Aile

Acquisition et interaction en langue étrangère

Aile... Lia 2 | 2009
Phonétique, bilinguisme et acquisition

Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners

Takeki Kamiyama and Jacqueline Vaissière

OpenEdition Journals

Electronic version
URL: http://journals.openedition.org/aile/4533
ISSN: 1778-7432

Publisher
Association Encrages

Printed version
Date of publication: 1 December 2009
Number of pages: 9-41
ISSN: 1243-969X

Electronic reference
Takeki Kamiyama and Jacqueline Vaissière, « Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners », *Acquisition et interaction en langue étrangère* [Online], Aile... Lia 2 | 2009, Online since 21 August 2012, connection on 19 April 2019. URL : http://journals.openedition.org/aile/4533

PERCEPTION AND PRODUCTION OF FRENCH CLOSE AND CLOSE-MID ROUNDED VOWELS BY JAPANESE-SPEAKING LEARNERS

Takeki KAMIYAMA et Jacqueline VAISSIÈRE

Abstract

A series of three experiments were conducted to examine the behaviour of Japanese-speaking learners of French (JSL) concerning the perception and production of French /u/, /y/ and /ø/, compared with English-speaking learners of French, in order to examine how L2 (second language) vowels which do not have phonemic and/or phonetic equivalents in L1 (native language) are acquired by learners.

The results indicate that JSL tend to produce French /u/ with a high F2 (> 1000 Hz), which is heard as /ø/ by native French (NF) listeners. They suggest that French /u/ is considered by JSL as phonemically similar to Japanese /u/ (as a high back vowel) and produced as such, while the phonetic realisation of French /u/ is new and different; and that /y/ is considered as both phonemically and phonetically new, and /ø/ as phonemically new but phonetically (acoustically) similar. Phonemically similar but phonetically new vowels seem to be most difficult to learn to produce accurately for foreign/second language learners.

Key words: perception, production, focal vowels, French, Japanese-speaking learners

1. Introduction

Learning the phonological system and its phonetic realisations in a second language is a challenge that learners face at some points of their learning process. As we all know empirically, and as many researchers have shown experimentally, some phonemes are more difficult to acquire than others. Unlike the acquisition of other components of language, such as word order of some constructions (English and German negation, for example, as summarised in Ellis 1994), the recent influential models of speech acquisition (Flege's Speech Learning Model, Best's Perceptual Assimilation Model, Kuhl's Magnet theory: *cf.* Flege 1995, Best 1995, Kuhl 2000) all grant an important role to prior linguistic experience of learners, rather than to a universal tendency observed with learners of different first languages.

Vowels, in particular, are known to exhibit more variability both synchronically and diachronically than consonants: in French, for instance, the consonant inventory is almost the same in practically all varieties, whereas vowel inventory largely depends on the dialect concerned (Vaissière 2006, among others). In addition, the continuous nature of the articulation of vowels makes it much more difficult for teachers to define and explain, and thus for learners to learn to reproduce the articulatory targets precisely, as compared to consonants. As for perception, vowels have been shown to present a less categorical perception than consonants (Schouten & van Hessen 1992 for a comparison of stop consonants and vowels, for example). It is also known that very slight deviations of vowel quality from the expected one may convey particular attitudes and emotions (Fónagy 1983, among others), since listeners are very sensitive to formant frequencies and fundamental frequency contours. The acquisition of L2 (second language) vowels is especially important for these reasons, and that is why it is particularly interesting to study cross-language perception and production of vowels.

A number of experimental studies have investigated the behaviour of native speakers of Japanese learning English vowels. Lambacher *et al.* (2000) showed that Japanese speakers had difficulty identifying back vowels /ʌ/ and /ɔ/, and open vowels /æ/ and /ɑ/ in American English. Strange *et al.* (1998) concluded that the «point vowels» (/i/ /ɑ/ /u/) of American English were considered as most similar to the corresponding Japanese vowels (/i/ /a/ /u/) in a relatively stable manner, but less so as long as other vowels are concerned.

The vowel systems of Japanese and English share some characteristics. First, neither of them has phonemically front rounded vowels as in French, German, Cantonese, and many other languages. The high front rounded /y/ in

French, which originated from /u/ in Latin, is characterised by close F2 and F3 around 1900 Hz in male voice (CALLIOPE 1989, Vaissière 2007, among others). A similar vowel timber may be produced as an allophone in Japanese (during the sequence /ju/ as a brief transition) and English (as a fronted /u/ in a coronal context as in «tune» /tun/), but does not correspond to any vowel phoneme in these languages. Close-mid /ø/ is characterised by evenly distributed formants with no grouping, and with F2 around 1500 Hz (CALLIOPE 1989, Vaissière 2007, among others). Note also that the open-mid /œ/ is another non-focal, "acoustically central" vowel similar to /ø/ (Vaissière 2006), with slightly higher formants, at least in citation form. These two acoustically very similar vowels are in partially limited distribution (no /œ/ is found in word-final open syllables, and the number of minimal pairs in closed syllables is quite limited: *e.g. jeûne* /ʒøn/ and *jeune* /ʒœn/).

French /u/ (transcribed [u]) is often referred to as a "fine" example of a high back rounded vowel, and indeed, shares the same acoustic characteristics as cardinal vowel number 8 pronounced by Daniel Jones and Peter Ladefoged, *i.e.* grouping of the first two formants under 1 kHz (Vaissière 2007). Japanese vowel /u/, which is phonemically high and back, usually does not correspond to phonetic realisations close to that of French /u/. It is often described phonetically as an unrounded vowel [ɯ]. F2 is typically located above 1 kHz (Sugito 1995, Mokhtari & Tanaka 2000). Figure 1 displays the formant values found by Sugito 1995. Note that F2 of /u/ ([ɯ]), which is slightly higher than 1000 Hz, is higher than that of /o/ ([o]), and that the F1 of /u/ is found between that of /i/ and that of /o/. Such values are fully explained by referring to articulation. The midsagittal profiles in Uemura (1990) show that the tongue is more fronted and the lips are less rounded[1] compared to a typical French [u], as illustrated in Bothorel *et al.* (1986). Note that a weaker degree of lip rounding (as observed in Japanese) contributes to raising the first formant (F1), and that both tongue fronting and lip spreading contribute to raising the first two formants (F2) in the velar region (see the nomograms of Fant 1960 for details). This articulatory configuration and the resulting acoustic properties are found in Chiba and Kajiyama (1955, first edition published in 1941), suggesting that it is not a recent phenomenon.

1. The X-ray movie «Cineradiograph of Japanese Pronunciation» produced by the Speech Laboratory of the National Institute for Japanese Language is available (in Japanese) on the website of the Institute: http://www6.kokken.go.jp/x-sen/

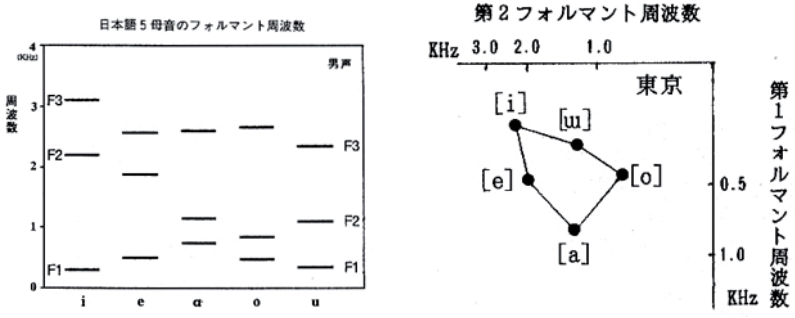


Figure 1

The first three formants (left: with phonemic transcription /u/ as the author put it) and the first two formants in two dimensions (right: the first formant – F1 – on the vertical axis, the second formant – F2 – on the horizontal axis; with phonetic transcription [ɯ] as the author put it) of the five vowels (male voice) of Tokyo Japanese (Sugito 1995).

English /u/, on the other hand, has undergone fronting in the course of the second half of the twentieth century, as exemplified in Hawkins and Midgley (2005) for British RP (Received Pronunciation), and in Hillenbrand *et al.* (1995) for American English. In Hawkins and Midgley's data, mean F2 of male speakers' /u/ ranges from 994 Hz (over 65 years old) to 1616 Hz (20-25 years old)[2]. Similar figures are found in Deterding (2006). Hillenbrand *et al.* compared their findings with the data of Peterson and Barney (1952), showing that the average values of the first two formants of /u/ pronounced in /hVd/ context were higher in their data both for male and female speakers.

The acoustic differences stated above among French /u/, Japanese /u/, and American English /u/ are illustrated in Figure 2 below. It can be seen that the first two formants of French /u/ are closely located together below 1000 Hz. In Japanese /u/, F2 is between 1000 and 1500 Hz, away both from F1 and F3. In American English, /u/ in the word "who" /hu/ is diphthongised, as is often the case even in isolation, with F2 lowering from above to below 1000 Hz and getting closer and closer to F1, as intensity goes down toward the end of the vowel.

2. The eleven monophthongs in RP were spoken in /hVd/ contexts by five men in each of the following age groups: 20–25, 35–40, 50–55, and 65–73 years in 2001. The test words in citation form were read four times in a randomised list.

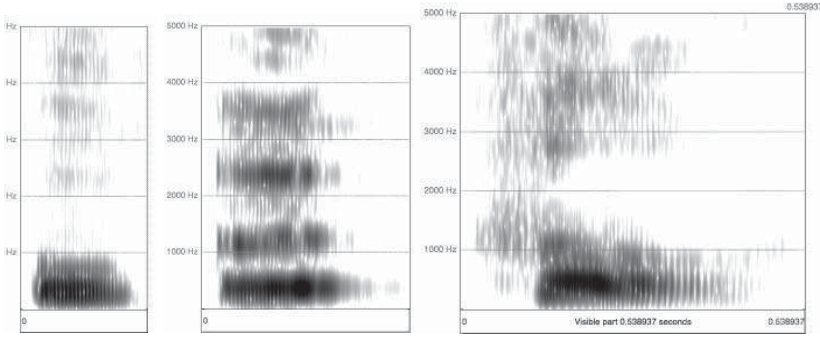


Figure 2

Wide-band spectrogram (on Praat; window length: 5 milliseconds) of French /u/ in isolation pronounced by a male French native speaker (left), of Japanese /u/ in isolation pronounced by a male native speaker of Tokyo Japanese (middle), of American English /u/ pronounced by a male speaker in the word "who" hu/ (right).

Indeed, the acoustic and perceptual similarity between Japanese /u/ [ɯ] (and /uR/ [ɯɯ]: /R/ represents the second half of a long vowel in Japanese) and American English (AE) /u/ was proven by Nishi *et al.* (2008). According to their acoustic data (four male speakers of each language pronounced nonsense words [hVba] three times each in citation and sentence forms), both /u/ [ɯ] and /uR/ [ɯɯ] in Japanese are acoustically (as far as F1 and F2 are concerned) closer to AE /uː/ than to any other AE vowel. Their perceptual assimilation experiment also showed that 12 native AE listeners assimilated both /u/ [ɯ] and /uR/ [ɯɯ] to AE /u/ in at least 89% of the cases.

How, then, do native speakers of Japanese and English acquire the front rounded series /y/, /ø/ (and /œ/), and the high back rounded /u/ in French as a second/foreign language?

Gottfried (1984) showed that American non-native speakers of French were significantly less accurate than native French in identifying and discriminating French vowels, especially front rounded ones. Strange *et al.*'s study (2004) on the perceptual assimilation patterns by American English speakers suggests that the front rounded vowels are perceived as more similar to back than front vowels. Levy and Strange (2008) conducted an AXB discrimination experiment using the vowels /u y œ i/ pronounced by three native speakers of

Parisian French in non-sense words /rabVp/ and /radVt/. The performances of two groups of 10 American English listeners (non French speakers – inexperienced – and those who had studied French extensively beginning after age 13 – experienced) were compared. The experienced group performed better than the inexperienced on /u-œ/, /i-y/, and /y-œ/ (mean errors: Exp. 5% vs. Inexp. 24%), while they did not differ for /u-y/ (Exp. 30% vs. Inexp. 24%). It may be argued that one of the factors that could explain the even poorer discrimination of the pair /u-y/ by the experienced group (learners of French) is orthography: grapheme <u> corresponds not to phoneme /u/ but to /y/ (/u/ corresponds to the sequence <ou>).

Regarding production, Rochet (1995) showed that French /y/ produced by 10 Canadian English speakers in monosyllables were judged by three native speakers of French to be [u] or [u]-like vowel, or a vowel between [y] and [u] 92% of the time (and as an [i]-like vowel or a vowel between [y] and [i] only 8% of the time). Lauret (1998) reported that native speakers of American English learning French tended to produce, in place of /y/, a sequence similar to American English /ju/, where F2 decreases gradually during the vowel. Flege (1987) investigated the production of more advanced American learners of French and showed that their /y/ was not far from native French speakers' target, but that their /u/ had a significantly higher F2 than that of native speakers. He then suggested that it might be more difficult to pronounce a "similar" L2 phone that has a phonemic equivalent in L1 but is phonetically different (/u/), than a "new" phone which does not have any equivalent in L1 (/y/).

How, then, do native speakers of Japanese learn to produce French /y/, a "new" phone, and /u/, a "similar" phone? Do they produce a native-like /y/ more easily than a native-like /u/? How do they cope with /ø/? A series of three experiments were conducted in order to investigate this issue.

2. Target French vowels /u/, /y/ and /ø/

Before investigating the production and perception of Japanese learners, native speakers of French were recorded and the data compared with that of previous studies (CALLIOPE 1989 in Figure 5, Gendrot & Adda-Decker 2004).

Four native speakers of French from the northern half of France (two male and two female: in the 25-35 age range at the time of the recording in July 2005) pronounced 10 French oral vowels in the carrier sentence "*Je dis* /V/ *comme dans …*" ('I say /V/ as in …'). Each vowel was presented in phonemic transcription with an example word written orthographically (*e.g.* "*loup"* /lu/

for /u/) on a computer screen, and the speaker was invited to say the sentence. Note that all speakers were familiar with phonemic transcription of French. The vowels were presented three times in a random order. The recording was carried out in a soundproof room through a headset microphone and the sound data were stored at 16 kHz/16 bits. It is indispensable to measure formants at several different time points within a vowel, firstly to check if native speakers really produce vowels with stable formants (as they are in French), and secondly to study the differences possibly observed between native and non-native speakers (See Lauret 1998 for American English speakers). The first four formants of each target vowel were measured in five different time zones (1. from the beginning up to the 20% point; 2. the second fifth; and so forth. See Figure 3) using Praat (Boersma & Weenink 2007), and a script written by Cédric Gendrot and modified for the present study by the first author. Unlikely values due to errors in automatic formant detection were discarded after checking the spectrogram and/or spectral slice (20 out of 2400 values, *i.e.* 0.8%; 12 cases for /u/ and 6 for /i/: note that /u/ is characterised by close F1/F2, /u/, close F3/F4). As expected for French, all vowel tokens had relatively stable formant values, showing that none of them were diphthongised. Table 1 indicates the mean formant values of each speaker.

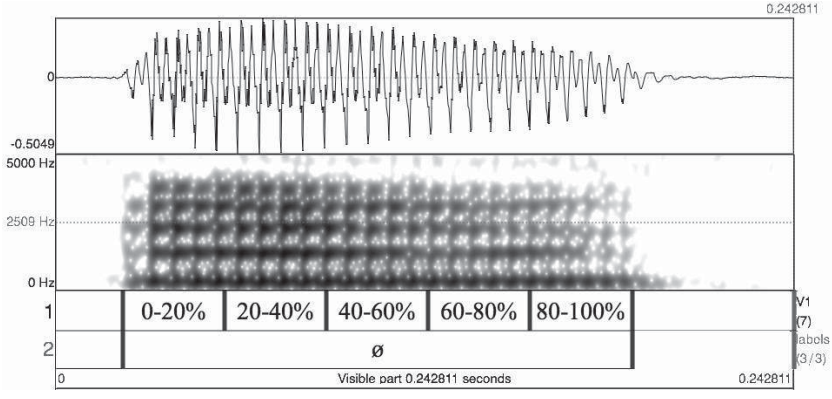


Figure 3

Illustration of the five time zones in which formant frequencies were measured. The vowel /ø/ pronounced by Speaker FR1 (male native French) in the carrier sentence "*Je dis* /ø/ *comme dans leu*" ('I say /ø/ as in *leu*'). Oscillogram and wide-band spectrogram (window length: 5 milliseconds) on Praat.

Note also that many acoustic studies on vowels, especially on English vowels, measure only the first two formants, but that F3 value is indispensable to represent vowel systems with front-rounded vs. front-unrounded opposition. For instance, some speakers may show similar F1 and F2 values for /i/ and /y/, with only F3 differentiating the two vowels. It is also known that F4 makes it possible to describe some cross-linguistic differences of similar phonemes: French /i/ is characterised by a grouping of F3 and F4, while English /i/ is characterised by a high F2 (higher than in French /i/, but F3 and F4 are generally not grouped: see Gendrot *et al.* 2008). In Japanese, the first two formants are largely sufficient to simply distinguish the five vowels, but it has been shown that the vocal tract during vowel production is modelled best with the first four formants (Mokhtari *et al*. 2007). In the present study, the first four formants are presented in tables and figures in order to better visualise the whole acoustic characteristics of vowels produced by native and non native speakers (e.g. F3 and F4 located away from the first two formants in French /u/ pronounced by native speakers vs. evenly distributed formants of the same vowel pronounced by JSL, as we shall see).

FR1 (M)	F1	F2	F3	F4	FR2 (M)	F1	F2	F3	F4	FR3 (F)	F1	F2	F3	F4	FR4 (F)	F1	F2	F3	F4
i	323	2194	3140	3829	i	312	2091	3160	3679	i	347	2519	3903	4439	i	343	2451	3536	3967
e	441	2048	2576	3479	e	332	2123	2691	3041	e	500	2488	3162	3985	e	558	2321	2966	3914
ɛ	512	1891	2556	3569	ɛ	505	1964	2478	2879	ɛ	686	2190	2972	3917	ɛ	696	2088	2667	3562
a	699	1261	2392	3434	a	654	1396	2444	2993	a	869	1238	3071	3977	a	845	1441	2612	3731
ɔ	495	951	2629	3401	ɔ	505	1053	2501	3336	ɔ	694	1051	2940	4078	ɔ	735	1108	2826	3827
o	372	688	2618	3496	o	330	715	2565	3317	o	481	776	2968	4024	o	515	912	2835	3878
u	289	631	2489	3503	u	304	637	2416	3514	u	302	798	2462	3435	u	385	744	2311	3457
y	284	1806	2100	3290	y	288	1749	2089	3128	y	306	2032	2471	3804	y	378	1989	2522	3733
ø	396	1276	2304	3271	ø	324	1372	2282	3144	ø	490	1570	2583	4012	ø	556	1431	2583	3954
œ	463	1366	2390	3386	œ	479	1461	2401	3251	œ	669	1733	2778	4149	œ	689	1551	2635	3987

Table 1

The first four formants (in Hertz) of the 10 oral vowels of French pronounced by four native speakers of French from the northern half of France (mean of 5 measuring points x 3 repetitions). Speakers FR1 and FR2 are male, FR3 and FR4 female speakers.

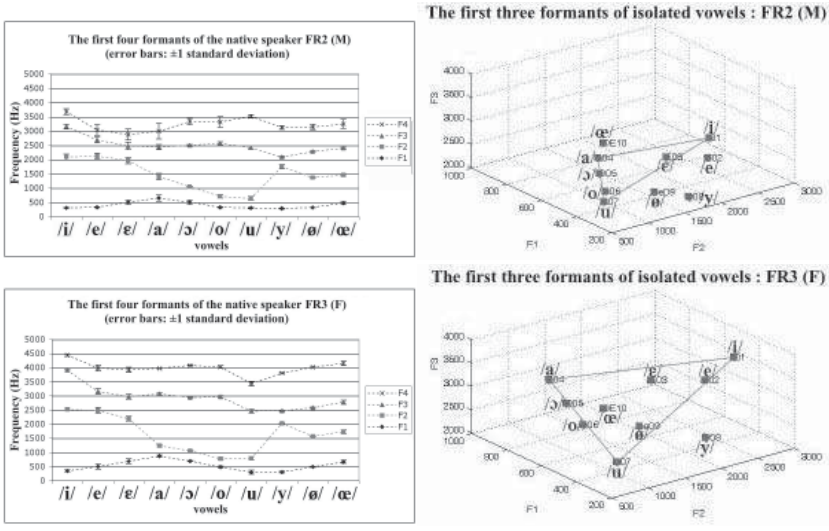


Figure 4

Graphic representation of the formant values of native French speakers FR2 (male: top) and FR3 (female: bottom) in Table 1. The first four formants (right), and the first three formants represented in a vowel triangle (right). Note the grouping of F3 and F4 for /i/ for both speakers. Note also that FR2's /i/ and /e/ have similar F1 and F2 values, but are distinguished by F3 and F4.

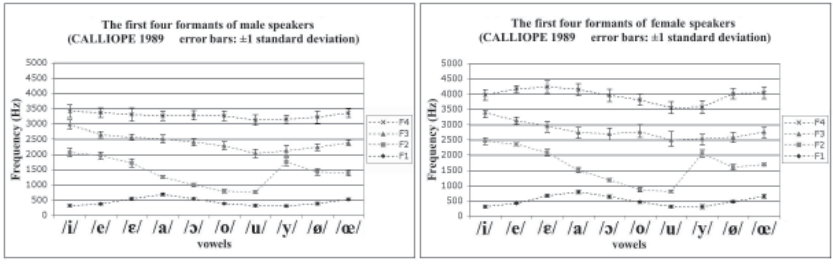


Figure 5

The first four formants (in Hertz) of the 10 oral vowels of French pronounced in [pV] and [pVʀ] by 10 male (left) and 9 female (right) native speakers of French. Two repetitions per subject (CALLIOPE 1989: 84).

Concerning vowels /u/, /y/, and /ø/, the following tendencies are found:

1) /u/, along with /o/, has very low first two formants, way below 1000 Hz both for male and female speakers, at least in citation form. These characteristics are also found in CALLIOPE. In Gendrot and Adda-Decker, F2 is above 1000 Hz; the difference probably lies in the fact that these data were collected from connected speech in radio broadcast corpora, in which vowels were generally pronounced in various (mainly anterior) consonantal and prosodic contexts. It follows that those tokens pronounced in a fronting context (*e.g.* between two coronals, such as in "*toute*" /tut/) contributed to making the mean value higher than that of /u/ in isolation. It is interesting to note that French listeners are mostly exposed to non focal realisations of /u/ in everyday speech.

2) /y/ shows close F2 and F3 between 1750 Hz and 2100 Hz for male speakers, approximately 2000 Hz and 2500 Hz for female speakers. CALLIOPE's data are quite similar to ours. In Gendrot and Adda-Decker, F2 is similar, but F3 is higher (2425 Hz for male, 2746 Hz for female speakers) than CALLIOPE's and ours. F3, which mainly depends on the front cavity in front vowels (particularly for /i/), becomes affiliated with the back cavity in hyperarticulated /y/ due to lip rounding and protrusion (for further details and references, Vaissière 2007): the high F3 in Gendrot and Adda-Decker's data might be due to weaker lip rounding or protrusion in many tokens in connected speech.

3) /ø/ is pronounced with no grouping of formants, with F2 around 1500 Hz (1276 and 1372 Hz for male, 1570 and 1431 for female speakers). F2 is lower than in CALLIOPE's (1417 Hz for male, 1605 Hz for female speakers) and Gendrot and Adda-Decker's data (1474 Hz and 1693 Hz for male and female speakers, respectively), but unlike /u/ and /y/, this vowel is characterised in all cases by the fact that F2 is located halfway between F1 and F3 (as is also the case with /œ/).

The values found and examined in the present section will be compared with those of Japanese learners' productions.

3. Preliminary study: Japanese learners' production

A first preliminary production experiment was carried out with a small group of Japanese speakers learning French (JSL) in a relatively well-controlled condition.

3.1. Method

The vowels /u/, /y/ and /ø/ were among the French vowels pronounced by the learners. They were embedded in the same carrier sentence as the previous experiment ("*Je dis* /V/ *comme dans* …"), but the sentences were presented visually. The target vowel was represented by its phonemic transcription, and the example word, which appeared at the end of the carrier sentence, orthographically. The speakers were invited to read the sentences one by one, twice in a row.

The speakers were three native speakers of Japanese (2 male, 1 female). They had studied French at the University of Tokyo for one semester and volunteered for the present study.

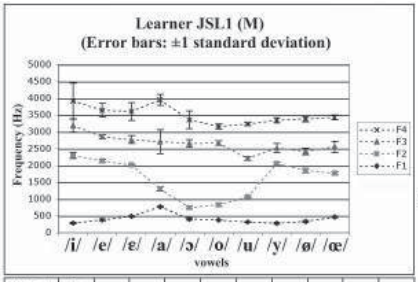
The oral data were recorded in a soundproof booth through a headset microphone and were stored at 16 kHz/16 bits. The first four formants were measured using Praat, as in the previous experiment. The mean of two repetitions for five time zones during the vowel (as for native speakers' data) was calculated.

3.2. Results

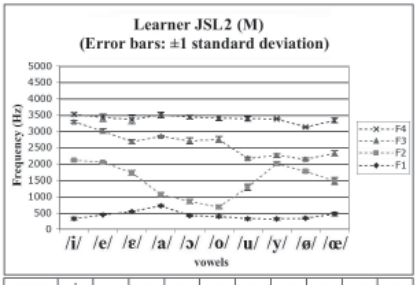
The results in Figure 6 indicate that none of the three learners pronounced /u/ with a low F2 (< 1000 Hz), unlike native speakers. Besides, the spectrograms show that it is not a mere difference of F2 value. Figure 9a (left and middle) illustrates a typical French /u/, in which the energy is highly concentrated around the first two formants located under 1000 Hz, and a typical token of /u/ pronounced by a Japanese learner, in which formants are distributed at similar distances, without any particular concentration of energy. This tendency is further illustrated in Figure 9b (right), which shows energy distribution in different frequency zones.

Concerning /y/, F2 and F3 are close in the productions of learners JSL1 and JSL2, even if F2 is higher (over 2000 Hz) than for the native speakers. By contrast, JSL3 produced a diphthong similar to the sequence /ju/ in Japanese, with F2 decreasing gradually toward the end of the vowel (see Figures 7 and 8).

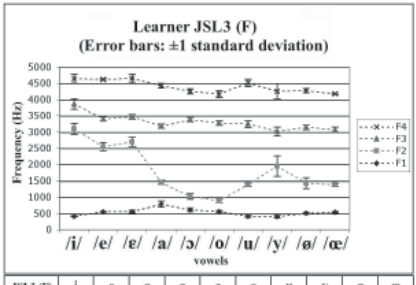
As for /ø/, learners JSL1 and JSL2 produced a vowel with close F2 and F3, probably because of the English spelling <eu> /ju/ which is interpreted as Japanese /juR/, which, in turn, is associated with French /y/. JSL 3 pronounced this vowel with an F2 around 1500 Hz, making it similar to native speakers' productions.



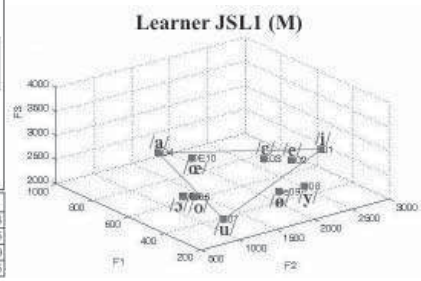
JSL1 (M)	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	3923	3643	3606	3955	3360	3164	3234	3346	3383	3433
F3	3188	2855	2770	2709	2656	2665	2212	2523	2414	2555
F2	2297	2135	2015	1304	749	830	1065	2062	1850	1770
F1	285	382	486	772	407	376	316	281	334	465

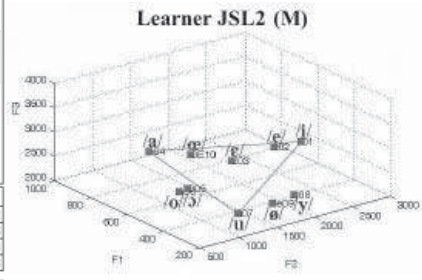


JSL2 (M)	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	3523	3420	3361	3502	3434	3403	3384	3375	3126	3332
F3	3290	3006	2680	2843	2703	2751	2168	2256	2144	2318
F2	2114	2045	1729	1073	858	685	1283	2005	1767	1474
F1	324	444	543	721	414	391	326	317	336	468



JSL3 (F)	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





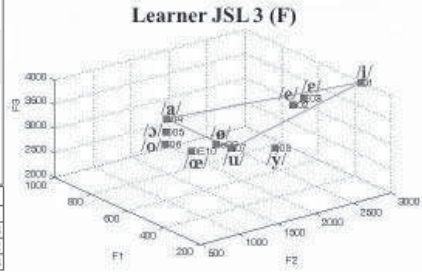


Figure 6

The first four formants (left) and the first three formants (in a 3D vowel triangle: right) of the ten oral vowels of French pronounced by three Japanese-speaking learners (JSL), Learners JSL1, JSL2 (male: top and middle) and JSL3 (female: bottom). Mean of 5 measures x 2 repetitions.

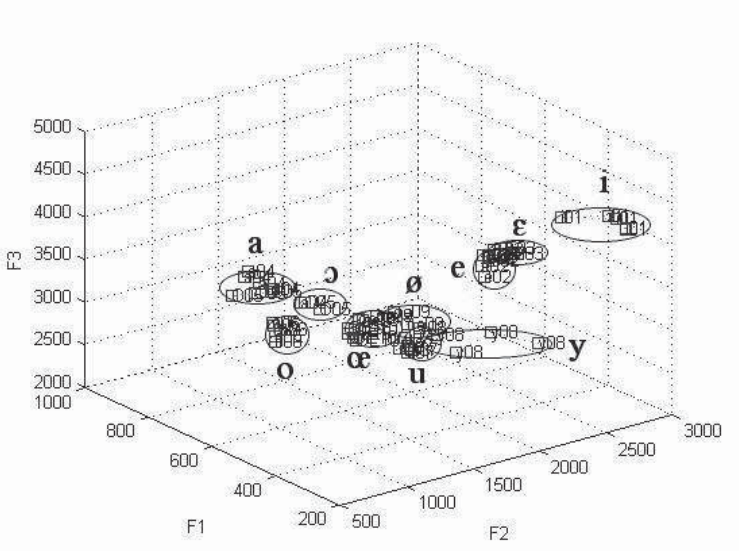


Figure 7

A 3D representation of the first three formants (Hz) of the vowels pronounced by Learner JSL3 (five measuring points during the vowel x one of the two repetitions).

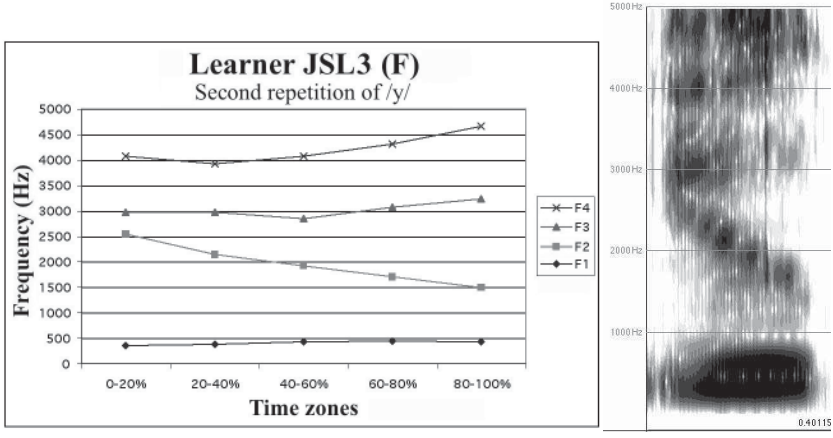


Figure 8

The formants values (left) and the spectrogram of one of the two repetitions of /y/ pronounced by Learner JSL3 (female).

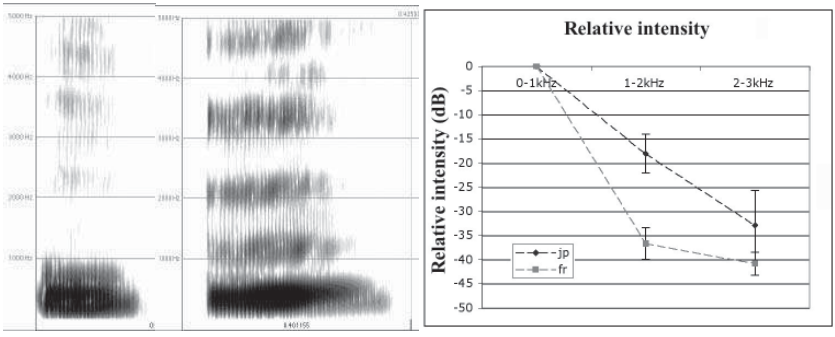


Figure 9a

The spectrograms of /u/ pronounced by the native speaker FR1 (left), and by the Japanese-speaking learner JSL1 (middle).

Figure 9b

Relative intensity in three frequency zones (0-1 kHz, 1-2 kHz, 2-3 kHz) of /u/ pronounced by the four native speakers (fr: mean of 12 tokens) and by the three Japanese-speaking learners (jp: mean of six tokens). The error bars represent ±1SD (Standard Deviation).

To sum up, /u/ was not pronounced in a native-like manner by any of the learners, while /y/ was, to some extent (in spite of slightly high F2), at least by two out of the three learners. /ø/ was produced in a native-like manner by one of them. Are these tendencies observed with a larger number of learners? Is there some more variability? How do they perceive these vowels? Are the tendencies observed in production explained by the learners' perceptual difficulties? A series of three experiments were conducted to tackle those issues: an AXB discrimination test including pairs /u-y/, /y-ø/, /u-ø/ with 14 JSL; the production of /u y ø/ by 47 JSL; the evaluation of JSL's /u y ø/ by 16 native French listeners.

4. Experiment 1: JSL's perception (AXB discrimination)

Experiment 1 consisted of an AXB discrimination task in which 14 Japanese-speaking learners listened to triplets composed of pairs including /u-y/, /y-ø/, and /u-ø/. The aim was to clarify the perceptual difficulties that JSL face and to relate the discrimination results to those of production tasks.

4.1. Method

Seventy-two triplets of vowels presented in isolation were composed of six pairs of vowels /u-y/, /y-ø/, /u-ø/, as well as /i-e/, /u-o/ and /ɛ-a/: 6 pairs x 4 combinations and orders of tokens x 3 speaker conditions: 1) all three stimuli in the triplet were pronounced by the same speaker (female FR3); 2) the second one was pronounced by another speaker of the same sex (female FR4); 3) the second was pronounced by another speaker of the opposite sex (male FR1). The vowel contrasts /i-e/ and /o-u/ were included since some cases of incorrect identification were observed in Kamiyama (2006). On the other hand, it is predicted from the results of the same study that /ɛ-a/ will be discriminated almost perfectly. The mean duration of the vowel tokens was 180 milliseconds (ms). The intra-stimulus interval (between the stimuli in each triplet) was set to one second, so that the stimuli would be processed as linguistic (phonemic) units rather than physical (acoustic) ones; the goal of the acquisition of L2 sound system is to establish phonemic categories of L2, not to develop the ability to discriminate any pairs of sounds based on their acoustic differences. The inter-stimulus interval (between triplets) was set to five seconds. The whole experiment lasted 19 minutes.

The subjects were 14 high school students studying French (as a second foreign language) at Musashi High School (Tokyo, Japan). Half of them had studied it for 15 months ("level 1"), the other half for 27 months or more ("level 2"). They listened to two repetitions of the triplets (144 trials). They were asked to choose either the first or the third sound in each triplet as being identical to the second one, and gave their answers on a sheet of paper. They were seated in a classroom and heard the stimuli through a pair of loudspeakers installed in the room. It should be noted that in the usual classroom situation learners listen to L2 sound material either directly from the instructor or through loudspeakers, except in language laboratories, where each workstation is equipped with a headset.

4.2. Results

The results of the experiment are shown in Figure 10. The mean score of all 14 JSL (the two levels included) was better for /u/-/y/ (94.9%) than for /y/-/ø/ (91.1%), and /u/-/ø/ (84.8% correct) was the most difficult of the three pairs. When we look at the difference between the two groups of learners, it is almost non-existent for /u-y/, but larger for /u-ø/, and still larger for /y-ø/ although statistically non significant. These results suggest that the distinction

between /u-y/ is the least difficult to acquire, that /u-ø/ is the most difficult, and that there is an effect of learning for /y-ø/.

We can also observe that /ɛ -a/ was discriminated almost perfectly. As for /o-u/, the score was as low as that of /u-ø/ for both learner groups, while the advanced group obtained a better score (as high as /u-y/) than the other group (as low as /u-ø/) for the pair /i-e/ (see Kamiyama 2009 for further details on these pairs).

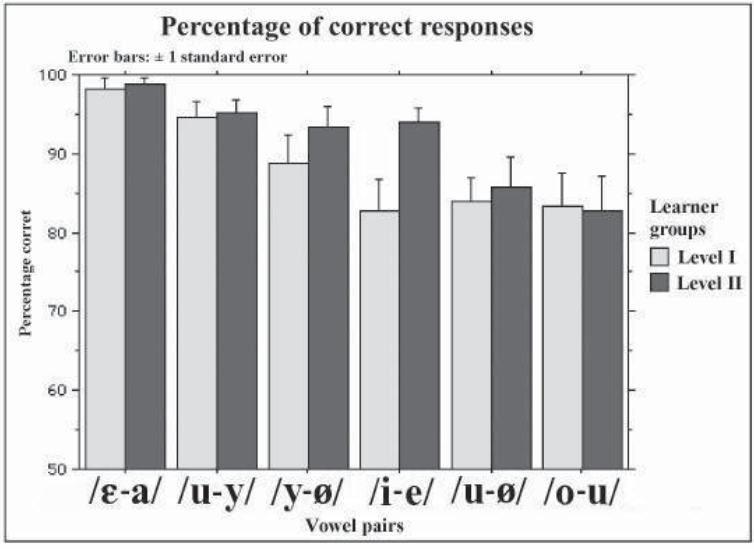


Figure 10

The results of the AXB discrimination task. Perception of French vowels by Japanese high school students. Mean score of each of the two groups (seven listeners in "level 1" and seven in "level 2" groups), two repetitions of 72 triplets. Error bars represent one standard error.

How, then, are these perception results related to the production of these vowels?

5. Experiment 2: JSL's production

Experiment 2 consisted of an acoustic analysis of the three French vowels /u/, /y/, and /ø/ pronounced by 50 JSL studying French in four different

classes. The aim of this experiment was to confirm the tendencies observed in our preliminary study, and to compare the results with the previous perception experiment.

5.1. Method

French oral and nasal vowels including /u/, /y/, and /ø/ were embedded in the carrier sentence "*Je dis* /V/ *comme dans …*" as in the preliminary study. The sentences were arranged in a random order. They were presented in the same experimental session and the subjects were invited to read the sentences one by one. At the end of the list of sentences, a dummy sentence was inserted so that all the other sentences would be read with a similar rhythm and intonation.

The 50 subjects were groups of students attending four different classes in French as a foreign language in universities of the Tokyo area. Group 1 (elementary) was mainly composed of learners who had learned French for one semester (3 hours a week), Group 2 (intermediate) of those who had studied it for 2 years (9 to 12 hours a week), Group 3 (intermediate) of those who had studied it for 1 to 1.5 years (9 hours a week), and Group 4 (upper intermediate) of those who had studied it for one semester (7.5 hours a week) at the university, with some prior learning experience (*e.g.* 2 to 8 hours a week for three years in high school).

The recording was carried out in a CALL (computer assisted language learning) laboratory. The subjects read the sentences at their own pace, using a headset microphone, and the sound was stored on the network at 22.5 kHz/16bits. They were asked to repeat the list five times. Some repeated it six times, others less than five times. The data of those who did not follow the instructions (*e.g.* those who did not pronounce the target vowel) were discarded. The data of 47 learners (out of 50; 40 female and 7 male) were thus retained.

The first four formants of /u/, /y/, and /ø/ were measured near the beginning, in the middle, and near the end of the vowel. The mean values throughout the vowel were also measured and compared with the mean of the three values measured at three points. Some parts of the vowels were excluded for reasons of recording quality or transitions from the preceding word in the carrier sentence. In order to determine quantitatively if the vowel is diphthongised (in particular, /y/ pronounced like Japanese /ju/), F2 was compared at the beginning and at the end of the vowel ("F2 change": F2 at the end divided by F2 at the beginning).

The two measures (1. mean of all values measured every 6 milliseconds during the whole vowel, and 2. mean of the three measures; toward the beginning, middle and end) were compared, and the error was calculated for each

formant and each speaker by dividing the difference between the two measures by the mean of the three-point measure. The maximum error between them was 6.8% (0.7% in average). In 78% of the cases, the error was smaller than 1%. Assuming that such error is negligible, the mean of the whole vowel was used as the formant value.

5.2. Results

Figures 10-12 show the mean first four formants of the three vowels /u y ø/ pronounced by 40 female learners of each group, as compared to those of the two female native speakers (FR3 and FR4). The data of the seven male learners were not included in the calculation of the mean, since the formant values are known to be different for male and female speakers except those due to Helmholtz resonance such as F1 of French /i y u/, F2 of French /u/ (Fant 1960). In spite of the large intra- and inter-speaker variability observed in the data (e.g. a given learner produced a [y]-like vowel three times and [jɯ]-like vowel twice for target phoneme /y/), the mean values of each group reveal some relevant tendencies.

Concerning /u/, only four learners out of 47 (40 female and 7 male) produced a low F2 below 1000 Hz. All other speakers produced an F2 higher than 1000 Hz. The mean first four formants of female learners (Figure 11) show that none of the four groups present values similar to those of the two female native speakers.

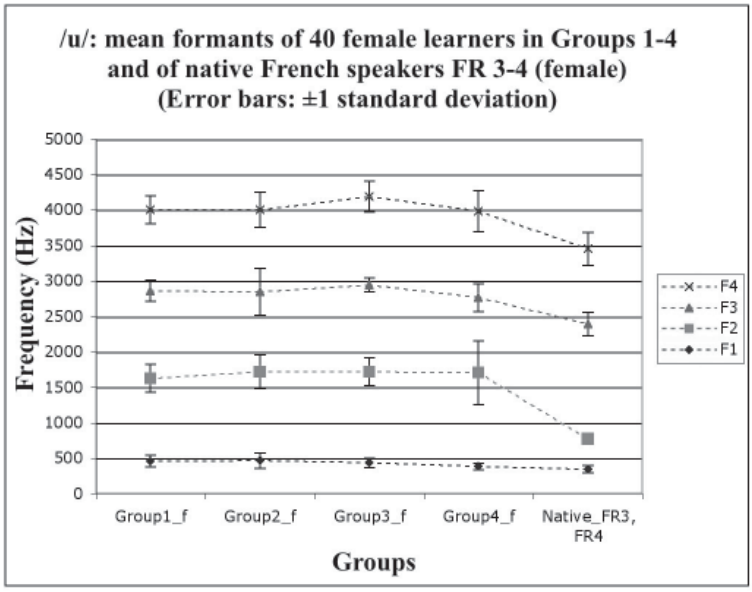


Figure 11

The mean first four formants of /u/ pronounced by 40 female JSL studying French in four different groups: Group 1 (elementary: 14 female learners, 70 repetitions), Group 2 (intermediate: 7 female learners, 34 repetitions), Group 3 (intermediate: 12 female learners, 63 repetitions), Group 4 (upper-intermediate: 7 female learners, 34 repetitions). The mean of the two female native speakers of French (FR3 and FR4, 6 repetitions) were added on the right-most column for comparison. The error bars represent the standard deviation.

By contrast, as many as 20 students our of 47, including elementary level learners (Group 1), succeeded in pronouncing a /y/ token with close F2 and F3 (with F2 not higher than 2400 Hz) without diphthongisation in at least one of the repetitions, even though their production is highly variable except for upper-intermediate learners (Group 4). The mean formant values of the 40 female learners show that the difference between the learners and the native speakers is smaller than that of /u/. Group 4 presents closer F2/F3 than the other JSL groups, making them similar to the native speakers' values in spite of higher F2/F3, probably due to insufficient lip protrusion compared to French native speakers (Figure 12).

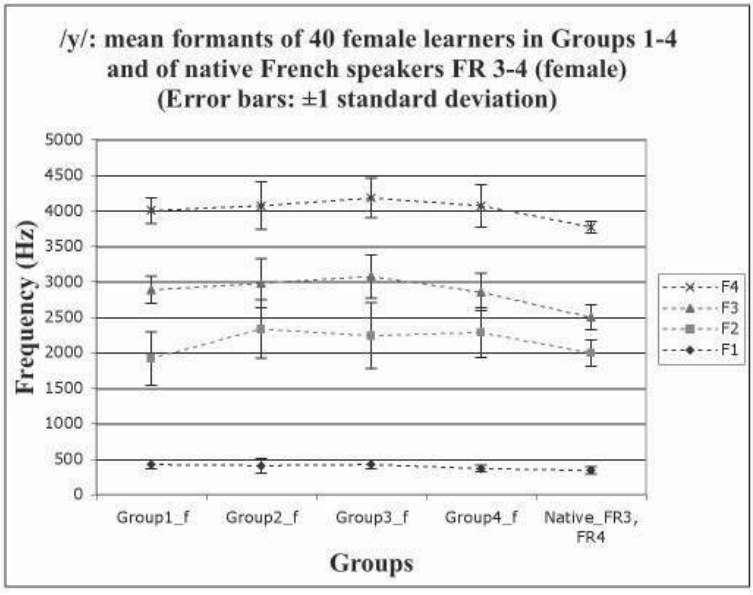


Figure 12

The mean first four formants of /y/ pronounced by 40 female JSL studying French in four different groups: Group 1 (elementary: 14 female learners, 69 repetitions), Group 2 (intermediate: 7 female learners, 34 repetitions), Group 3 (intermediate: 12 female learners, 63 repetitions), Group 4 (upper-intermediate: 7 female learners, 34 repetitions). The mean of the two female native speakers of French (Fr3 and FR4, 6 repetitions) were added on the right-most column for comparison. The error bars represent the standard deviation.

The vowel /ø/ was pronounced by 31 learners out of 47 in a native like manner: F2 was close neither to F1 nor to F3, and located around 1500 Hz. The mean values of female learners indicate that all four JSL groups produced formant values similar to those of the two native speakers, except higher F3 (Figure 13).

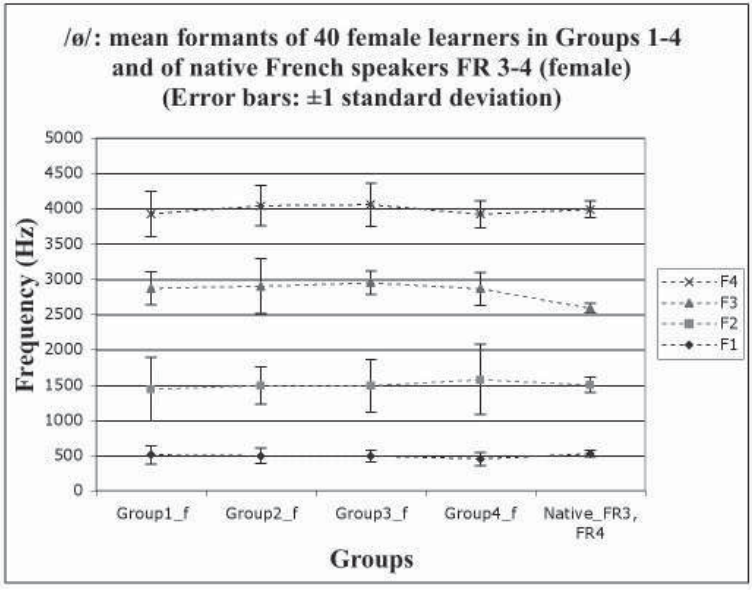


Figure 13

The mean first four formants of /ø/ pronounced by 40 female JSL studying French in four different groups: Group 1 (elementary: 14 female learners, 68 repetitions), Group 2 (intermediate: 7 female learners, 34 repetitions), Group 3 (intermediate: 12 female learners, 63 repetitions), Group 4 (upper-intermediate: 7 female learners, 34 repetitions). The mean of the two female native speakers of French (Fr3 and FR4, 6 repetitions) were added on the right-most column for comparison. The error bars represent the standard deviation.

These findings suggest that JSL learn to produce /ø/ in a native like manner more easily than /y/, which in turn is easier than /u/.

How, then, are these vowels actually perceived by native speakers of French?

6. Experiment 3: native speakers' perception of vowels /u y ø/ produced by JSL

Experiment 3 consisted of a test of identification and rating of JSL's /u y ø/ by native listeners of French. The aim of this experiment was to examine how

the acoustic properties of these vowel phonemes produced by JSL and observed in the previous experiment are perceived by native speakers of French.

6.1. Method

First, six tokens of /u/ in isolation pronounced by the three JSL (2 male, 1 female) in the preliminary study were chosen. Then, 10 tokens of the same vowel produced by the same three JSL in a task of immediate repetition (Kamiyama 2006) were added. Finally, two tokens of /u/ pronounced by two (1 male, 1 female) of the four JSL who produced it with low first two formants (< 1000 Hz) in Experiment 2 were added so that the F2 of the stimuli would vary in a balanced manner. A total of 18 tokens of /u/ were thus selected (Figure 14, top left). As for /y/ and /ø/, the tokens produced in the same series of experiments as each one of /u/ were selected. The mean duration of the stimuli was 225 ms.

Sixteen native speakers of French (NF) aged between 23 and 34 listened to these 54 stimuli (3 vowels x 18 tokens) four times, choosing one of the French vowels they had heard and giving a rating corresponding to its goodness (1 to 5). They were seated in a quite, closed room with a computer. The stimuli were presented through a headphone, and the listeners used a mouse to choose their answers on a computer screen. The interface used for the experiment was Experiment MFC 4 on Praat (Boersma & Weenink 2007). There was no time limit for the answers: the listeners responded at their own pace. The whole experiment took approximately 20 to 25 minutes.

Each stimulus was thus identified and rated 64 times (4 repetitions x 16 NF listeners). The number of times each stimulus was identified as a given vowel was calculated along with the mean goodness score ranging from 0 (no response for that category) to 5 (given 5 out of 5 in all the 64 responses).

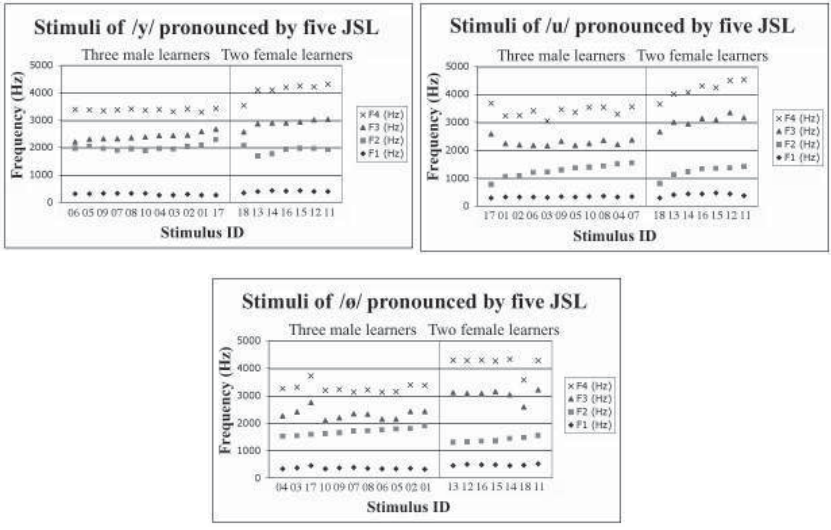


Figure 14

The first four formants of the stimuli used in the experiment (18 tokens of each of the three vowels /u y ø/ pronounced by three male and two female JSL). Each column represents the mean first four formants of all values measured every six milliseconds during the whole vowel stimulus (thus 18 columns for each vowel, 11 for male learners' stimuli, 7 for female learners' stimuli). The stimuli are arranged here in the order of their F2 value for /u/ and /ø/, of F3 for /y/. Note that the tokens of /y/ pronounced by the female learner in the preliminary study were mostly diphthongised, as in Japanese /ju/.

6.2. Results

The results are shown in Figures 16-18. The findings include the following:

i) /u/ (Figure 15): the stimuli with an F2 below 1000 Hz were almost unanimously identified as /u/ with a good rating score. Those with an F2 between 1000 and 1100 Hz were categorised as /ø/ almost as often as /u/. With an F2 above 1100 Hz, the stimuli were mainly identified as /ø/ with a good rating score. F2 shows a very good correlation (0.89 for the stimuli of male learners, 0.83 for those of female learners) with the index "/u/ score - /ø/ score" (ranging from -5 to 5: see Figure 16).

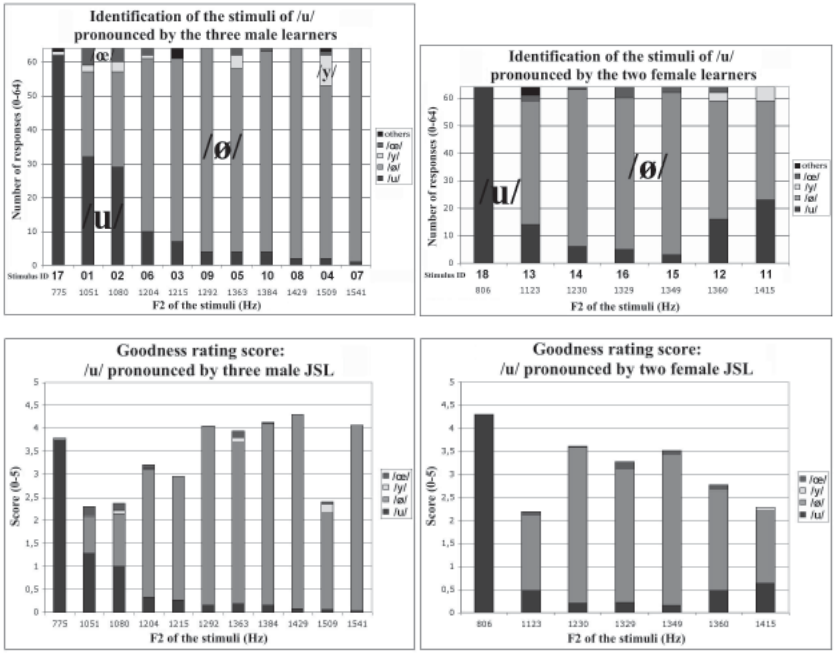


Figure 15

Results of the identification task (above: 64 responses) and mean rating (below: 0 to 5). Eighteen tokens of /u/ pronounced by 5 JSL (3 male and 2 female learners), 64 responses (4 repetitions x 16 NF listeners) for each stimulus.

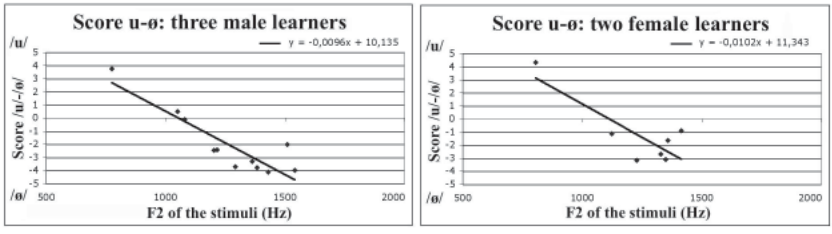


Figure 16

F2 of the stimuli (x-axis) and the index "/u/ score - /ø/ score" (y-axis) for stimuli pronounced by three male learners (left) and two female learners (right). 64 responses (4 repetitions x 16 NF listeners for each stimulus).

ii) /y/ (Figure 17): the stimuli with close F2/F3, and with a not too high F2, were mainly perceived as /y/ with a good rating score. Those with diphthongisation, as in Japanese /ju/ [jɯ] (produced by the female learner examined in the preliminary study: see Figure 14) were identified either mainly as /ø/, or equally as /u/, /y/, or /ø/, but the rating score was low in all cases.

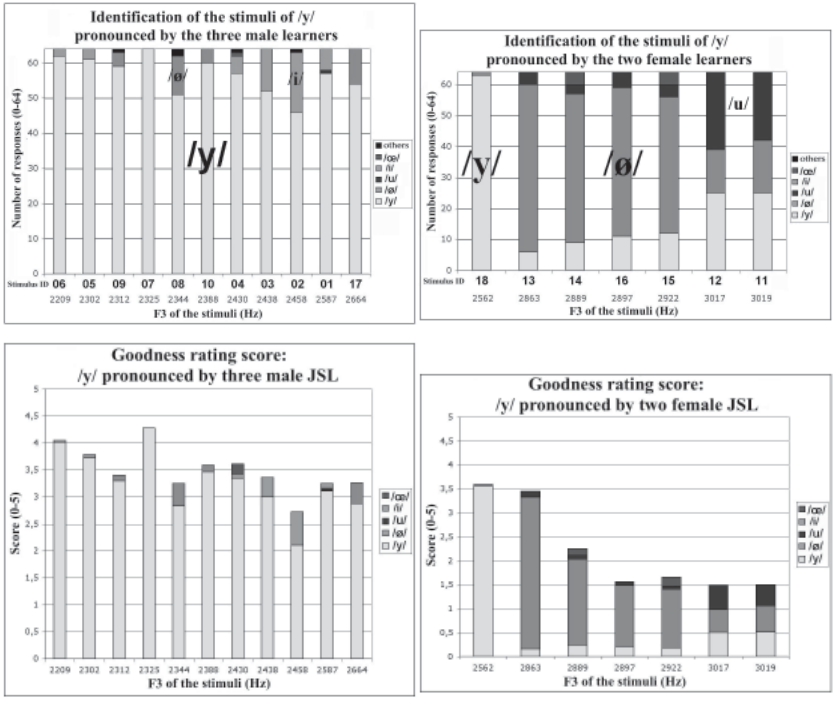


Figure 17

Results of the identification task (above: 64 responses) and mean rating (below: 0 to 5). Eighteen tokens of /y/ pronounced by 5 JSL (3 male and 2 female learners), 64 responses (4 repetitions x 16 NF listeners) for each stimulus.

iii) /ø/ (Figure 18): the stimuli with F2 around 1500 Hz and not close to F3 were identified mainly as /ø/ and obtained good rating scores.

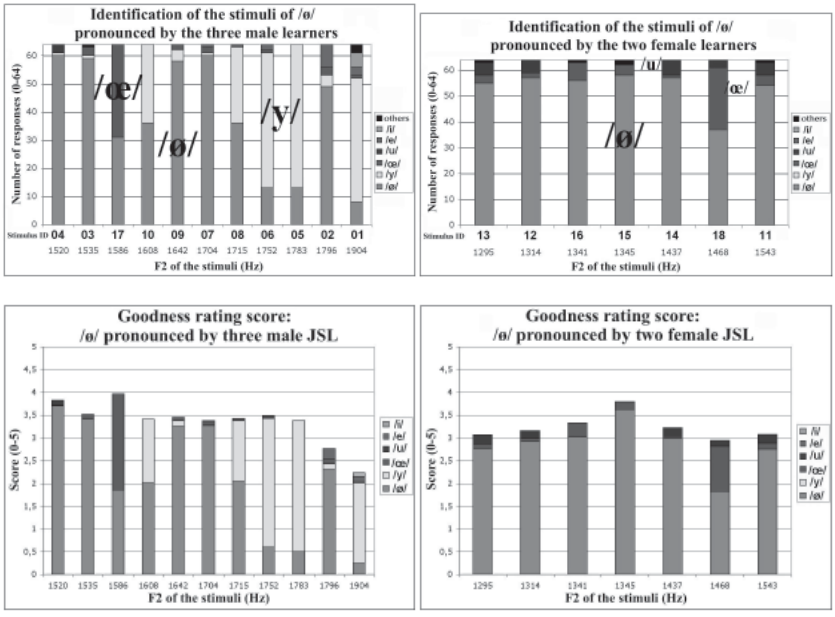


Figure 18

Results of the identification task (above: 64 responses) and mean rating (below: 0 to 5). Eighteen tokens of /ø/ pronounced by 5 JSL (3 male and 2 female learners), 64 responses (4 repetitions x 16 NF listeners) for each stimulus.

7. Discussion and conclusion

The present series of experiments has shown and confirmed the following points:

1. JSL have difficulty distinguishing /u-ø/ perceptually, less so for /y-ø/, and still less for /u-y/ (Experiment 1).

2. Out of the 47 JSL examined, only four produced a native-like /u/, whereas 20 produced a more or less native-like /y/ at least once, and 31 JSL did so for /ø/ (Experiment 2).

3. NF listeners mainly perceived /ø/ when they heard stimuli intended as /u/ by JSL (but produced with a high F2 between 1100 and 1600 Hz), giving a mean rating of 2 to 4.5 out of 5 for /ø/, but less than 1 for /u/ (Experiment 3).

Production

The findings in 2 above confirm the tendency observed in the preliminary study conducted with a smaller number of learners, suggesting that it is not only those three individuals but a larger number of JSL who had difficulty producing /u/, less so with /y/, and still less with /ø/.

It should also be noted that the four learners who pronounced a native-like /u/ in Experiment 2 all had a considerably richer learning experience: three of them had lived in a French speaking country, and the other had learned French in a language school for three years (two hours a week).

The phonetic differences between French /u/ and Japanese /u/ (or between French /u/ and English /u/) are usually known, at least explicitly, only to phoneticians and linguists as well as a small number of students who study phonetics. Most non-native teachers and learners tend to consider French and Japanese /u/ as practically the same sounds and do not pay particular attention to their articulatory and acoustic differences. On the other hand, /y/ is considered to be a "new" vowel that does not exist in L1 from the beginning, and so relatively much attention is paid to the difficulty of its perception and production. It is also easier to show the articulatory configuration of /y/ than that of /u/. Lip rounding and protrusion, which are shared characteristics of /u/ and /y/, are relatively easy to show and perceive visually. However, it is much more difficult for learners to perceive the tongue position of /u/, which is further back compared to Japanese /u/, and to articulate it correctly, simply because it is invisible from outside. It is not easy for them to associate mid-sagittal profiles and position of their own tongue either. This is why researchers and practitioners have devised such methods as moving from low back vowels to high back [u] or using the help of pitch (association of low pitch and [u]).

It is interesting to note that in French both /u/ and /y/ are "quantal" vowels in which two formants are found close together and enhance the amplitude of each other, making them perceptually salient. However, when they are perceived and produced as L2 phones, learners behave differently due to the L1 vowel system.

Another category should be taken into account for the case of /ø/ learned by JSL, who had less difficulty acquiring it. We saw earlier that French /ø/ is characterised by equally distributed formants, and these characteristics are also observed in Japanese /u/. In Experiment 3 the tokens of French /u/ pronounced by JSL, which may be considered to be close to Japanese /u/, were mainly perceived as /ø/ by NF listeners. This finding also supports the acoustic similarity of French /ø/ and Japanese /u/, even if they differ in terms of articulation (front

rounded French /ø/ vs. not fully back, not fully rounded Japanese /u/). The difficulty probably lies in the fact that French /ø/ is considered to be a new sound to JSL (and taught as such), transcribed with an unfamiliar symbol, which might lead to some confusion, also due to its spelling (association of <eu> with English /ju/, leading JSL to pronounce a sequence like Japanese /ju/ or French /y/, as we observed in the preliminary study on production). The case of /u/ is totally different; on the one hand, it is transcribed with a familiar symbol (letter *u* is used to represent Japanese phoneme /u/ in roman transcription *rômaji*, as in "*Ueno*" /ueno/: a place name in Tokyo). On the other hand, the large phonetic difference between the target sound (French /u/) and the L1 sound (Japanese /u/) is not consciously noticed by most JSL. This lack of consciousness might make its acquisition harder than that of /y/, which is phonetically "new", but also phonemically in the sense that there is no equivalent phoneme in the L1 (Japanese) vowel inventory; a fact that might help learners to notice the deviation from the target form more easily. We suggest a hierarchy of difficulty in production represented in Table 2: L2 phonemes could be considered as phonemically or phonetically, similar or new, as compared to L1 phonemes.

	Difficulty in production → /a/	/ø/	/y/	/u/
Phonemically similar	+	-	-	+
Phonetically similar	+	+	-	-

Table 2

Difficulty of acquisition of production in terms of phonemic and phonetic similarity.

Perception

Another interesting issue is the relative perceptual difficulty of pairs /u-ø/ and /u-y/ by JSL and native speakers of American English learning French (AESL), as reported in Gottfried (1984), Levy and Strange (2008), and

mentioned in the introduction of the present paper. Experiment 1 resulted in the following hierarchy of difficulty: /u-ø/ > /y-ø/ > /u-y/, whereas it is with the pair /u-y/ that experienced AESL were found to have difficulty, according to Levy and Strange (2008).

The fact that the pair /u-y/ is not as difficult as /u-ø/ for JSL could be explained by the existence of the sequence /ju/ in Japanese. In our preliminary study on JSL's production, we observed cases in which French /y/ was pronounced as a sequence resembling Japanese /ju/. In loanwords from French, for instance, /y/ is usually interpreted as /ju/, while /u/ is interpreted as /u/ (e.g. "*Lumière*" /lymjɛʀ/ > /rjumieRru/, "*Louvre*" /luvʀ/ > /ruRburu/: /R/ represents the second half of a long vowel. For more data and detailed phonological analysis on loanword adaptation in Japanese from French and English, see Shinohara 1997). Note that /u/ and /ju/ are in opposition in Japanese (ex. /uku/ 'to float' vs. /juku/ 'to go'; /'umi/ 'sea' vs. /'jumi/ (female first name)). In terms of PAM, the French pair /u-y/ could thus be considered as "two-category assimilation", in which case discrimination is expected to be excellent.

Sequences /u/ and /ju/ are also found in English, but the opposition could be considered as partially neutralised because of "j-dropping", even more so in American English (see Wells 2000, for example). There exist some minimal pair words in contexts that did not undergo the phenomenon (e.g. "cue" /kju/ vs. "coo" /ku/), but the two sequences are neutralised in other contexts (e.g. "tune" /tun/, "toon" /tun/). It is not easy to decide whether j-dropping was caused by the similarity between /u/ and /ju/, or whether j-dropping favoured the similarity between these two sequences. In any case, it is likely that AESL cannot easily assimilate French /u-y/ to English /u-ju/, unlike JSL (note also that English /u/ is diphthongised, with its F2 lowering during the vowel, unlike Japanese /u/). This case could thus be considered as an example of "single-category assimilation" (or "category-goodness difference") rather than of "two-category assimilation" in terms of PAM, making it difficult for native speakers of American English to distinguish this pair.

One of the reasons for the relative ease with which JSL learn to distinguish French /u/ and /y/ perceptually may be due to the fact that both /u/ and /y/ have the advantage of being focal vowels (with a high concentration of energy). Besides, the focal nature of these vowel phonemes may and should be included in their symbolic representation to better characterise them (Vaissière, 2006). In addition, /u/ has a much lower centre of gravity than /y/ in all cases, which may explain in part why /u/-/y/ is easier to discriminate than the other pairs.

The present article only dealt with vowels in isolation. Levy and Strange (2008) report the effect of consonantal context (bilabial and alveolar contexts) on American English speaking learners (AESL)'s perception of French vowel pairs including /u/, /y/ and /œ/. It will be interesting to examine how the consonantal context influences the perception of the three vowels /u y ø/ by JSL.

Also, the findings of the present study could be applied to the development of a training system of these three vowels, and eventually, of others: users-learners would be able to pronounce the target vowel and get feedback on the acceptability as a token of the intended vowel, based on the correspondence between the real-time acoustic analysis of the vowel produced, and the expected values by native listeners. Learners are thus led to associate the mental representation of the target vowel and the articulatory gestures needed for its phonetic realisation. Further experiments collecting more detailed and ample data will be needed to specify the parameters required for such a system.

Acknowledgments: The authors wish to express their gratitude to Miho Inoue for kindly accepting to record her students during precious class time, and to the two anonymous reviewers for reading an earlier draft of the paper and for providing valuable comments.

REFERENCES

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

Boersma, P. & Weenink, D. (2007). *Praat: doing phonetics by computer* (Version 4.6.13) [Computer program]. Retrieved in August 2007, from http://www.praat.org/.

Bothorel, A., Simon, P., Wioland, F. & Zerling, J.-P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Publications de l'Institut de Phonétique de Strasbourg, Strasbourg.

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

Chiba, T. & Kajiyama, M. (1955). *The Vowel: Its Nature and Structure*. Tokyo: Tokyo-Kaiseikan Publishing (first edition published in 1941).

Deterding, D. (2006). The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association* n° 36 (2), 187-196.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

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

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* n° 15, 47-65.

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

Fónagy, I. (1983). *La vive voix : essais de psycho-phonétique*. Payot, Paris.

Gendrot, C. & Adda-Decker, M. (2004). Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande. *Proceedings of the Workshop MIDL (Modélisations pour l'IDentification des Langues),* 7-12.

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

Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics* n° 12, 91-114.

Hawkins, S. & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* n° 35 (2), 183-199.

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

Kamiyama, T. (2006). La production et la perception des voyelles orales françaises par les apprenants japonophones : une étude préliminaire. *Actes des XXVIèmes Journées d'étude sur la Parole,* 49-52.

Kamiyama, T. (2009). *Apprentissage phonétique des voyelles orales du français langue étrangère chez des apprenants japonophones*. Unpublished PhD. Dissertation. Université de la Sorbonne Nouvelle, Paris.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences* n° 97, 11850-7.

Lambacher, S., Martens, W. & Molholt, G. (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.

Lauret, B. (1998). *Aspects de Phonétique expérimentale contrastive : «l'accent» anglo-américain en français*. Unpublished PhD. Dissertation. Université de la Sorbonne Nouvelle, Paris.

Levy, E. S. & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetic,* n° 36, 141–157.

Martin, P. (2002). Le système vocalique du français du Québec. De l'acoustique à la phonologie. *La linguistiqu* n° 38 (2), 71-88.

Mokhtari, P. & Tanaka K. (2000). A Corpus of Japanese Vowel Formant Patterns. *Bulletin of Electrotechnical Laboratory* n° 64 (special issue), 57-66.

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

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

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

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 379-410. York Press, Baltimore.

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

Shinohara, S. (1997). Analyse phonologique de l'adaptation japonaise de mots étrangers. Unpublished PhD. Dissertation. Université de la Sorbonne Nouvelle, Paris.

Strange, W. (ed.) (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Speech Research*. Baltimore: York Press.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K. & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics* n° 26, 311-344.

Strange, W., Weber, A., Levy, E., Shafiro, V. & Nishi, K. (2004). 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* n° 115 (5), 2606-2606.

Sugito, M. (1995). *Oosaka - Toukyou akusento onsei jiten CD-ROM: kaisetsuhen* [CD-ROM Accent dictionary of Spoken Osaka and Tokyo Japanese]. Maruzen, Tokyo.

Uemura, Y. (Kokuritsu Kokugo Kenkyûjo [National Institute for the Japanese Language]) (1990). *Nihongo no boin, shiin, onsetu: chouon undou no jikken-onseigakuteki kenkyû* [Vowels, consonants and syllables in Japanese: an exprimental phonetic study on articulatory movements]. Shûei shuppan, Tokyo.

Vaissière, J. (2006). *La phonétique*. Presses universitaires de France, Paris.

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 & M. Ohala (ed.), *Experimental Approaches to Phonology*, 54-71. Oxford: Oxford University Press.

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

Résumé

Deux expériences de perception et une expérience de production ont été effectuées afin d'examiner les comportements des apprenants japonophones (AJ) concernant les trois voyelles /u/, /y/ et /ø/, et de les comparer avec le cas des apprenants anglophones. Les résultats de ces expériences suggèrent les points suivants : i) les AJ ont tendance à produire le /u/ français avec un F2 élevé (> 1000 Hz), ce qui est perçu comme /ø/ par les auditeurs natifs francophones (NF) ; ii) pour les AJ, le /u/ français pourrait être considéré comme similaire sur le plan phonémique (en tant que voyelle postérieure fermée) mais nouveau sur le plan phonétique, /y/ comme nouveau sur les plans phonémique et phonétique, /ø/ comme nouveau sur le plan phonémique mais similaire sur le plan phonétique (acoustique). Les voyelles phonémiquement similaires mais phonétiquement nouvelles semblent être les plus difficiles à apprendre à produire correctement pour les apprenants des langues étrangères/secondes.