ que pour /u/ français. En revanche, nous avons observé des cas où ils confondaient avec d'autres voyelles, notamment avec /y/ (ce qui a donné lieu à un [y] ou un [jɯ]), dans la tâche de lecture, ce qui suggère qu'ils n'ont pas nécessairement bien associé la valeur phonétique à la catégorie phonémique.

C'est pour ces raisons que nous proposons de considérer la similitude phonétique et la conscience phonémique des apprenants comme deux facteurs différents. Nous examinons dans ce travail de thèse l'apprentissage de la prononciation en milieu institutionnel, et nous supposons que le rôle des notions métalinguistiques peut jouer un rôle plus important que dans l'acquisition en milieu naturel. Les apprenants sont confrontés quotidiennement, dans le cadre de leur apprentissage, à l'orthographe, qui reflète au moins en partie des suites de phonèmes, à la transcription phonémique de mots dans des dictionnaires (même s'ils ont tendance à se concentrer

---

[432] J. E. Flege. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15 : 47-65, 1987.

[433] J. E. Flege. "Second language speech learning: Theory, findings, and problems". in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 233-277, 1995.

[434] « *H2 A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds*. » Flege *op. cit.* (p. 239)

[435] « H3 *The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned*. » Flege *op. cit.* (p. 239)

plutôt sur le sens ou la traduction), ou à des remarques de leurs enseignants en classe. Même s'ils ne sont pas toujours conscients des catégories phonémiques des voyelles de la langue cible, la plupart des apprenants auraient appris de l'existence des phonèmes vocaliques qui sont transcrits avec des symboles qui ne sont pas utilisés pour transcrire les 5 voyelles du japonais, dans des manuels d'apprentissage (voir la quatrième partie) ou en classe. La conscience des phonèmes qui n'ont pas d'équivalent dans la L1 (ex. /y/ et /ø/, représentés avec des symboles autres que les 5 utilisés pour le système du japonais : /i e a o u/), pourrait faciliter l'apprentissage de ces voyelles dans une certaine mesure, malgré l'écart phonétique qui pourrait éventuellement se trouver entre la voyelle cible et la voyelle la plus proche dans le système de L1 ; les voyelles qui sont phonétiquement proches, comme le /ø/, seraient naturellement plus faciles que celles qui sont phonétiquement différentes, comme le /y/. En revanche, concernant les voyelles qui sont transcrites avec les mêmes symboles que celles de L1 (/i e a o u/), si l'écart phonétique entre la cible en L2 (/u/ français) et le phonème vocalique correspondant, transcrit avec le même symbole (/u/ japonais), la prise de conscience serait plus difficile, et cela rendrait l'apprentissage de ces voyelles encore plus difficile que celles qui sont pas phonémiquement nouvelles (ex. /y/ et /ø/). Il reste un autre cas de figure : les voyelles qui sont similaires phonémiquement, et la cible en L2 n'est pas très écartée du phonème de L1 (/o/, et éventuellement le /i/, si l'apprenant prononce le /i/ japonais comme un [i] avec un rapprochement de F3 et F4), les apprenants parviendraient facilement à produire ces voyelles avec les cibles des locuteurs natifs. La Figure 129 résume les quatre cas de figure décrits ci-dessus.

Difficulté de production →

| | /a/ | /ø/ | /y/ | /u/ |
|---|---|---|---|---|
| Phonémiquement similaire | + | - | - | + |
| Phonétiquement similaire | + | + | - | - |

**Figure 129 :** La hiérarchie de difficulté de production observée chez les apprenants japonophones étudiés dans la présente étude, selon deux caractéristiques : similarité phonémique et similarité phonétique.

Cette hiérarchie de difficulté permet d'expliquer les tendances observées dans notre série d'expériences, mais cette proposition reste évidemment à vérifier avec d'autres cas d'études qui portent sur l'apprentissage des langues étrangères ou secondes en milieu institutionnel. Dans les cas d'autres langues cibles, la difficulté que les apprenants japonophones rencontrent dépendra naturellement du système vocalique cible. Les apprenants japonophones seraient confrontés à la même difficulté du /u/ lors de l'apprentissage de l'allemand (/u/, /ʊ/ et /ø/ sont distincts), ou du coréen (/u/ et /ɯ/ sont en opposition). En revanche, en ce qui concerne des langues comme l'espagnol (5 voyelles) ou l'italien (7 ou 5 voyelles selon la région), qui n'ont pas d'autres voyelles qui ressemblent acoustiquement au /u/ [ɯ] japonais, le /u/ prononcé à la japonaise serait probablement interprété correctement comme /u/, même s'il ne correspond pas à un bon exemplaire de la catégorie.

Il nous restera également à examiner la conscience phonémique explicite des apprenants afin de confirmer véritablement cette classification proposée. Dans la quatrième partie, une analyse de méthodes d'apprentissage de français langue étrangère publiées au Japon sera présentée pour éclaircir des descriptions et des explications que les apprenants rencontrent typiquement dans leur apprentissage.

Nous développerons par la suite trois autres sujets de discussion concernant ces trois voyelles /u y ø/.

**/u/ produit par les apprenants japonophones et le [ʊ] relâché canadien** La voyelle /u/ produite par des apprenants japonais dans notre série d'expériences a été identifiée par des auditeurs francophones français principalement comme /u/ si le F2 était inférieur à 1000 Hz, comme /u/ ou /ø/ mais avec une mauvaise évaluation si le F2 se trouvait entre 1000 Hz et 1100 Hz, comme /ø/ si le F2 était plus élevé (Expérience 8, perception de francophones : Figure 87, p 210 et Figure 88, p 211). Or, les francophones québécois disposent de l'allophone [ʊ], qui apparaît dans les syllabes fermées (ex. « soupe » /sup/ [sʊp] : voir Ostiguy *et al.* 1996[436]), et qui est prononcé avec F1 et F2 plus élevés que ceux du [u] qui apparaît dans les syllabes ouvertes (Martin 2002[437] : Figure 121 ci-dessous). Nous pouvons donc supposer que les locuteurs natifs du français québécois accepteraient plus facilement le /u/ prononcé par les apprenants japonophones avec F1 et F2 plus élevés que les valeurs attendues chez les locuteurs du français parisien. Cette hypothèse reste à vérifier dans une expérience de perception effectuée auprès d'auditeurs québécois dans divers contextes syllabiques, car nous ne pouvons savoir si les auditeurs acceptent un vocoïde qui ressemblerait à un [ʊ] dans une syllabe ouverte (ou en contexte isolé). Le résultat d'une étude préliminaire (Tremblay et Kamiyama 2009)[438] montre que les auditeurs québécois donnent de meilleures notes que les auditeurs français aux stimuli du /u/ prononcés par des apprenants japonophones avec un F2 entre 1000 Hz et 1100 Hz.

---

[436] L. Ostiguy, R. Sarrasin, G. Irons. *Introduction à la phonétique comparée : les sons : le français et l'anglais nord-américains*. Sainte-Foy, Les Presses de l'Université Laval, 1996.

[437] P. Martin. Le système vocalique du français du Québec. De l'acoustique à la phonologie. *La linguistique* 38(2) : 71-88, 2002.

[438] M.-C. Tremblay, T. Kamiyama. Perception of L2 production by L1 speakers of different dialectal backgrounds: the case of Japanese-speaking learners' /u/ perceived by French and Quebec native speakers. *The Journal of the Acoustical Society of America* 125(4) : 2776, 2009.

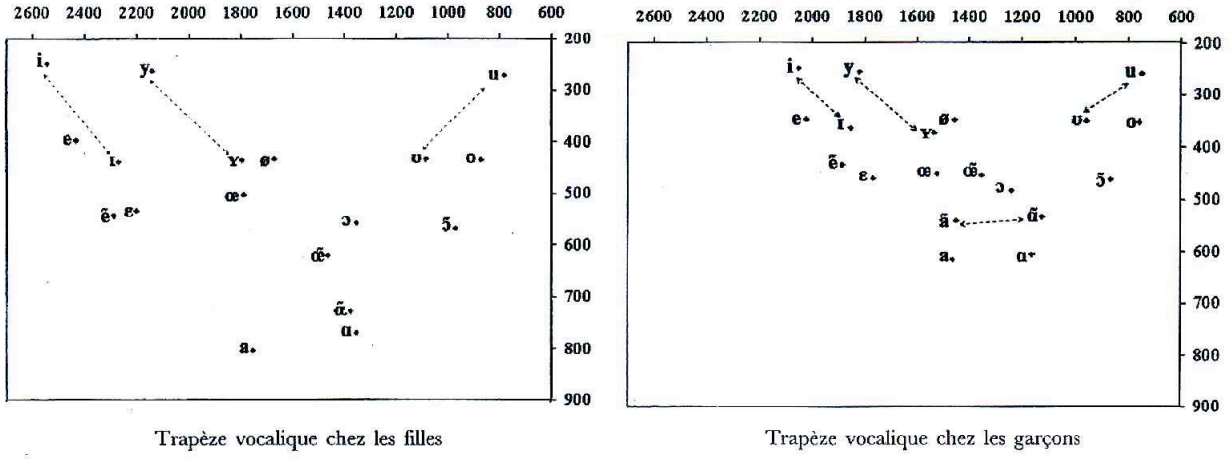



**Figure 130 :** Les valeurs moyennes des deux premiers formants en Hertz des voyelles du français québécois dans 159 mots prononcés par 6 femmes (à gauche) et 6 hommes (à droite). Les flèches réunissent deux variantes d'un même phonème (Martin 2002 : 84).

**Le /y/ français et le /ju/ [jɯ] japonais** Dans les mots d'emprunt en japonais, le /y/ français est adapté systématiquement, quel que soit le contexte consonantique, à la séquence /ju/ (ou /juR/ : /R/ représente la deuxième more d'une voyelle longue), ainsi que la séquence /jø/ du français : Hugo /'jugoR/ /'juRgoR/, Camus /'kamju/, Dumas /'djuma/, Curie /'kjuriR/ ou /'kjuRriR/), et cela correspond à ce que l'apprenante 3 a produit dans l'Expérience 4 (production de 3 apprenants : Figure 55, p. 139). Nous pouvons interpréter cette adaptation comme combinant 2 éléments (ou traits) du /y/ français (antériorité et labialité), en cumulant l'antériorité du premier élément /j/, et la labialité du second élément /u/ (même si le /u/ [ɯ] japonais n'est pas aussi labialisé que le /u/ français).

Le japonais n'est d'ailleurs pas la seule langue qui interprète le /y/ comme séquence de deux phonèmes. Le cas particulièrement intéressant est celui du coréen. Dans cette langue, /y/ est adapté comme /wi/ (ex. "buffet" : /pwip$^{h}$e/[439] en coréen, /'bjuQhwe/[440] en japonais, "Bruxelles" : /pɯrwiser/ en coréen, /bu'rjuQseru/ en japonais). Nous retrouvons, dans tous les deux cas, une combinaison de deux éléments /i/ (antériorité) et /u/ (labialité), mais avec des ordres différents, et des poids (syllabiques) différents, /i/ étant le noyau en coréen, et le /u/ en japonais. Nous pourrions représenter ce cas de figure comme à la Figure 131 ci-dessous. Cependant, il n'est pas facile de trouver une raison plausible pour cette divergence. Ce serait peut-être dû à la différence du système vocalique : le /u/ du coréen est arrondi, et il s'oppose au non-arrondi /ɯ/ (transcrit

[439] En coréen le phonème /p/ (lenis) se réalise non-voisé en début de syntagme, et que le /p$^{h}$/ est aspiré. Il est également intéressant de noter que les deux langues présentent des comportements différents concernant l'adaptation du /f/, qui n'existe dans ni l'une ni l'autre des langues : /p$^{h}$/ [p$^{h}$] en coréen, /h(w)/ [ɸ] en japonais.

[440] La raison de l'insertion du /Q/ (première more d'une consonne géminée) reste encore à éclaircir, mais selon Kubozono (2006 : 53-54), la langue préfère des syllabes HL (*Heavy-Light* : lourde-légère) à la fin du mot, et l'insertion du /Q/ dans l'avant-dernière more crée cette structure. Ce principe explique le cas de /'bjuQhwe/, et bien d'autres.

H. Kubozono. *Akusento no housoku* [Règles de l'accent]. Tokyo, Iwanami shoten, 2006.

également /ɨ/), tandis qu'en japonais il n'y a pas d'opposition arrondi/non-arrondi pour les voyelles postérieures fermées. Certes, il existe des langues qui interprètent /y/ comme /u/ en ce qui concerne la production L2 (italien, espagnol, entre autres : Lauret 2007 : 76[441], études expérimentales sur les apprenants anglophones américains de Rochet 1995[442]), mais ces langues ne possèdent pas d'opposition arrondi/non-arrondi pour les voyelles postérieures. En coréen, la séquence /j/ + /ɯ/, qui ressemblerait phonétiquement au /ju/ japonais, serait une adaptation possible, mais les contraintes phonotactiques de la langue ne la permettent pas[443]. Les exemples que nous venons de considérer ont été tirés de mots d'emprunts, et ne correspondent peut-être pas exactement à la perception réelle chez les Coréens. Il sera donc intéressant d'effectuer des expériences de production et de perception chez les Coréens, ce qui permettra de comparer les comportements des apprenants japonais et coréens.

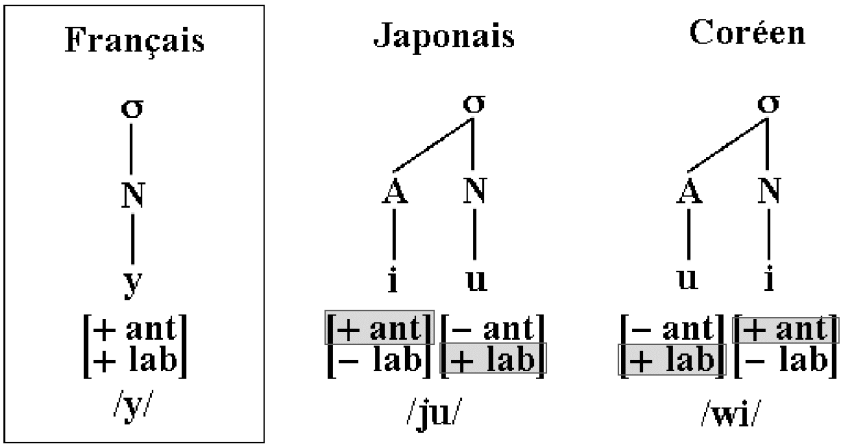



**Figure 131 :** Adaptation du /y/ français (à gauche) en japonais (au milieu) et en coréen (à droite). σ représente une syllabe, A et N, attaque et noyau de syllabe, [ant] et [lab], antériorité et labialité, respectivement.

Il serait également possible d'expliquer ce pattern d'adaptation par la ressemblance acoustique d'une portion de la séquence /ju/ [jɯ] japonaise au /y/ français. Dans la séquence /ju/ [jɯ] japonaise, on observe un croisement de F2 et F3 durant le passage du /j/ au /u/, qui serait dû au recul de la langue : la fréquence de la résonance de la cavité antérieure baisse au fur et

---

[441] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français*. Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[442] B. L. Rochet. "Perception and production of second-language speech sounds by adults", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 379-410, 1995.

[443] La semi-voyelle /w/ correspond à /u/, et il existe une autre semi-voyelle /ɯ/ qui correspond à /ɯ/ (transcrit également /ɨ/), et qui se combine uniquement avec /i/. En revanche, /ɯi/ est simplifié à /i/ quand cette séquence est précédée par une autre consonne d'attaque (ex. /hɯi/ > /hi/). Par conséquent, l'opposition entre /w/ et /ɯ/ est neutralisée dans ce contexte (Lee *et al*. 2004, entre autres). C'est probablement pour cette raison que l'on observe une réalisation antériorisée de /w/ dans /Cwi/, ce qui pourrait expliquer la similitude entre /Cwi/ en coréen et /Cy/ en français. Notons également que la prononciation historique de /wi/ est [y] (Lee *et al*. 2004).

I.-S. Lee, S.-O. Lee, W. Chai. *Kankokugo gaisetsu* [Aperçu de la langue coréenne], traduit en japonais par H. Umeda, T. Maeda. Tokyo, Taishûkan, 2004.

à mesure que la longueur de cette cavité augmente (*cf.* passage du /i/ au /y/ en français : Vaissière 2007 : Figure 6[444] pour la modélisation). Les deux formants sont ainsi proches sur une portion de la séquence /ju/ [jɯ] même si les formants ne sont pas stables (Figure 132).

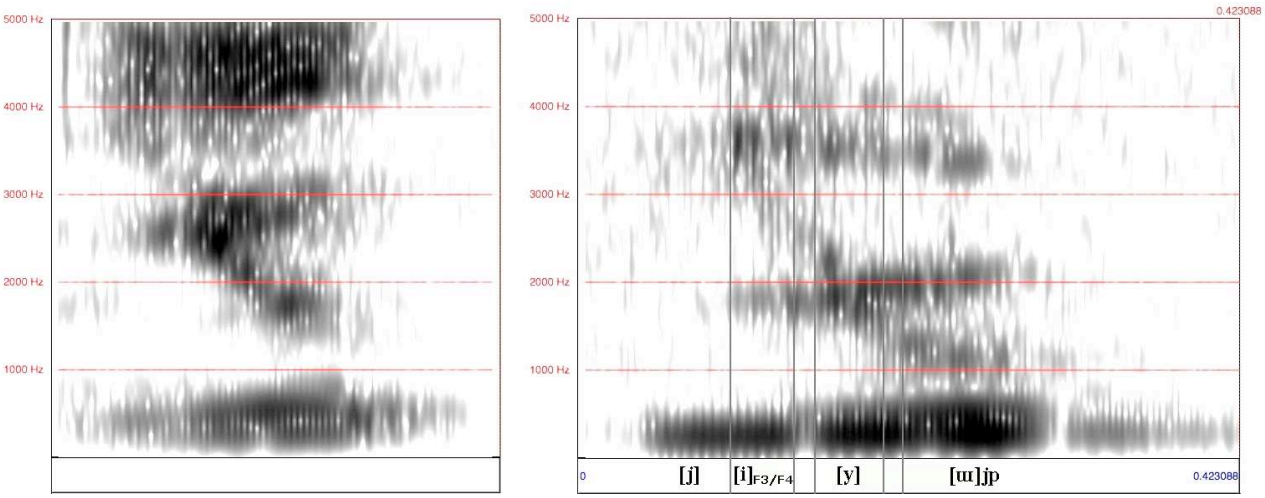



**Figure 132 :** Spectrogrammes à bande large (Praat ; largeur de la fenêtre : 5 millisecondes) du /ju/ [jɯ] japonais prononcé par une locutrice native (JP3 : à gauche) et un locuteur natif (JP1 : à droite) du japonais de Tokyo dans la phrase 「それは『ユ』といいます」 *Sorewa* /ju/ *to iimasu* « on appelle cela /ju/ ». Dans le spectrogramme de droite, six phases acoustiques peuvent être distinguées : 1) [j] ; 2) [i] F3/F4 ; 3) transition (labialisation) ; 4) [y] ; 5) transition (postériorisation) ; 6) [ɯ] japonais.

Nous pouvons voir dans les spectrogrammes de la Figure 132 que le F2 et le F3 se rapprochent autour de 2000-3000 Hz pour la femme (Figure 132 : à gauche) et de 2000 Hz pour l'homme (Figure 132 : à droite) pendant la transition du /j/ vers le /u/ [ɯ], puis s'écartent vers la partie stable du /u/ [ɯ]. Dans ce dernier exemple, six phases acoustiques peuvent être distinguées : [j], [i] F3/F4, transition (labialisation), [y], transition (postériorisation), [ɯ] japonais.

[444] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

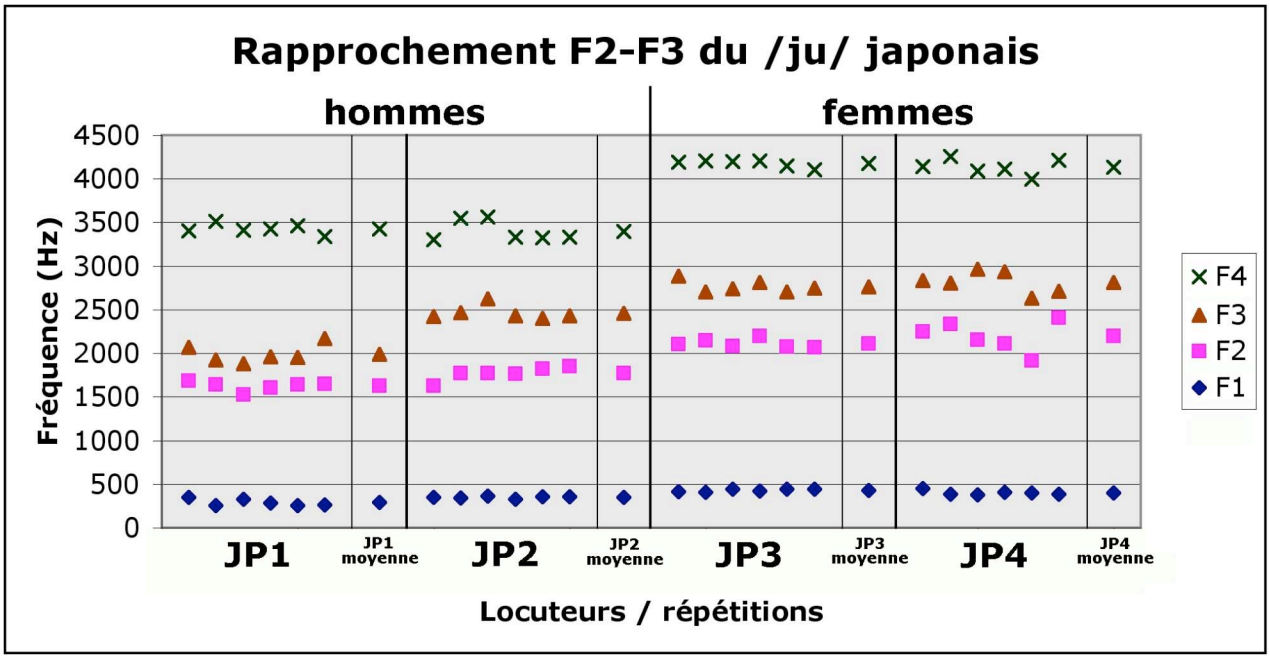



**Figure 133 :** Les valeurs des 4 premiers formants durant la transition du /j/ au /u/ (où F2 et F3 sont proches) dans la séquence /ju/ japonaise prononcée par 4 locuteurs natifs (2 femmes et 2 hommes) du japonais de Tokyo dans la phrase 「それは『ユ』といいます」 *Sorewa* /ju/ *to iimasu* « on appelle cela /ju/ ». Les valeurs moyennes de la portion de transition sélectionnée manuellement (variant de 18 millisecondes à 63 millisecondes, la moyenne étant de 35 millisecondes).

La Figure 133 montre les valeurs des quatre premiers formants, mesurés sur la portion du /ju/ [jɯ] où les F2 et F3 sont proches (Phase 4 ci-dessus). Même s'il ne s'agit que d'une petite portion de transition manuellement sélectionnée (variant de 18 à 63 millisecondes, la moyenne étant de 35 millisecondes), et que les valeurs ne sont pas stables, les locuteurs natifs du japonais produisent ainsi un vocoïde avec un rapprochement F2/F3 lors de la Phase 4.

**<u>La difficulté de la paire /u/-/ø/ : comparaison avec les auditeurs anglophones américains</u>**

La difficulté de distinction perceptive de l'opposition /u/-/ø/ que nous avons observée dans nos expériences est particulièrement intéressante à noter quand nous considérons le cas des auditeurs anglophones américains, qui manifeste plus de difficulté pour la paire /u/-/y/ (Gottfried 1984[445], Levy et Strange 2002[446], Levy 2006[447], Levy et Strange 2008[448] pour la perception, Rochet 1995[449],

[445] T. L. Gottfried. Effects of consonant context on the perception of French vowels. *Journal of Phonetics* 12 : 91-114, 1984.

[446] E. S. Levy, W. Strange. Effects of consonantal context on perception of French rounded vowels by American English adults with and without French language experience. *Journal of the Acoustical Society of America* 111(5) : 2361-2362, 2002.

[447] E. S. Levy. Consonantal context and experience effects on American L2-learners' perception of French vowels. *Journal of the Acoustical Society of America* 119(5) : 3424, 2006.

[448] E. S. Levy, W. Strange. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36 : 141–157, 2008.

Lauret 1998[450] pour la production). Entre autres, les résultats de Levy et Strange (2008 : discrimination AXB) ont montré que les auditeurs expérimentés (en français) ont mieux distingué la paire /u/-/œ/[451] (taux d'erreurs de 3%) que les auditeurs non-apprenants (taux d'erreurs de 27%), alors qu'aucun effet d'expérience d'apprentissage n'a été trouvé pour la paire /u/-/y/, qui n'a pas été bien distinguée même par les auditeurs expérimentés (taux d'erreurs de 30% contre 24% pour les auditeurs non-expérimentés). Ce cas de figure contraste très nettement avec les résultats de notre test AXB (Expérience 6, Figure 66, p. 182 et Figure 69, p. 187) : c'est la paire /u/-/ø/ qui était plus difficile à distinguer que /y/-/ø/ et /u/-/y/, quelle que soit l'expérience d'apprentissage des auditeurs ; en outre, la paire /u/-/y/ a été relativement bien distinguée même par les non-apprenants. Quelles sont les similitudes et les différences entre les deux langues, le japonais et l'anglais américain, qui nous permettraient d'expliquer ces différences de comportements ?

Comme nous l'avons vu dans la première partie, le /u/ [ɯ] japonais et le /u:/ anglais sont caractérisés tous les deux par une valeur élevée de F2 par rapport à celle du /u/ français, et les deux langues ne possèdent pas d'équivalent phonétique du /u/ français. Les locuteurs natifs des deux langues ont effectivement tendance à prononcer le /u/ français avec un F2 plus élevé que les locuteurs natifs (Expériences 4 et 7 pour les apprenants japonais, Flege 1987[452] pour les apprenants américains).

La réalisation diphtonguée du /y/ est un phénomène également observé chez les apprenants anglophones américains. Lauret (1998, *op. cit.*) a montré que les locuteurs natifs de l'anglais américain (GA : *General American*) apprenant le français avaient tendance à produire pour le /y/ français un vocoïde diphtongué, avec un F2 descendant, (Figure 38, p. 113, à comparer avec la Figure 55, p. 139) comme le /u:/ américain. Nous observons néanmoins une différence entre le /y/ français envisagé par les apprenants américains et le /u:/ américain prononcés par les mêmes locuteurs : dans le premier, le rapprochement de F2/F3 est plus nettement marqué au début de la voyelle, et la durée est plus courte, sans la partie stable observée à la fin du /u:/ américain.

Il est d'ailleurs intéressant de noter que le cas inverse n'est pas inexistant, c'est-à-dire que la séquence /ju:/ de l'anglais pourrait se réaliser comme [y] par certains francophones. La chanteuse Céline Dion, native du français canadien, prononce occasionnellement le mot <*you*> de l'anglais avec un [y] prolongé, même s'il est suivi d'une voyelle plus relâchée, comme le témoigne le spectrogramme de la Figure 134 ci-dessous. Ce n'est probablement pas un phénomène très répandu, mais ce serait au moins considéré comme une illustration de la similitude entre le /ju:/ anglais et le /y/ français.

---

[449] B. L. Rochet. "Perception and production of second-language speech sounds by adults", in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 379-410, 1995.

[450] B. Lauret. *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français.* Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle, 1998.

[451] Dans les logatomes testés dans cette étude, la voyelle /œ/ se trouvait dans les syllabes /b_p/ et /d_t/.

[452] J. E. Flege. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15 : 47-65, 1987.

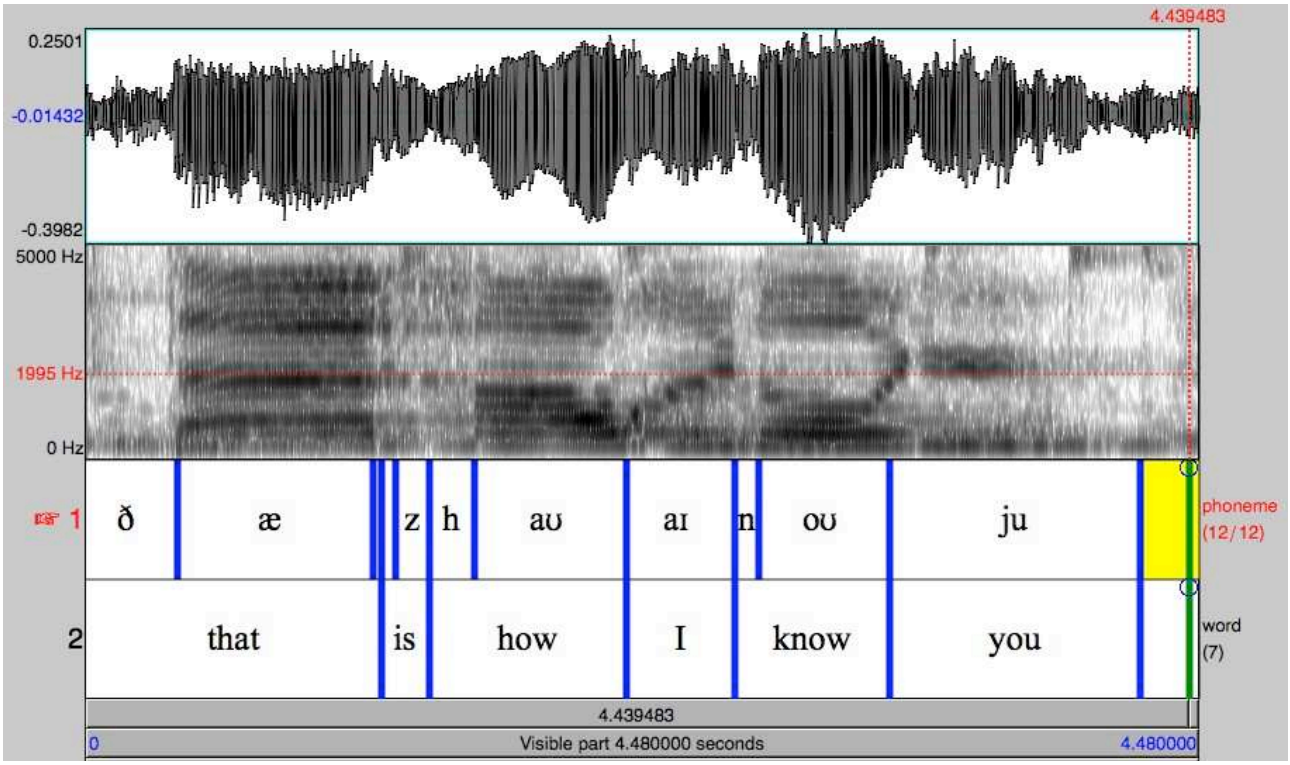



**Figure 134 :** Spectrogramme à bande large sous Praat (au milieu ; longueur de la fenêtre : 5 millisecondes) de la phrase « *that is how I know you (go on)* », dans la chanson « *My Heart Will Go On* », chantée par Céline Dion, et enregistrée dans l'album « *Au cœur du stade* » (enregistrement live du concert qui a eu lieu les 19 et 20 juin 1999 au Stade de France), Columbia, 1999. Nous observons un [y] prolongé, avec un regroupement de F2 et F3 pour le mot <you> /juː/[453].

Le /y/ diphtongué prononcé par l'apprenante japonaise (apprenante 3) dans notre étude ressemblerait au /y/ français prononcé par les apprenants américains. Ces deux types de réalisation, un par les japonophones, l'autre par les anglophones américains, sont ressemblants sur le plan phonétique, mais du point de vue phonémique, nous pourrions leur donner deux interprétations différentes. Pour les japonophones, c'est la seule voyelle cible pour laquelle certains apprenants ont produit une diphtongaison, et ce serait dû à l'existence de la séquence /ju/ du japonais, qui est constituée de deux éléments qui comportent respectivement l'antériorité et l'arrondissement (même si le /u/ [ɯ] japonais n'est pas aussi labialisé que le /u/ français). En revanche, pour ce qui est des apprenants américains, il ne s'agit qu'une des réalisations diphtonguées des cibles monophtongues (/e/ /o/ du français prononcés [eɪ] [oʊ], entre autres), comme le montre Lauret (1998, *op. cit.*), à cause du transfert des tendances diphtonguées des phonèmes correspondants en L1 (/eɪ/ /oʊ/).

Cette tendance générale de diphtongaison constitue en effet une différence importante qui se trouve entre l'anglais et le japonais. En anglais américain, c'est non seulement la séquence /juː/ mais aussi la voyelle isolée /uː/ qui présentent un changement de formants au cours de la séquence ou de la voyelle, alors qu'en japonais, c'est uniquement les séquences qui contiennent /j/

[453] C'est la forme forte /juː/, au lieu de la forme faible /ju/, qui est utilisée ici, car la note est prolongée, et suivie par une pause.

(dont /ju/ [jɯ]), le /u/ se réalisant comme un vocoïde monophtongue, comme nous l'avons observé dans l'Expérience 1 (voyelles du japonais).

En outre, l'opposition entre le /juː/ et le /uː/ de l'anglais (américain) pourrait être considérée comme étant partiellement neutralisée suite à la chute du /j/ (« *jod dropping* ») de <u> /juː/. Wells (2000)[454] décrit les cas de la chute du /j/ comme suit[455] :

- *après les consonnes* /tʃ, dʒ, ʃ, r, j/
- *quelques fois en anglais britannique, et toujours en anglais américain, après* /l, □, s, z/
- *généralement en anglais américain, mais pas en anglais britannique, après* /t, d, n/[456]

Nous pouvons trouver des paires minimales entre <u> /juː/ qui n'a pas subit la chute du /j/ et <oo> /uː/ (ex. *queue, cue, Kew* /kjuː/ vs. *coo* /kuː/). Cependant, dans les contextes où la chute du /j/ a eu lieu, nous observons une neutralisation (ex. *tune* /tuːn/ en anglais américain, *toon* /tuːn/ ; cf. *tune* /tjuːn/ ou /tʃuːn/ en anglais britannique).

Nous pouvons donc supposer que ce phénomène de neutralisation partielle a créé une similitude entre /uː/ et /juː/, ou inversement, que la chute du /j/ a été provoquée par la similitude entre /uː/ et /juː/ dans les contextes palatal et coronal (le cas du contexte palatal semble être particulièrement naturel). Quelle que soit la relation cause-effet entre ces deux phénomènes, le contexte consonantique semble avoir une influence considérable sur la distinction perceptive de la paire /u/-/y/ du français. Les auditeurs non expérimentés américains qui ont participé à l'étude de Levy et Strange (2008[457] : discrimination AXB) ont discriminé cette paire plus facilement dans le contexte labial (/rab_p/ : taux d'erreurs de 9%) que dans le contexte dental (/rad_t/ : taux d'erreurs de 39%).

D'ailleurs, il nous restera à étudier l'influence du contexte consonantique dans le cas des apprenants (et non-apprenants) japonophones. La paire /u/-/y/, relativement bien distinguée dans le contexte isolé, serait plus difficile dans le contexte palatal, si l'on juge du pattern d'emprunt. Le recours au /j/ n'est pas disponible, car le /j/ est déjà utilisé pour interpréter les consonnes palatales (palato-alvéolaires, ou post-alvéolaires). Les syllabes /Cu/ et /Cy/ sont ainsi interprétées toutes les deux comme /Cju/ en japonais : bonjour /bõʒuʁ/ > /boNzjuRru/[458] [boõʑuːrɯ] / [bondʑuːrɯ] ; Jule /ʒyl/ > /zjuRru/ [dʑɯːrɯ] / [ʑɯːrɯ]. Les résultats préliminaires

---

[454] J. C. Wells. *Longman Pronunciation Dictionary* (Second edition). Harlow, Pearson Education Limited, 2000.

[455] « *In the case of expected* juː, jʊə, jʊ, *the* j *drops out as follows:*

*- after the consonant sounds* tʃ, dʒ, ʃ, r, j, *as in jury* ˈdʒʊər i ‖ ˈdʒʊr i, *rude* ruːd;

*- sometimes in BrE, and always in AmE, after* l, □, s, z, *as in assume* ə ˈs(j)uːm ‖ ə ˈsuːm;

*- usually in AmE, but not in BrE, after* t, d, n, *as in tune* tjuːn ‖ tuːn *(see also ASSIMILATION for the BrE possibility of* tʃuːn*).* » (p. 805)

[456] L'auteur mentionne la possibilité de la coalescence de /j/ (« *jod coalescence* ») pour t et d en anglais britannique (> /tʃ/, /dʒ/).

[457] E. S. Levy, W. Strange. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36 : 141–157, 2008.

[458] Rappelons que /R/ représente la deuxième moitié d'une voyelle longue en japonais.

d'une étude non publiée semblent confirmer cette difficulté : nous avons observé plus de cas de confusion perceptive entre les trois voyelles /u/, /y/, et /ø/ dans le contexte palatal (/ʃ, ʒ, j/ dans des logatomes /CVCVCV/) que dans d'autres contextes. Le contexte palatal poserait des difficultés également aux apprenants anglophones, si nous considérons le fait que la neutralisation partielle entre /juː/ et /uː/ due à la chute du /j/ a eu lieu après les consonnes palatales.

### 3.6. CONCLUSION PARTIELLE

Dans cette partie, nous avons examiné la perception et la production des voyelles /u/, /y/, et /ø/ par les apprenants japonophones, et remarqué les faits suivants :

1) Sur le plan perceptif, le /y/, qui serait assimilé à la séquence /ju/ du japonais à la différence des deux autres voyelles /u/ et /ø/, est le plus facile à distinguer des deux autres, alors que la paire /u/-/ø/ est plus difficile à distinguer.

2) Sur le plan de la production, le /ø/, qui est phonétiquement (acoustiquement) ressemblant au /u/ [ɯ] japonais, a été mieux prononcé que les deux autres voyelles.

3) Le /y/, qui est phonétiquement non similaire à la séquence la plus proche de la L1, /ju/ [jɯ] japonais, et phonémiquement nouveau (pas d'équivalent phonémique en japonais), a été prononcé mieux que le /u/, qui est différent phonétiquement du /u/ [ɯ] japonais malgré la similiarité phonémique (en tant que voyelle postérieure fermée).

4) Cette difficulté du /u/ serait due au fait que les apprenants ont des difficultés à prendre conscience de la différence phonétique entre le /u/ français et le /u/ [ɯ] japonais, transcrits tous les deux avec le même symbole phonémique.

Il nous reste maintenant à étudier comment les apprenants japonophones apprennent ces trois phonèmes vocaliques, /u/, /y/, et /ø/. Comment sont-ils traités dans l'enseignement de la prononciation au Japon ? Comment, si possible, pourrait-on l'améliorer ?

# Quatrième partie : l'enseignement de la prononciation des voyelles /u/, /y/ et /ø/ aux apprenants japonophones


**RÉSUMÉ**

Le but de cette partie est de d'apporter une réflexion sur la didactique de la prononciation tenant compte des résultats obtenus dans les parties précédentes.

i) **Analyse de manuels existants** : nous avons analysé 71 manuels d'apprentissage du français langue étrangère destinés aux apprenants japonophones afin d'étudier la place accordée à l'enseignement et l'apprentissage explicites des voyelles /u y ø/. Le résultat montre que : 1) les différences entre le /u/ français et le /u/ [ɯ] japonais, et la difficulté du /u/ français, en particulier, sont souvent négligées, même si la labialisation est mentionnée plus souvent que la position de la langue ; 2) la ressemblance acoustique entre le /ø/ et le /u/ japonais n'est presque jamais traitée. Cela suggère que les apprenants et les enseignants (japonophones, en particulier) sont rarement conscients de ces difficultés au cours de l'enseignement et l'apprentissage du français langue étrangère au Japon.

ii) **Illustration de quelques méthodes d'enseignement de la prononciation** : nous explorons des pistes pour favoriser la prise de conscience des apprenants concernant les difficultés de prononciation. Nous prenons comme exemple le /u/ français et le /u/ [ɯ] japonais, et illustrons quelques méthodes qui faciliteraient l'apprentissage du /u/ français : 1) méthode verbo-tonale, déjà largement utilisée (combinaison de la gravité du /u/ avec une F0 basse, de l'articulation postérieure avec des consonnes postérieures) ; 2) utilisation d'un continuum de voyelles cardinales (parvenir au [u] français graduellement en partant de [ɑ] et passant par [ɔ] et [o]) ; 3) images acoustiques (combinées avec des gestes, des couleurs) ; 4) synthétiseurs articulatoires (VTCalcs, TractSyn) ; 5) représentation spectrale ; 6) spectrogramme en temps réel (WaveSurfer).


**L'objectif de la présente partie** Les savoirs savants des spécialistes ne sont pas toujours directement applicables à l'enseignement et l'apprentissage. Ils doivent être adaptés à des objectifs de divers niveaux. Selon Cuq (dir. : 2003 :240)[459], la notion de transposition didactique a été utilisée d'abord par le sociologue Michel Verret (1975)[460], et reprise par Yves Chevallard (1985)[461]. Selon ce dernier auteur, une notion issue de l'extérieur de la sphère didactique est transformée d'abord en « *un objet enseignable (sélection, programmation), puis en un objet enseigné (présentation, explication, évaluation) et enfin éventuellement, par extension, en un objet d'apprentissage* » (cité dans Cuq dir. 2003 *op. cit.*).

Dans le cas de la phonétique (expérimentale) et l'enseignement et l'apprentissage de la prononciation, nous pourrons penser aux étapes suivantes :

---

[459] J.-P. Cuq (dir.). *Dictionnaire de didactique du français langue étrangère et seconde*. Paris, Clé International/Asdifle, 2003.

[460] M. Verret. *Le temps des études*. Paris, Champion Honore, 1975.

[461] Y. Chevallard. *La transposition didactique*. Grenoble, La Pensée sauvage, 1985.

i) Les connaissances acquises par des phonéticiens, des phonologues, des psycholinguistes, des ingénieurs, des médecins *etc*. sur les systèmes de production et de perception de la parole par des natifs et non natifs, et sur le système phonémique et la réalisation phonétique des langues concernées ;

ii) Les connaissances des enseignants : parmi i), celles qui pourraient être utiles aux enseignants de langues ;

iii) Les connaissances sur les activités d'enseignement et le processus d'apprentissage chez les apprenants) ;

iv) Les connaissances explicites transmises aux apprenants ;

v) Les connaissances explicites appropriées éventuellement par les apprenants ;

vi) Les connaissances implicites, ou compétences pratiques de prononciation chez les apprenants ;

Dans les parties précédentes, nous avons traité de la première et de la dernière (et l'avant-dernière, en partie) étape dans une série d'études expérimentales. Nous étudierons maintenant les étapes intermédiaires.

Nous nous intéressons d'abord au traitement des voyelles /u/, /y/, et /ø/ dans des manuels de français. Pour cela, nous avons analysé 71 manuels généralistes publiés au Japon (les étapes ii, iv, et v).

Par la suite, nous considérons diverses manières de présenter les caractéristiques articulatoires et acoustiques des voyelles cibles qui faciliteraient la prise de conscience des apprenants (les étapes iii, iv et v).

## 4.1. ANALYSE DE MANUELS GÉNÉRALISTES

### *4.1.1. Les manuels généralistes analysés*

Nous avons recueilli 71 manuels généralistes d'apprentissage de français langue étrangère rédigés en japonais et publiés au Japon (voir Tableau 63). Ces manuels ont été conçus principalement pour des apprenants adultes. Il existe également un petit nombre de manuels pour un public adolescent (lycéens et collégiens), mais comme la grande majorité des apprenants sont des étudiants universitaires, c'est à ce public que la grande majorité des manuels sont adressés.

La quasi-totalité des 71 manuels généralistes que nous avons recueillis sont des livres de classe, conçus pour un usage dans une classe de langues. Par conséquent, les explications ne sont pas nécessairement très détaillées, car celles de l'enseignant sont censées compléter le contenu. Nous considérons que les points importants à retenir sont néanmoins expliqués explicitement, et que le contenu de ces manuels reflète le poids accordé à ces points par les auteurs, et sur connaissances auxquelles les enseignants et les apprenants ont accès dans une situation typique de

l'enseignement et de l'apprentissage du français langue étrangère au Japon. Il existe également des manuels de prononciation, dont un sera présenté dans le chapitre suivant, mais ces manuels spécialisés ne sont malheureusement pas très largement utilisés, principalement à cause du manque de temps à consacrer aux activités de la prononciation, donc leur contenu ne correspondrait pas aux connaissanees typiquement partagées par les enseignants et les apprenants.

Le courant méthodologique le plus fréquemment adopté dans les manuels généralistes japonais est la méthode grammaire-traduction. Dans le cadre de l'enseignement d'une deuxième langue étrangère (第二外国語 *daini-gaikokugo*), les étudiants suivent typiquement un ou deux cours de grammaire-traduction assurés par un enseignant non-natif (japonophone), et éventuellement un cours ou deux de « conversation » assurés par un enseignant francophone natif. C'est dans le premier type de cours que la plupart des manuels que nous avons recueillis sont utilisés.

Le style le plus courant du traitement de la prononciation dans ces livres est de présenter la correspondance graphie-phonie, accompagnée éventuellement de tableaux représentant le système phonémique, au tout début du volume. Un petit nombre de manuels (16 sur 71) proposent des exercices de graphie-phonie ou de prononciation dans les leçons qui suivent, mais la majorité se contente d'une introduction succincte au début du volume.

**Tableau 63** : Les 71 manuels de français langues étrangères recueillis pour l'analyse, ordonnés par l'année de publication, puis par l'ordre alphabétique du premier auteur. Les maisons d'édition : AS (Éditions Asahi), DS (Daisan-Shobô), GK (Gakushû-Kenkyûsha), HS (Hakusuisha), NHSK (Nihon Hôsô Shuppan Kyôkai), ONF (Ôsaka Nichi-Futsu Bunka Center), SB (Sôbi Shuppansha), SG (Srugadai Shuppansha), SS (Sanshûsha).

| | Référence | Premier auteur | Année | Titre en français [titre original en japonais traduit en français] | Maison d'édition |
|---|---|---|---|---|---|
| 1 | Udagawa 1991 | Hiroshi Udagawa | 1991 | Mon ami Vincent | DS |
| 2 | Wada et al. 1991 | Toshimasa Wada | 1991 | Clarté -lecture- | HS |
| 3 | Hiraga et Matsuda 1992 | Isao Hiraga | 1992 | En un an | HS |
| 4 | Ichikawa 1993 | Shin'ichi Ichikawa | 1993 | Le français dans le monde | AS |
| 5 | Kawai et al. 1994 | Georgette Kawai | 1994 | Mosaïques | ONF |
| 6 | Naitô et Tamada 1994 | Yôya Naitô | 1994 | Passeport pour le français | HS |
| 7 | Mizubayashi et Mizubayashi 1999 | Akira Mizubayashi | 1999 | 106 bis, boulevard du Montparnasse : une approche globale du français quotidien | DS |
| 8 | Chikushi 2001 | Fumiteru Chikushi | 2001 | Premier vol en français | AS |
| 9 | Ishino 2001 | Kôichi Ishino | 2001 | [Introduction au français avec un CD] | HS |
| 10 | Nouchi 2001 | Nouchi Ryôzô | 2001 | [Lecture rapide : initiation ludique au français] | DS |
| 11 | Satô et Satô 2001 | Kumiko Satô | 2001 | abCD français | HS |
| 12 | Hisamatsu 2002 | Ken'ichi Hisamatsu | 2002 | Félicitations : futsuken, chat chat chat (troisième édition) | SG |
| 13 | Kazue 2002 | Jôji Kazue | 2002 | ABC français | HS |
| 14 | Nakai 2002 | Tamako Nakai | 2002 | [CD-livre Français avec la grand-mère Hana pour se faire comprendre] | NHSK |
| 15 | Nakamura 2002 | Atsuko Nakamura | 2002 | Le français avec des amis | HS |
| 16 | Sekura et al. 2002 | Masakatsu Sekura | 2002 | Grasys : nouvelle grammaire systématique du français | SG |
| 17 | Shimaoka 2002 | Shigeru Shimaoka | 2002 | Mécanisme du français | HS |
| 18 | Takahashi et Giunta 2002 | Hideo Takahashi | 2002 | Nouveau Coup de Cœur | DS |
| 19 | Birmann et al. 2003 | Olivier Birmann | 2003 | Dialogues (deuxième édition) | DS |
| 20 | Matsumoto et al 2003 | Etsuji Matsumoto | 2003 | [Le français maîtrisé sans peine] | GK |
| 21 | Ôta et al. 2003 | Kôichi Ôta | 2003 | Grammaire française de base (nouvelle édition) | SG |
| 22 | Takizawa et Ôiwa 2003 | Takayuki Takizawa | 2003 | [Maîtrisons solidement le français : avec un CD] | SS |
| 23 | Abrial 2004 | Jean-Pierre Abrial | 2004 | [Se faire des amis avec des Français : conversation facile en français] | SS |
| 24 | Akashi 2004 | Nobuko Akashi | 2004 | RDV@Paris | HS |
| 25 | Anan et Mollé 2004 | Fumiyo Anan | 2004 | Le français Passionnément 1 (nouvelle édition) | DS |
| 26 | Fujita et al. 2004 | Yûji Fujita | 2004 | Mon Premier Vol Tokyo-Paris (nouvelle édition) | SG |
| 27 | Hamana 2004 | Hélène Hamana | 2004 | À la carte | SG |
| 28 | Kitayama et Mehl 2004 | Kenji Kitayama | 2004 | Paris seizième | SG |
| 29 | Satô et al. 2004 | Natsuo Satô | 2004 | Promenade en français -communication et grammaire- | SG |
| 30 | Shôda2004 | Yasuko Shôda | 2004 | Brise-Glace | SG |
| 31 | Tanabe et al. 2004 | Yasuko Tanabe | 2004 | Salut ! | SG |
| 32 | Watanabe 2004 | Kimiko Watanabe | 2004 | Harmonie | SG |
| 33 | Anan et Mollé 2005 | Fumiyo Anan | 2005 | Bon plaisir avec En Marche | SG |
| 34 | Higuchi et al. 2005 | Atsushi Higuchi | 2005 | Premier contact avec le français | HS |
| 35 | Katô et Birmann 2005 | Kazutaka Katô | 2005 | Grammaire française élastique | SG |
| 36 | Katô 2005 | Musafumi Katô | 2005 | Rapide en français I | SG |
| 37 | Kawaguchi et al. 2005 | Junji Kawaguchi | 2005 | Premier montage du français | SG |
| 38 | Kumamoto et Tsujino 2005 | Tetsuya Kumamoto | 2005 | CaDRage | SG |
| 39 | Nakamura et al. 2005 | Eiko Nakamura | 2005 | À la française (troisième édition) | HS |
| 40 | Nakamura et Katô 2005 | Atsuko Nakamura | 2005 | Clique ! Cliquez ! | DS |
| 41 | Nakayama 2005 | Masahiko Nakayama | 2005 | Bonjour, Paris | HS |
| 42 | Roussel et Marukawa | François Roussel | 2005 | Tome 1 | DS |
| 43 | Shiotani et Gonthier-Hosoi | Mayumi Shiotani | 2005 | À la ligne | SG |
| 44 | Teruki et al. 2005 | Takeshi Teruki | 2005 | Ma première rencontre avec Paris | DS |
| 45 | Denda et al. 2006 | Kuniko Denda | 2006 | Manuel pratique de grammaire française | DS |
| 46 | Fujita 2006 | Yuji Fujita | 2006 | Pascal au Japon | HS |
| 47 | Iwata et al. 2006 | Yoshinori Iwata | 2006 | Oreille tendue | SG |
| 48 | Kitayama 2006 | Kenji Kitayama | 2006 | Récemment | AS |
| 49 | Nakajima et Delmont-Hosaka 2006 | Hiroji Nakajima | 2006 | Parcours | HS |
| 50 | Ogura et Ogura 2006 | Kôsei Ogura | 2006 | Aimez-vous la France ? (nouvelle édition) | DS |
| 51 | Sugiyama 2006 | Rieko Sugiyama | 2006 | Les Rencontres 1 | DS |
| 52 | Yasuda 2006 | Etsuko Yasuda | 2006 | Vite fait, bien fait ! | HS |
| 53 | Yoshino 2006 | Satoko Yoshino | 2006 | Petit à Petit | DS |
| 54 | Itô et al. 2007 | Yôji Itô | 2007 | La Règle du jeu | SG |
| 55 | Kurakata 2007 | Hidenori Kurakata | 2007 | Grammaire systématique du français | SB |
| 56 | Naitô 2007 | Yôya Naitô | 2007 | Nouveau Cahier de Français | SG |
| 57 | Sagaz et Nakazato 2007 | Michel Sagaz | 2007 | Civi-Langue | SG |
| 58 | Sawada et al. 2007 | Nao Sawada | 2007 | Amicalement bis : invitation à la langue et à la culture françaises (nouvelle édition) | SG |
| 59 | Suzuki et Nakano 2007 | Takayoshi Suzuki | 2007 | L'unité de la grammaire française | SG |
| 60 | Yamasaki et al. 2007 | Yôichiro Yamasaki | 2007 | Quelques scènes à Paris (nouvelle édition) | DS |
| 61 | Itô 2008 | Madoka Itô | 2008 | Le Français sans Maître : Grammaire française élémentaire pour ceux qui étudient seuls | SG |
| 62 | Kai 2008 | Motofumi Kai | 2008 | Möbius | DS |
| 63 | Kaneko et al. 2008 | Mitsuko Kaneko | 2008 | Petit Poisson | HS |
| 64 | Lorrillard et al. 2008 | Olivier Lorrillard | 2008 | Monsieur Pourquoi : méthode de français pour l'université | SG |
| 65 | Nakamura et al. 2008 | Kazu Nakamura | 2008 | Bonne Chance ! | SG |
| 66 | Ono et Muramatsu 2008 | Yuriko Ono | 2008 | Ma grammaire | HS |
| 67 | Ôtsu 2008 | Toshikatsu Ôtsu | 2008 | La France au Patrimoine Mondial | AS |
| 68 | Sauvegrain et al. 2008 | Sylvain Sauvegrain | 2008 | En Grammaire ! | SG |
| 69 | Tanabe 2008 | Yasuko Tanabe | 2008 | Amie-amie (nouvelle édition) | DS |
| 70 | Teruki et al. 2008a | Takeshi Teruki | 2008 | Nouveau chemin d'accès à la langue française | DS |
| 71 | Teruki et al. 2008b | Takeshi Teruki | 2008 | Votre navigateur pour le français | DS |

### *4.1.2. Les points étudiés*

Afin d'éclaircir les caractéristiques du traitement de la prononciation (tout particulièrement celle des voyelles /u/, /y/ et /ø/) de chaque manuel, nous nous somme posé les 12 questions suivantes :

1. La transcription phonémique des trois voyelles /u/, /y/, et /ø/ figure-t-elle ?

2. La transcription approximative en syllabaire *kana* figure-t-elle ? Si oui, quelles sont les correspondances établies ?

3. Des explications détaillées sur l'articulation de /u/ figurent-elles ? Si oui, portent-elles sur la labialisation, la position de la langue, ou les deux ?

4. Des explications détaillées sur l'articulation de /y/ figurent-elles ?

5. Des explications détaillées sur l'articulation de /ø/ figurent-elles ?

6. La différence phonétique entre le /u/ français et le /u/ japonais est-elle bien explicitée ?

7. La différence phonétique entre le /y/ français et le /ju/ japonais est-elle bien explicitée ?

8. La ressemblance phonétique (acoustique) entre le /ø/ français et le /u/ japonais est-elle mentionnée ?

9. Un triangle ou un trapèze vocalique figure-t-il dans le manuel ?

10. /a/ et /ɑ/ sont-ils représentés comme deux phonèmes différents ?

11. /ɛ̃/ et /œ̃/ sont-ils représentés comme deux phonèmes différents ?

12. Des activités de prononciation ou de graphie-phonie sont-elles proposées dans les leçons (non seulement au début du volume) ?

### *4.1.3. Les résultats*

Tout d'abord, parmi les 71 manuels analysés, 3 ne contenaient aucune description ou activité sur la prononciation ou les correspondances graphie-phonie. Nous allons donc exclure ces 3 manuels. Les 12 questions posées dans la section précédente sont étudiées une par une dans ce qui suit.

**Q1. La transcription phonémique des trois voyelles /u/, /y/, et /ø/ figure-t-elle ?**

Parmi les 68 manuels qui traitent de la prononciation, 60 font figurer le symbole /u/, 62 /y/, et 59 /ø/ (Figure 135). Cela signifie qu'il existe 6 manuels qui n'utilisent pas du tout la transcription en API. Un utilise uniquement /y/, deux /u/ et /y/, et un /y/ et /ø/, alors que les 58 autres utilisent tous les trois symboles.

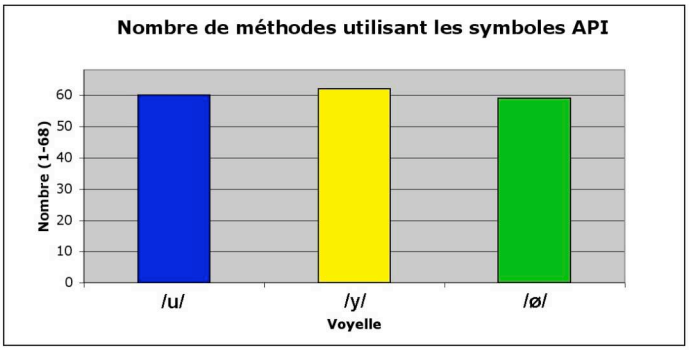



**Figure 135 :** Nombre de manuels qui font figurer les symboles API des voyelles /u/, /y/ et /ø/ (parmi les 68 qui traitent de la prononciation).

**Q2. La transcription approximative en syllabaire *kana* figure-t-elle ? Si oui, quelles sont les correspondances établies ?**

Presque la moitié des manuels ont adopté une transcription approximative en écriture syllabaire *kana* pour /u/ et /y/ (32 sur 68), et moins pour /ø/ (23 sur 68 : Figure 136)[462].

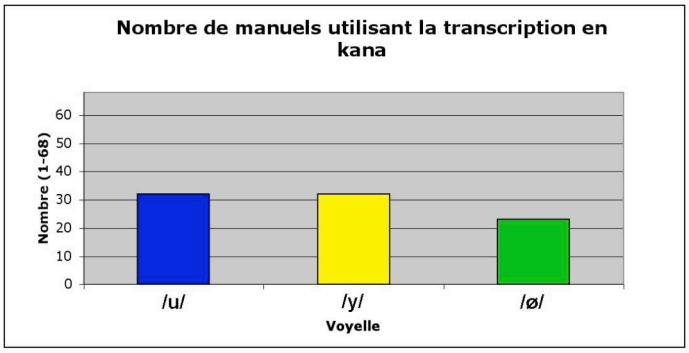



**Figure 136 :** Nombre de manuels qui ont adopté la transcription approximative en écriture japonaise (*kana*).

Nous avons trouvé 23 manuels qui ont utilisé le *kana* à toutes les trois voyelles, et la manière la plus courante (15 manuels) est la suivante : ウ (/u/ [ɯ] japonais) pour /u/, ユ (/ju/ [jɯ] japonais) pour /y/, et de nouveau ウ (/u/ [ɯ] japonais) pour /ø/ (Tableau 64). Si l'on inclut des variations de longueur vocalique (ex. ウー (/uR/ [ɯː] japonais) pour /u/), le nombre de manuels qui tombent dans cette catégorie sera de 17. Six autres manuels ont des types similaires de représentation, mais différencient /u/ et /ø/ d'une manière ou autre ;

1) ウー (/uR/ [ɯː] japonais) pour /u/, ウ (/u/ [ɯ] japonais) pour /ø/ ;
2) ウ (/u/ [ɯ] japonais) pour /u/, ゥ (/u/ [ɯ] japonais : lettre minuscule) pour /ø/ ;
3) ウ (/u/ [ɯ] japonais) « fort » (強い［ウ］) pour /u/, ウ (/u/ [ɯ] japonais) « faible » (弱い［ウ］) pour /ø/ ;
4) ウ (/u/ [ɯ] japonais) « fort » (（強い）ウ) pour /u/, ユ (/ju/ japonais) « aigu » (（鋭い）ユ) pour /y/, ウ (/u/ [ɯ] japonais) « ouvert » (（開いた）ウ) pour /ø/ ;
5) ウ (/u/ [ɯ] japonais), mais une « voyelle postérieure la plus fermée » (「ウ」ですが、最も狭い後母音) pour /u/.

[462] Notons que le choix d'utilisation du *kana* ne dépend pas nécessairement de l'intention de l'auteur. Selon Kuroda (2008), les auteurs de manuels de langues sont parfois amenés par la maison d'édition à rajouter de la transcription en *kana* (p. 189).

R. Kuroda. *Gogaku wa yarinaoseru!* [On peut réapprendre les langues étrangères]. Tokyo, Kadokawa, 2008.

**Tableau 64** : Patterns de transcription approximative en *kana*.

| Phonème français | /ø/ | /u/ | /y/ |
|---|---|---|---|
| Kana | ウ | | ユ |
| Prononciation japonaise | [ɯ] | | [jɯ] |

| | | | |
|---|---|---|---|
| /u/ | ウ /u/ | 2 | |
| /y/ | ユ /ju/ | 2 | |
| /y/ | bouche plus tendue que ユ /ju/ | 1 | 5 |
| /u y/ | ウ /u/. ユ /ju/ | 4 | |
| /u y/ | ウ /u/ dit/prononcé avec la bouche pointue, ユ /ju/ | 3 | |
| /u ø/ | ウ /u/, ゥ /u/ (minuscule) | 1 | 8 |
| /u y ø/ | ウ /u/, ユ /ju/, ウ /u/ | 15 | |
| /u y ø/ | ウ(ー) /u(R)/, ユ /ju/, ウ(ー) /u(R)/ | 1 | |
| /u y ø/ | ウ /u/, ユー /juR/, ウ /u/ | 1 | |
| /u y ø/ | ウー /uR/, ユ /ju/, ウ /u/ | 2 | |
| /u y ø/ | ウ /u/, ユ /ju/, ゥ /u/ (minuscule) | 1 | |
| /u y ø/ | ウ /u/ fort, ユ /ju/, ウ /u/ faible | 1 | |
| /u y ø/ | ウ /u/ fort, ユ /ju/ aigu, ウ /u/ ouvert | 1 | |
| /u y ø/ | ウ /u/ mais une V post. la plus fermée, ユ /ju/, ウ /u/ | 1 | 23 |

Huit livres qui fournissent du *kana* à deux voyelles seulement, sept pour /u/ et /y/, et un seul pour /u/ et /ø/. Dans les cinq autres, une seule des trois voyelles est transcrite en *kana* : /u/ dans deux manuels, /y/ dans les trois autres.

Nous voyons ainsi que parmi les 3 voyelles /ø/ est le moins souvent associé à la transcription en *kana* (toujours à ウ: /u/ [ɯ] japonais, avec une précision éventuelle, comme « faible » ou « ouvert »). Cette tendance pourrait s'expliquer par le fait que les auteurs auraient trouvé qu'il était moins convenable d'associer au caractère ウ (/u/ japonais) le /ø/ français que le /u/ français, celui-ci étant transcrit habituellement avec le même symbole que le /u/ [ɯ] japonais malgré un écart phonétique important. Le symbole phonémique /ø/ aurait fait penser que cette voyelle soit tout à fait différente du /u/ [ɯ] japonais.

**Q3. Des explications détaillées sur l'articulation de /u/ figurent-elles ? Si oui, portent-elles sur la labialisation, la position de la langue, ou les deux ?**

Parmi les 68 manuels, 17 comportent des explications explicites sur l'articulation de /u/, dont 4 mentionnent les lèvres (arrondissement et/ou protrusion des lèvres) et la position de la langue, 4 autres les lèvres et la langue, avec une description impressionniste de la position de la langue (ex. 口の奥で « *au fond de la boche* », のどの辺りで « *aux alentours de la gorge* », ou mention à la pointe de la langue), et les 9 autres uniquement les lèvres (Figure 137, gauche). Ceux qui explicitent la labialisation utilisent presque tous des expressions qui mentionnent les gestes de l'arrondissement (16 sur 17), et de la protrusion (16 sur 17).

**Q4. Des explications détaillées sur l'articulation de /y/ figurent-elles ?**

Vingt manuels qui fournissent des explications explicites sur l'articulation de /y/ (Figure 137, milieu), dont 5 qui n'avaient pas de description précise (ex. mention seulement à la pointe de

la langue ; 「ウ」の口で「イ」と発音する « *prononcer* イ (/i/ japonais) *avec la bouche de* ウ (/u/ [ɯ] japonais) », ce dernier son n'étant pas tout à fait arrondi).

**Q5. Des explications détaillées sur l'articulation de /ø/ figurent-elles ?**

Nous avons observé des explications explicites sur l'articulation de /ø/ dans 19 manuels parmi les 68 étudiés. La grande majorité d'entre eux (16 sur 19) mentionnent la combinaison de /o/ et /e/ d'une manière ou autre (ex. position de la langue de /e/ et la forme des lèvres de /o/).

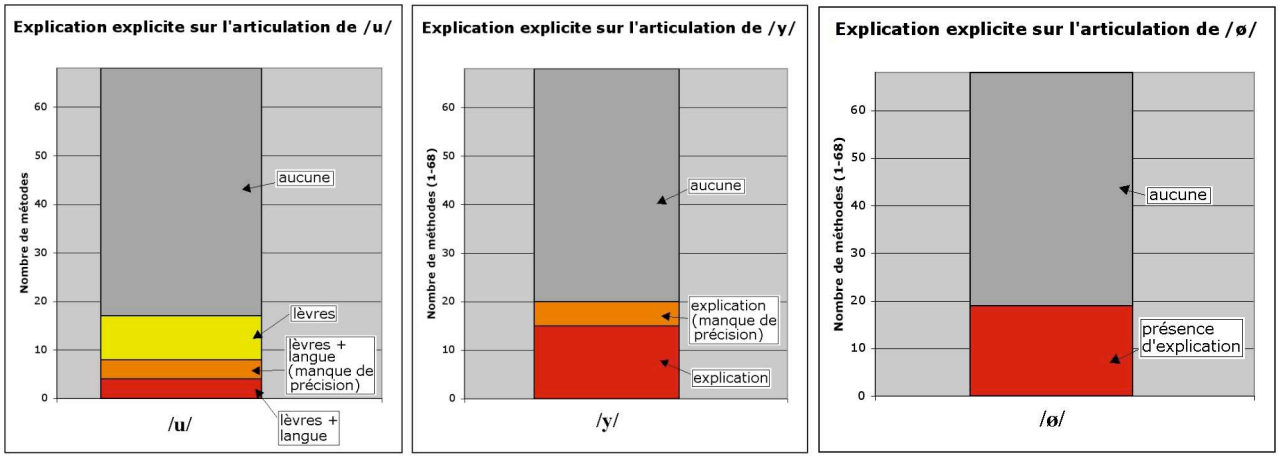



**Figure 137 :** Présence d'explication explicite sur l'articulation de /u/ (gauche), de /y/ (milieu), et de /ø/ (droite) dans les 68 manuels qui avaient du contenu sur la prononciation.

**Q6. La différence phonétique entre le /u/ français et le /u/ [ɯ] japonais est-elle bien explicitée ?**

La différence phonétique (articulatoire, en particulier) entre le /u/ français et le /u/ [ɯ] japonais est explicitement expliquée dans 7 manuels, et 10 autres décrivent l'articulation du /u/ français sans mentionner la différence avec celle du /u/ [ɯ] japonais (Figure 138, gauche). Les explications sur la différence incluent les expressions suivantes : « *le* /u/ [ɯ] *japonais, prononcé sans arrondissement des lèvres, est un son complètement différent* [du /u/ français] » (日本語の「ウ」は唇を丸めずに発音する、まったく別の音です : Nakai 2002 : 36[463]) ; « *le son* [ウ], *prononcé avec les lèvres arrondies et protruses, n'existe pas en japonais* » (唇を丸く尖らせて発音する［ウ］の音は日本語にはありません : Kawaguchi *et al*. 2005 : 3[464]).

**Q7. La différence phonétique entre le /y/ français et le /ju/ [jɯ] japonais est-elle bien explicitée ?**

Le /y/ français est décrit comme étant différent du /ju/ [jɯ] japonais dans 5 manuels. Un affirme que l'on peut remplacer le /y/ français par le /ju/ [jɯ] japonais, tout en soulignant la différence considérable entre les deux sons. Il existe 15 autres manuels qui décrivent l'articulation de /y/, sans mentionner la différence entre le /y/ français et le /ju/ [jɯ] japonais (Figure 138 : milieu).

---

[463] T. Nakai. *CD bukku : korede tsuujiru ! Hana-baachan no furansugo* [CD-livre Français avec la grand-mère Hana pour se faire comprendre]. Tokyo, Nihon Hôsô Shuppan Kyôkai, 2002.

[464] J. Kawaguchi, R. Miyashita, A. Shioda. *Premier montage du français*. Tokyo, Surugadai Shuppansha, 2005.

**Q8. La ressemblance phonétique (acoustique) entre le /ø/ français et le /u/ [ɯ] japonais est-elle mentionnée ?**

Nous n'avons trouvé que 2 manuels, écrits par le même premier auteur, qui mentionnaient la ressemblance phonétique (impression acoustique) entre le /ø/ français et le /u/ japonais, et cela, pour expliquer que le /ø/ est plus proche du /u/ [ɯ] japonais que du /ju/ [jɯ] japonais (Figure 138, droite).

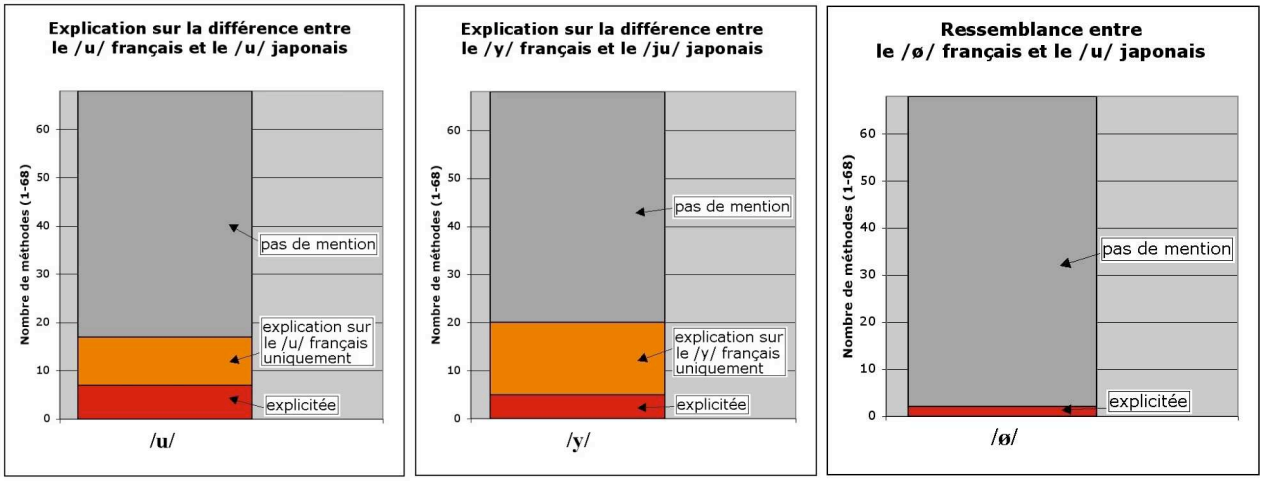



**Figure 138 :** Nombre de manuels qui expliquent explicitement : 1) la différence phonétique entre le /u/ français et le /u/ [ɯ] japonais (gauche) ; 2) la différence phonétique entre le /y/ français et /ju/ [jɯ] japonais ; 3) la ressemblance entre le /ø/ français et le /u/ japonais (71 manuels étudiés, dont 68 traitent de la prononciation).

**Q9. Un triangle ou un trapèze vocalique figure-t-il dans le manuel ?**

**Q10. /ɛ̃/ et /œ̃/ sont-ils représentés comme deux phonèmes différents ?**

**Q11. /a/ et /ɑ/ sont-ils représentés comme deux phonèmes différents ?**

Trente manuels font figurer un triangle ou un trapèze vocalique. Parmi ces 30 titres, 17 montraient les deux voyelles nasales /ɛ̃/ et /œ̃/ comme deux sons différents sans explications, 11 les traitaient comme deux sons, mais avec une explication sur la neutralisation des deux phonèmes, et les 2 autres indiquaient qu'un seul son /ɛ̃/. Quant à la paire /a/-/ɑ/, 22 manuels faisaient figurer les deux symboles sans explication, 7 avec une mention à la neutralisation, et un représentait qu'une voyelle ouverte /a/ (Figure 139).

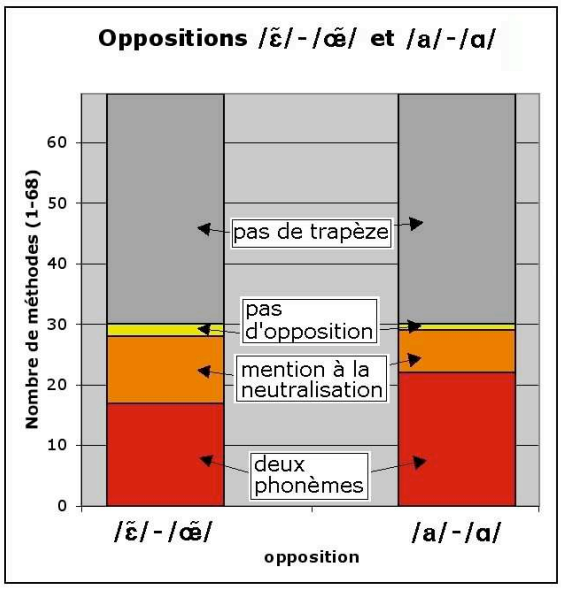



**Figure 139 :** Manuels qui faisaient figurer les paires /ɛ̃/-/œ̃/ et /a/-/ɑ/.

**Q12. Des activités de prononciation ou de graphie-phonie sont-elles proposées dans les leçons ?**

Seize manuels proposent des activités (y compris un simple tableau de récapitulation) de graphie-phonie et/ou de prononciation dans les leçons (c'est-à-dire, non seulement au début du volume), dont 9 traitaient plutôt la graphie-phonie, 3 plutôt la prononciation et la perception, et 4 les deux aspects. Un manuel présentait uniquement le cas de liaison avec les chiffres.

Concernant le contenu des activités, 13 manuels traitent des voyelles nasales (dont 7 les correspondances graphie-phonie, 6 la prononciation). Par contraste, 4 seulement proposent des activités de prononciation sur l'opposition /u/-/y/ (dont 2, dans le triplet /i u y/), un seul pour la graphie-phonie de /y/-/ø/ (parmi d'autres combinaisons de graphie). Quant à l'opposition /u/-/ø/ (ou /u/-/œ/), 3 titres seulement traitent de cette difficulté : un dans le triplet /Œ u O/ (c'est-à-dire /ø u o/ et /œ u ɔ/) ; un, l'opposition /u/-/œ/ ; un qui traitent de la graphie-phonie parmi d'autres combinaison de graphie.

Les 51 autres ne proposaient aucune activité de graphie-phonie ou de prononciation, sauf une présentation du système phonémique du français ou un récapitulatif de graphie-phonie au début du volume, ou à la fin pour certains (Figure 140).

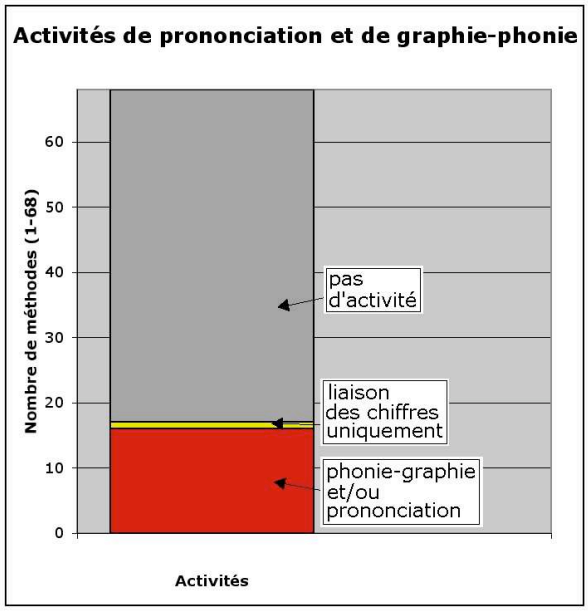



**Figure 140 :** Manuels qui proposent des activités de prononciation et/ou de graphie-phonie.

### *4.1.4. Synthèse et discussion*

Nous allons maintenant relever des points qui semblent importants parmi les observations faites ci-dessus :

1. La grande majorité des manuels étudiés font figurer les symboles API des voyelles /u/, /y/, et /ø/ (59-62 sur 68 : Q1. Figure 135), mais presque la moitié d'entre eux utilisent une transcription approximative en *kana* (23-32 sur 68 : Q2. Figure 136). Le pattern le plus courant de transcription en *kana* est le suivant : ウ (/u/ [ɯ] japonais) pour /u/, ユ (/ju/ [jɯ] japonais) pour /y/, et ウ (/u/ [ɯ] japonais) pour /ø/, malgré un petit nombre de manuels qui différencient /u/ et /ø/ d'une manière ou autre (majuscule et minuscule, « fort » et « faible », *etc.* : Q2. Tableau 64).

2. Seul un quart des manuels étudiés expliquent l'articulation du /u/ français (17 sur 68 : Q3. Figure 137), et moins de la moitié d'entre eux (7 sur 17) explicitent la différence phonétique entre le /u/ français et le /u/ [ɯ] japonais (Q6. Figure 138).

3. Quant à /y/, son articulation est explicitée dans plus d'un quart de manuels (de manière complète ou incomplète : 20 sur 68, Q4. Figure 137), et la différence entre le /y/ français et le /ju/ [jɯ] japonais est explicitée dans 5 manuels seulement (Q7. Figure 138).

4. La ressemblance acoustique entre le /ø/ français et le /u/ [ɯ] japonais est traitée dans 2 manuels seulement, mais d'une manière indirecte (ex. /ø/ est plus proche du /u/ [ɯ] japonais que du /ju/ [jɯ] japonais : Q8. Figure 138).

5. En revanche, les paires /ɛ̃/-/œ̃/ et /a/-/ɑ/ sont représentées dans la quasi-totalité des manuels qui font figurer un triangle ou un trapèze vocalique (28 sur 30 pour /ɛ̃/-/œ̃/, et 29 sur 30 pour /a/-/ɑ/), avec une mention à la neutralisation dans une minorité d'entre eux (11 sur 28 pour /ɛ̃/-/œ̃/, et 7 sur 29 pour /a/-/ɑ/ : Q10-11. Figure 139).

6. Seule une petite minorité de manuels (7 sur 68) proposent des activités de prononciation de sorte que les apprenants puissent s'exercer au fur et à mesure que l'apprentissage progresse ; 16 sur 68 même en incluant des activités de graphie-phonie (Q12. Figure 140)

Nous pouvons supposer que la différence phonétique entre le /u/ français et le /u/ [ɯ] japonais, et la difficulté d'apprentissage de cette voyelle française ne seraient pas suffisamment bien traitées dans les manuels généralistes japonais. Cette situation ne faciliterait pas la prise de conscience des apprenants japonophones concernant la particularité du /u/ français par rapport au /u/ [ɯ] japonais, les deux phonèmes étant transcrits avec le même symbole /u/.

/y/ ne semble pas être mieux traité que /u/, mais nous avons vu que les apprenants japonophones identifiaient mieux cette voyelle (Tableau 22, p. 130), et la distinguaient des autres voyelles (/u/ et /ø/) mieux que /u/ (Figure 66, p. 182). En outre, le symbole /y/, qui n'est pas utilisé dans le système vocalique du japonais, permettrait aux apprenants de prendre conscience du fait que cette voyelle n'est pas identique à /ju/ [jɯ] japonais. Ce facteur-là rendrait l'apprentissage de /y/ moins difficile que celui de /u/.

Le fait que /ø/ ressemble acoustiquement à /u/ [ɯ] japonais malgré la différence de lieux d'articulation ne semble pas être suffisamment traité dans les manuels généralistes japonais. Comme les japonophones ont des difficultés à distinguer les deux voyelles /u/-/ø/, qui seraient toutes les deux assimilées au /u/ [ɯ] japonais, des connaissances sur la ressemblance acoustique entre /ø/ et le /u/ [ɯ] japonais, avec celles de la différence acoustique (et articulatoire) entre le /u/ français et le /u/ [ɯ] japonais, pourraient faciliter la distinction de cette paire.

En revanche, les paires /ɛ̃/-/œ̃/ et /a/-/ɑ/ sont traitées dans un grand nombre de manuels, malgré la neutralisation observée auprès de la majorité des locuteurs natifs français de nos jours. Cela contraste avec le traitement insuffisant des voyelles /u/, /y/, et /ø/, qui sont distinguées par tous les locuteurs natifs, et doivent se prononcer de manière distincte pour se faire comprendre correctement en français.

En outre, il est très dommage que les manuels généralistes japonais ne soient pas très nombreux à proposer des activités de prononciation. Certes, les enseignants sont très souvent confrontés au problème de manque de temps, et ont donc du mal à trouver suffisamment de temps à consacrer à la prononciation. Néanmoins, s'ils trouvent des activités proposées dans les manuels, ils seront plus facilement amenés à les effectuer en classe.

Nous avons étudié dans ce chapitre comment la prononciation, notamment celle des voyelles /u/, /y/ et /ø/, est traitée dans des manuels généralistes de français langue étrangère publiés au Japon. Nous avons observé que ces trois voyelles en question, notamment la difficulté de /u/, n'étaient pas suffisamment bien traitées, ce qui ne facilite pas la prise de conscience des apprenants. Que pourrait-on faire pour soulever la conscience des apprenants ? Comment pourrait-on améliorer, si possible, la pratique de l'enseignement de la prononciation du français langue étrangère auprès des apprenants japonophones ? Dans le chapitre 4.3., nous verrons des outils qui permettraient de le faire.

## 4.2. PRÉSENTATION D'UN MANUEL DE PRONONCIATION PUBLIÉ AU JAPON

Nous présenterons dans ce chapitre le manuel d'Anan et Guimbretière (1995)[465], écrit en japonais pour les apprenants japonophones.

### *4.2.1. Les phénomènes traités*

Le manuel commence par une partie consacrée à la prosodie (syllabe, groupe rythmique, accent, intonation, émotions ; voir Tableau 65), comme le font Abry et Chalaron (1994)[466], Kaneman-Pougatch et Pedoya-Guimbretière (2001)[467], et Charliac *et al.* (2003)[468]. Les voyelles et les semi-voyelles, et les consonnes sont traitées dans la deuxième partie, suivie de la troisième partie qui traite des phénomènes observés à l'intérieur de groupe rythmiques (enchaînement, liaison, « e caduc », assimilation de consonnes.

Concernant les voyelles /u y ø/, les voyelles fermées /i u y/ et moyennes /e ɛ o ɔ ø œ/ sont traitées séparéments (Dossiers 1 et 2, respectivement). Des activités qui portent sur l'opposition /u/-/ø/ ou /y/-/ø/ ne sont pas proposées. C'est en opposition avec /o/ et /e/ que /ø/ est présenté ; son articulation est expliquée par une combinaison de [e] (position de la langue) et de [o] (lèvres arrondies). La labialité de /y/ et de /u/, ainsi que la postériorité de la langue pour /u/, est explicitée dans le manuel (notamment la labialité). /u y ø/ se trouvent parmi les huit voyelles (/i y u/, /ø œ/, /õ ɑ̃ ɛ̃/) que les auteures décrivent comme des voyelles qui nécessite un entraînement pour les japonophones.

**Tableau 65** : Table des matières d'Anan et Guimbretière (1995).

*Partie 1 : Rythme et mélodie du français*
*1. Syllabe*
*2. Groupe rythmique et accent*
*3. Éléments d'intonation*
*4. Expressions des émotions*
*Chanson : Sur le pont d'Avignon*

*Partie 2 : Acquisition des sons du français*
*Chapitre 1 : Voyelles et semi-voyelles*

*Dossier 1 : Voyelles* [i] [y] [u]

*Dossier 2 : Voyelles* [e] [ɛ] [o] [ɔ] [ø] [œ]

*Dossier 3 : Voyelles nasales* [õ] [ɑ̃] [ɛ̃] [œ̃]

*Dossier 4 : Semi-voyelles* [j] [ɥ] [w]

---

[465] F. Anan, E. Guimbretière. *Komyunikêshon no tameno furansugo hatsuon-hou: hatsuon no kisoku to renshuu* [Plaisir de communiquer : règles et exercices de phonétique pour bien communiquer]. Tokyo, Sugugadai Shuppansha, 1995.

[466] D. Abry, M.-L. Chalaron. *350 exercices de phonétique*. Paris, Hachette, 1994.

[467] M. Kaneman-Pougatch, E. Pedoya-Guimbretière. *Plaisir des sons : enseignement des sons du français*. Paris, Hatier/Didier, 1991.

[468] L. Charliac, A. C. Motron, J.-T. Le Bougnec, B. Loreil. *Phonétique progressive du français (niveau débutant)*. Paris, CLE International, 2003.

*Chapitre 2 : Consonnes*

*Dossier 5 : Consonnes* [ʀ] [l]

*Dossier 6 : Consonnes* [f] [v] [b]

*Dossier 7 : Consonnes* [ʃ] [ʒ] [s] [z]

*Chanson : Le temps des cerises*

*Partie 3 : Règles à l'intérieur de groupe rythmique*
*1. Enchaînement*
*2. Liaison*
*3. e caduc*
*4. Assimilation de consonnes*
*Chanson : Auprès de ma blonde*

*Tests*
*Test 1*
*Test 2*

### *4.2.2. L'organisation des dossiers*

Chaque dossier est composé de cinq types d'activité. Les auteures expliquent ces activités dans la préface (page *ix*) :

> *1) « Écoutez » vous permettra d'entendre et de connaître les sons,*
> *2) « Ressentez » vous permettra de vous familiariser avec les traits de chaque son,*
> *3) « Parlez » vous permettra de vous exercer à produire les sons en contexte facilitant d'abord, puis dans des dialogues,*
> *4) « Écrivez » vous permettra de faire la liaison entre les sons et leurs graphies,*
> *5) « Amusez-vous » vous détendre en continuant à prononcer les sons.*

Avant l'activité 1) « *écoutez* » de chaque dossier, les symboles phonémiques, l'orthographe représentatif, les caractéristiques articulatoires (tension, labialité) et acoustiques (gravité), les profils articulatoires, et la forme des lèvres des sons étudiés sont montrés, accompagnés d'un paragraphe qui explique l'articulation.

Dans la phase 1) « *Écoutez* », des exercices de discrimination (/i/-/y/, /y/-/u/, /i/-/y/ dans Dossier 1) et d'identification sont proposés.

L'activité 2) « *Ressentez* » propose une intégration corporelle des sons étudiés (Figure 141 ci-dessous. Voir 4.3.3.2., Figure 150, p. 287).

L'activité 3) « *Parlez* » propose des exercices de répétition, de transformation grammaticale, et de jeu de dialogue.

L'activité 4) « *Écrivez* » porte sur les correspondances graphie-phonie (repérage de graphie correspondant aux sons entendus ; dictée partielle sur les voyelles concernées).

La dernière activité 5) « *Amusez-vous* » propose d'utiliser les sons travaillés dans des contextes linguistiques plus authentiques (quiz, écoute et lecture de poème, *etc*.).

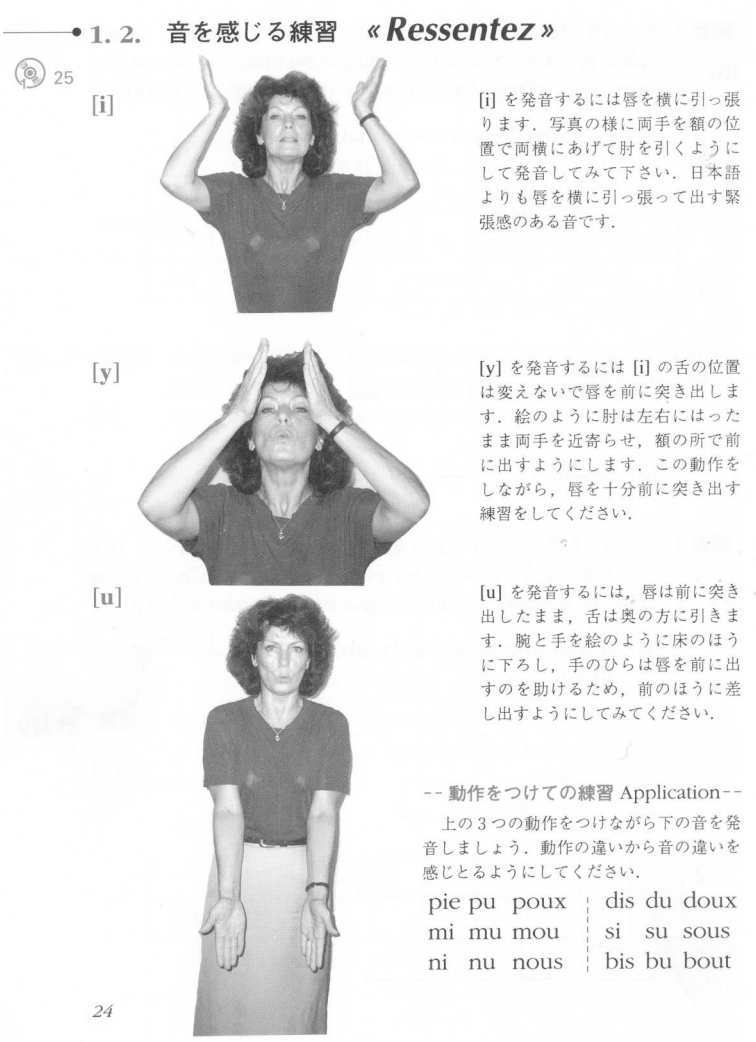



**Figure 141** : Activité d'intégration corporelle de /i y u/ dans la section « *Ressentez* » (Anan et Guimbretière 1995 : 26).

L'ordre des activités présentées dans ce manuel correspond bien à la progression que Lauret (2007 : 139) propose :

> « *de la perception vers la production* » ;
> « *de l'auditif au linguistique* » ;
> « *de l'oral vers l'écrit* » ;
> « *du plus contraint au moins contraint (la répétition simple ne sollicite pas les mêmes compétences que la transformation grammaticale)* »
> « *du plus favorisant* [contexte facilitant] *au moins favorisant* ».

### 4.3. OUTILS FACILITANT LA PRISE DE CONSCIENCE DES APPRENANTS

Nous allons présenter dans ce chapitre un survol des outils qui permettraient de faciliter la prise de conscience des apprenants. Nous nous concentrons notamment sur la voyelle /u/, qui serait la plus difficile à produire parmi les voyelles /u y ø/ pour les apprenants japonophones, et qui n'est très souvent pas traitée comme une voyelle difficile dans les manuels généralistes japonais de français langue étrangère. La difficulté de production de cette voyelle se trouve dans les deux

facteurs suivants : la labialité (l'arrondissement et la protrusion des lèvres) et le recul de la langue. C'est notamment ce dernier facteur qui ne peut être illustré facilement, à cause de l'invisibilité de la langue. Nous présenterons ici les outils qui permettraient de compenser particulièrement cette difficulté.

Les méthodes suivantes seront présentées dans ce chapitre :

1. La méthode verbo-tonale
2. Continuum de vocoïdes postérieurs (proposé par Bertrand Lauret)
3. Image acoustique, pour illustrer la gravité du /u/ français par rapport au /u/ [ɯ] japonais
4. Synthèse articulatoire (VTCalc et TractSyn), qui permet aux apprenants de prendre conscience de la relation entre l'articulation et l'acoustique.
5. Spectrogramme en temps réel (WaveSurfer), pour visualiser les formants en temps réel.

### *4.3.1. La méthode verbo-tonale*

L'approche structuro-globale audiovisuelle (SGAV) a développé la méthode verbo-tonale de correction phonétique (Renard 1971[469], Renard 2002[470]). Cette méthode propose de « rendre plus saillantes les informations mal perçues ou non-perçues » par les apprenants (présentation de cette méthode par Lauret 2007 : 108[471]). Les sons clairs ou aigus (ex. /i/ par rapport à /y/) peuvent être renforcés par une intonation montante, et les sons graves (ex. /y/ par rapport à /i/) par une intonation descendante (Figure 142). De la même manière, les sons aigus sont renforcés par des consonnes aiguës, les sons graves par des consonnes graves (Figure 143).

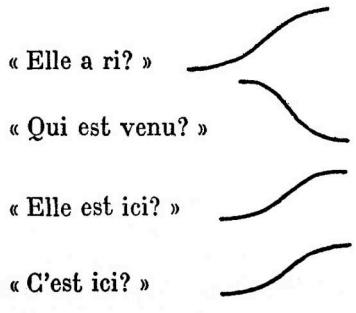



**Figure 142 :** /i/ placé à la fin d'une intonation montante, et au début d'une intonation descendante, une solution proposée aux apprenants qui prononcent cette voyelle de manière trop relâchée, en émettant un son proche de [e] (Renard 1971 : 73-74).

[469] R. Renard. *Introduction à la méthode verbo-tonale de correction phonétique*. Paris, Didier, 1971.

[470] R. Renard. "Une phonétique immergée", in R. Renard, *Apprentissage d'une langue étrangère/seconde 2 : la phonétique verbo-tonale*. Bruxelles, De Boeck Université, 11-24, 2002.

[471] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

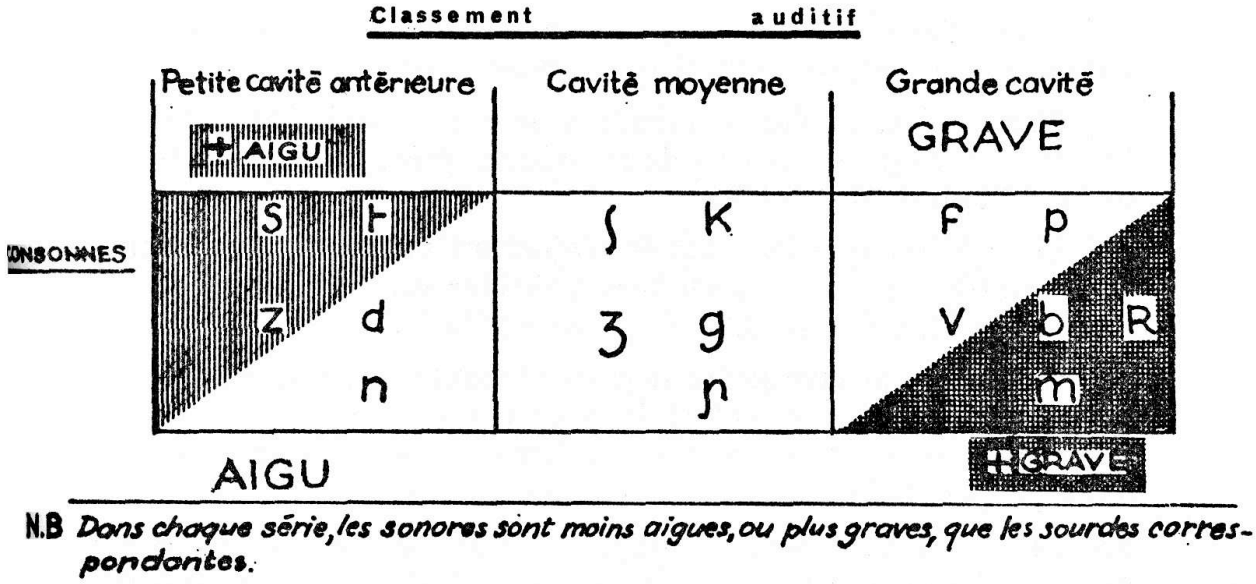



**Figure 143 :** Classement auditif des consonnes du français, proposé par Renard (1971 : 93)[472]

Étant donné que les apprenants japonophones ont du mal à distinguer perceptivement l'opposition /u/-/ø/, que c'est /u/ plutôt que /ø/ qui est phonétiquement plus éloigné du /u/ japonais, et qu'ils ont des difficultés à prononcer le /u/ français, nous pouvons renforcer les caractéristiques graves de cette voyelle. Pour ce faire, la méthode propose comme facteurs favorables un débit lent, le registre bas, une intonation descendante, et les consonnes « graves » (/f v p b m ʀ/ : Renard 1971 *op. cit.* ; ex. « jusqu'au bout » /ʒyskobu/).

Callamand (1981)[473] propose 28 fiches pédagogiques établies par voyelle, et par le type de l'écart entre la cible et la production des apprenants. Chaque fiche précise les traits à renforcer, les contextes facilitants, et présente 12 phrases contenant la voyelle cible dans la dernière syllabe (c'est donc une syllabe « accentuée » allongée). Ces phrases sont ordonnées de sorte que la difficulté augmente progressivement, des contextes facilitants aux contextes non-facilitants, et par conséquent, du plus facile au plus difficile.

La Figure 144 montre le cas des apprenants qui prononcent [ɯ] au lieu de [u] (l'auteure précise que c'est une « erreur fréquente chez les locuteurs japonais ». Callamand 1981 : 119). Les traits à renforcer sont « plus de labialité » et « moins d'antériorité », et les consonnes qui forment les contextes facilitants pour [u] sont [p] [b] [m] [f] [v] [w]. On commence par des phrases qui se terminent par une syllabe ouverte contenant un [u] et une des consonnes facilitantes (*fou* /fu/, *debout* /dəbu/, *vous* /vu/, *boue* /bu/, *mou* /mu/), puis une avec une syllabe fermée (*pour* /puʀ/). Ensuite, nous y trouvons deux phrases avec des consonnes pré-palatales (facteur non facilitant pour [u]) et arrondies (facteur facilitant) /ʃ/ et /ʒ/ (*joues* /ʒu/, *chou* /ʃu/), deux avec des

[472] Notons que les consonnes /f v p b m ʀ/ ne correspondent pas à une « grande cavité » comme indiqué par Renard, mais à un manque de cavité (Fant 1960).

G. Fant. *Acoustic Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*. The Hague, Mouton, 1960.

[473] M. Callamand. *Méthodologie de l'enseignement de la prononciation : organisation de la matière phonique du français et correction phonétique*. Paris, Création Loisirs Enseignement International, 1981.

consonnes « neutres » à l'attaque (*rouge* /ʀuʒ/, *coupe* /kup/), et pour finir, deux avec la labialité « neutre » (*tout* /tu/, *sourd* /suʀ/).

Nous y observons que la difficulté augmente progressivement, sauf que les consonnes [f] et [v] n'aideront pas nécessairement les apprenants japonophones. Ces consonnes n'existent pas dans leur langue maternelle japonaise, et certains apprenants ont tendance à s'efforcer de former un contact solide (parfois même plus que nécessaire) entre les dents supérieures et la lèvre inférieure (notamment vers le sommet de la lèvre, pour ceux qui ne sont pas appris que le contact peut se former également sur le côté intérieur de la lèvre). Ces gestes risquent d'empêcher la protrusion des lèvres.

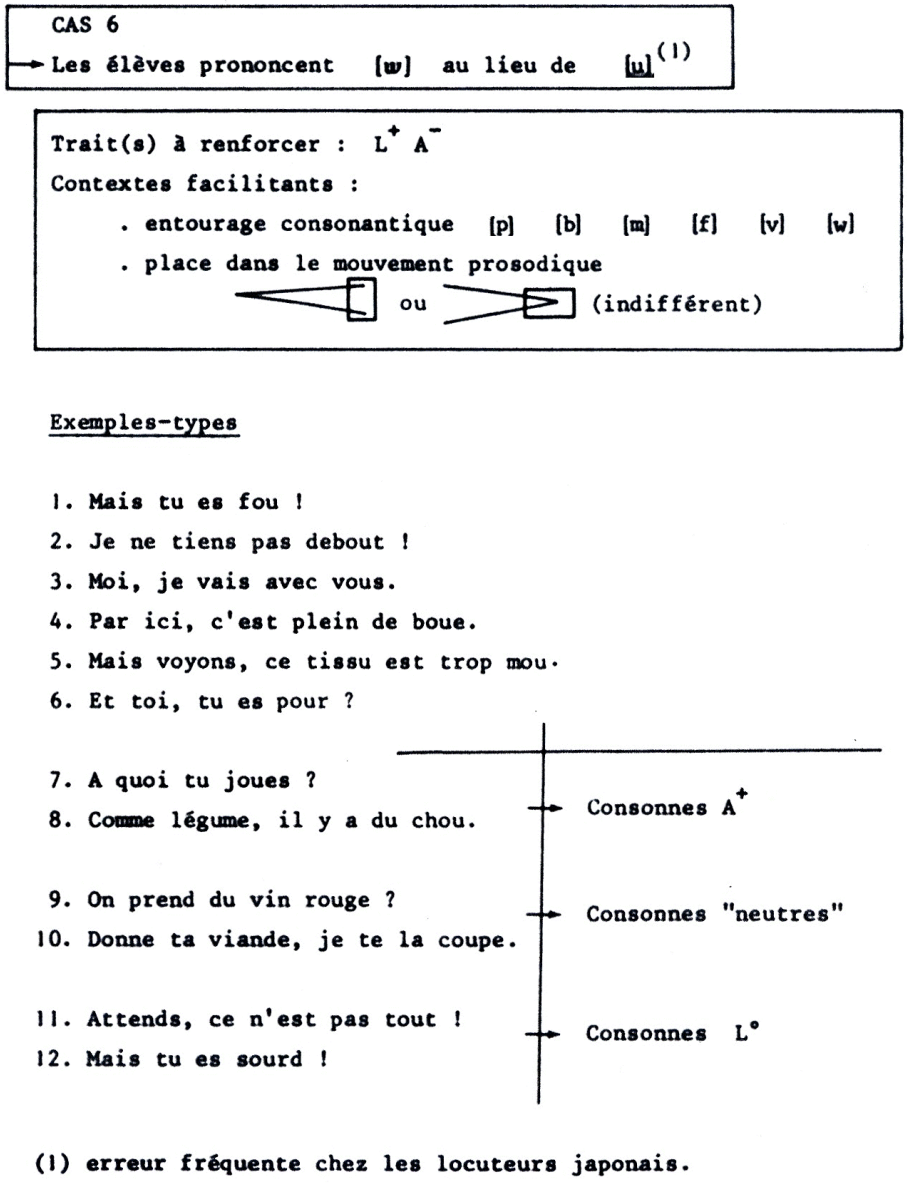
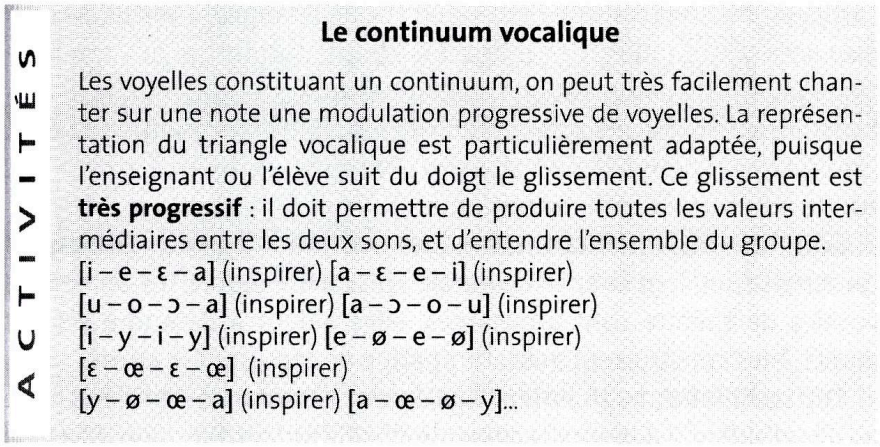



**Figure 144 :** Fiche pédagogique proposée pour les élèves prononçant [ɯ] au lieu de [u] (Callamand 1981 : 119).

### *4.3.2. Continuum de vocoïdes postérieurs (Lauret)*

Lauret (2007 : 123)[474] propose une activité de production de continua de vocoïdes (Figure 145). Cette activité est conçue pour l'échauffement vocal qui prépare les apprenants à d'autres tâches de production, mais selon l'auteur (communication personnelle), la série de vocoïdes postérieurs (progressivement arrondis), en allant de [ɑ] vers [u], en passant par [ɔ] et [o], est utile pour amener les apprenants japonophones à prononcer le /u/ français. Les apprenants sont invités à produire ces vocoïdes en les modulant très progressivement et très lentement, en permettant de produire toutes les valeurs intermédiaires, et d'entendre l'ensemble du continuum. La Figure 146 illustre un exemple de réalisation de ce continuum. Nous y voyons que les deux premiers formants baissent progressivement, un mouvement comparable à la série des voyelles postérieures du français, prononcées de manière discrète (Figure 147).


ACTIVITÉS

**Le continuum vocalique**

Les voyelles constituant un continuum, on peut très facilement chanter sur une note une modulation progressive de voyelles. La représentation du triangle vocalique est particulièrement adaptée, puisque l'enseignant ou l'élève suit du doigt le glissement. Ce glissement est **très progressif** : il doit permettre de produire toutes les valeurs intermédiaires entre les deux sons, et d'entendre l'ensemble du groupe.
[i – e – ɛ – a] (inspirer) [a – ɛ – e – i] (inspirer)
[u – o – ɔ – a] (inspirer) [a – ɔ – o – u] (inspirer)
[i – y – i – y] (inspirer) [e – ø – e – ø] (inspirer)
[ɛ – œ – ɛ – œ] (inspirer)
[y – ø – œ – a] (inspirer) [a – œ – ø – y]...

**Figure 145 :** Activité d'échauffement vocal « continuum vocalique » proposée par Lauret (2007 : 123).

[474] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

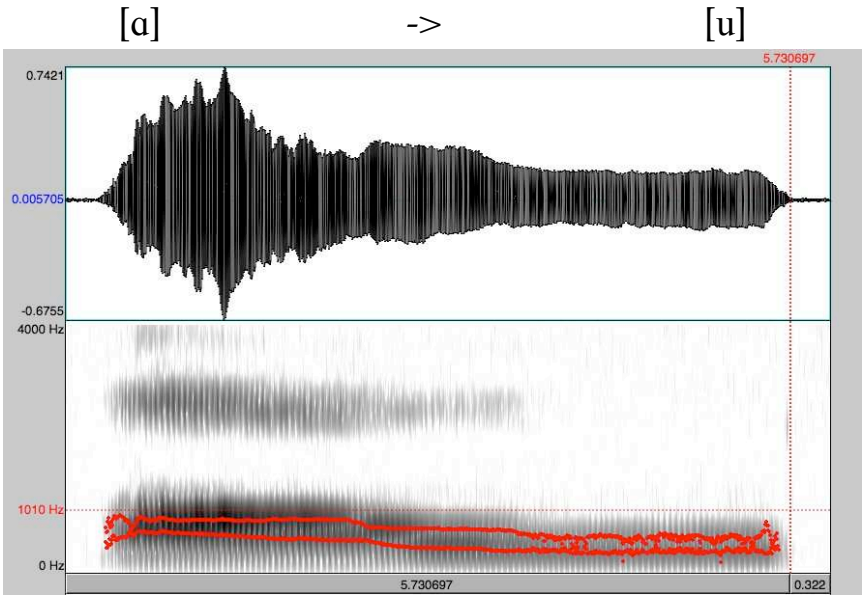
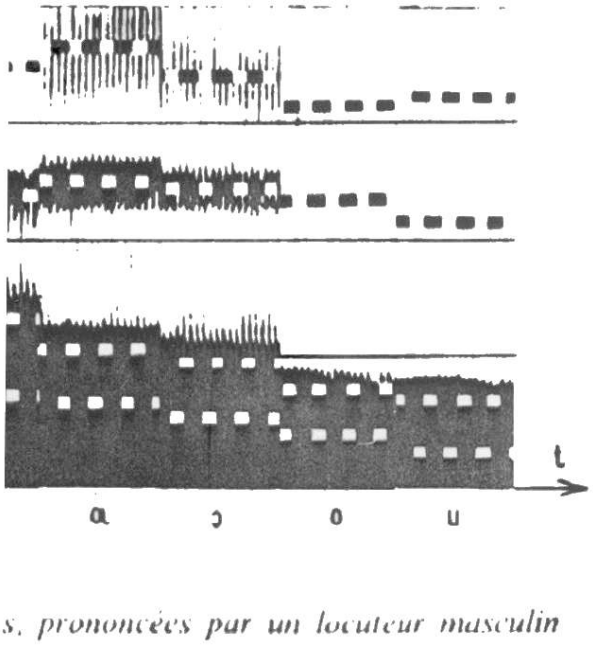



**Figure 146 :** Continuum de vocoïdes postérieurs (progressivement arrondis) variant entre [ɑ] et [u], en passant par [ɔ] et [o]. L'oscillogramme (en haut), le spectrogramme à bande large (largeur de la bande : 5 millisecondes), et les deux premiers formants détectés par Praat (paramètres de détection : 2 formants à détecter en dessous de 1800 Hz).




*s, prononcées par un locuteur masculin*

**Figure 147 :** Spectrogramme des voyelles postérieures du français prononcées par un locuteur masculin (Liénard 1977[475]).

Cette méthode permet aux apprenants d'aborder le [u] français en travaillant sur des entités plutôt non-linguistiques, sans avoir à traiter consciemment du /u/ français ou du /u/ [ɯ] japonais. Si l'on adopte cette activité comme échauffement, le passage à d'autres exercices de

[475] J.-S. Liénard. *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole*. Paris, Masson, 1977.

productions qui portent plus explicitement sur le phonème /u/ sera facilité. D'ailleurs, cela correspond bien à la progression des activités en classe que Lauret propose : « de l'auditif au linguistique » (Lauret 2007 : 139).

### *4.3.3. Image acoustique : la gravité de /u/*

Les propriétés acoustiques des sons sont souvent difficiles à saisir intuitivement si l'on cherche à les expliquer en utilisant des termes spécialistes de la physique et de l'acoustique. Comme les apprenants ne sont bien évidemment pas tous des phonéticiens ni physiciens, de diverses méthodes sont employées par les enseignants pour que les apprenants puissent associer le phonème et son impression acoustique à une « image » acoustique, qui leur facilite d'identifier le phonème en question et de le distinguer des autres. En ce qui concerne le /u/ français, comme nous l'avons déjà vu, il s'agit du caractère « grave », dû à la concentration d'énergie sur les basses fréquences (< 1000 Hz), autour des deux premiers formants regroupés et donc renforcés (Vaissière 2008[476]). Que peut-on associer à ce caractère grave du /u/ français ?

#### 4.3.3.1. Des icônes qui représentent la gravité (vs. acuité) et la tension musculaire (Charliac *et al.*)

Charliac *et al.* (2003)[477] utilisent une série d'icônes qui représentent des propriétés articulatoires et acoustiques, dont ceux qui illustrent l'acuité (avec une note de musique aiguë et grave) et la tension musculaire (un bonhomme qui tend les bras et une jambe, et qui a les bras baissés : Figure 148).

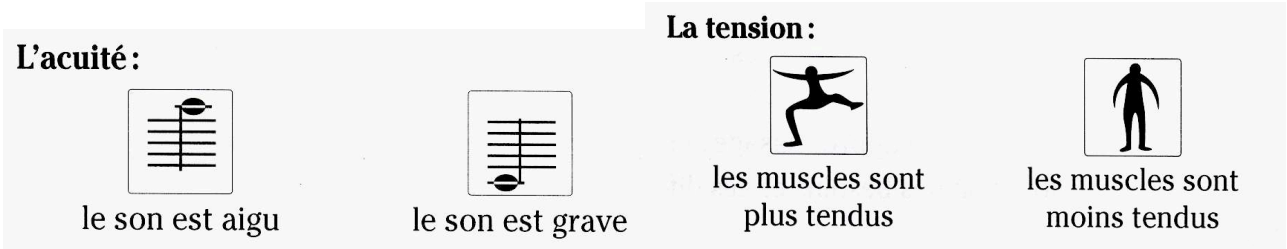



**Figure 148 :** Icônes représentant l'acuité (gauche) et la tension musculaire (droite). Charliac *et al.* (2003 : 4).

Les auteurs présentent les oppositions /y/-/u/, /y/-/Œ/, et /Œ/-/u/ en utilisant ces icônes en combinaison avec ceux qui représentent le lieu d'articulation : l'acuité et l'antériorité de la langue pour /u/-/y/, la tension musculaire et l'aperture pour /y/-/Œ/, et l'acuité et l'antériorité de la langue pour /Œ/-/u/ (Figure 149).

---

[476] J. Vaissière. Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal* 2, sous presse.

[477] L. Charliac, A. C. Motron, J.-T. Le Bougnec, B. Loreil. *Phonétique progressive du français (niveau débutant)*. Paris, CLE International, 2003.

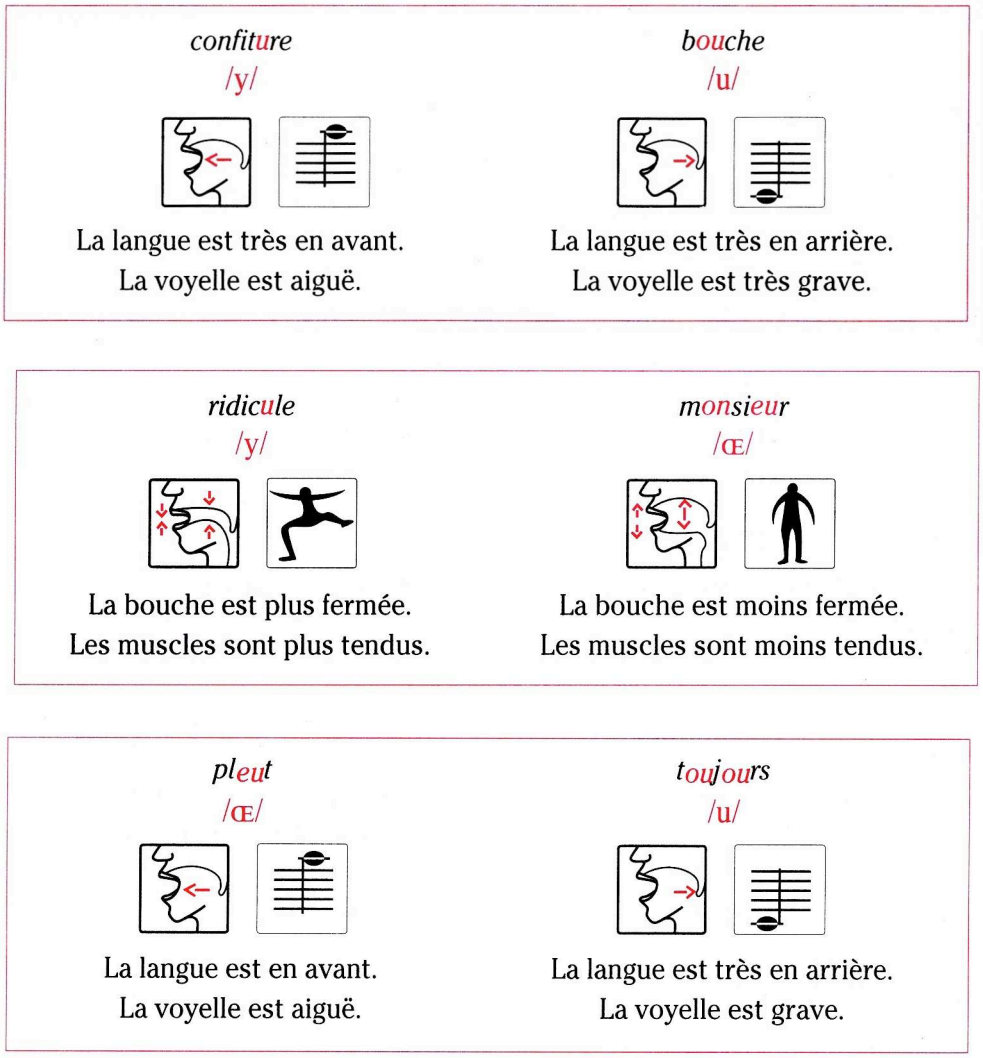



**Figure 149 :** Représentation par icônes des paires de voyelles /y/-/u/ (en haut), /y/-/œ/ (au milieu), et /œ/-/u/ (en bas). Charliac *et al.* (2003 : 48, 50, 68)

L'avantage de cette méthode consiste en une simplicité de la notation par note de musique, que les non-linguistes pourraient saisir assez facilement, même si l'association du timbre (l'acuité ou la gravité) à la fréquence fondamentale (la note de musique aiguë ou grave) n'est pas fondée scientifiquement : ce serait plutôt à la différence d'instruments de musique (clavecin et piano, par exemple) que la différence de timbre vocalique devrait être comparée, mais l'association serait probablement beaucoup plus difficile, et ne faciliterait donc pas l'apprentissage.

#### 4.3.3.2. Des images sonores et l'intégration corporelle (Kaneman-Pougatch et Pedoya-Guimbretière)

Kaneman-Pougatch et Pedoya-Guimbretière (1991)[478] proposent une activité corporelle dans laquelle des phones (voyelles et consonnes) sont associés à une image sonore représentée par

[478] M. Kaneman-Pougatch, E. Pedoya-Guimbretière. *Plaisir des sons : enseignement des sons du français*. Paris, Hatier/Didier, 1991.

des gestes corporels. L'objectif de cette activité est de « *solliciter le corps pendant la production vocale, pour favoriser l'acquisition et renforcer la mémoire* » (résumé de Lauret 2007 : 123[479]).

La gravité de /u/ est associée aux gestes suivants : « *Le son* [u] *qui est grave trouve son chemin le long des jambes et finit par sortir par une petite bouche imaginaire qui, elle, se trouve sous la plante des pieds* » (Kaneman-Pougatch et Pedoya-Guimbretière 1991 : 16. Nous avons reproduit les « images » des trois voyelles fermées à la Figure 150, celle de /ø/ à la Figure 151).




**Figure 150 :** « Images » des voyelles fermées /y/, /i/ et /u/ (Kaneman-Pougatch et Pedoya-Guimbretière 1991 : 16).

[479] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.



**[ɸ]**

Inspirez, puis sur expiration, prononcez **[ɸ]** en continu en le faisant sortir par le petit creux qui se situe au centre, entre la base du nez et la lèvre supérieure, comme si vous aviez à cet endroit un petit trou tout rond, aux bords tendus, puisque **[ɸ]** est un son labial. Exercez-vous à sentir des vibrations à cet emplacement quand s'échappe votre **[ɸ]**.

**Figure 151 :** « Image » de la voyelle /ø/ (Kaneman-Pougatch et Pedoya-Guimbretière 1991 : 41).

Ces images sonores et les gestes associés ne correspondent pas nécessairement à des propriétés articulatoires ou acoustiques, mais comme les auteures le soulignent, les images sont fondées sur trois propriétés articulatoires et acoustiques : la tension (musculaire), l'acuité, et la labialité. En outre, les auteures précisent que ces idées ne sont « *que des suggestions, des pistes, des exemples parmi d'autres* », en ajoutant qu'« *Il est bon que l'élève construise son propre système d'images à travers son vécu et son appréhension du monde* » (p. 10)[480]. L'importance attachée à l'initiative des apprenants se rejoint à la position prise par Lauret (2007 : 131) :

> « […] *on pourrait presque dire, sans provocation, qu'il n'y a de méthode que celle que se constitue l'apprenant. Celui qui progresse est ici celui qui s'investit dans son apprentissage et qui prend goût à entendre, à reconnaître, à imiter, à jouer avec la nouvelle langue. L'approche pédagogique de l'enseignant, qui doit susciter l'implication personnelle (narcissique) de l'élève, est plus importante que les techniques d'entraînement employées.* »

#### 4.3.3.3. Les couleurs des voyelles

Les divers timbres vocaliques peuvent être associés à des couleurs différentes. Fónagy (1983 : 59-75)[481] présente une série d'enquêtes effectuées sur les métaphores en phonétique auprès de locuteurs natifs du hongrois (de diverses générations), du français, et de l'anglais américain. Une des questions posées était de savoir laquelle des deux voyelles /i/ et /u/ correspond mieux aux qualifications proposées, dont « sombre (*dark*) », « gros (*bigger*) », *etc*. Les résultats montrent que la grande majorité de tous les groupes de sujets (trois langues et différentes générations) ont répondu que le /u/ (/uː/ pour l'anglais) était le plus sombre (92 Hongrois sur 95, 33 Français sur 34, 89 Américains sur 97). Ce résultat suggère que l'association entre l'image acoustique et certaines images métaphoriques ne dépendrait pas de la langue native du locuteur. Si c'est le cas, ces images métaphoriques seront utilisables pour représenter des phones des langues étrangères et secondes.

---

[480] Notons que les images adoptées dans Anan et Guimbretière (1995 : voir Figure 141, p. 279) ne sont pas identiques, tout en gardant les images acoustiques de base (acuité pour [i], gravité pou [u]).

[481] I. Fónagy. *La vive voix : essais de psycho-phonétique*. Paris, Payot, 1983.

Gattegno a proposé un tableau de couleurs associées aux phonèmes de l'anglais américain (Figure 152) dans sa méthode d'enseignement « *the Silent Way* » (Gattegno 1963[482] ; présentée également dans Celce-Murcia *et al.* 1996[483]). Les apprenants peuvent associer ainsi les timbres vocaliques à des diverses couleurs sans avoir à recourir ni à l'orthographe (notons que le rapport graphie-phonie est beaucoup moins transparent que celui du français) ni aux symboles phonétiques. En revanche, l'association entre les voyelles et les couleurs dans le système de Gattegno ne semble pas être fondée sur les propriétés acoustiques. Les voyelles postérieures, par exemple, qui sont caractérisées par un timbre grave dû à la concentration des deux premiers formants dans les basses fréquences, sont représentées tantôt par des couleurs plutôt sombres (vert pour /uː/ dans *noon*, gris pour /ʊ/ dans *put*), tantôt par des couleurs plutôt claires (orange pour /ɔː/ dans *all*). Nous pouvons également poser la question sur le choix des voyelles : nous n'y trouvons pas toutes les voyelles colorées de /r/ (*r-colored vowels* ; il manque /ɪr/ et les diphtongues), tandis qu'un petit nombre de séquences « semi-consonne + voyelle » sont incluses (/juː/ dans *you*, /jʊr/ dans *your*, /wʌn/ dans *one*, et /wɑːr/ dans *memoire*, mais cette dernière est rare en anglais). Il ne semble pas approprié non plus de donner comme mots clés des mots grammaticaux qui sont prononcés plus souvent dans leur forme faible (ex. la forme faible /ət/ au lieu de la forme forte /æt/ pour *at*).

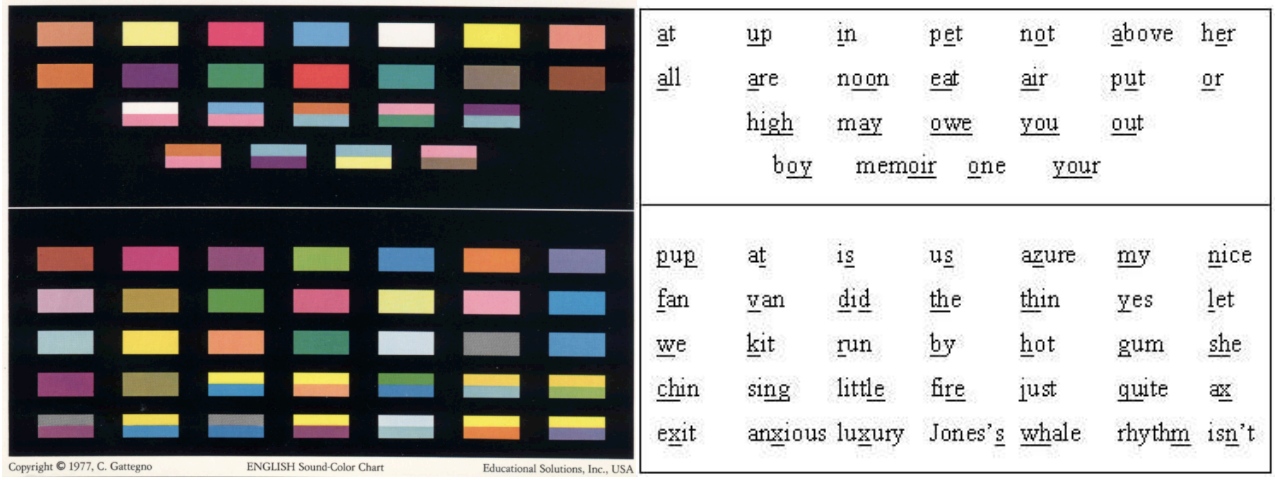



**Figure 152 :** Tableau original de sons et couleurs (*Sound Color Chart* : gauche) de « *the Silent Way* » de Gattegno et les mots clés (droite). Les voyelles sont disposées en haut, et les consonnes en bas.

[482] C. Gattegno. *Teaching foreign languages in schools: the silent way*. Reading (England), Educational Explorers, 1963.

[483] M. Celce-Murcia, D. M. Brinton, J. M. Goodwin. *Teaching Pronunciation: a Reference for Teachers of English to Speakers of Other Languages*. Cambridge, Cambridge University Press, 1996.




**Figure 153 :** Exemple d'application[484] de la table de couleurs des voyelles de l'anglais de Gattegno.

Wrembel (2007)[485] a proposé un tableau coloré de voyelles pour le polonais et l'anglais (Figure 154) en se fondant sur des études perceptives effectuées auprès des locuteurs natifs des deux langues. Dans ces expériences, les tâches des sujets consistaient à écouter les voyelles de leur langue maternelle (en contexte isolé et dans une syllabe /b_d/), et de choisir une des 11 couleurs de base présentées sur une palette. Les réponses des anglophones et des Polonais ont montré des tendances similaires, qui sont les suivantes : les voyelles antérieures fermées (/i/ du polonais, /i:/ de l'anglais) ont été associées principalement au jaune et au vert, les voyelles ouvertes (/a/ du polonais, /æ/ et /ɑ:/ de l'anglais) au rouge, et les voyelles postérieures aux couleurs sombres. Nous y retrouvons donc une association entre l'acuité des voyelles antérieures fermées, due à la concentration d'énergie sur les hautes fréquences (F2 et F3, ou F3 et F4), et les couleurs claires, d'une part, et entre la gravité des voyelles postérieures arrondies, due à la concentration d'énergie sur les basses fréquences (F1 et F2), et les couleurs sombres, de l'autre. Les tableaux de Wrembel sont donc fondés sur des propriétés acoustiques et sur la perception réelle des auditeurs.

---

[484] Tiré du site « *English Raven* » créé par Jason D. Renshaw. Les mots et les dessins ont été ajoutés par l'auteur du site. http://www.englishraven.com/Phonics_soundcolor.html

[485] M. Wrembel. "Still sounds like a rainbow" - a proposal for a coloured vowel chart. *Proceedings of Phonetics Teaching and Learning Conference 2007*, Paper ID 14, 4 pages, 2007.

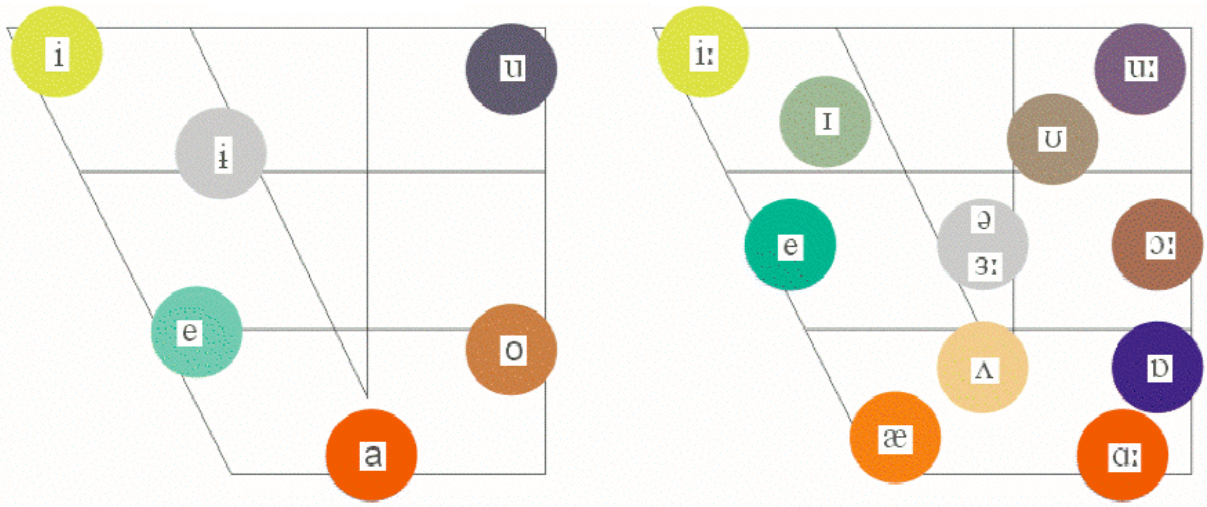



Figure 1. Coloured vowel charts for Polish and English

**Figure 154 :** Tableaux colorés de voyelles pour le polonais (gauche) et pour l'anglais RP (droite) proposés par Wrembel (2007).

#### 4.3.3.4. Des coupes spectrales schématisées

Les propriétés physiques sont en général difficilement applicables directement aux matériels pédagogiques. Les coupes spectrales du /u/ français prononcé par un locuteur natif et par un apprenant japonophone sont représentées à la Figure 155. Les phonéticiens y trouvent des différences remarquables entre les deux. La concentration d'énergie sur les basses fréquences (< 1000 Hz) due au regroupement des deux premiers formants, qui contraste avec des pics très bas des formants supérieurs, donnant lieu à une sonorité grave, est très nettement visible dans la coupe spectrale du /u/ français prononcé par le locuteur natif, alors que celle de l'apprenant présente une distribution plus équilibrée d'énergie, ce qui caractérise une voyelle « acoustiquement centrale » (Vaissière 2006 : 73). En revanche, l'association entre ces coupes spectrales et la sonorité plus ou moins grave ne serait pas évidente pour les apprenants[486].

[486] Des spectres similaires sont utilisés pour la rééducation des enfants malentendants (fonction « *Spectra Patterning* » de SpeechViewer III, IBM).

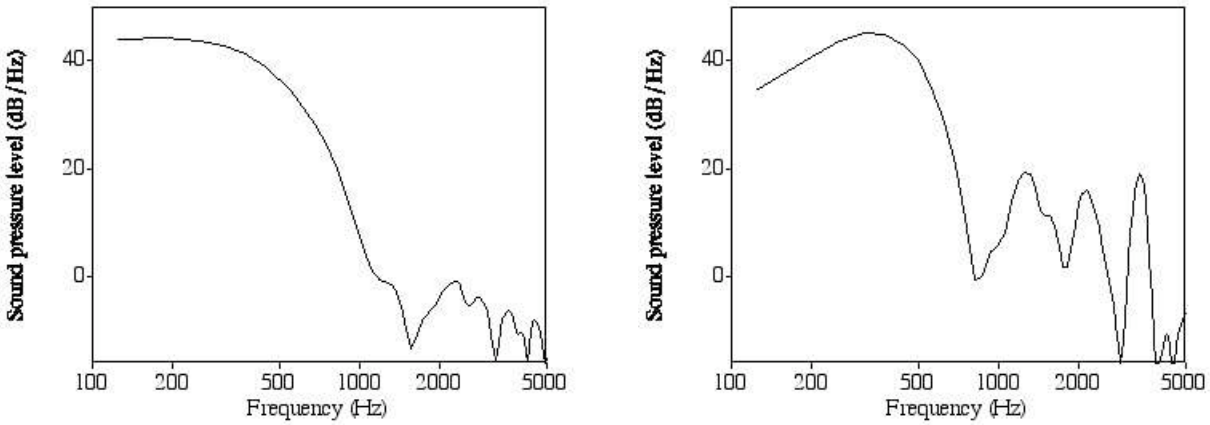



**Figure 155 :** Coupes spectrales (Praat ; échelle logarithmique) du /u/ français prononcé par un locuteur natif (gauche) et par un apprenant japonophone (droite).

Les images de la Figure 156 sont fondées sur les coupes spectrales de la Figure 155. Nous avons collé ces dernières, tournées de 90 degrés à gauche, avec leurs images en miroir. Ensuite, nous avons coloré celle du /u/ français du natif en bleu foncé, et celle du /u/ de l'apprenant en bleu clair, en considérant que la gravité serait associée à une couleur sombre, comme nous l'avons vu dans la section précédente. La hauteur géométrique correspond donc à la hauteur de la fréquence et à l'acuité. La concentration d'énergie sur les basses fréquences du /u/ français est ainsi représentée par la grande masse de la partie inférieure, qui contraste avec la structure légère de la partie supérieure, alors que la partie supérieure de l'image du /u/ de l'apprenant est plus imposante.

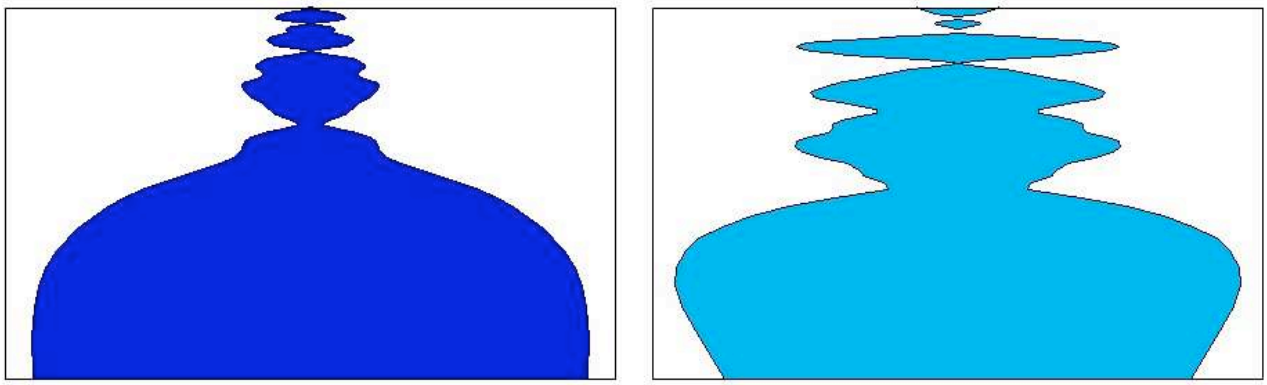


**Figure 156 :** Coupes spectrales schématisées du /u/ français prononcé par un locuteur francophone natif (gauche) et par un apprenant japonophone (droite).

L'avantage de cette représentation consiste en une association des propriétés acoustiques à une image plus abstraite que nous pourrions avoir instinctivement sur les voyelles concernées.

Il serait possible d'automatiser la procédure en fondant la couleur en fonction d'un certain nombre de propriétés acoustiques, comme le centre de gravité, le rapport d'énergie entre différentes bandes de fréquences.

### *4.3.4. Synthèse articulatoire (VTCalc et TractSyn)*

Les synthétiseurs articulatoires nous permettent de comprendre les rapports entre la configuration articulatoire du conduit vocal et le produit acoustique. Les synthétiseurs ayant une interface visuelle du conduit vocal seraient particulièrement utiles pour la didactique de la prononciation, pour que les apprenants puissent prendre conscience de la relation entre la forme du conduit vocal et les gestes articulatoires. Nous présentons ici deux synthétiseurs, VTCalc (Maeda 1982)[487], utilisé pour générer les stimuli de l'Expérience 9 (perception de stimuli synthétisés), et TractSyn (Birkholz *et al.* 2006)[488].

Nous présentons dans ce qui suit un essai de simulation d'un [u] français, d'un [ɯ] (/u/) japonais, et d'un [ø] français, faite à l'aide des deux synthétiseurs en imitant les profils articulatoires de Wioland (1991)[489] et d'Uemura (1990)[490] représentés à la Figure 157. Le [u] français typique peut être modélisé par un conduit vocal avec deux fortes constrictions, une au milieu (région vélaire) et une aux lèvres (comme le montre le profil articulatoire de Wioland à la Figure 157), ce qui forme deux résonateurs Helmholtz (Fant 1960[491], Vaissière 2007[492]).

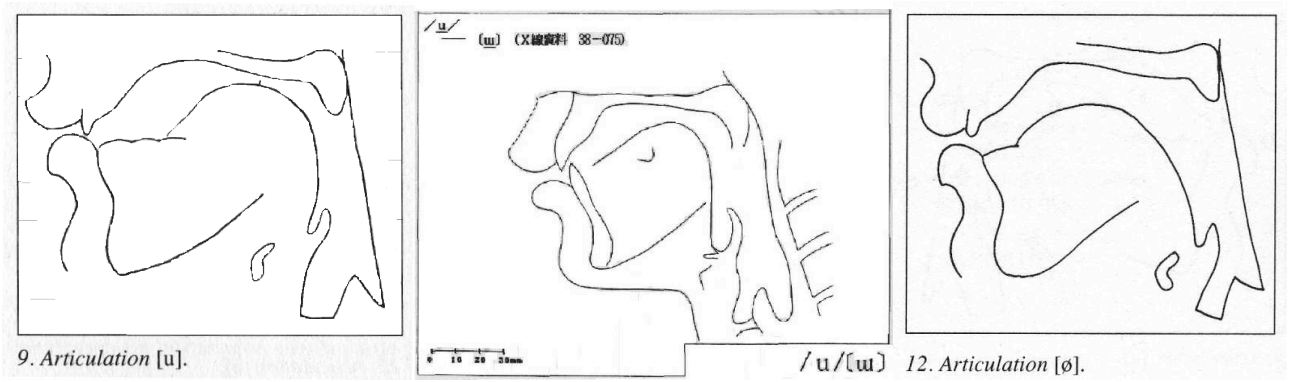



**Figure 157 :** Profils articulatoires d'un [u] français (Wioland 1991 : 17), d'un [ɯ] (/u/) japonais (Uemura 1990 : 59), et d'un [ø] français (Wioland 1991 : 18).

---

[487] S. Maeda. A digital simulation method of the vocal tract system. *Speech Communication* 1(3-4) : 199-229, 1982.

[488] P. Birkholz, D. Jackèl, B. J. Kröger. Construction and control of a three-dimensional vocal tract model. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 2006*, I : 873-876, 2006.

[489] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

[490] 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]. Tokyo, Shûei shuppan, 1990.

[491] G. Fant. *Acoustic Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*. The Hague, Mouton, 1960.

[492] J. Vaissière. "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", in M. J. Solé, P. S. Beddor, M. Ohala, *Experimental Approaches to Phonology*. Oxford, Oxford University Press, 54-71, 2007.

#### 4.3.4.1. VTCalc

VTCalc (*Vocal Tract Calculations*) a été développé par Shinji Maeda, sur la base de son modèle articulatoire. Le logiciel calcule les caractéristiques acoustiques de l'ensemble d'un conduit vocal, de la glotte jusqu'aux lèvres, spécifié par les 7 paramètres articulatoires suivants (Maeda 1982, *op. cit.*) : 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-*.

L'interface graphique de VTCalc est représentée à la Figure 158. Les données d'entrée sont les 7 paramètres articulatoires mentionnés ci-dessus, à régler en utilisant les boutons se trouvant en haut, à droite de la fenêtre. Les utilisateurs peuvent également régler le nombre de sections du conduit vocal (une section : 1 cm de longueur) et des deux paramètres concernant la nasalité (le lieu du branchement de la cavité nasale, l'aire de couplage nasal). En fonction de ces paramètres articulatoires, le profil articulatoire (en haut, au milieu) et la fonction d'aire (en haut, à gauche) sont modifiés. Par la suite, la fonction de transfert (en bas, au milieu) et la fréquence, la largeur de la bande, et l'amplitude des 5 premiers formants (en bas, à droite) sont obtenues en cliquant sur « *calculate* » de la barre de menu. Les utilisateurs peuvent également écouter le son en cliquant sur « *synthetize* » de la barre de menu.

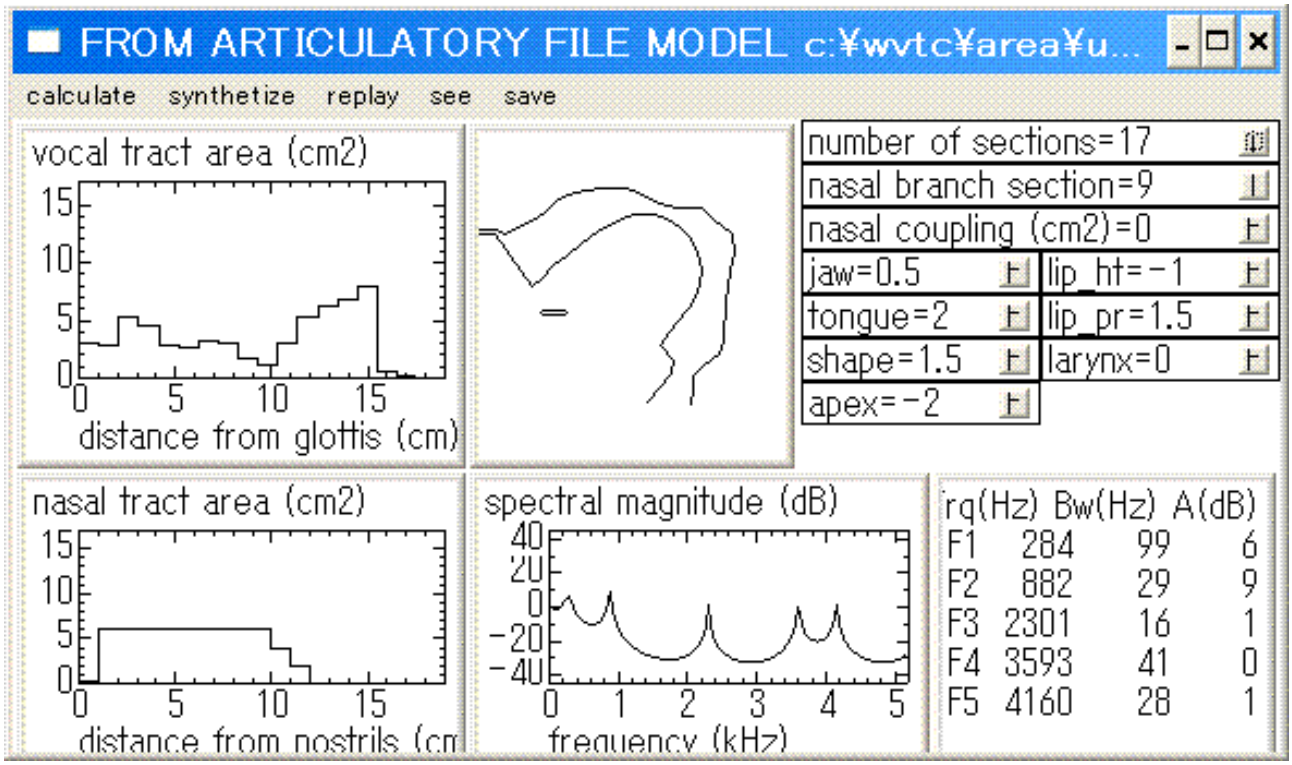



**Figure 158 :** Interface graphique de VTCalc. La fonction d'aire du conduit vocal (en haut, à gauche), le profil articulatoire simulé (en haut, au milieu), les paramètres articulatoires (en haut, à droite), la fonction d'aire du conduit nasal (en bas, à gauche), la fonction de transfert (en bas, au milieu), la fréquence (Hz), la largeur de bande (Hz) et l'amplitude (dB) des 5 premiers formants (en bas, à droite).

Nous avons simulé un [u] français, un [ɯ] (/u/) japonais, et un [ø] français en imitant les profils articulatoires de la Figure 157. La Figure 159 montre les trois voyelles simulées en utilisant VTCalc, et les spectrogrammes des voyelles de sortie. Nous pouvons y observer des différences articulatoires et acoustiques entre le [u] français et le [ɯ] (/u/) japonais, et des ressemblances acoustiques entre le [ɯ] (/u/) japonais et le [ø] français.

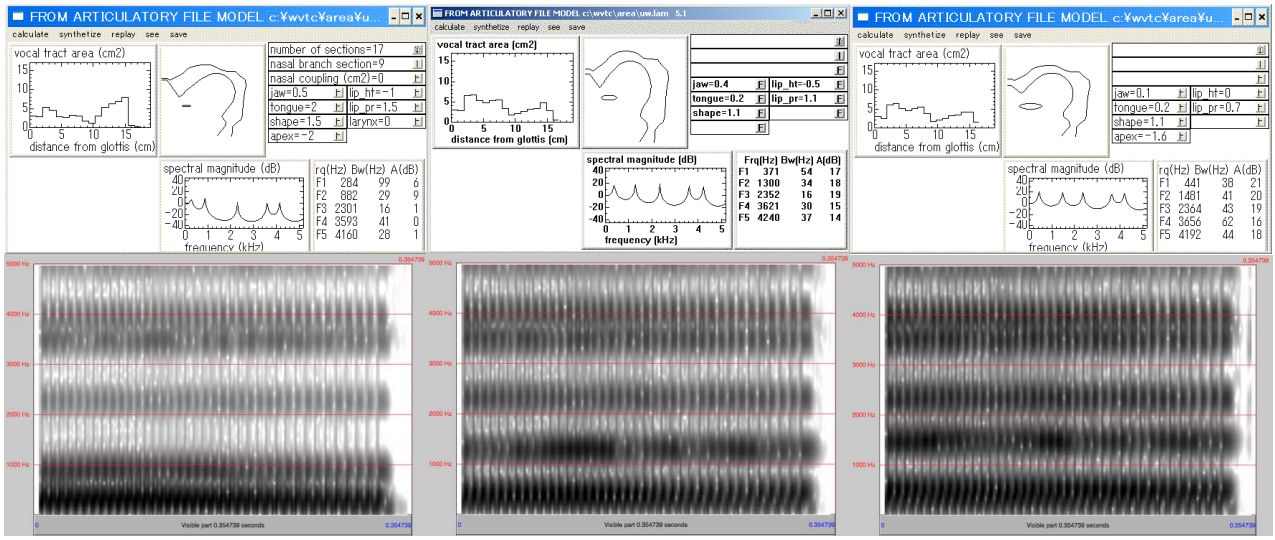



**Figure 159 :** Simulation d'un [u] français (gauche), d'un [ɯ] (/u/) japonais (milieu), et d'un [ø] français (droite) réalisée sous VTCalc. Les fenêtres de VTCalc (en haut) : les fonctions d'aire (en haut, à gauche), les profils articulatoires (en haut, au milieu), les paramètres articulatoires d'entrée (en haut, à droite), les fonctions de transfert (en bas, au milieu). Les spectrogrammes à bande large (sous Praat : la largeur de la fenêtre est de 5 millisecondes) des vocoïdes synthétisés (en bas).

### 4.3.4.2. TractSyn

Le synthétiseur TractSyn (version 1.0)[493] est un autre synthétiseur articulatoire développé par Peter Birkholz (*Institut für Informatik*, *Universität Rostok*, Allemagne), utilisant un modèle du conduit vocal à 3 dimensions (Birkholz *et al.* 2006)[494]. Ce synthétiseur permet aux utilisateurs de modifier la forme du conduit vocal simulé avec une interface graphique, en déplaçant des petits carreaux (points de contrôle) qui sont liés aux organes tels que les lèvres, la langue, le bout de la langue, la mâchoire, *etc.* (Figure 160). La fonction d'aire et la fonction de transfert qui correspondent au conduit vocal sont calculées et affichées sur le même écran. Les utilisateurs peuvent également écouter le son synthétisé. Ils peuvent ainsi voir et entendre les conséquences acoustiques de la modification du conduit vocal.

[493] http://www.vocaltractlab.de/index.htm?/vocaltractlab_download_en.htm

[494] P. Birkholz, D. Jackèl, B. J. Kröger. Construction and control of a three-dimensional vocal tract model. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 2006*, I : 873-876, 2006.

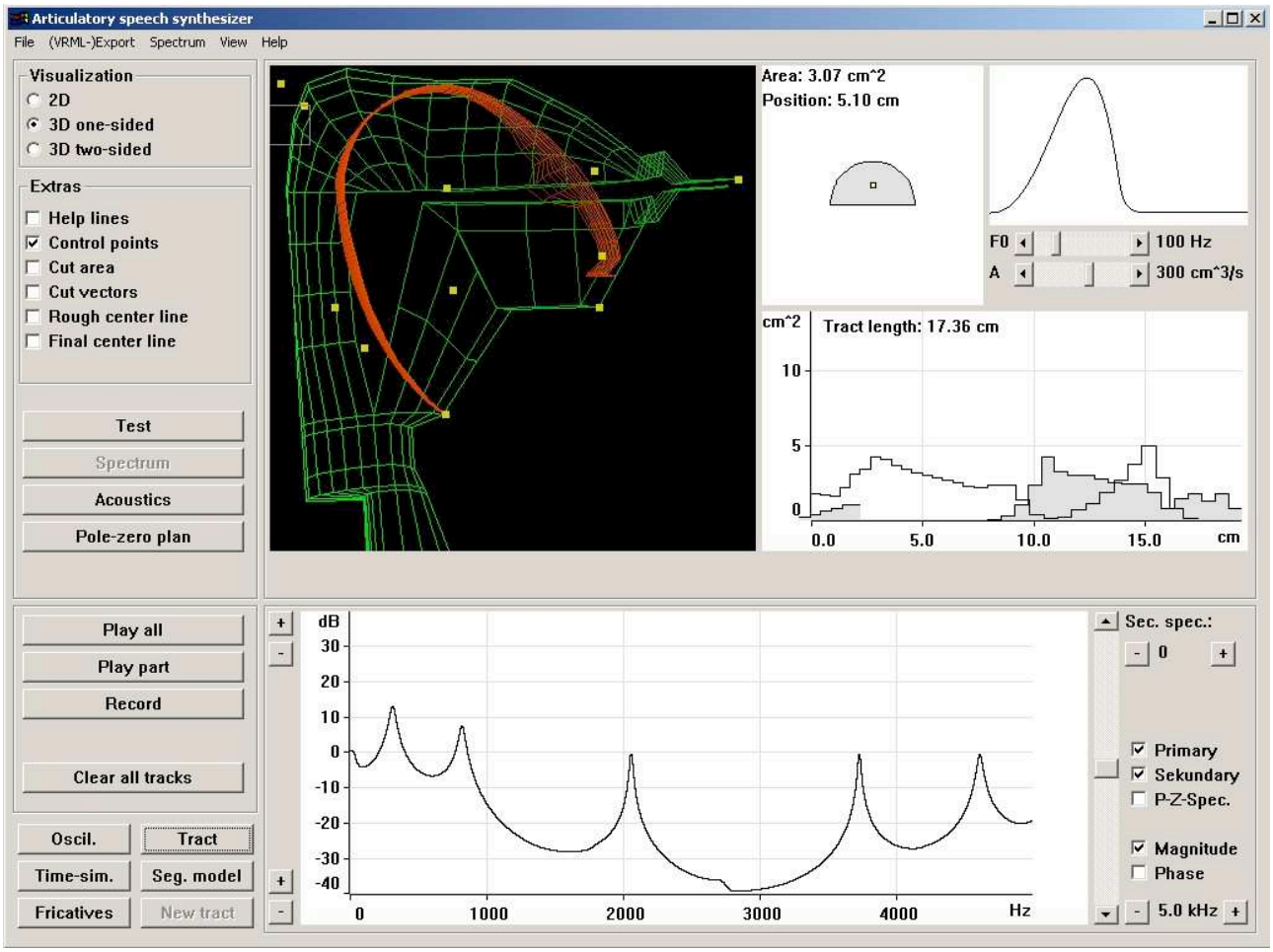



**Figure 160 :** Interface de TractSyn 1.0. Simulation d'un [u] français. L'utilisateur peut modifier le conduit vocal en déplaçant les petits carreaux. La fonction d'aire est affichée à droite, et la fonction de transfert en bas.

La Figure 161 illustre un essai de simulation d'un [u] français, d'un [ɯ] (/u/) japonais, et d'un [ø] français, faite en imitant les profils articulatoires de Wioland (1991)[495] et d'Uemura (1990)[496] représentés à la Figure 157. La Figure 161 illustre bien les différences articulatoires (la position des constrictions, l'aire des constrictions) et acoustiques (la focalisation F1/F2 et la baisse d'énergie des formants supérieurs *vs.* la répartition plus équilibrée d'énergie) entre le [u] français et le [ɯ] (/u/) japonais, et les différences articulatoires et les ressemblances acoustiques entre le [ɯ] (/u/) japonais et le [ø] français.

[495] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

[496] 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]. Tokyo, Shûei shuppan, 1990.

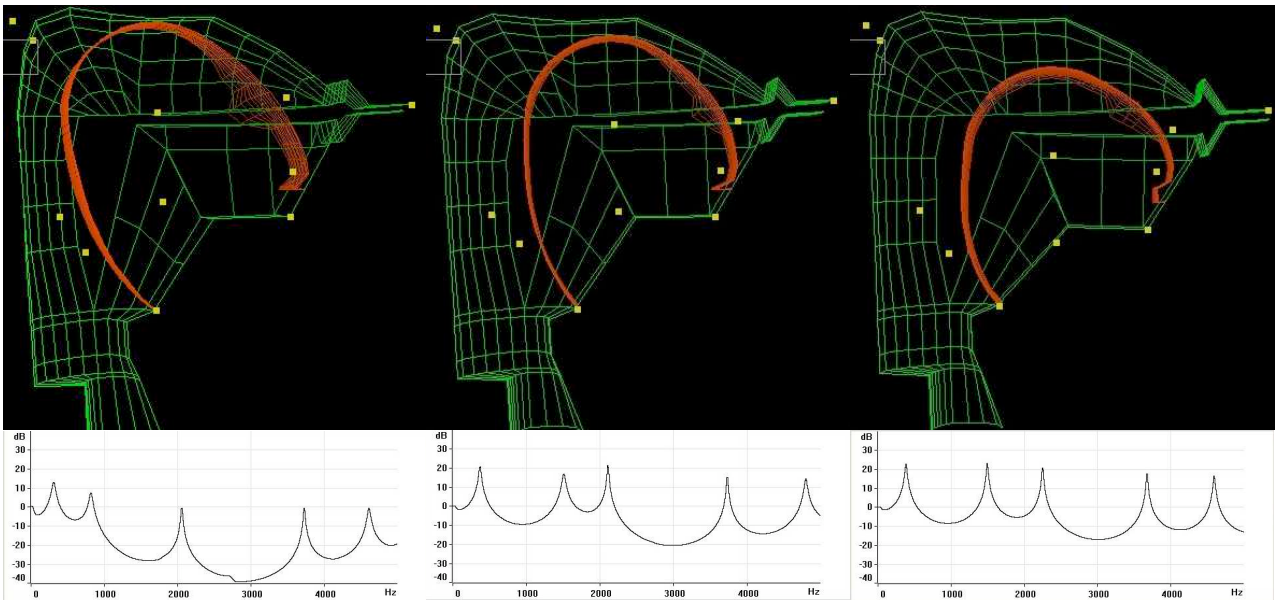



**Figure 161 :** Simulation d'un [u] français (gauche), d'un [ɯ] (/u/) japonais (milieu), et d'un [ø] français (droite) effectuée sous TractSyn. Les profils articulatoires d'entrée (haut) et les fonctions de transfert (bas) des voyelles synthétisées.

L'avantage d'utilisation des synthétiseurs articulatoires consiste en la possibilité d'illustrer la configuration articulatoire (notamment la position et la forme de la langue, qui est invisible de l'extérieur) et ses conséquences acoustiques. TractSyn, en particulier, permet de le faire de manière interactive.

### *4.3.5. Spectrogramme en temps réel (WaveSurfer)*

Les caractéristiques acoustiques du /u/ français, ainsi que les différences entre le /u/ français et le /u/ [ɯ] japonais, peuvent être montrées par des spectrogrammes. En revanche, les apprenants ne sont pas tous des phonéticiens qui savent lire les spectrogrammes facilement. L'enseignant pourrait attirer leur attention sur un point précis, tel que le regroupement des deux premiers formants observé pour le /u/ français. Ce travail sera facilité si l'on peut illustrer les différences entre les voyelles en temps réel, de manière interactive, afin que les apprenants puissent comparer leur écoute (en écoutant les sons prononcés par l'enseignant), l'articulation (la partie visible comme les lèvres, en regardant leur enseignant), et les caractéristiques acoustiques visualisées (spectrogramme)[497].

[497] Concernant la F0, la représentation en temps réel est disponible sous WinPitch (Martin 1996). Ce logiciel permet également de modifier des contours de F0 dans une interface facile à manipuler. Il existe aussi une version développée pour l'enseignement de la prosodie (WinPitch LTL : Martin 2005).

Ph. Martin. WinPitch : un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows. *Actes des XXIV Journées d'Étude sur la Parole*, Avignon, 224-227, 1996.

Ph. Martin. WinPitch LTL, un logiciel multimédia d'enseignement de la prosodie. *Apprentissage des langues et systèmes d'information et de communication* (ALSIC) 8(2) : spécial Atala : 95-108, 2005.

La Figure 162 montre un exemple de spectrogramme du logiciel d'analyse phonétique WaveSurfer (développé par Kåre Sjölander et Jonas Beskow, KTH, Suède)[498]. Le spectrogramme représente une alternance entre un [ɯ] (/u/) japonais et un [u] français. Nous y observons bien que lors du passage du [ɯ] (/u/) japonais au [u] français, le F2 plonge en dessous de 1000 Hz et rejoint le F1, et que les formants supérieurs s'affaiblissent considérablement, ce qui contraste nettement avec la concentration d'énergie sur les basses fréquences (< 1000 Hz) due au regroupement des deux premiers formants.

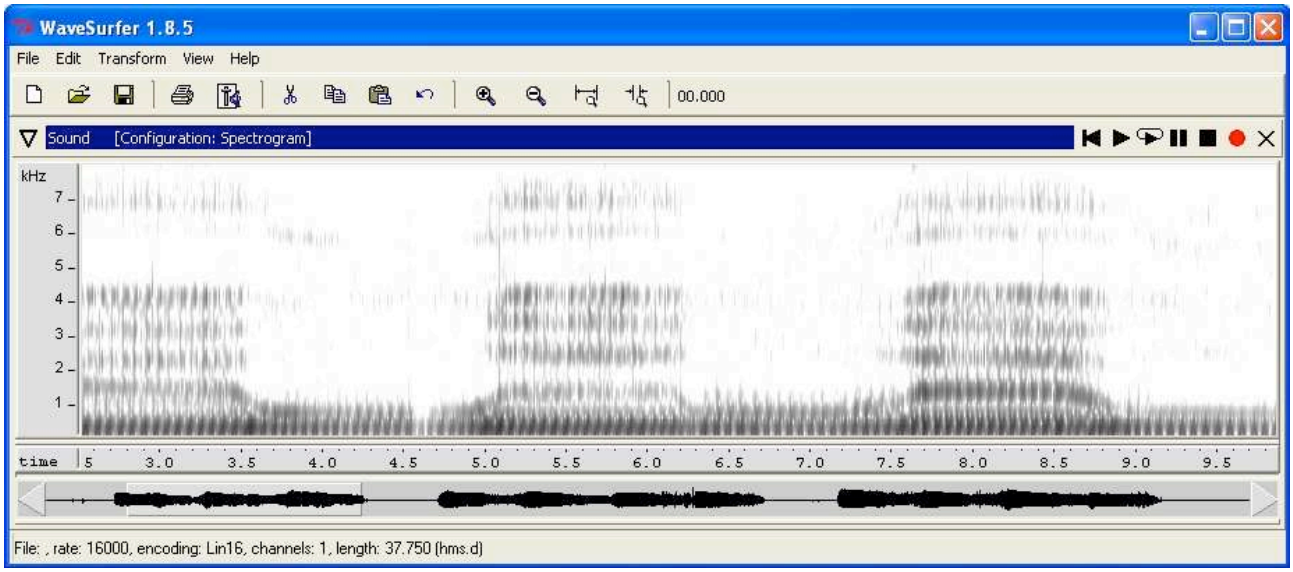



**Figure 162 :** Exemple de spectrogramme de WaveSurfer. Une alternance entre un [ɯ] (/u/) japonais et un [u] français y est illustrée. Le spectrogramme correspond à la partie sélectionnée de l'oscillogramme en bas (vers le début de la séquence affichée). Visualisation en temps réel.

Nous pouvons utiliser cet outil dans un premier temps pour illustrer les différences entre les deux phones des deux langues, en même temps que l'enseignant montre l'articulation aux apprenants. Par la suite, le logiciel peut être utilisé pour donner un feedback en temps réel aux apprenants : ils seront invités à prononcer /u/ de sorte que les deux « bandes » se rejoignent en bas du spectrogramme, le plus bas possible.

### 4.4. CONCLUSION PARTIELLE

Dans la présente partie, nous avons commencé par étudier 71 manuels généralistes de français langue étrangère publiés au Japon. La difficulté relative de /u/ et la différence entre le /u/ français et le /u/ [ɯ] japonais ne sont pas suffisamment traitées dans les manuels généralistes, voire dans l'enseignement du français langue étrangère au Japon. D'ailleurs, le nombre de manuels proposant des activités ou des exercices de prononciation est limité (16 sur 71).

[498] http://www.speech.kth.se/wavesurfer/index.html

Nous avons présenté par la suite divers outils qui permettraient de faciliter la prise de conscience des différences entre les phones de L2 et ceux de L1. Ces outils ont été intégrés dans la proposition de fiche pédagogique d'un cours qui porte sur l'opposition /u/-/ø/, et est destiné aux apprenants japonophones.

Nous avons ainsi montré comment les résultats des expériences de perception et de production présentées dans les deuxième et troisième parties de notre étude pourraient être appliqués à l'enseignement de la prononciation du français langue étrangère, en recourant à divers outils et méthodes. Ce qui est important à retenir est l'importance de la diversité des méthodes à adopter selon les apprenants et la situation, comme le souligne Lauret (2007 : 170)[499] :

> « *Enfin, il semble essentiel de se méfier des tenants de la solution phonétique : à chaque apprenant son évolution et sa solution adaptée. C'est plutôt à cette permanente attention sur les moyens variés de l'entraînement et au plaisir à produire la nouvelle langue que l'enseignant doit inviter l'apprenant, plutôt que d'imposer une méthode phonétique, qu'il n'aura souvent pas le loisir d'exploiter pleinement en classe*. »

Il est important de retenir que c'est à chaque apprenant de travailler sur la prononciation, en écoutant régulièrement la langue cible, et en prononçant, dans la mesure du possible, avec des points à améliorer en tête. Le rôle de l'enseignant est d'animer et d'encourager ce processus d'apprentissage.

[499] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

# Conclusion générale

Nous avons présenté dans cette thèse une série de réflexions et des expériences acoustiques et perceptives portant sur l'apprentissage des voyelles du français langue étrangère chez des apprenants japonophones. Nous avons mis en évidence ou confirmé les facteurs suivants :

1) Un certain nombre d'erreurs semblent être induites ou favorisées par les correspondances phonie-graphie.[500] [501] Par exemple, nous avons vu que <eu> /ø/, associé incorrectement (en français) au /juː/ anglais, peut être associé au /juR/ [jɯː] (et /ju/ [jɯ]) japonais.

2) Le système phonémique et la représentation des phonèmes chez les apprenants joue un rôle[502]. Le /u/ français (voyelle postérieure fermée) est phonémiquement similaire au /u/ [ɯ] japonais (voyelle postérieure fermée). Les apprenants japonophones ont tendance à prononcer le /u/ français (voyelle focale F1/F2) avec un timbre proche de celui du /u/ [ɯ] japonais (voyelle acoustiquement centrale).

3) Les caractéristiques acoustiques intrinsèques de la réalisation des voyelles[503] : la confusion perceptive entre /u/-/o/ français serait due au fait que F2 du /o/ japonais est plus bas que celui du /u/ [ɯ] japonais.

4) L'acoustique semble primer sur l'articulatoire[504]: les voyelles antérieures arrondies ressemblent acoustiquement aux voyelles postérieures non arrondies, qui sont perçues comme des voyelles antérieures arrondies pour les locuteurs de langues qui n'ont pas de voyelles postérieures non arrondies (francophones). Plus concrètement, /u/ prononcé à la japonaise (langue centrale-postérieure, lèvres légèrement arrondies) est identifié comme /ø/ par des francophones ; des vocoïdes simulant une voyelle postérieure non arrondie (par synthèse articulatoire) sont perçus comme /œ/ par des francophones.

5) La réalisation des traits d'antériorité et de labialité est simultanée pour le /y/ français par les locuteurs francophones natifs, mais n'est pas simultanée en ce qui concerne son l'adaptation en japonais et en anglais, où l'arrondissement est réalisé postérieurement à l'antériorité ; ou en coréen, où l'antériorité est réalisée postérieurement à la labialité.[505]

6) L'influence de l'anglais comme deuxième langue[506]: <eu> /ø/, associé incorrectement (en français) à /juː/ anglais, peut être associé à /juR/ [jɯː] (et /ju/ [jɯ]) japonais ; l'acquisition de [u] quantique, avec F1/F2 bas, n'est pas facilitée par l'anglais L2, car /uː/ anglais ressemble à /u/ [ɯ] japonais, du moins sur le plan F1-F2.

---

[500] Chapitre 3 de la partie 2 ; chapitre 2 de la partie 3.
[501] Figure 59, p. 142
[502] Chapitre 3 de la partie 2 ; chapitre 2 de la partie 3.
[503] Chapitres 1 et 2 de la partie 2; chapitre 1 de la partie 3.
[504] Chapitres 3 et 4 de la partie 3.
[505] Chapitre 5 de la partie 3.
[506] Chapitre 3 de la partie 2; chapitre 3 de la partie 3.

Ce travail de thèse souhaite contribuer, sur le plan théorique, à l'éclaircissement des processus complexes d'apprentissage et d'enseignement de la prononciation des langues étrangères, et sur le plan pratique, à l'amélioration de l'apprentissage et de l'enseignement de la prononciation du français langue étrangère par les apprenants japonophones. Notamment, il permet d'illustrer des cas de phones qui sont similaires ou non sur le plan phonémique et/ou phonétique.

L'image acoustique précise d'un phonème dépend d'un grand nombre de facteurs, du locuteur, de la région, du débit, du style, de sa position par rapport aux frontières de syllabe, de mot, de sa position par rapport à l'accent lexical, des phonèmes environnants, du contexte prosodique, et aussi de l'attitude du locuteur, ou de son état émotif. La réalisation de /u/ est très variable en contexte non expérimental, comme le montrent le corpus ESTER et les résultats de Gendrot et Adda (2004), mais il existe bien une cible relativement fixe, qui se révèle quand les locuteurs francophones natifs le prononce de façon isolée. En outre, l'anticipation de l'arrondissement et de la protrusion des lèvres pour le /u/ français influence les voyelles qui se trouvent dans les syllabes précédentes (sur l'anticipation, voir Sock *et al*. 1999, Vaxelaire *et al*. 2002). /œ/ dans les syllabes (accentuées) fermées serait plus éloigné de /u/ [ɯ] japonais que /ø/. La présente étude a examiné les voyelles en contexte isolé . Il est important de travailler les voyelles dans des syllabes, et dans des contextes prosodiques divers lors des activités dans la classe de langues, comme illustré dans Wioland (1991)[507], Lauret (2007)[508], ou dans des manuels de prononciation tels que Abry et Chalaron (1994)[509], Kaneman-Pougatch et Pedoya-Guimbretière (2001)[510], ou Charliac *et al.* (2003)[511].

Il semble également possible de nos jours d'intégrer des nouveaux outils informatiques qui puissent aider l'apprentissage et l'enseignement de la prononciation, comme la représentation spectrographique en temps réel, ou l'utilisation de synthétiseurs articulatoires, par exemple. La représentation visuelle de diverses caractéristiques phonétiques, tels que les formants et la concentration d'énergie sur une zone fréquentielle (pour /u/ et /y/ français, par exemple), ou les conséquences sur le spectre de gestes articulatoires précises (comme l'arrondissement et/ou la protrusion des lèvres) devrait faciliter la prise de conscience des apprenants de la relation entre gestes articulatoires, image acoustique et son perçu. Un travail ultérieur consistera à tester ces apports.

---

[507] F. Wioland. *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette, 1991.

[508] B. Lauret. *Enseigner la prononciation du français : questions et outils*. Paris, Hachette, 2007.

[509] D. Abry, M.-L. Chalaron. *350 exercices de phonétique*. Paris, Hachette, 1994.

[510] M. Kaneman-Pougatch, E. Pedoya-Guimbretière. *Plaisir des sons : enseignement des sons du français*. Paris, Hatier/Didier, 1991.

[511] L. Charliac, A. C. Motron, J.-T. Le Bougnec, B. Loreil. *Phonétique progressive du français (niveau débutant)*. Paris, CLE International, 2003.

# Bibliographie


Abry, D., M.-L. Chalaron (1994). 350 exercices de phonétique. Paris, Hachette.

Akinaga, K. (2001). "Toukyou akusento no shuutoku housoku [Règles pour l'acquisition de l'accent de Tokyo]," in H. Kindaichi, K. Akinaga, *Shin-Meikai nihongo akusento jiten* [Dictionnaire Shin-Meikai de l'accent lexical du japonais]. Tokyo, Sanseidô, pages annexes 10-106.

Amanuma, Y., K. Ôtsubo, O. Mizutani (1978). *Nihongo Onseigaku* [Phonétique du japonais]. Tokyo, Kuroshio.

Amino, K., S.-R. Ji, S. Shinohara (2007). Perception of Japanese plosives by Korean speakers. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1789-1792.

Anan, F., E. Guimbretière (1995). *Komyunikêshon no tameno furansugo hatsuon-hou: hatsuon no kisoku to renshuu* [Plaisir de communiquer : règles et exercices de phonétique pour bien communiquer]. Tokyo, Sugugadai Shuppansha.

Beckman, M. (1982). Segment Duration and the 'Mora' in Japanese. *Phonetica* 39 : 113-135.

Best, C. T. (1995). "A Direct Realist View of Cross-Language Speech Perception," in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 171-204.

Best, C. T., G. W. McRoberts, N. N. Sithole (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click perception by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14 : 345-60.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning* 28 : 69-84.

Birkholz, D., Jackèl, B. J. Kröger (2006). Construction and control of a three-dimensional vocal tract model. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 2006* I : 873-876.

Boersma, P., D. Weenink. *Praat: doing phonetics by computer* (logiciel). http://www.praat.org/

Bôle-Richard, R. (2009). *Les mots du Saugeais : dictionnaire du patois sauget, d'après les travaux de Joseph Bobillier*. Besançon, Cêtre.

Bôle-Richard, R. (2009). *Note sur les voyelles du français local de Franche-Comté* (communication personnelle).

Boula de Mareüil, P., B. Vieru-Dimulescu (2006). The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica* 63(4) : 247-267.

Callamand, M. (1981). *Méthodologie de l'enseignement de la prononciation : organisation de la matière phonique du français et correction phonétique*. Paris, Création Loisirs Enseignement International.

CALLIOPE (1989). *La parole et son traitement automatique*. Paris, Milano, Barcelona, Mexico, Masson.

Celce-Murcia, M., D. M. Brinton, J. M. Goodwin (1996). *Teaching Pronunciation: a Reference for Teachers of English to Speakers of Other Languages*. Cambridge, Cambridge University Press.

Chang, S., M. C. Plauché, J. J. Ohala (2001). "Markedness and consonant confusion asymmetries," in E. Hume, K. Johnson, *The role of speech perception in phonology*. San Diego, Academic Press, 79-101.

Charliac, L., A. C. Motron, J.-T. Le Bougnec, B. Loreil (2003). *Phonétique progressive du français (niveau débutant)*. Paris, CLE International.

Chevallard, Y. (1985). *La transposition didactique*. Grenoble, La Pensée sauvage.

Clarke, C. M., M. F. Garrett (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America* 116(6) : 3647-3658.

Clopper, C. G., A. R. Bradlow (2006). Effects of dialect variation on speech intelligibility in noise. *The Journal of the Acoustical Society of America* 119(5) : 3424.

Collado, D. (1632). *Ars Grammaticae Iaponicae Linguae*. Rome.

Cruttenden, A. (2001). *Gimson's Pronunciation of English: 6th edn. of A. C. Gimson, An Introduction to the Pronunciation of English*. London, Edward Arnold.

Cruttenden, A. (1997). *Intonation* (2nd edition). Cambridge, Cambridge University Press.

Cuq, J.-P. (dir.) (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. Paris, Clé International/Asdifle.

Cutler, A., J. Mehler, D. Norris, J. Segui (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25(4) : 385-400.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11 : 51-62.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22 : 499-534.

Delattre, P. (1965a). "Comparing the Prosodic Features of English, German, Spanish and French," in P. Delattre, *Comparing the Phonetic Features of English, French, German and Spanish: an Interim Report*. Heidelberg, Julius Groos Verlag, 23-43.

Delattre, P. (1965b). Comparing the vocalic features of English, German, Spanish and French. *International Review of Applied Linguistics* 2 : 71-97.

Delattre, P. (1966a). "Accent de mot et accent de groupe," in P. Delattre, *Studies in French and Comparative Phonetics*. The Hague, Mouton & co., 69-72.

Delattre, P. (1966b). "La leçon d'intonation de Simone de Beauvoir : étude d'intonation déclarative comparée," in P. Delattre, *Studies in French and Comparative Phonetics*. The Hague, Mouton & co., 75-82.

Delattre, P. (1966c). Les dix intonations de base du français. *The French Review* 40(1) : 1-14.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review / La Revue canadienne des langues vivantes* 59(4) : 547-567.

Deterding, D. H. (1990). *Speaker Normalisation for Automatic Speech Recognition*. Ph.D. dissertation, University of Cambridge.

Deterding, D. H. (2006). The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association* 36 : 187-196.

Di Cristo, A. (1975). Présentation d'un test de niveau destiné à évaluer la prononciation des anglophones. *Revue de Phonétique Appliquée* 33-34 : 9-35.

Durand, J., B. Laks, C. Lyche (2003). Le projet 'Phonologie du français contemporain'. *La tribune*

*internationale des langues vivantes* 33 : 3-9.

Dutoit, T., V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken (1996). The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of the International Conference on Spoken Language Processing (ICSLP) '96*, 1393-1396.

Ellis, R. (1994). *The study of second language acquisition*. Oxford, Oxford University Press.

Esposito, A., M. G. Di Benedetto (1999). Acoustic and Perceptual Study of Gemination in Italian Stops. *The Journal of the Acoustical Society of America* 106(4 Pt 1) : 2051-2062.

Evans, B. G., P. Adank (2008). Differences in the time-course of accent adaptation: a comparison of adaptation to foreign-accented and unfamiliar regionally-accented speech. *The Journal of the Acoustical Society of America* 123(5) : 3073.

Fant, G. (1960). *Acoustic Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*. The Hague, Mouton.

Flege, J. E. (1984). The detection of French accent by American listeners. *The Journal of the Acoustical Society of America* 76(3) : 692-707.

Flege, J. E. (1987). The production of new and similar phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15 : 47-65.

Flege, J. E. (1995). "Second language speech learning: Theory, findings, and problems." in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 233-277.

Flege, J. E., E. M. Frieda, T. Nozawa (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25 : 169–186.

Flege, J. E., J. Hillenbrand (1984). Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76(3) : 708-721.

Flege, J. E., I. R. A. MacKay, D. Meador (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America* 106(5) : 2973-2987.

Flege, J. E., M. J. Munro, I. R. A. Mackay (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16 : 1-26.

Fónagy, I. (1983). *La vive voix : essais de psycho-phonétique*. Paris, Payot.

Fougeron, C., C. L. Smith (1999). "Illustrations of the IPA: French," in International Phonetic Association, *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge, Cambridge University Press, 78–81.

Fujimoto, M. (2006). Is vowel devoicing in Japanese phonological or phonetic?: studies using PGG, MRI and EMG techniques. Poster présenté à la Dixième Conférence de Phonologie de Laboratoire, Paris, 29 juin – 1er juillet.

Fujimoto, M., E. Murano, S. Niimi, S. Kiritani (2002). Differences in glottal opening pattern between Tokyo and Osaka dialect speakers: factors contributing to vowel devoicing. *Folia Phoniatrica et Logopaedica* 54(3) : 133-143.

Fujisaki, H., K. Hirose (1984). Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustic Society of Japan* 5(4) : 233–242.

Gaillard-Corvaglia, A., T. Kamiyama (2008). La /b/ forte in salentino (Puglia): uno studio acustico,

percettivo e fisiologico. *Actes du congrès international "la Comunicazione Parlata 2006,"* Nâples.

Galazzi-Matasci, E., E. Pedoya (1983). Et la pédagogie de la prononciation ? *Le français dans le monde*, 180, 39-44.

Gattegno, C. (1963). *Teaching foreign languages in schools: the silent way*. Reading (England), Educational Explorers.

Gendrot, C., M. Adda-Decker (2004). Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande. *Proceedings of the Workshop MIDL*, Paris, France, 7-12.

Gendrot, C., M. Adda-Decker, J. Vaissière (2008). Les voyelles /i/ et /y/ du français : focalisation et variations formantiques. *XXVIIèmes Journées d'Étude sur la Parole*, juin 2008 Avignon, 205-208.

Gottfried, T. L., (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics* 12 : 91-114.

Gottfried, T. L., W. Strange (1980). Identification of coarticulated vowels. *The Journal of the Acoustical Society of America* 68 : 1626-1635.

Gougenheim, G. (1935). *Éléments de phonologie française*. Paris, Belles Lettres.

Guion, S. G., J. E. Flege, R. Akahane-Yamada, J. C. Pruitt (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America* 107(5) : 2711–2724.

Hallé, P. A., B. de Boysson-Bardies, M. M. Vihman (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* 34 : 299-318.

Hattori, K., P. Iverson (2008). English /r/-/l/ pronunciation training for Japanese speakers. *The Journal of the Acoustical Society of America* 123(5) : 3327.

Hattori, K., P. Iverson (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *The Journal of the Acoustical Society of America* 125(1) : 469–479.

Hattori, S. (1984). Onseigaku [Phonétique]. Tokyo, Iwanami shoten.

Hawkins, S., J. Midgley (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35 : 183-199.

Hazan, V., A. Sennema, M. Iba, A. Faulkner (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication* 47(3) : 360-378.

Hillenbrand, J., L. A. Getty, M. J. Clark, K. Wheeler (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America* 97(5) : 3099-3111.

Hoequist, C. J. (1983). Syllable Duration in Stress-, Syllable- and Mora-Timed Languages. *Phonetica* 40 : 203-237.

Hughes, A., P. Trudgill, D. Watt (2005). *English Accents and Dialects*. London, Hodder Arnold.

Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31(2-3) : 213-238.

Inoue, F. (1997). "Intonêshon no shakaisei [Aspect social de l'intonation]," in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 143-168.

Itô, T., A. Kagawa (2000). Moji kakutoku-mae no on'in tan'i eno jikakuteki bunsetsuka no hattatsu [Développement de la segmentation consciente en unités phonologiques avant l'acquisition de la langue écrite]. *Rapport annuel du projet "Kokoro no hattatsu: ninchiteki seichou no kikou* [Développement mental : organisation du développement cognitif]" (année 1999-2000), 141-148.

Iverson, P., V. Hazan, K. Bannister (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, 118(5) : 3267-3278.

Jespersen, O., (1904). *How to teach a foreign language*. London, George Allen and Unwin.

Jones, D., (1956). *Cardinal Vowels*. London, Linguaphone Institute (livret d'explication accompagné de 2 disques audio).

Jones, D., P. Roach, J. Hartman, J. Setter (2006). *Cambridge English Pronouncing Dictionary* (17th edition). Cambridge, Cambridge University Press.

Kaiki, N., K. Takeda, Y. Sagisaka (1992). Gengojouhou o riyou shita boin keizoku jikanchou no seigyo [Contrôle de la durée des voyelles par des informations linguistiques]. *Denshi jouhou tsuushin gakkai ronbun-shi* J-75A(3) : 467-473.

Kamiyama, T. (2002). *Les aspects suprasegmentaux des phrases déclaratives courtes en français lues par les apprenants japanophones*. Mémoire de DEA (Diplôme d'Études Approfondies) de phonétique, Université de la Sorbonne Nouvelle.

Kamiyama, T. (2003). Initial pitch in words beginning with a CVV syllable with a long vowel in Tokyo Japanese. *Proceedings of the 16th International Congress of Phonetic Sciences*, 543-546.

Kamiyama, T. (2004a). L1 nihongo gakushuusha no furansugo hatsuon ni okeru tsuyosa to takasa: tanbun yomiage kara mita sono tokuchou [Des caractéristiques sur l'intensité et la F0 dans de courtes phrases en français lues par des apprenants japonophones]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2003 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2003-2004], Graduate School of Arts and Sciences, Université de Tokyo, 69-84.

Kamiyama, T. (2004b). Perception of foreign accentedness in L2 prosody and segments: L1 Japanese speakers learning L2 French. *Proceedings of the International Congress: Speech Prosody 2004*, 721-724.

Kamiyama, T. (2004c). Tokyo and Osaka Japanese: is it possible to distinguish them by prosody alone? *Proceedings of the Workshop MIDL*, Paris, France, 167-172.

Kamiyama, T. (2005a). Does Explicit Knowledge Help L2 Comprehension? The Case of Determiners du and deux in French Learned by Japanese-speaking Learners. *Proceedings of the Phonetics Teaching and Learning Conference 2005*, Department of Phonetics and Linguistics, University College London, July 2005 Paper ID 5 (4 pages).

Kamiyama, T. (2005b). L1 nihongo washa no furansugo hatsuon ni okeru purosodjî: sono keikou no shikaku-ka to suuchi-ka no kokoromi [La prosodie de phrases lues en français par des apprenants japonophones : essai de représentation graphique et de quantification de quelques tendances observées]. *Toukyou daigaku gaikokugo kyouikugaku kenkyuukai 2004 nendo kiyou* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2004-2005], Graduate School of Arts and Sciences, Université de Tokyo, 72-79.

Kamiyama, T. (2005c). Perception of foreign accentedness by native speakers and L2 learners: a preliminary study. *Actes des 7èmes Rencontres Jeunes Chercheurs de l'École Doctorale 268 'Langage et langues', Université Paris III*, 37-41.

Kamiyama, T. (2006). La production et la perception des voyelles orales françaises par les apprenants japonophones. *Actes des XXVIèmes Journées d'Étude sur la Parole*, 49-52.

Kamiyama, T. (2007a). Acquisition of French vowels by Japanese-speaking learners: close and close-mid rounded vowels. Phonetics and Phonology in Third Language Acquisition (réunion satellite du XVIème Congrès International des Sciences Phonétiques), Université de Fribourg (en Brisgau), Allemagne, 3-avr août 2007.

Kamiyama, T. (2007b). Les occlusives sourdes et sonores en japonais et en français sont-elles phonétiquement similaires ? Communication orale présentée aux XXIèmes Journées de Linguistique d'Asie Orientale, EHESS, Paris, 28-30 juin 2007.

Kamiyama, T. (2008). Acquisition of the production of 'new' and 'similar' vowels: the case of /u/ and /y/ in French by Japanese-speaking learners. *The Journal of the Acoustical Society of America* 123(5) : 3741.

Kamiyama, T. (2009). Acoustic similarities between front rounded and back unrounded vowels as evidenced by French /ø/ and /u/ produced by Japanese-speaking learners. *The Journal of the Acoustical Society of America* 125(4) : 2756.

Kamiyama, T. (à paraître). Hi-senmon katei no sentaku jugyou ni okeru onsei imputto-busoku o oginau tameno intânetto fukushuu kyouzai [Matériels pédagogiques de révision pour compenser le manque d'input sonore dans un cours optionnel de japonais]. *Furansu Nihongo Kyouiku* [Enseignement du Japonais en France] 4 Association des enseignants de japonais en France, 9 pages.

Kamiyama, T., A. Gaillard-Corvaglia (2006a). Le /b/ fort en salentin (sous-région des Pouilles, Italie du Sud) : une étude acoustique et perceptive. *Actes des 8èmes Rencontres Jeunes Chercheurs de l'École Doctorale 268 'Langage et langue', Université Paris III*, 39-42.

Kamiyama, T., A. Gaillard-Corvaglia (2006b). Le occlusive bilabiali in salentino (Puglia): uno studio acustico e percettivo. *Actes du second Congrès National de l'AISV* (Associazione Italiana di Scienze della Voce), Salerne, 683-694.

Kamiyama, T., M. Sakamoto (en préparation). Intonation of yes-no questions in French: is it really enough to raise pitch at the end?

Kamiyama, T., Y. Yamamoto (2007). Visual representation of prosody for tactful communication skills: the case of request in Japanese as a Foreign Language taught to French university students. *Proceedings of the Phonetics Teaching and Learning Conference 2007*, University College London, Paper ID 37 (4 pages).

Kaneman-Pougatch, M., E. Pedoya-Guimbretière (1991). *Plaisir des sons : enseignement des sons du français*. Paris, Hatier/Didier.

Kataoka, R. (2009). A production study on phonologization of /u/-fronting in alveolar context. *The Journal of the Acoustical Society of America* 125(4) : 2570.

Kawaguchi, J., R. Miyashita, A. Shioda (2005). *Premier montage du français*. Tokyo, Surugadai Shuppansha.

Kawakami, S. (1956). Buntou no intonêshon [Intonation du début de phrase]. *Kokugogaku* [Philologie japonaise] 25 : 21-30.

Kawakami, S. (1973a). *Nihongo akusento-hou* [Méthode de l'accent en japonais]. Tokyo, Gaku shobô

shuppan.

Kawakami, S. (1973b). "Nihongo onsuuritsu no saikentou / Reexamination of the rhythm of Japanese verse," in *Ônishi hakushi kiju kinen onseigaku sekai ronbunshuu / Worldwide anthology of papers in phonetics in honor of Dr. Onishi's seventy-seventh birthday*. Tokyo, Phonetic Society of Japan, 665-671.

Keating, P. A., M. K. Huffman (1984). Vowel variation in Japanese. *Phonetica* 41 : 191-207.

Kindaichi, H. (1972). "Gagyô bion-ron [Essai sur les nasales de la ligne ga]" (1942), reproduit dans H. Kindaichi, *Nihongo on'in no kenkyuu* [Recherches sur la phonologie du japonais]. Tokyo, Tôkyôdô Shuppan, 168-197.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *The Journal of the Acoustical Society of America* 67(3) : 971-95.

Kong, E.-J., M. Beckman (2006). Fine-grained phonetics and developmental universals for glottal features. Travaux présentés à la Dixième Conférence de Phonologie de Laboratoire, Paris, 29 juin – 1er juillet.

Kôri, S. (1997). "Nihongo no intonêshon: kata to kinou [Intonation du japonais: ses patrons et fonctions]" in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 169-202.

Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford, Pergamon.

Kubozono, H. (1985). Speech errors and syllable structure. *Linguistics and Philology* 6 : 220-243.

Kubozono, H. (1993). *The Organization of Japanese Prosody*. Tokyo, Kuroshio.

Kubozono, H. (1996). Syllable and Accent in Japanese: Evidence from Loanword Accentuation. Onseigakkai kaihou / *Journal of the Phonetic Society of Japan* 211 : 71-82.

Kubozono, H. (1998). "*On'inron* [Phonologie]," in Onsei [Sons de la parole]. Tokyo, Iwanami shoten, 53-91.

Kubozono, H. (1999). *Nihongo no onsei* [Système des sons du japonais]. Iwanami shoten, Tokyo.

Kubozono, H. (2006). *Akusento no housoku* [Règles de l'accent]. Tokyo, Iwanami shoten.

Kubozono, H., S. Ôta (1998). *On'in kouzou to akusento* [Structure phonologique et accent]. Tokyo, Kenkyûsha.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science USA* 97 : 11850-7.

Kuhl, P. K., P. Iverson (1995). "Linguistic experience and the 'Perceptual Magnet Effect'," in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 121-154.

Kuhl, P. K., K. A. Williams, F. Lacerda, K. N. Stevens, B. Lindblom (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255 : 606 – 608.

Kurematsu, A. (1997). "Inritsu jouhou no jouhou kagaku eno tekiyou [Application des informations prosodiques aux sciences de l'information]," in M. Sugitô, T. Kunihiro, H. Hirose, M. Kôno, *Nihongo onsei 2 : akusento, intonêshon, rizumu to pôzu* [Phonétique du japonais 2 : accent, intonation, rythme et pause]. Tokyo, Sanseidô, 303-318.

Kuroda, R. (2008). *Gogaku wa yarinaoseru!* [On peut réapprendre les langues étrangères]. Tokyo, Kadokawa.

Labrune, L. (2000). Transcrire le japonais. *Daruma Revue d'Études Japonaises* 6/7 (printemps 2000), Philippe Picquier : 340-356.

Labrune, L. (2006). *La phonologie du japonais*. Leuven ; Paris, Peeters.

Ladefoged, P. (1993). *A Course in Phonetics* (troisième édition). Fort Worth, Texas, Harcourt Brace Jovanovich.

Lado, R. (1964). *Language Teaching: A Scientific Approach*. New York, San Francisco, Toronto, London, McGraw-Hill. Inc.

Lambacher, S., W. Martens, G. Molholt (2000). Comparison of identification of American English vowels by native speakers of Japanese and English. *Proceedings of the Phonetic Society of Japan* 2000, 213-217.

Laroy, C. (1995). *Pronunciation*. Oxford, Oxford University Press.

Lauret, B. (1998). *Aspect de Phonétique Expérimentale Contrastive : l'accent anglo-américain en français*. Thèse de doctorat de phonétique, Université de la Sorbonne Nouvelle.

Lauret, B. (2007). *Enseigner la prononciation du français : questions et outils*. Paris, Hachette.

LeBel, J.-G. (1990). *Traité de correction phonétique ponctuelle*. Québec, Centre international de Recherche en aménagement linguistique.

Lee, I.-S., S.-O. Lee, W. Chai (2004). *Kankokugo gaisetsu* [Aperçu de la langue coréenne], traduit en japonais par H. Umeda, T. Maeda. Tokyo, Taishûkan.

Léon, M., P. Léon (1997). *La prononciation du français*. Paris, Nathan Université.

Léon, P. R. (2000). *Phonétisme et prononciations du français* (4ème édition). Paris, Nathan.

Levitt, A., Q. Wang (1991). Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech* 34 : 235-249.

Levy, E. S. (2006). Consonantal context and experience effects on American L2-learners' perception of French vowels. *The Journal of the Acoustical Society of America* 119(5) : 3424.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *The Journal of the Acoustical Society of America* 125(2) : 1138-1152.

Levy, E. S., W. Strange (2002). Effects of consonantal context on perception of French rounded vowels by American English adults with and without French language experience. *The Journal of the Acoustical Society of America* 111(5) : 2361-2362.

Levy, E. S., W. Strange (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36 : 141–157.

Liénard, J.-S. (1977). *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole*. Paris, Masson.

Lisker, L., A. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20 : 384–422.

Maddieson, I., K. Precoda (1990). Updating UPSID. *UCLA Working Papers in Phonetics* 74 : 104–111.

Maeda, S. (1976). *A characterization of American English intonation*. Ph.D. dissertation, Massachusetts Institute of Technology.

Maeda, S. (1982). A digital simulation method of the vocal tract system. *Speech Communication* 1(3-4) : 199-229.

Martin, Ph. (1982). Phonetic Realisations of Prosodic Contours in French. *Speech Communication* 1 : 283-294.

Martin, Ph. (1996). WinPitch : un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows. *Actes des XXIV Journées d'Étude sur la Parole*, Avignon, 224-227.

Martin, Ph. (2005). WinPitch LTL, un logiciel multimédia d'enseignement de la prosodie. *Apprentissage des langues et systèmes d'information et de communication (ALSIC)* 8(2) : spécial Atala : 95-108.

Martin, P. (2002). Le système vocalique du français du Québec. De l'acoustique à la phonologie. *La linguistique* 38(2) : 71-88.

Martinet, A. (1958). C'est jeuli le Mareuc ! *Romance Philology* 11 : 345-355. (Réédité dans A. Martinet, Le Français sans Fard. Paris, Presses Universitaires de France, 1969.)

Martinet, A. (1969). *Le Français sans Fard*. Paris, Presses Universitaires de France.

Matthews, S., V. Yip (1994). *Cantonese : A Comprehensive Grammar*. London, Routledge.

Mehler, J., J.-Y. Dommergues, U. Frauenfelder, J. Segui (1981). The Syllable's Role in Speech Segmentation. *Journal of Verbal Learning and Verbal Behavior* 20 : 298-305.

Mokhtari, P., T. Kitamura, H. Takemoto, K. Honda (2007). Principal components of vocal-tract area functions and inversion of vowels by linear regression of cepstrum coefficients. *Journal of Phonetics* 35(1) : 20-39.

Mokhtari, P., K. Tanaka (2000). A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* 64 (special issue) : 57-66.

Montagu, J. (2004). Les sons sous-jacents aux voyelles nasales en français parisien : indices perceptifs des changements. *Actes des XXIVèmes Journées d'Étude sur la Parole*, 385-388.

Munro, M. J. (1995). Nonsegmental Factors in Foreign Accent. *Studies in Second Language Acquisition* 17 : 17-34.

Nakai, T. (2002). *CD bukku : korede tsuujiru ! Hana-baachan no furansugo* [CD-livre Français avec la grand-mère Hana pour se faire comprendre]. Tokyo, Nihon Hôsô Shuppan Kyôkai.

Nakamichi, M. (1980). Nihongo no rizumu no tan'i ni tsuite [Sur l'unité de rythme en japonais]. *Onsei, gengo no kenkyuu* [Recherches en sons de la parole et langage], Laboratoire de phonétique de l'Université des langues étrangères de Tokyo 1 : 15-22.

Neufeld, G. G. (1979). Vers une théorie de la capacité d'apprentissage linguistique. *Encrages*, numéro spécial de linguistique appliquée, Université de Paris VIII : 9-15.

Nishi, K., W. Strange, R. Akahane-Yamada, R. Kubo, S. A. Trent-Brown (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America* 124(1) : 576-588.

Ohala, J. J. (1983). "The origin of sound patterns in vocal tract constraints," in P. F. MacNeilage, *The production of speech*. New York, Springer-Verlag, 189-216.

Ooigawa, T. (2008). Nihon-jin, Amerika-jin, Furansu-jin choushusha ni yoru eigo, futsugo, seigo ni okeru boinkan no ryuuon no benbetsu : yobi kenkyuu [Discrimination de liquides intervocaliques en anglais, en français, et en espagnol chez des auditeurs japonais, américains et français : étude préliminaire]. *Sophia University Working Papers in Phonetics 2008,* Sophia University Phonetics Laboratory, Tokyo : 2-20.

O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. *Journal of Phonetics* 9 : 385-406.

Ostiguy, L., R. Sarrasin, G. Irons (1996). *Introduction à la phonétique comparée : les sons : le français et l'anglais nord-américains*. Sainte-Foy, Les Presses de l'Université Laval.

Ôtake, T. G. Hatano, A. Cutler, J. Mehler (1993). Mora or Syllable? Speech Segmentation in Japanese. *Journal of Memory and Language* 32(2) : 258-278.

Palmer, H. E. (1939). *A grammar of spoken English on a strictly phonetic basis*. Cambridge, England, W. Heffer and sons, ltd.

Passy, P. (1886a). *Le français parlé. Morceaux choisis à l'usage des étrangers avec la prononciation figurée*. Heilbronn, Henninger.

Passy, P. (1886b). *Les éléments d'anglais parlé*. Paris, Firmin-Didot.

Patkowski, M. (1990). Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics* 11 : 73-89.

Payne, M. E. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2) : 153-181.

Peterson, G. E., H. L. Barney (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America* 24(2) : 175-184.

Pickett, E. R., S. E. Blumstein, M. W. Burton (1999). Effects of Speaking Rate on the Singleton/Geminate Consonant Contrast in Italian. *Phonetica* 56(3-4) : 135-157.

Pierrehumbert, J. B., M. E. Beckman (1988). *Japanese Tone Structure*. Cambridge, Massachusetts ; London, England, The MIT Press.

Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, University of Michigan Press.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague* 4 : 79-96.

Port, R. F., J. Dalby, M. O'Dell (1987). Evidence for mora timing in Japanese. *The Jounal of the acoustical Society of America* 81 : 1574-1585.

Poser, W. J. (1990). Evidence for Foot Structure in Japanese. *Language* 66(1) : 78-105.

Renard, R. (1971). *Introduction à la méthode verbo-tonale de correction phonétique*. Paris, Didier.

Renard, R. (2002). "Une phonétique immergée," in R. Renard, *Apprentissage d'une langue étrangère/seconde 2 : la phonétique verbo-tonale*. Bruxelles, De Boeck Université, 11-24.

Riney, T. J., N. Takagi, K. Ôta, Y. Uchida (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35 : 439–443.

Rochet, B. L. (1995). "Perception and production of second-language speech sounds by adults," in W. Strange, *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press, 379-410.

Sadanobu, T. (2004). A natural history of Japanese pressed voice. *Journal of the Phonetic Society of Japan* 8(1) : 29-44.

Sadanobu, T. (2005). *Sasayaku koibito, rikimu repôtâ: kuchi no naka no bunka* [Murmures des amoureux, voix pressée des reporters : culture dans la bouche]. Tokyo, Iwanami.

Sagisaka, Y. (1999). Nihongo on'in no jikanchou seigyo to chikaku [Contrôle de durée et perception des phonèmes en japonais]. *Gengo* [Langage] 28(9) : 51-56, 1999.

Saitô, Y. (1997). *Nihongo onseigaku nyuumon* [Introduction à la phonétique du japonais]. Tokyo, Sanseidô.

Sawashima, M., S. Miyazaki (1973). Glottal Opening for Japanese Voiceless Consonants. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo* 7 : 1-10.

Sawashima, M., S. Niimi (1974). Laryngeal Conditions in Articulations of Japanese Voiceless Consonants. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo* 8 : 13-18.

Schouten, M. E. H., A. J. van Hessen (1992). Modeling phoneme perception. I: Categorical perception. Journal of the Acoustical Society of America 92(4) : 1841-1855.

Scovel, T. (2000). A Critical Review of The Critical Period Research. *Annual Review of Applied Linguistics* 20 : 213-223.

Shibata, T. (1994). "Gairaigo ni okeru akusento-kaku no ichi [Position du noyau d'accent dans les mots d'emprunt]," in K. Satô, *Gendaigo, hougen no kenkyuu* [Recherches sur le japonais contemporain et des dialectes]. Tokyo, Meiji shoin, 388-418.

Shibatani, M. (1990). *The Languages of Japan*. Cambridge, Cambridge University Press.

Shimizu, K. (1996). *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Tokyo, Seibidô Publishing.

Shinohara, S. (1997). *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Thèse de doctorat, Université de la Sorbonne Nouvelle.

Shinohara, S. (2000). Default Accentuation and Foot Structure in Japanese: Evidence from Japanese Adaptations of French Words. *Journal of East Asian Linguistics* 9 : 55-96.

Shôchi, T., V. Aubergé, A. Rilliard (2006). Comment les attitudes prosodiques sont parfois de « faux-amis » : les affects sociaux du japonais vs. français. *Actes des XXVIèmes Journées d'études sur la parole*, Dinard, 259-262.

Shôchi, T., V. Aubergé, A. Rilliard (à paraître). Nihongo bogo-washa wa hatsuwa no ichibu o kikudakede taido o chikaku dekirunoka? – Gating paradaimu niyoru jikken o tôshite [Peut-on percevoir les attitudes japonaises avant la fin de phrase ? – test perceptif selon le paradigme de gating]. *Furansu Nihongo Kyouiku* [Enseignement du Japonais en France] 4 Association des enseignants de japonais en France, 9 pages.

Sidaras, S. K., J. E. Alexander, L. C. Nygaard (2008). Talker-specific and accent-general perceptual adaptation to Korean-accented English. *The Journal of the Acoustical Society of America* 123(5) : 3887.

Sock, R., V. Hecker, M.-A. Cathiard (1999). The perceptual effects of anticipatory labial activity in French. The proceedings of the XVIth International Congress of Phonetic Sciences, Vol. 5 2057-2060.

Straka, G. (1965). Album phonétique. Sainte-Foy, Les Presses de l'Université Laval.

Strange, W., R. Akahane-Yamada, R. Kubo, S. A. Trent, K. Nishi, J. J. Jenkins (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics* 26 : 311-344.

Strange, W, E. S. Levy, R. Lehnholf (2004). Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. *The Journal of the Acoustical Society of America* 115(5) : 2606.

Sugitô, M. (2001). "Bunpou to nihongo no akusento oyobi intonêshon: Toukyou to Ôsaka no baai" [Grammaire, et accent et intonation du japonais : les cas de Tokyo et d'Osaka]", in Onsei bunpou kenkyuukai / Spoken Language Working Group, *Bunpou to onsei III / Speech and grammar III.* Tokyo, Kuroshio, 197-210.

Sugitô, M. (1997). "Nihongo no akusento to intonêshon: toukyou hougen no 'hana' to 'hana' no soui [Accent et intonation en japonais : différence entre 'hana' (fleur) et 'hana' (nez) en dialecte de Tokyo]", in M. Sugitô, *Hana to hana* ['Hana' (fleur) et 'hana' (nez)]. Osaka, Izumi shoin : 3-22.

Sugitô, M. (1995). *Ôsaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [Dictionnaire CD-ROM sonore d'accent des japonais d'Osaka et de Tokyo : commentaires]. Tokyo, Maruzen.

Sweet, H. (1900). *Practical study of languages*. New York, Henry Holt and Company.

Takebayashi, S. (1996). *Eigo onseigaku* [Phonétique de l'anglais]. Tokyo, Kenkyûsha.

Tanaka, S., H. Kubozono (1999). *Nihongo no hatsuon kyoushitsu* [Classe de prononciation du japonais]. Tokyo, Kuroshio.

Thorsen, N. G. (1980). A study of the perception of sentence intonation – evidence from Danish. *The Journal of the Acoustical Society of America* 67 : 1014-1030.

Tremblay, M.-C., T. Kamiyama (2009). Perception of L2 production by L1 speakers of different dialectal backgrounds: the case of Japanese-speaking learners' /u/ perceived by French and Quebec native speakers. *The Journal of the Acoustical Society of America* 125(4) : 2776.

Troubetzkoy, N. S. (1938/2005). *Principes de phonologie* (traduction de Jean Cantineau, revue et corrigée par Luis Jorge Prieto). Paris, Klincksiek.

Tsujimura, N. (1996). *An Introduction to Japanese Linguistics*. Oxford, Blackwell.

Uemura, Y. (Kokuritsu Kokugo Kenkyûjo [Institut national de la langue japonaise]) (1990). *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]. Tokyo, Shûei shuppan.

Uwano, Z. (1993). "On no kouzou [Structure des sons de la parole]," in K. Kazama, *Gengogaku* [Linguistique]. Tokyo, Tokyo daigaku shuppankai [Presses de l'Université de Tokyo], 193-249.

Uwano, Z. (1989). "Nihongo no akusento [Accent en japonais]," in *Kouza nihongo to nihongo-kyouiku 2 : Nihongo no onsei, on'in I* [Série : Langue japonaise et enseignement du japonais 2 : Phonétique et phonologie du japonais, tome I]. Tokyo, Meiji Shoin : 178-205.

Vaissière, J. (1983). "Language Independent Prosodic Features," in A. Cutler, R. Ladd, *Prosody: Models and Measurements*. Berlin, Springer Verlag, 53-65.

Vaissière, J. (1991). "Rhythm, accentuation and final lengthening in French," J. Sundberg, L. Nord,

R. Carlson, *Music, Language, Speech and Brain*. London, Macmillan Press, 108-120.

Vaissière, J. (1997). Langues, prosodies et syntaxe. *Revue Traitement Automatique des Langues, ATALA* 38(1) : 53-82.

Vaissière, J. (2002). "Cross-Linguistic Prosodic Transcription: French versus English," in N. B. Volskaya, N. D. Svetozarova, P. A. Skrelin, *Problems and Methods of experimental Phonetics: in honour of the 70th anniversary of Prof. L. V. Bondarko*. St.-Petersburg, St.-Petersburg State University, 147-164.

Vaissière, J. (2006a). *La phonétique*. Paris, Presses Universitaires de France.

Vaissière, J. (2006b). Using an articulatory model as an integrated tool for a better understanding of the combined articulatory, acoustic and perceptual aspects of speech. Communication orale présentée au INAE Workshop on Image and Speech Processing, Chennai, Inde.

Vaissière, J. (2007). "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages," in M. J. Solé, P. S. Beddor, and M. Ohala, *Experimental Approaches to Phonology*. Oxford University Press, Oxford, 54-71.

Vaissière, J. (sous presse). Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal* 2.

Vance, T. J. (2008). *The sounds of Japanese*. Cambridge, Cambridge University Press.

Vaxelaire, B., V. Ferbach-Hecker, R. Sock (2002). La perception auditive de gestes vocaliques anticipatoires. *Les Actes des XXIVèmes Journées d'étude sur la Parole*, 49-52.

Verret, M. (1975). *Le temps des études*. Paris, Champion Honore.

Viala, A., H. Ôta, M.-N. Vacheron, P. Martin, F. Caroli (2004). Les Japonais en voyage pathologique à Paris : un modèle original de prise en charge transculturelle. *Nervure* 12(5) : 31-34.

Weinreich, U. (1953/1968). *Languages in Contact*. The Hague, Mouton.

Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication* 49 : 28–48.

Wells, J. C. (1962). *A Study of the Formants of the Pure Vowels of British English*. MA dissertation, University College London.
<http://www.phon.ucl.ac.uk/home/wells/formants/index.htm>

Wells, J. C. (1982). *Accents of English*. Cambridge, Cambridge University Press.

Wells, J. C. (2000). *Longman Pronunciation Dictionary* (Second edition). Harlow, Pearson Education Limited.

Wells, J. C., J. House (1995). *The sounds of the IPA* (CD audio). London, University College London.

Wenk, B. J., F. Wioland (1982). Is French really syllable-timed? *Journal of Phonetics* 10 : 193-216.

Wioland, F. (1985). *Les structures syllabiques du français : fréquence et distribution des phonèmes consonantiques ; contraintes idiomatiques dans les séquences consonantiques*. Genève, Slatkine / Paris, Champion.

Wioland, F. (1991). *Prononcer les mots du français : des sons et des rythmes*. Paris, Hachette.

Wioland, F. (2001). Que faire de la graphie « e » dans le cadre de la didactique de la

prononciation ? *Le français dans le monde* 318 : 31-33, 2001.

Wioland, F. (2005). *La vie sociale des sons du français*. Paris, l'Harmattan.

Wioland, F. (2007). "Préface", in B. Lauret, *Enseigner la prononciation du français : questions et outils*. Hachette, Paris.

Woehrling, C., P. Boula de Mareüil (2006). Identification d'accents régionaux en français : perception et catégorisation. *Bulletin PFC (Phonologie du Français Contemporain : Usages, Variétés et Structure)* numéro 6 (Prosodie du français contemporain) : 89-102.

Wrembel, M. (2007). "Still sounds like a rainbow" - a proposal for a coloured vowel chart. *Proceedings of Phonetics Teaching and Learning Conference 2007*, Paper ID 14 (4 pages).

Yamasaki, H., P. A. Hallé (1999). How do native speakers of Japanese discriminate and categorize French /r/ and /l/? *Proceedings of the XIVth International Congress of Phonetic Sciences*, 909-912.

Yoshikawa, M. (2001). *Honkon etsugo: hatsuon / An introduction to Hong Kong Cantonese: Pronunciation*. Tokyo, Hakuteisha.

Yuan, J., C. Shih, G. P. Kochanski (2002). Comparison of declarative and interrogative intonation in Chinese. *Proceedings of the International Congress: Speech Prosody 2002*, 711-714.

Yukawa, Y. (1999). *Gengogaku* [Linguistique]. Tokyo, Hituzi syobô.

**Apprentissage phonétique des voyelles du français langue étrangère chez des apprenants japonophones**

**Résumé**

Ce travail s'intéresse à l'application de la phonétique expérimentale (acoustique et perceptive) à la didactique de la prononciation des langues étrangères. Le propos est illustré par les difficultés d'apprentissage par des japonophones des voyelles du français ; les expériences portent spécifiquement sur les voyelles /u y ø/. Le but est d'élucider les difficultés que présentent ces phones selon que leur statut phonémique et leur réalisation phonétique diffèrent ou non entre la langue maternelle et la langue apprise.

Le /u/ français diffère phonétiquement de son équivalent phonémique, le /u/ japonais. L'étude confirme que le /u/ français, phonémiquement « similaire » au /u/ japonais, est plus difficile que la voyelle « nouvelle » /y/, qui n'a pas d'équivalent ni phonémique ni phonétique en japonais. La production du /ø/, qui est « nouveau » phonémiquement mais proche du /u/ japonais au plan acoustique, semble présenter encore moins de difficulté.

La thèse apporte également une réflexion sur la didactique de la prononciation. L'analyse de manuels généralistes de français publiés au Japon suggère que les apprenants et les enseignants sont rarement conscients de la différence de difficultés des /u y ø/. Quelques méthodes d'enseignement de la prononciation – certaines traditionnelles, d'autres innovantes – sont proposées, dans l'idée de favoriser la conscientisation de ces difficultés.

Le but de cette thèse est une contribution à l'éclaircissement des processus d'apprentissage de la prononciation des langues étrangères, et à l'amélioration de son apprentissage et de son enseignement.

**Mots clés : apprentissage phonétique, français langue étrangère, voyelles, perception et production, niveaux phonétique et phonémique, apprenants japonophones, didactique de la prononciation.**

**Pronunciation of French vowels by Japanese speakers learning French as a foreign language**

**Summary**

This dissertation deals with the application of experimental (acoustic and perceptual) phonetics to the teaching of pronunciation. The issue is illustrated by Japanese speakers' difficulties in learning French vowels in general. Experiments were specifically conducted on the vowels /u y ø/. The objective is to elucidate the case of individual phones depending on whether or not their phonemic status and their phonetic realisation differ in the two languages under study.

French /u/ differs phonetically from its phonemic counterpart, Japanese /u/. The present study confirms that French /u/, which is phonemically "similar" (to Japanese /u/), turns out to be more difficult than the "new" vowel /y/, which has no phonemic or phonetic counterpart in Japanese. The production of /ø/, which is phonemically "new" but acoustically close to Japanese /u/, seems to present still less difficulty.

The dissertation also brings a reflection on the teaching of pronunciation. The analysis of general French textbooks published in Japan suggests that learners and teachers are seldom aware of the difference in the difficulties caused by the three vowels /u y ø/. Also, some methods of pronunciation teaching (some traditional, others new) are presented in terms of how they foster learners' awareness of these difficulties.

The goal of this dissertation is to help to shed light on the learning processes of the pronunciation of foreign languages, and to improve its learning and teaching.

**Keywords: pronunciation learning, French as a foreign language, vowels, perception and production, phonetic and phonemic levels, Japanese-speaking learners, pronunciation teaching.**

Discipline : Phonétique
Laboratoire de Phonétique et Phonologie (UMR 7018)
Institut de Linguistique et Phonétique Générales et Appliquées
19, rue des Bernardins 75005