Espaces Linguistiques

Perception and production of French oral vowels by Japanese-speaking learners: does dialectal variation play a role?

Perception et production des voyelles orales du français par des apprenants japonophones : la variation dialectale a-t-elle un rôle à jouer ?

Takeki KAMIYAMA

UR TransCrit, Université Paris 8 Vincennes - Saint-Denis

takeki.kamiyama@univ-paris8.fr

Jumpei KANEDA

ATR Advanced Telecommunications Research Institute International

Ryoko HAYASHI

Kobe University

Abstract: This paper presents a series of studies on the impact of regional variation in L1 (source language) on L2 (target language) pronunciation, focusing on the high back vowel /u/ in French perceived or produced by native speakers of Japanese from the Kanto area (around Tokyo) and the Kansai area (around Osaka). 1) To serve as a base line, the 5 Japanese vowels were pronounced by 11 speakers (5 women and 6 men) from Kansai with no learning experience of French and their first 4 formants were measured to examine their acoustic properties. 2) 25 naïve listeners from Kansai took part in an AXB auditory discrimination task for French vowels including pairs /u/-/y/, /y/-/ø/ and /u/-/ø/; the results were compared with those of Kanto speakers in previous studies. 3) Some Japanese-speaking learners of French (JSL) from Kansai produced the French /u/ with a lower F2 (around 1,000 Hz for men) than typical values observed for learners from the Tokyo area. These results show limited differences for some of the speakers from Kansai compared to the tendency commonly observed with Kanto speakers.
Keywords: French vowels, Japanese speakers, regional variation, perception, production

Résumé : Cet article présente une série d'études portant sur l'impact de la variation régionale en L1 (langue source) sur la prononciation en L2 (langue cible), en se focalisant sur la voyelle fermée postérieure /u/ en français perçue ou produite par des locuteurs natifs du japonais du Kanto (région de Tokyo) et du Kansaï (région d'Osaka). 1) Pour servir de référence, les 5 voyelles du japonais ont été prononcées par 11 locuteurs (5 femmes et 6 hommes) du Kansaï (Osaka) sans expérience d'apprentissage du français et les 4 premiers formants ont été mesurés afin d'examiner leurs propriétés acoustiques. 2) 25 auditeurs originaires du Kansaï et non-apprenants du français ont effectué une tâche de discrimination auditive AXB de voyelles françaises, comprenant les paires /u/-/y/, /y/-/ø/ et /u/-/ø/ : les résultats ont été comparés avec ceux de locuteurs du Kanto dans des études antérieures. 3) Certains apprenants japonophones du français langue étrangère du Kansaï ont produit le /u/ français avec un F2 plus bas (autour de 1000 Hz pour les hommes) que les valeurs typiques observées chez des apprenants du Kanto. Ces résultats montrent des différences limitées chez certains des locuteurs du Kansaï par rapport à la tendance communément observée chez des locuteurs du Kanto.
Mots clés : voyelles françaises, japonophones, variation régionale, perception, production

Introduction

In the acquisition of second language (L2) pronunciation, the learner's source language (L1) is generally considered to have a crucial impact, as seen in influential models of L2 speech learning such as PAM-L2 (Perceptual Assimilation Model of Second Language Speech Learning: Best & Tyler, 2007) and SLM-r (revised Speech Learning Model: Flege & Bohn, 2021), as well as in more traditional accounts of Polivanov (1931), Trubetzkoy's phonological 'sieve' (1939/1969), or in contrastive analysis (Weinreich, 1953/1968; Lado, 1964). A large number of studies on L2 pronunciation or cross-language speech perception focus on a single L1 variety, often widely considered as a standard (e.g. Seoul as opposed to Gyungsang Korean or other accents).

Some, however, report cross-language studies comparing the impact of different L1 varieties on the classification of L2 vowels. Morrison (2008) ran a perceptual classification experiment using a series of synthesized stimuli covering the acoustic space acceptable to L1 English listeners as the English /ɪ/. Three groups of monolingual listeners (19 Western-Canadian English, 17 Peninsular-Spanish and 20 Mexican-Spanish speakers) were asked to classify the stimuli into phonemic categories of their L1. The stimuli for which Western-Canadian listeners' modal response was /ɪ/ were almost all identified as Spanish /e/ by Peninsular-Spanish listeners, while three-quarters of the portion of the same stimulus space was identified as Spanish /i/ and one quarter as Spanish /e/ by Mexican listeners. Chládková & Podlipský (2011) conducted a perceptual assimilation experiment of Dutch vowels by Czech listeners from Bohemia (BC: n=19) and Moravia (MC: n=22) with no knowledge of or no previous exposure to Dutch. In Bohemia, long and short high front vowels /iː/ and /i/ have spectrally different phonetic realizations ([i] and [ɪ], respectively), whereas in Moravia the spectral difference is much smaller, if not totally absent. The result of the experiment shows that BC and MC listeners perceive the Dutch vowels /i/ et /ɪ/ differently: BC labelled the tense vowel /i/ more often (61.8%) as the Czech long vowel /iː/ than MC did (34.4%), and the difference turns out to be significant (p < .001).

These results, obtained in cross-language studies with non-learners, suggest that the variety of L1 could also have a considerable impact on the acquisition of phonemic contrasts amongst L2 learners.

Based on the above-mentioned literature, the goal of the present paper is to present a series of studies enabling a comparison of the influence of two L1 regional varieties of Japanese on the perception and production of the high back vowel /u/ in French by native speakers of Japanese,

contrasted by neighbouring vowels. The L1 varieties under scrutiny are 1) from the greater Tokyo area (Kanto region), often considered to be a "national standard" (Shibatani, 1990), and 2) from the Kansai region, including Osaka (Figure 1).

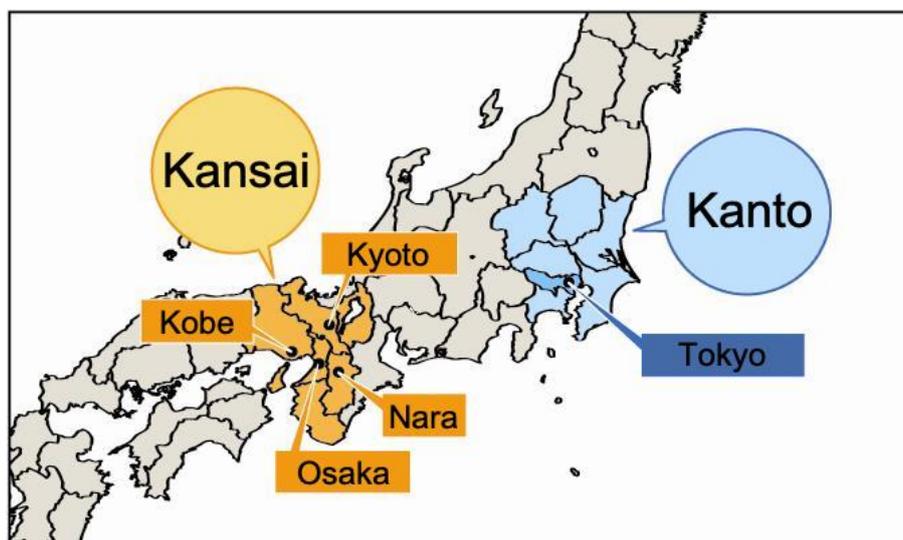Figure 1: Kansai (including cities of Osaka, Kyoto, Kobe and Nara) and Kanto (including Tokyo) regions.

Figure adapted from Tanaka, Tetsuya "*Nihongo kyouiku tsûshin: jugyou no hinto: hougen (kansaiben) ni fureru* (Japanese Language Education Newsletter: Lesson tips: Exposure to dialects - Kansai dialect)", Japan Foundation[1].

Figure 2: Vowel inventories of Reference French (left: based on Vaissière, 2006, *inter alia*) and of Tokyo (Kanto) Japanese (right: based on Shibatani, 1990, Sugitô, 1995, Vance, 2008, *inter alia*).


---

[1] https://www.jpf.go.jp/j/project/japanese/teach/tsushin/hint/201107.html
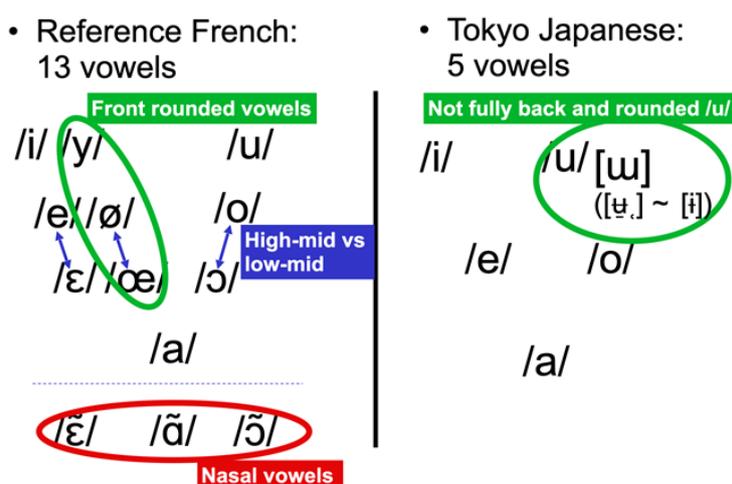
The target variety, Reference French (RF), as described by Detey, Lyche *et al.* (2016)[2], was chosen for two reasons. First, the learners who participated in the experiment resided in Japan and had very little experience abroad in a French-speaking area. Second, the RF variety is widely adopted in teaching material used in foreign language settings. RF has 3 nasal vowels, and 10 oral vowels, which include a series of front rounded vowels /y/, /ø/ and /œ/. By contrast, the 5-vowel system of Tokyo and Osaka Japanese does not include such front rounded vowels (Figure 2). These vowels, together with /u/, can be a challenge for Japanese L1 learners of French (JSL) (Kamiyama, 2011). While both languages have a phoneme commonly transcribed as /u/, the high back vowel in RF is realized as a focal vowel (Schwartz et al., 1997) with F1 and F2 being close together and under 1,000 Hz, thereby forming a frequency zone of high concentration of energy (Liénard, 1977; Vaissière, 2007, *inter alia*). It has been shown that native speakers of Tokyo (Kanto) Japanese who learn French as a foreign language (FFL) tend to produce this target vowel /u/ with a higher F2 without a high concentration of energy for F1 and F2 (Figure 3). This corresponds to a tongue position typically more fronted than that of native speakers (Kocjančič Antolík, Pillot-Loiseau & Kamiyama, 2019). The vowel tokens thus produced by JSL may be perceived by native listeners of French as the mid-high front rounded /ø/, characterized by evenly distributed formants without a zone of high concentration of energy (Kamiyama & Vaissière, 2009). In perception, the vowel contrast /u/-/ø/ is among the most difficult for JSL to distinguish (Kamiyama & Vaissière, 2009), as expected from French loanwords in Japanese, where both vowels are usually adapted to /u/ in Japanese ("Strasbourg" /stʀasbuʀ/ -> ストラスブール /sutorasubuRru/; "Périgueux" /peʀigø/ -> ペリグー /periguR/[3]), suggesting the case of Single-Category (SC) assimilation (or Category-Goodness assimilation) in PAM (Perceptual Assimilation Model: Best, 1995). A perceptual categorization experiment using stimuli made by articulatory synthesis with Maeda's articulatory parameters (Maeda, 1982) corroborates the fact that the acoustic-articulatory space of French native listeners' /u/ and /ø/ corresponds roughly to the acoustic-articulatory space of /u/ for Japanese native listeners (Kamiyama, 2011).

---

[2] The authors prefer this term to that of *Standard French*, "seen by many as too prescriptive", but at the same time, they acknowledge that "unfortunately, all too often, Reference French represents exactly the same reality: an abstraction, a set of features attributed to a certain social class (educated people) and geographical area (Paris and surroundings)" (Detey, Lyche *et al.*, 2016, p. 58-59).

[3] /R/ is a phoneme that is realized as the lengthening of the preceding vowel in Japanese (cf. Labrune, 2012; some authors use /H/ instead of /R/: Vance, 2008).

Figure 3: Wide-band spectrogram (window length: 5 milliseconds), obtained using Praat (Boersma & Weenink, 2007), of French /u/ in isolation pronounced by a male native speaker of French (left), and by a male JSL from Tokyo (right).
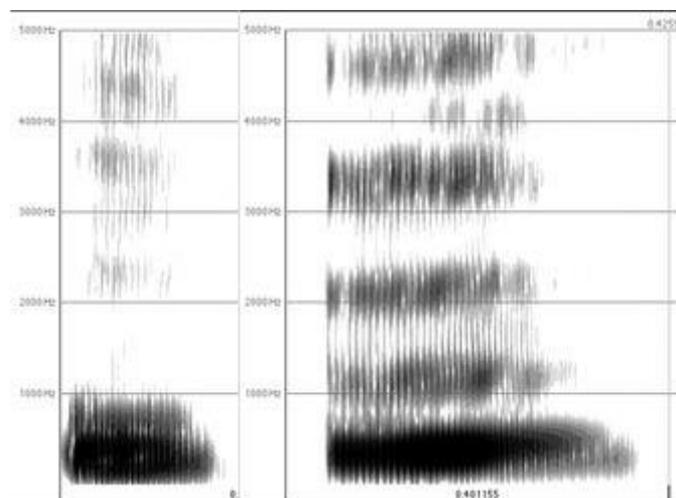
The trend described above stems from the fact that the high non-front vowel in Kanto (Tokyo) Japanese /u/ shows a higher F2 (> 1,000 Hz) than that of the French /u/. By contrast, the vowel /u/ in Kansai Japanese is usually described as more rounded (Shibatani, 1990) and its F2, as well as F1, is lower (Sugitô, 1995) than that of Kanto (Figure 4).

Figure 4: F1 (vertical axis) and F2 (horizontal axis) of the 5 vowels in Osaka (Kansai: left) and Tokyo (Kanto: right) Japanese pronounced by male speakers.
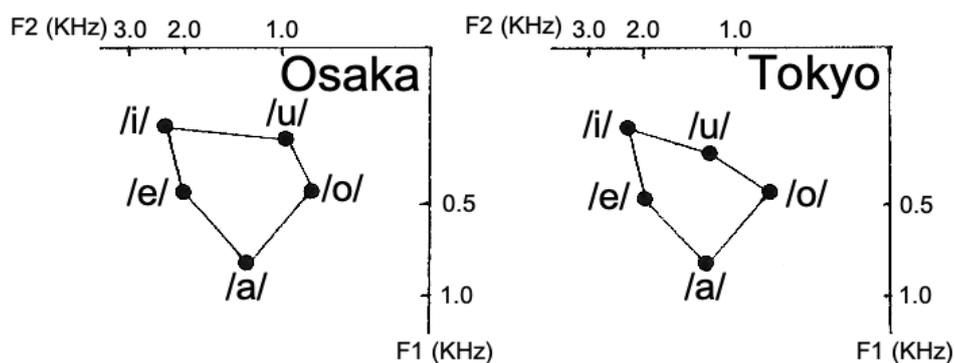
Figures adapted from Sugito (1995)

The articulatory and acoustic characteristics of the vowel /u/ in Tokyo (Kanto) Japanese are described as follows: /u/ in isolation shows a more front, and especially lower tongue position, with less protruded lips (Kokuritsu Kokugo Kenkyûjo, 1990; Takebayashi, 1996) compared to the French /u/ (Bothorel *et al.*, 1986); acoustically, F2 of /u/ varies mostly between 1,000 and 1,500 Hz

for male speakers in various consonantal contexts, which is comparable to F2 of /a/ among other vowels (Figure 5, adapted from Mokhtari & Tanaka, 2000). Little is known, however, about articulatory and acoustic properties of Kansai Japanese vowels, except for the illustrative descriptions mentioned above. It is therefore unclear whether Japanese listeners from Kansai may be influenced by possibly low values of F2 in /u/ and categorize non-native vowels differently from Kanto listeners, or whether they also benefit from these articulatory and acoustic properties when they acquire second or foreign language vowels (PAM and PAM-L2: Best, 1995; Best & Tyler, 2007).

Figure 5 : F1 (horizontal axis) and F2 (vertical axis) of the 5 vowels in Tokyo (Kanto) Japanese pronounced by 5 male speakers. 22 words containing either a long vowel /VR/ or a hiatus /VV/
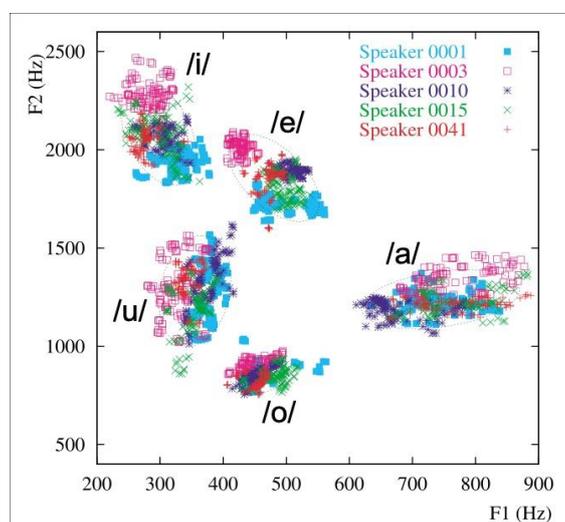
Figure adapted from Mokhtari & Tanaka (2000)

The findings presented above lead us to the following hypotheses:

1) Kansai-Japanese speakers produce /u/ with lower F2 than Kanto speakers in their L1 Japanese.

2) Kansai-Japanese listeners better distinguish the French /u/ from other French vowels which may sound similar for Japanese listeners in general, than Kanto listeners.

3) Kansai-Japanese speakers learning French as a foreign language learn to produce more native-like tokens of the French /u/ than learners from Kanto.

These hypotheses motivated three experiments, presented in the following sections:

1) Production of the 5 vowels in Japanese by native speakers from Kansai

2) AXB discrimination task for French vowel pairs by naive listeners (non-learners of French) from Kansai

3) Production of French vowels by Japanese-speaking learners of French from Kansai.

1. Experiment 1: acoustic analysis of Kansai Japanese vowels

In the first experiment, the acoustic properties of the 5 Japanese vowels produced by Kansai speakers were examined to observe the actual tendency amongst speakers and to serve as a baseline.

1.1. Method

11 speakers (5 women and 6 men) from Kansai (Hyogo, Osaka, Nara, Shiga, Wakayama prefectures), aged 18 to 23, all students at Kobe University, participated in this experiment. One of them spent 3 years from age 6 to 8 in Ibaraki Prefecture in Kanto, and another lived in Germany for a year at the age of 21, but none of the others had lived outside Kansai. Their self-assessed degree of use of *kyoutsûgo* (literally, "common language", a term referring to non-dialectal variety, often considered close to Tokyo Japanese) ranged from 1 (not frequent) to 5 (frequent). The 5 Japanese vowels in isolation /i/ /e/ /a/ /o/ /u/, as well as the sequences /ja/ /ju/ /jo/ were embedded in the carrier sentence /sore o _____ to iu/ ("one calls that …") presented one by one on a screen in a semi-random order. The test items were preceded by 3 training sentences in colloquial Kansai (Osaka) Japanese taken from Sugito (1995). The list of carrier sentences was read 5 times. The participants wore a head-set microphone and the audio data were recorded at 44.1 kHz, 16 bits, using ROCme! (Ferragne *et al.*, 2012). The procedures were conducted by a male researcher (from Tokyo) the participants met for the first time. The first 4 formants were measured in three zones, as mean values during the first, second, and last third of the vowel portion. It should be noted that no clear diphthongization tendency has been reported about Japanese vowels in isolation. This was also the case in the present data set.

1.2. Results

The mean F2 of /u/, shown in Figure 6, mostly ranges between 1,000 Hz and 1,500 Hz for women and 1,000 Hz and 1,300 Hz for men, which is comparable to that of /a/.

Figure 6: The mean values and Standard Deviation (SD) of the first 4 formants of the 5 Japanese vowels /i/ /e/ /a/ /o/ /u/ pronounced by 5 women (left) and 6 men (right) from Kansai (3 measures per token, 5 repetitions). The error bars represent ±1SD
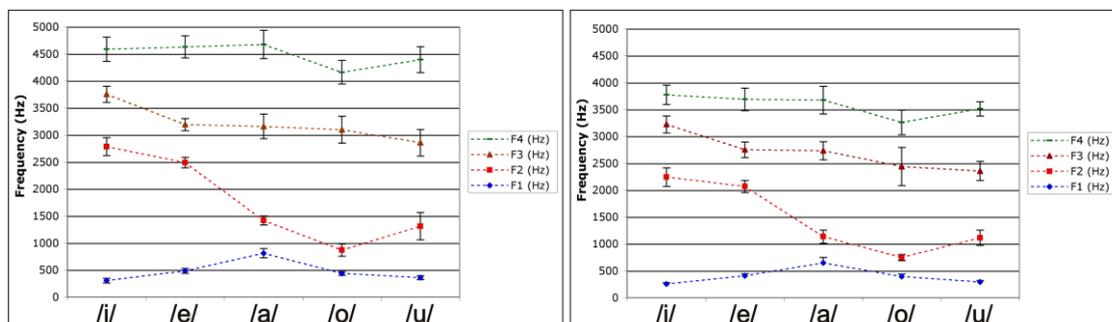
Taking a closer look at the distance between F1 and F2, which is known to be larger in the Kanto (Tokyo) Japanese /u/ than the French /u/, reveals a higher inter-speaker variability than for the other non-front vowels /o/ and /a/, as shown in Figure 7. The relatively low F1-F2 distance for the female speaker sp08 and the male speakers sp02, sp04, sp05 and sp06 is caused by F2 ranging between 1,000 and 1,100 Hz. Some speakers showed a higher intra-speaker variability than others: for sp04, for example, mean F2 ranged between 1,150 and 1,200 Hz in the first two repetitions, but then dropped below 1,000 Hz in the other three repetitions.

F1-F2 distance tends to be higher for speakers who gave a higher rating to the self-assessed degree of use of the non-local form, "common language" (*kyoutsûgo*), even though this tendency is not systematic.

Figure 7: The mean values and Standard Deviation (SD) of the distance between F1 and F2 of the Japanese vowel /u/ pronounced by 5 women (left) and 6 men (right) from Kansai (3 measures per token, 5 repetitions). The number after Speaker ID represents the self-assessed degree of use of the "common language" (k*youtsûgo).* The error bars represent ±1SD
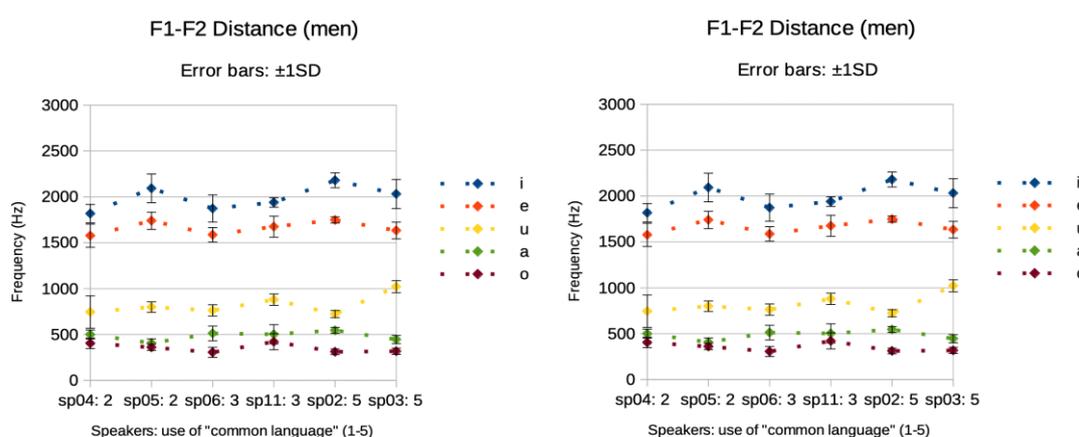
The findings suggest that Kansai listeners are indeed exposed to /u/ tokens with relatively low F2 to some extent at least (compared to Kanto listeners), in spite of variability, which might impact the perception of the French /u/, produced typically with even lower F2 in isolation.

2. Experiment 2: Auditory discrimination of French vowel pairs by naïve listeners from Kansai

Following the results of Experiment 1, an AXB auditory discrimination task was set up to examine the perceptual distinction of French vowel pairs including /u/ by naïve listeners of Kansai Japanese.

2.1. Method

The participants were 25 students enrolled in either of two different universities located in the cities of Osaka and Kobe. None of them had learnt French or any other language characterized by having a series of front rounded vowels in its vowel inventory.

Six vowel pairs /u/-/y/, /y/-/ø/, /u/-/ø/, as well as /i/-/e/, /u/-/o/ and /ɛ/-/a/, were used to compose 72 triplets of vowels in isolation: 6 pairs x 4 combinations and orders of tokens (/u u y/, /u y y/, /y u u/ and /y y u/ for the vowel pair /u/-/y/) x 3 speaker conditions: 1) all 3 stimuli in the AXB triplet were pronounced by the same female native speaker of Reference French; 2) "A" (the first stimulus in the triplet) and "B" (the last) were pronounced by the same female native speaker as in the previous condition, but "X" (the second one) was pronounced by another RF female native speaker; 3) "X" was pronounced by a male RF native speaker, while "A" and "B" were produced by the same female speaker as in the first speaker condition. The vowel contrasts /i/-/e/ and /o/-/u/ were included since some cases of incorrect identification were observed in a perceptual identification test administered for FFL in Kamiyama (2011). By contrast, 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 1 second, so that the stimuli would be processed as linguistic (phonemic) units rather than physical (acoustic) ones. The list of 72 triplets were played 2 times in different orders. The experiment was self-paced and conducted using Praat (Boersma & Weenink, 2007).

2.2. Results

The percentage correct is represented in Figure 8. Among the pairs including the vowels /u/, /y/ and /ø/, the contrast /u/-/ø/ was the most difficult to distinguish perceptually (78.3 % correct), followed by /y/-/ø/ (83,4% correct) and /u/-/y/ (93.4% correct). This order is the same as for listeners from Kantô: for 7 non-learners of French, 76.2% correct for /u/-/ø/, 79.8% correct for /y/-/ø/, 85,7% correct for /u/-/y/; for 14 learners of French as a foreign language, 84.8% correct for /u/-/ø/, 91.1% correct for /y/-/ø/, 94.9% correct for /u/-/y/ (Kamiyama & Vaissière, 2009).

Figure 8: Mean percentage correct of discrimination for French vowel pairs perceived by 25 non-learners of French from Kansai. 2 repetitions x 12 triplets x 6 vowel pairs. Chance level: 50%.
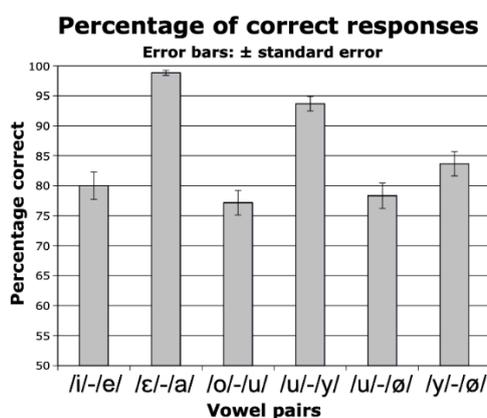
3. Experiment 3: Production of French vowels by JSL from Kansai

The previous experiment shows that the French phonemic contrast /u/-/ø/ is as difficult for Kansai listeners to distinguish perceptually as for Kanto listeners. The present study sheds light on the production of these two vowels among other French vowels by learners of French as a foreign language from Kansai in comparison with those from Kanto.

3.1. Method

The speech material (part of the corpus reported in Landron *et al.*, 2010) consists of the 13 RF vowels (10 oral /i e ɛ a ɔ o u y ø œ/ and 3 nasal (/ɛ̃ ɑ̃ ɔ̃/), embedded in carrier sentences such as "*Bébé, je dis « é » comme dans bébé*" (Baby, I say "é" as in baby). In the present study only the 10 oral vowels are analyzed. The target vowel was elicited with the help of a keyword (e.g. "*bébé*", "*pouce*", "*deux*" containing /e/, /u/, /ø/, respectively) and a typical spelling pattern (e.g. « é » /e/, « ou » /u/, « eu » /ø/).
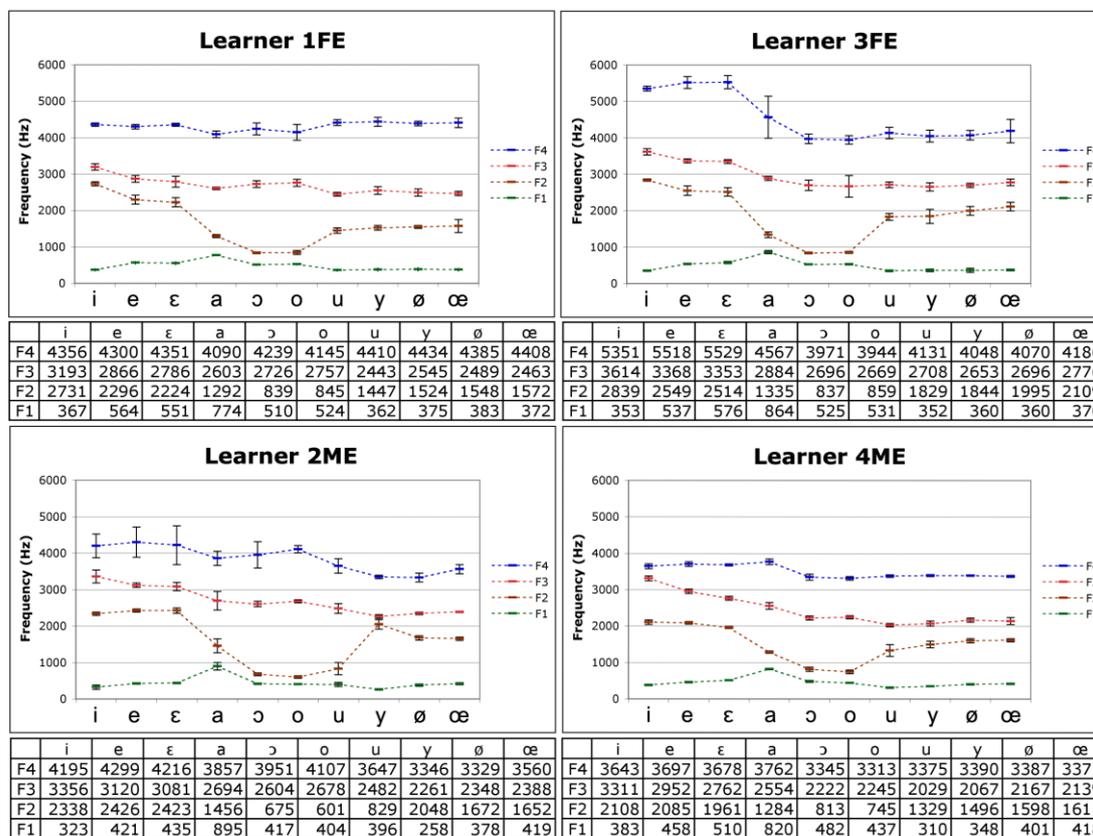
Two subgroups of JSL of FFL participated in the experiment. The first consisted of 4 students (2 women and 2 men) enrolled at Tokyo University of Foreign Studies. They had all lived in the Greater Tokyo area (Kanto) for at least 3 or 4 years, but originally came from various regions, including one participant from Tokyo, one from Kagawa, located across the Seto Inland Sea from Kansai. Two of the participants had started learning French at 17-18 years old and the others earlier (13 and 15 years old), which makes 3 to 9 years of learning experience. The second subgroup was composed of 4 second- and third-year students (2 women and 2 men) at Kobe University (Hyogo Prefecture), approximately 30 km West of Osaka. They all grew up mainly in Kansai. They had all started learning French at university (approximately at 18 years old), which makes 1.5 to 2.5 years of learning experience. All 8 participants were in the age range of 18-24 years.

The carrier sentences were presented one by one in a semi-random order on a computer screen. The sentence list was repeated 4 times without a break. The participants were invited to separate the target vowel from the rest of the carrier sentence so as to avoid formant transitions due to the adjacent consonants as far as possible. A training session preceded the recording session. The recording took place in recording studios of Tokyo University of Foreign Studies and Kobe University (Faculty of Global Human Sciences: named "Faculty of Intercultural Studies" at the time of the recording) using a headset microphone. The learners' production was recorded at 44.1 kHz, 16 bits. The target vowels were annotated manually using Praat (Boersma & Weenink, 2007). Portions where F2 and higher formants are not clearly observable, or where irregular periods are found, were excluded. The formant values were extracted every 6.25 milliseconds and the mean value was taken for the first, second, and last thirds of the vowel portion. The automatic formant detection was checked and the parameters were modified when necessary.

3.2. Results

The formant values of the vowels produced by the 4 learners from Kanto are presented in Figure 9.
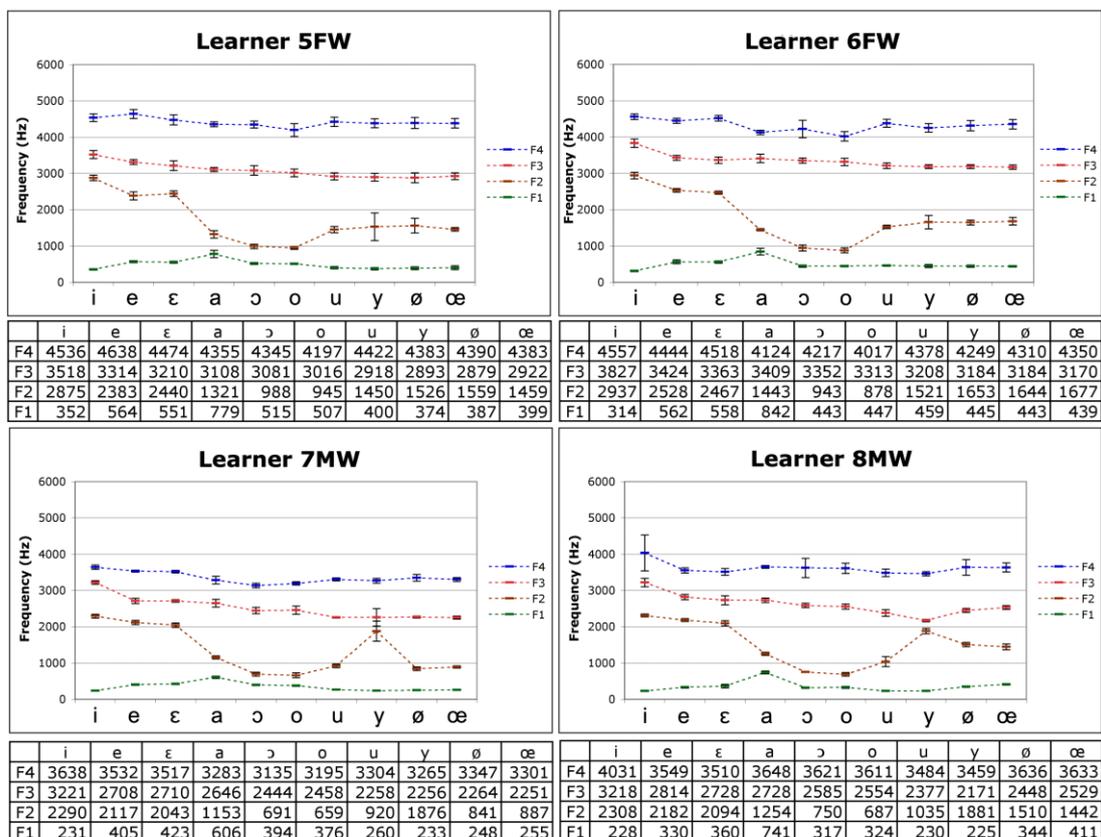
Figure 9: Learners from Kanto: mean F1, F2, F3 and F4 of the French oral vowels
(3 measures x 4 repetitions), by 4 learners (2 women above; 2 men below). /e/ produced 3 times
by the learner 2ME. Error bars: ± 1SD

Learner 1FE

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	4356	4300	4351	4090	4239	4145	4410	4434	4385	4408
F3	3193	2866	2786	2603	2726	2757	2443	2545	2489	2463
F2	2731	2296	2224	1292	839	845	1447	1524	1548	1572
F1	367	564	551	774	510	524	362	375	383	372

Learner 3FE

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	5351	5518	5529	4567	3971	3944	4131	4048	4070	4186
F3	3614	3368	3353	2884	2696	2669	2708	2653	2696	2776
F2	2839	2549	2514	1335	837	859	1829	1844	1995	2109
F1	353	537	576	864	525	531	352	360	360	370

Learner 2ME

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	4195	4299	4216	3857	3951	4107	3647	3346	3329	3560
F3	3356	3120	3081	2694	2604	2678	2482	2261	2348	2388
F2	2338	2426	2423	1456	675	601	829	2048	1672	1652
F1	323	421	435	895	417	404	396	258	378	419

Learner 4ME

	i	e	ε	a	ɔ	o	u	y	ø	œ
F4	3643	3697	3678	3762	3345	3313	3375	3390	3387	3371
F3	3311	2952	2762	2554	2222	2245	2029	2067	2167	2139
F2	2108	2085	1961	1284	813	745	1329	1496	1598	1611
F1	383	458	510	820	482	437	310	348	401	418

It was decided not to limit the graphic representation to rounded vowels but to include all oral vowels, for it is important to observe vowels as a whole system of contrasts. Three of the learners (1FE, 3FE and 4ME) show F2 values of /u/ higher than 1,400 Hz and far from F1. This tendency corroborates the observation made in previous studies on Kanto learners (Kamiyama & Vaissière, 2009). The only participant who produced the first two formants close together under 1,000 Hz (2ME) had started learning French earlier than the other participants, at the age of 13. His /y/ is also realized with close F2/F3 (Liénard, 1977, among others) with a small Standard Deviation (SD), which indicates that this vowel is not diphthongized either, unlike cases observed for some JSL (falling F2 due to an onglide [j], like in the Japanese sequence /ju/: Kamiyama & Vaissière, 2009).

Let us note that formant values are generally higher for women than men, but to a lesser extent for the formants essentially due to the Helmholtz resonance (F1 of /i/ /y/ /u/ and F2 of /u/ in French: Tubach, 1989).

Figure 10: Learners from Kansai: mean F1, F2, F3 and F4 of the French oral vowels
(3 measures x 4 repetitions), by 4 learners (2 women above; 2 men below). Error bars: ± 1SD

Learner 5FW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4536	4638	4474	4355	4345	4197	4422	4383	4390	4383
F3	3518	3314	3210	3108	3081	3016	2918	2893	2879	2922
F2	2875	2383	2440	1321	988	945	1450	1526	1559	1459
F1	352	564	551	779	515	507	400	374	387	399

Learner 6FW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4557	4444	4518	4124	4217	4017	4378	4249	4310	4350
F3	3827	3424	3363	3409	3352	3313	3208	3184	3184	3170
F2	2937	2528	2467	1443	943	878	1521	1653	1644	1677
F1	314	562	558	842	443	447	459	445	443	439

Learner 7MW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	3638	3532	3517	3283	3135	3195	3304	3265	3347	3301
F3	3221	2708	2710	2646	2444	2458	2258	2256	2264	2251
F2	2290	2117	2043	1153	691	659	920	1876	841	887
F1	231	405	423	606	394	376	260	233	248	255

Learner 8MW

	i	e	ɛ	a	ɔ	o	u	y	ø	œ
F4	4031	3549	3510	3648	3621	3611	3484	3459	3636	3633
F3	3218	2814	2728	2728	2585	2554	2377	2171	2448	2529
F2	2308	2182	2094	1254	750	687	1035	1881	1510	1442
F1	228	330	360	741	317	324	230	225	344	411

The results of the Kansai learners are presented in Figure 10. Concerning /u/, the two women (5FW and 6FW) show F2 around 1,500 Hz like the learners from Kanto. By contrast, the two men (7MW and 8MW) present F2 around 1,000 Hz (920 Hz for 7MW and 1,035 Hz for 8MW; still significantly higher than the native speakers of French reported in Kamiyama & Vaissière, 2009: Wilcoxon rank sum test, $p < .0001$), even if the F1-F2 distance is relatively large compared to native speakers: 660 Hz for 7MW, 805 Hz for 8MW, against 342 Hz and 333 Hz for two male native speakers in Kamiyama & Vaissière (2009), the difference being significant (Wilcoxon rank sum test, $p < .0001$).

In addition to the phonetic realization of /u/, the distinction of this vowel from other target vowels was also examined by means of Euclidean distance. This measure was already applied to the data of French rounded vowels pronounced by JSL in order to estimate the degree of dispersion depending on the speaking tasks adopted (Marushima *et al.*, 2010: calculation based on the first two formants). In the present study, the Euclidean distance was obtained from the first four formants in Bark scale ($[26.81/(1+1960/f)]-0.53$: Traunmüller, 1990) so that the perceptual aspect would be taken into account [1]:

[1] Euclidean distance between vowels A and B

$$= \sqrt{\left(F1_{(A)} - F1_{(B)}\right)^2 + \left(F2_{(A)} - F2_{(B)}\right)^2 + \left(F3_{(A)} - F3_{(B)}\right)^2 + \left(F4_{(A)} - F4_{(B)}\right)^2}$$

Figure 11 represents the Euclidean distances between /u/-/o/, /u/-/y/, /y/-/ø/ et /u/-/ø/, in comparison with the values observed for native speakers of French who produced vowels in isolation in a similar carrier sentence (Kamiyama & Vaissière, 2009). The learners' (from Kanto and Kansai) formant values distinguish /u/ and /o/ as clearly as native speakers, which is not always the case with the other contrasts. It was noted earlier that the learner 7MW from Kansai pronounced /u/ with F2 lower than 1,000 Hz, but the small Euclidean distance between /u/-/ø/ seen in this figure shows that these two target vowels were produced in a similar manner.

Figure 11: Euclidean distance (based on mean F1, F2, F3 and F4 in Bark scale) between /o/-/u/, /u/-/y/, /y/-/ø/ and /u/-/ø/: women (left) and men (right): learners from Kanto (FE, ME) and Kansai (FW, MW); native speakers of French (fr) in Kamiyama & Vaissière (2009). 4 repetitions for learners, 3 for native speakers.
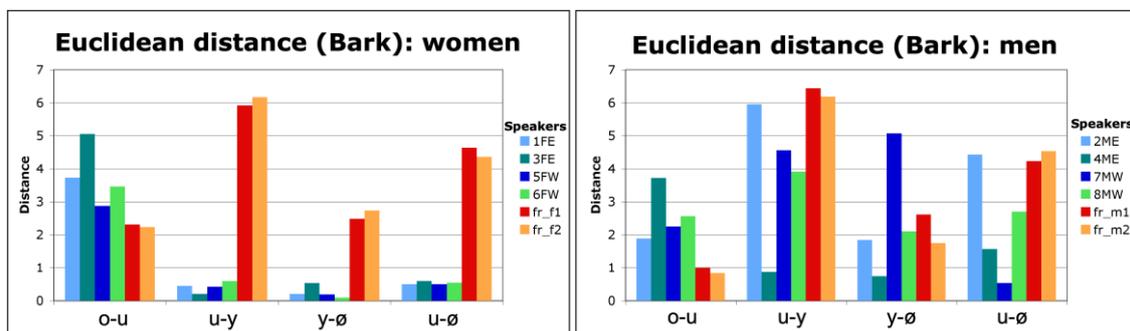
3.3. Discussion

The results of the series of experiments presented in the previous sections lead us to reconsider the hypotheses suggested earlier.

For Experiment 1, it was hypothesized that Kansai-Japanese speakers do indeed produce /u/ with lower F2 than Kanto speakers in their L1 Japanese. This hypothesis was partly adopted: for 5 of the 11 speakers, mean F2 ranged between 1,000 and 1,100 Hz, and an another showed a high intra-speaker variability: between 1,150 and 1,200 Hz in the first two repetitions, and then below 1,000 Hz in the three other repetitions. In examining the inter- and intra-speaker variability observed in this local accent not considered as a national standard, sociolinguistic factors need to be taken into account. The participants were university students and the recording took place in

their university, with an investigator coming from Tokyo, whom they met for the first time. All of these factors might have favoured a more formal speaking style than, for example, that used in an informal and casual speech with family members or childhood friends. Different tasks and settings could elicit context-dependent variability, as designed for the PFC (*Phonologie du français contemporain*: Phonology of Contemporary French) protocol, with word lists, text reading, formal and informal conversations (Detey, Durand *et al.*, 2016). The self-assessed degree of the use of *kyoutsûgo* (common language) showed a slight tendency to favour higher F2 for those who use it more frequently: other sociolinguistic factors such as the age of the speakers, the gender, the socio-economic status, the interlocutor (e.g. peer speaking in the same accent variety or external investigator using a stardardized variety) and the setting of the data elicitation (speaking task) might help clarify further the tendency observed in the present dataset.

The current study and the previous ones cited refer to articulation but only indirectly or with simple observations: to the best of our knowledge, there is no articulatory data available on the production of Japanese vowels in other varieties than Tokyo (Kanto) Japanese (e.g. Kokuristu Kokugo Kenkyûjo, 1990), except for the data from 5 Kansai speakers in the recently published Real-time MRI Articulatory Database (Version 1)[4] presented in Maekawa (2023). Amongst the 5 Kansai speakers, born between 1952 and 1970, who produced the sequence /uR/ (long /u/) in this database, 4 of them present relatively low F2 (< 1,100 Hz), corroborating previous descriptions and our findings. Furthermore, the female speaker born in 1952 produced F2 values as low as those of French native speakers' /u/ (c. 600 Hz). A qualitative observation of their tongue position and the lip configuration seems to reveal the following: 3 speakers (5 tokens) out of 5 (7 tokens) from Kansai are characterized by a considerable degree of lip protrusion, whereas only 3 "standard" speakers (5 tokens) out of 15 (21 tokens) showed the same articulatory configuration; the tongue position is considerably back for 3 speakers (4 tokens) out of 5 (7 tokens) from Kansai and is approximately central for the other 2 speakers (3 tokens), while it is considerably back for 1 "standard" speaker (1 token) only out of 15 (21 tokens), relatively back for 7 speakers (9 tokens), and approximately central for the other 7 speakers (10 tokens); all 5 Kansai speakers tend to share narrower constriction areas both for the lips and the tongue compared to Kanto speakers, contributing to lower both F1 and F2. Quantitative analyses of these articulatory data would provide further insights into the production of the Kansai /u/, among other vowels, in comparison with Kanto speakers.

---

[4] https://rtmridb.ninjal.ac.jp

The AXB auditory discrimination test in Experiment 2 revealed that the order of difficulty for the contrasts /u/-/ø/ (the most difficult), /y/-/ø/ and /u/-/y/ (least difficult) was the same for non-learners of French from Kansai as for those from Kanto and learners from Kanto. This finding does not support Hypothesis 2: Kansai-Japanese listeners distinguish the French /u/ perceptually from other vowels sounding similar to Japanese listeners more easily than Kanto listeners, thanks to their exposure to the Kansai /u/, which is acoustically closer to the French /u/ than the Kanto (Tokyo) Japanese /u/ is. Even though the Kansai listeners showed a slightly higher percentage correct for the contrasts /u/-/y/ and /y/-/ø/ than the Kanto listeners, a more comparable dataset will be needed to estimate the impact of the differences observed.

Compared with the previous cross-language studies on classification to L1 categories (Morrison, 2008; Chládková & Podlipský, 2011), it turns out that the vowel contrast /u/-/ø/ is not facilitated by a different dialectal background in the auditory discrimination task. As mentioned earlier, this contrast could be considered as a case of Single-Category (CS) in PAM (Best, 1995). In this model, Category-Goodness (CG) pattern, in which one of the L2 phonemes corresponds to a better token of an L1 category than the other L2 phoneme, perceived as a less good token of the same L1 category, predicts a better discrimination than in CS. If the French /ø/ were perceived as a less good token of the Japanese /u/ than the French /u/ were by Kansai listeners, unlike Kanto listeners, then the discrimination of the contrast /u/-/ø/ would be better for the former group than the latter. The fact that it is not the case suggests not only that there is no phonetic category (allophone) for the Kansai-type [u] with lower F2, apart from the Kanto-type [u] with higher F2, but also that these two types form a continuum with neither of them as a prototype, probably due to the intra- and inter-speaker variability, which, together with exposure to the Kanto variety through media and communication with speakers from Kanto and other regions, makes the Kanto-type [u] a frequently observed type of token.

It was shown in Experiment 3 that the two male learners from Kansai produced the French /u/ with F2 around 1,000 Hz (one of them lower than 1,000 Hz), but with a larger F1-F2 distance than the native speakers, and that the participant who produced F2 lower than 1,000 Hz also produced /ø/ in a similar manner. Let us note that the small distance between F1 and F2 characterizes the French /u/ (Gendrot *et al.*, 2008) and that the higher formants, including F3, of the French back vowels /u/ and /o/ are not perceptible (F'2 close to F2: Vaissière, 2011). This

result suggests that the phonetic realization of /u/ in Kansai Japanese facilitates the phonetic acquisition of the French /u/ to some extent, even if the first two formants are not as close as in native speakers' tokens. An auditory assessment test with native-speaking listeners from two dialectal backgrounds (RF speakers from Paris and speakers from Québec) shows that the RF listeners examined perceived those stimuli of /u/ (produced by JSL from Kanto) with F2 between 1,000 and 1,100 Hz as /ø/ and /u/ almost equally often, but considered as very poor exemplars of either of them, while the Québec French listeners tested identified the same stimuli as /u/ in the majority of the cases, with a better goodness rating than RF listeners (Tremblay & Kamiyama, 2009). This result suggests that the tokens produced by the two male learners from Kansai are also accepted quite well as /u/ by Québec listeners, since their F2 is located between 1,000 and 1,100 Hz. A more systematic study including Kansai learners would clarify the perceptual impact of this type of phonetic realization on native-speaking listeners' assessment. By contrast, the phonemic acquisition of the production of /u/, namely learning to pronounce it distinctly from neighbouring vowels, especially /ø/, is not necessarily facilitated, as shown by the Euclidean distance observed for the vowel contrast /u/-/ø/.

Considering these findings, Hypothesis 3, stating that Kansai-Japanese speakers learning French as a foreign language learn to produce more native-like tokens of the French /u/ than learners from Kanto, is accepted literally, but cannot be extended to phonemic acquisition of the vowel phoneme: the phonemic contrast between /u/ and other vowels, especially /ø/, is still part of the difficulties. A further study with a larger number of participants would enable an observation of more general tendencies.

The present series of experiments examined the vowels in isolation, but it will be also necessary to study different consonantal and prosodic contexts in further studies. Indeed, Gendrot and Adda-Decker's (2005) data present higher F2 for the French /u/ (1,153 Hz for women), which is probably due to the effect of coarticulation in continuous speech, where vowels are located in various consonantal and prosodic contexts. Likewise, F2 of /u/ in (Kanto - Tokyo) Japanese is also variable in various consonantal contexts, as shown in the data presented by Mokhtari and Tanaka (2000), cited earlier in Figure 5.

Conclusion

This paper reported a series of studies on the possible impact of the L1 varieties, namely, the Japanese varieties of Kansai and Kanto, on the acquisition of the French high back vowel /u/ in contrast with some neighbouring vowels. In spite of intra- and inter-speaker variability, some speakers from Kansai produce the Japanese /u/ with lower F2 than typical Kanto Japanese values, but this tendency does not seem to help them to better distinguish the French /u/-/ø/ perceptually. When learners produce the French vowel /u/, the Kansai-type [u] with lower F2 may facilitate the phonetic realization of the target vowel, but it does not necessarily mean that the phonemic contrast /u/-/ø/ is also acquired. These findings suggest that being aware of the regional or individual differences of learners may be useful in teaching foreign or second language pronunciation.

References

BEST Catherine T., 1995, "A direct realist view of cross-language speech perception", *in* Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore, York Press, p. 171-204.

BEST Catherine T. & TYLER Michael D., 2007, "Nonnative and second-language speechperception: Commonalities and complementarities", *in* Murray J. Munro & Ocke-Schwen Bohn (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*, Amsterdam, John Benjamins, p. 13-34.

BOERSMA Paul & WEENINK David, 2007, *Praat: doing phonetics by computer* (Version 4.6.13) [Computer program]. Retrieved in August 2007, from http://www.praat.org/.

BOTHOREL André, SIMON Péla, WIOLAND François & ZERLING Jean-Pierre, 1986, *Cinéradiographie des voyelles et consonnes du français*, Strasbourg, Institut de Phonétique de Strasbourg.

CHLÁDKOVÁ Kateřina & PODLIPSKÝ Václav Jonáš, 2011, "Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners", *The Journal of the Acoustical Society of America*, vol. 130(4), p. EL186–EL192.

DETEY Sylvain, DURAND Jacques, LAKS Bernard & CHANTAL Chantal, 2016, "The PFC programme and its methodological framework", *in* Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Varieties of Spoken French*, Oxford, Oxford University Press, p. 13-23.

DETEY Sylvain, LYCHE Chantal, RACINE Isabelle, SCHWAB Sandra & LE GAC David, 2016, "The notion of norm in spoken French: Production and perception", *in* Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (eds.), *Varieties of Spoken French*, Oxford, Oxford University Press, p. 55-67.

FERRAGNE Emmanuel, FLAVIER Sébastien & FRESSARD Christian, 2012, *ROCme!* (Version 2.0) [Computer software], retrieved June 4, 2012, Available from www.ddl.cnrs.fr/rocme.

FLEGE James Emil & BOHN Ocke-Schwen, 2021, "The Revised Speech Learning Model (SLM-r)", *in* Ratree Wayland (ed.), *Second Language Speech Learning: Theoretical and Empirical Progress*, Cambridge, Cambridge University Press, p. 3-83. doi.org/10.1017/9781108886901.002.

GENDROT Cédric & ADDA-DECKER Martine, 2005, "Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German", *in Proceeding of Interspeech 2005*, Lisbon, Portugal, p. 2453-2456.

GENDROT Cédric, ADDA-DECKER Martine & VAISSIÈRE Jacqueline, 2008, « Les voyelles /i/ et /y/ du français : focalisation et variations formantiques », *in Actes des 27èmes Journées d'Études sur la Parole*, Avignon, p. 205-208.

KAMIYAMA Takeki, 2011, "Pronunciation of French vowels by Japanese speakers learning French as a foreign language: back and front rounded vowels /u y ø/", *Phonological Studies* (Phonological Society of Japan), vol. 14, p. 97-108.

KAMIYAMA Takeki, 2012, « Production des voyelles du français par des apprenants japonophones : effet du dialecte d'origine », *in* Laurent Besacier, Benjamin Lecouteux & Gilles Sérasset (dir.), *Actes des Journées d'Études sur la Parole 2012*, Grenoble, p. 771–778.

KAMIYAMA Takeki, KANEDA Jumpei & HAYASHI Ryôko, 2013, « Perception et production des voyelles orales du français par des japonophones avec des dialectes différents », *19ème Congrès International des linguistes / 19th International Congress of Linguists* (juillet 2013), Genève, Suisse.

KAMIYAMA Takeki & VAISSIÈRE Jacqueline, 2009, "Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners", *AILE … LIA*, n° 2, p. 9-41.

KOCJANČIČ ANTOLÍK Tanja, PILLOT-LOISEAU Claire & KAMIYAM Takeki, 2019, "The effectiveness of real-time ultrasound visual feedback on tongue movements in L2 pronunciation training", *Journal of Second Language Pronunciation*, vol. 5(1), p. 72-97.

KOKURITSU KOKUGO KENKYÛJO [National Institute for Japanese Language and Linguistics] (UEMURA Yukio & TAKADA Shôji), 1990, *Nihongo no boin, shiin, onsetsu: chouon undou no jikken-onseigakuteki kenkyû* [Vowels, consonants and syllables in Japanese: an experimental phonetic study on articulatory gestures], Tokyo, Shûei shuppan.

LABRUNE Laurence, 2012, *The Phonology of Japanese*, Oxford, Oxford University Press.

LADO Robert, 1964, *Language Teaching: A Scientific Approach*, New York, McGraw-Hill.

LANDRON Simon, PAILLEREAU Nikola, NAWAFLEH Ahmad, EXARE Christelle, ANDO Hirofumi & GAO Jiayin, 2010, "The PhoDiFLE Corpus: a Joint Corpus of French Speech, for a Phonetic Study of L1-L2 Contrasts", *Les cahiers de praxématique,* vol. 54-55, p. 73-86.

LIÉNARD Jean-Sylvain, 1977, *Les processus de la communication parlée : introduction à l'analyse et la synthèse de la parole* (Vols. 1-1), Paris ; New York ; Barcelone [etc.], Masson.

MAEDA Shinji, 1982, "A digital simulation method of the vocal tract system", *Speech Communication* vol. 1(3-4), p. 199-229.

MAEKAWA Kikuo, 2023, "Production of the utterance-final moraic nasal in Japanese: A real-time MRI study", *Journal of the International Phonetic Association*, n° 53(1), p. 189-212.

MARUSHIMA Naoki, DETEY Sylvain & KAWAGUCHI Yuiji, 2010, « Caractéristiques phonétiques des voyelles orales arrondies du français chez des apprenants japonophones », *Flambeau* (Revue annuelle de la section française, Université des Langues Étrangères de Tokyo) vol. 36, p 53–72.

MOKHTARI Parham & TANAKA Kazuyo, 2000, "A Corpus of Japanese Vowel Formant Patterns", *Bulletin of Electrotechnical Laboratory*, n° 64 (special issue), p. 57-66.

MORRISON Geoffrey Stewart, 2008, "Perception of synthetic vowels by monolingual Canadian-English, Mexican-Spanish, and Peninsular-Spanish listeners", *Canadian Acoustics*, n° 36(4), p. 17-23.

POLIVANOV Evgenij, 1931, « La perception des sons d'une langue étrangère », *Travaux du Cercle linguistique de Prague*, vol. 4, p. 79-96.

SCHWARTZ Jean-Luc, BOË Louis-Jean, VALLÉE Nathalie & ABRY Christian, 1997, "The Dispersion-Focalization Theory of vowel systems", *Journal of Phonetics,* vol. 25(3), p. 255-286.

SHIBATANI Masayoshi, 1990, *The Languages of Japan*, Cambridge [England], Cambridge University Press.

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

TAKEBAYASHI Shigeru, 1996, *Eigo onseigaku* [English Phonetics], Tokyo, Kenkyûsha.

TRAUNMÜLLER Hartmut, 1990, "Analytical expressions for the tonotopic sensory scale", *Journal of the Acoustical Society of America,* vol. 88(1), p. 97-100.

TREMBLAY Marie-Claude & KAMIYAMA Takeki, 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. Cross-Language Speech Perception and Variations in Linguistic Experience" (2nd ASA Special Workshop on Speech, May 2009, Portland, Oregon, United States), *Journal of the Acoustical Society of America,* vol. 125(4), p. 2776.

TRUBETZKOY Nikolai S., [1939] 1969, *Principles of Phonology* (*Grundzüge der Phonologie,* translated by Christiane A. M. Baltaxe), Berkeley/Los Angeles, University of California Press.

TUBACH Jean-Pierre (dir.) (CALLIOPE), 1989, *La parole et son traitement automatique*, Paris, Masson.

VAISSIÈRE Jacqueline, 2006, *La phonétique*, Paris, Presses Universitaires de France.

VAISSIÈRE Jacqueline, 2007, "Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages", *in* Maria-Josep Solé, Patrice Speeter Beddor & Manjari Ohala (eds.), *Experimental Approaches to Phonology*, Oxford, Oxford University Press, p. 54-71.

VAISSIÈRE Jacqueline, 2011, "On the acoustic and perceptual characterization of reference vowels in a cross-language perspective", *in* Wai Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, City University of Hong Kong, p 52-59.

VANCE Timothy J., 2008, *The Sounds of Japanese,* Cambridge, Cambridge University Press.

WEINREICH Uriel, [1953] 1968, *Languages in Contact,* The Hague, Mouton.