

Formant Analysis of the English Alphabet Read by Japanese Female Students and Native Speakers

Giido Izuta

*Yamagata Prefectural Yonezawa Women's Junior College
6-15-1 Toori Machi, Yonezawa City, Yamagata, 992-0025 Japan*

Abstract—In order to understand and characterize the sounds of the letters of the alphabet as well as to figure out whether the students show some sort of affinity with spoken English of a particular region/country, the utterances of three groups (26 Japanese female students, 10 British and 20 American women) were recorded, analyzed, and statistically compared. It turned out that only a few letters had both F2 and F1 statistically similar to those of the native speakers; and preference to a specific English sound pattern was not statistically established.

Keywords—Phonetics, applied Linguistics, analysis of formants, characteristics of English utterances by Japanese female students.

I. INTRODUCTION

LINGUISTICS issues concerned with the difficulties that Japanese students face when learning English language have been reported (see for example [1]-[3]) and many theoretical as well as practical approaches have been taken into account to deal with this problem [4]-[5]. Nevertheless representative frameworks as the contrastive analysis have shed light on the differences between the sounds made by the native speakers and Japanese learners, teachers have struggled to find out an efficient methodology to teach the Japanese students pronunciation [6]-[7].

Motivated by this scientific background, this work has attempted to contribute to the field by handling this matter from the phonetic analysis standpoint. In fact, taking for granted some previous pilot investigations [8], it delved into the analysis of the formants composing the sounds of the English alphabet made by the Japanese students and the native speakers. More specifically, the aims of this study are two-folded: to characterize the formants of the sounds accordingly

to the category of the letters, and to find out which kind of spoken English, whether Received Pronunciation or Standard American English, is easier for Japanese students to utter.

II. EXPERIMENTAL PROCEDURE

A. Subjects and Data Recording

Twenty six young Japanese female college students, aged 19 to 20 years old, were randomly recruited to participate in this experiment. The subjects were asked to recite (not the phonics) the English alphabet. The utterances were recorded with the freeware “sound engine” running on a personal computer equipped with “Windows 8”. In addition, sounds of ten female speakers of Received Pronunciation (RP) and twenty female individuals of Standard American English (SAE) were acquired by means of street interviewing sessions. However, 3 out of 10 for RP, and 6 out of 20 for SAE were carefully scrutinized and their sounds downloaded through the Internet. All the native speakers were in age range between late 20s and 30s and reportedly college graduated healthy native speakers of English.

B. Data Processing

The digital sounds were pre-processed for noise filtering and analyzed with freeware “Praat”, which also generated the text format files with numerical values of the voice signals. These text files were imported into the software “Microsoft Excel” in order to carry out the statistical data processing.

III. RESULTS

In what follows, a graph depicting the values of formants F1 and F2 and two other graphs of vertical bars representing the

statistical testing of the formants are presented for each group of letters. It is worth noting that, hereafter, this paper is concerned only with the tests on the pairs JP and US, and JP and UK.

Fig.1 shows that the group of Japanese students had relatively higher F2 frequencies for nearly almost all the letters in the category [i:/i]. In fact, F2 frequencies had values greater than 2400 Hz whereas the US frequencies were roughly in-between the range from 2200 through 2500 Hz, and UK in the range from 2000 to 2200 Hz. As for the F1 frequencies, contrary to the clustery gathering seen in the US, which located in the interval from 450 to 650 Hz, and UK, in the range from 650 to 800 Hz, the group of students scattered widely from 500 to 800 Hz. Looking closely to each letter and statistically comparing them across the groups, the results turned out to be as shown in figs. 2 and 3. Fig.2 indicates that the students had mean values of F2s - mean and standard deviation of letter b: 2985 ± 125 Hz, c: 2133 ± 60 Hz, d: 2100 ± 120 Hz, e: 2177 ± 144 Hz, g: 2125 ± 105 Hz, p: 2127 ± 102 Hz, t: 2116 ± 125 Hz, and v: 2060 ± 75 Hz - statistically different from those of the group of RP speakers (b: 2467 ± 184 Hz, c: 2492 ± 158 Hz, d: 2504 ± 210 Hz, e: 2610 ± 174 Hz, g: 2512 ± 207 Hz, p: 2540 ± 182 Hz, t: 2511 ± 163 Hz, and v: 2411 ± 231 Hz) . On the other hand, comparing with US (b: 2271 ± 260 Hz, c: 2251 ± 223 Hz, d: 2267 ± 372 Hz, e: 2275 ± 336 Hz, g: 2465 ± 186 Hz, p: 2322 ± 341 Hz, t: 2351 ± 302 Hz, and v: 2220 ± 332 Hz), statistical similarity was verified for letters 'c', 'd', 'e', 'p' and 'v'. For the F1s in Fig. 3, JP (b: 678 ± 82 Hz, c: 789 ± 59 Hz, d: 696 ± 79 Hz, e: 659 ± 77 Hz, g: 709 ± 56 Hz, p: 693 ± 60 Hz, t: 719 ± 78 Hz, v: 714 ± 78 Hz.) was different from UK (b: 501 ± 116 Hz, c: 781 ± 198 Hz, d: 556 ± 163 Hz, e: 562 ± 192 Hz, g: 668 ± 284 Hz, p: 653 ± 186 Hz, t: 623 ± 196 Hz, v: 582 ± 205 Hz.) only for 'b', and different from US (b: 484 ± 84 Hz, c: 663 ± 93 Hz, d: 510 ± 72 Hz, e: 487 ± 84 Hz, g: 652 ± 93 Hz, p: 589 ± 125 Hz, t: 605 ± 77 Hz, v: 502 ± 124 Hz.) for all letters, but 'p'.

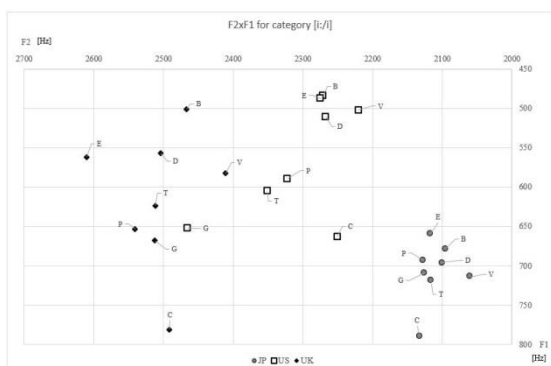


Fig. 1 Graph of the formants for category [i:/i] (letters b, c, d, e, g, p, t, and v.) ●: Japanese students. □: native speakers of American English. ◆: native speakers of RP.

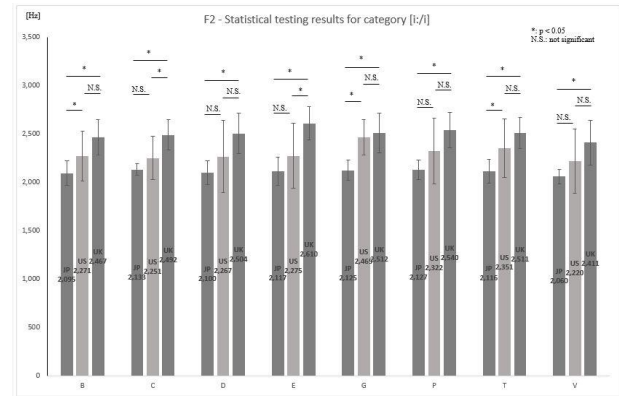


Fig. 2 Statistical testing of the formants F2s for category [i:/i] (letters b, c, d, e, g, p, t, and v.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

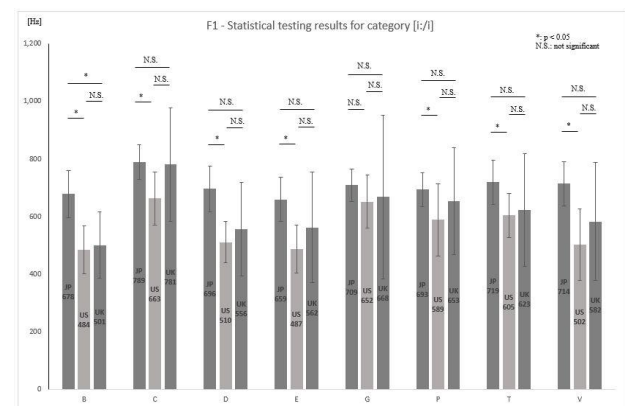


Fig. 3 Statistical testing of the formants F1s for category [i:/i] (letters b, c, d, e, g, p, t, and v.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

Category [e] had letters with F2 and F1 distributed as plotted in Fig. 4, in which JP had F2 from 1900 to 2100 Hz, and F1 from 750 to 1000 Hz whereas US varied from 1550 to 2100 Hz for F2 and from 600 to 1000 Hz for F1; and UK from 1600 to 2250 Hz for F2 and from 700 to 1250 Hz for F1. In fact, as for F2, the means and standard deviations read: f: 2033 ± 56 Hz, l: 1944 ± 133 Hz, m: 1992 ± 99 Hz, n: 1996 ± 82 Hz, s: 2042 ± 64 Hz, and x: 2032 ± 51 Hz for JP; f: 2171 ± 179 Hz, l: 1626 ± 127 Hz, m: 1880 ± 152 Hz, n: 1988 ± 197 Hz, s: 2219 ± 201 Hz, x: 2208 ± 222 Hz for UK; and f: 1957 ± 98 Hz, l: 1568 ± 107 Hz, m: 1779 ± 115 Hz, n: 1853 ± 170 Hz, s: 2006 ± 177 Hz, x: 2054 ± 96 Hz for US. Fig. 5 says that JP and UK were statistically related for 'f', 'n', 's' and 'x', and JP and US were not different for 's', and 'x'. As shown in Fig. 6, the same result was obtained for F1 when JP (f: 974 ± 112 Hz, l: 804 ± 51 Hz, m: 820 ± 98 Hz, n: 789 ± 81 Hz, s: 958 ± 89 Hz, and x: 981 ± 68 Hz) and UK (f: 1157 ± 220 Hz, l: 795 ± 87 Hz, m: 681 ± 96 Hz, n: 749 ± 131 Hz, s: $1174 \pm$

245 Hz, x: 1206 ± 249 Hz) were compared with each other. Slightly differently though, JP and US (f: 934 ± 98 Hz, l: 741 ± 65 Hz, m: 648 ± 97 Hz, n: 623 ± 101 Hz, s: 921 ± 133 Hz, and x: 955 ± 99 Hz) were similar for 'f', 's', and 'x'.

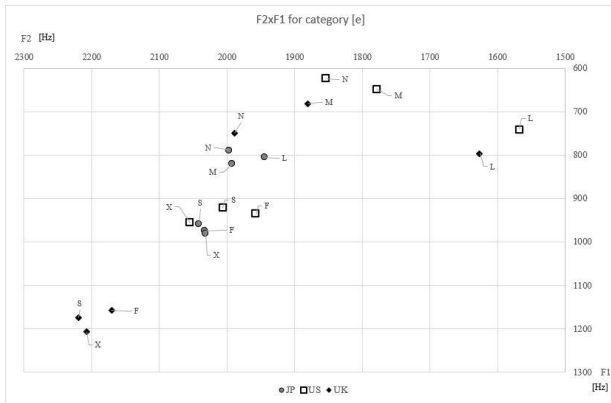


Fig. 4 Graph of the formants for category [e] (letters f, l, m, n, s and x.) ●: Japanese students. □: native speakers of American English. ◆: native speakers of RP.

speakers of RP. *: p<0.05, two tails. N.S.: not significant.

Fig.7 gives that the values of F2 for JP within the range from 2050 to 2150 Hz were smaller than the other groups with US (in the interval from 2150 to 2350 Hz) mid-way between JP and UK, which ranged from 2350 to 2600 Hz; and as for F1s, ruling out the letter 'h' close to the lower left corner the points stayed in the range between 600 and 1000 Hz. Now statistically comparing these points, Fig. 8 reveals that JP (a: 2058 ± 106 Hz, h: 2094 ± 71 Hz, j: 2091 ± 75 Hz, and k: 2104 ± 59 Hz) and UK (a: 2341 ± 135 Hz, h: 2544 ± 119 Hz, j: 2350 ± 197 Hz, and k: 2375 ± 146 Hz) were statistically different as far as F2s are concerned; and as for JP and US (a: 2161 ± 196 Hz, h: 2356 ± 168 Hz, j: 2254 ± 159 Hz, and k: 2246 ± 172 Hz), they were similar for only 'a'. Focusing on the comparison results of F1s given in Fig. 9, JP (a: 766 ± 50 Hz, h: 903 ± 57 Hz, j: 791 ± 48 Hz, and k: 785 ± 59 Hz) and UK (a: 632 ± 69 Hz, h: 1182 ± 186 Hz, j: 78 ± 291 Hz, and k: 784 ± 136 Hz) were similar for 'j' and 'k' whereas JP and US (a: 638 ± 81 Hz, h: 922 ± 168 Hz, j: 717 ± 169 Hz, and k: 676 ± 90 Hz) were not different for 'h' and 'j'

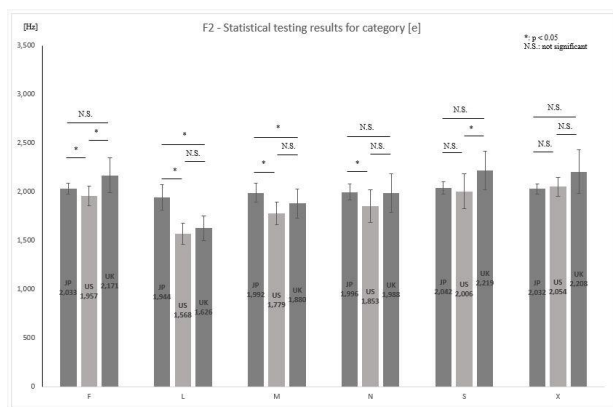


Fig. 5 Statistical testing of the formants F2s for category [e] (letters f, l, m, n, s and x.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: p<0.05, two tails. N.S.: not significant.

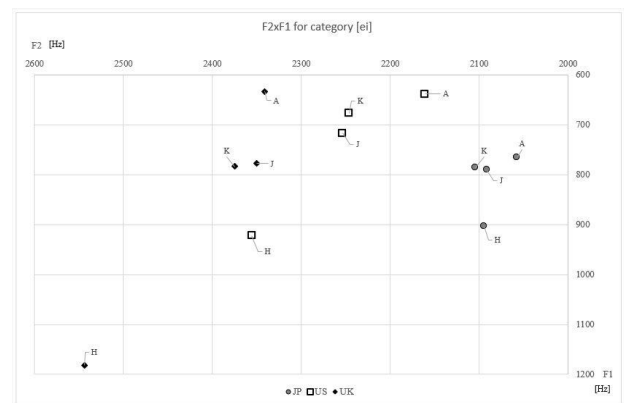


Fig. 7 Graph of the formants for category [ei] (letters a, h, j and k.) ●: Japanese students. □: native speakers of American English. ◆: native speakers of RP.

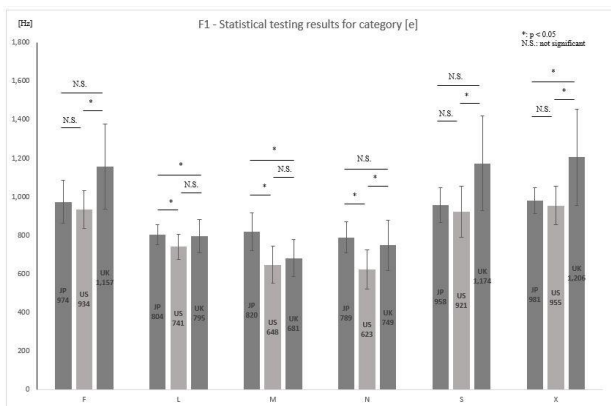


Fig. 6 Statistical testing of the formants F1s for category [e] (letters f, l, m, n and s.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native

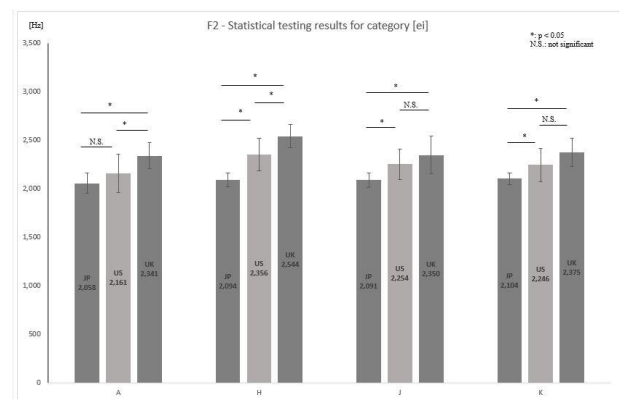


Fig. 8 Statistical testing of the formants F2s for category [ei] (letters a, h, j and k.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: p<0.05, two tails. N.S.: not significant.

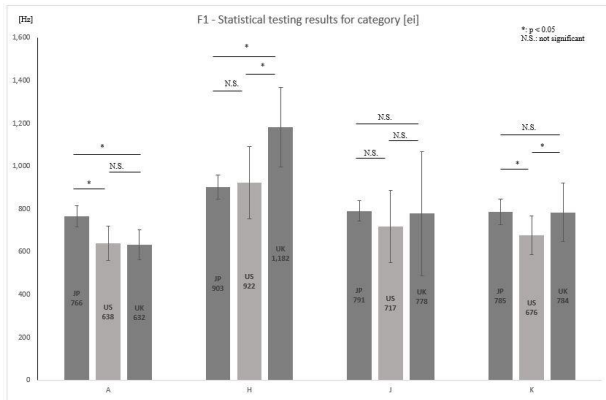


Fig. 9 Statistical testing of the formants F1s for category [ei] (letters a, h, j and k.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

The letters in the category [u:] are displayed in Fig. 10. The sounds of ‘q’s were in the scope defined along the F2 axis running from 1950 to 2100 Hz with US having the smallest values and UK the largest. Yet, this order was also verified for F1s which varied from 600 to 800 Hz. For ‘u’, unlike the “US, UK, and JP” frequency increasing order seen on the F2 axis which ranged from 1850 to 2000 Hz, the order along the F1 axis, which was within the interval from 400 to 650 Hz, was “UK, US, and JP”, from smallest to largest. Actually, apart from the frequency values, the same patterns were measured for ‘w’. Fig. 11 illustrates the results of the statistical testing for F2. JP (q: 1999 ± 82 Hz, u: 1985 ± 83 Hz, and w: 1923 ± 74 Hz) and UK (q: 2076 ± 202 Hz, u: 1968 ± 183 Hz, and w: 1844 ± 203 Hz) were positively correlated to each other for all the tests whereas JP and US (q: 1967 ± 154 Hz, u: 1887 ± 167 Hz, and w: 1734 ± 139 Hz) were similar for ‘q’ and ‘u’, but not for ‘w’. As far as F1 is concerned, Fig. 12 unveils that JP (q: 731 ± 59 Hz, u: 640 ± 56 Hz, and w: 684 ± 66 Hz) and UK (q: 785 ± 176 Hz, u: 433 ± 74 Hz, and w: 526 ± 121 Hz) were statistically similar for ‘q’. Actually, this was also true for JP and US (q: 648 ± 148 Hz, u: 500 ± 110 Hz, and w: 562 ± 64 Hz.)

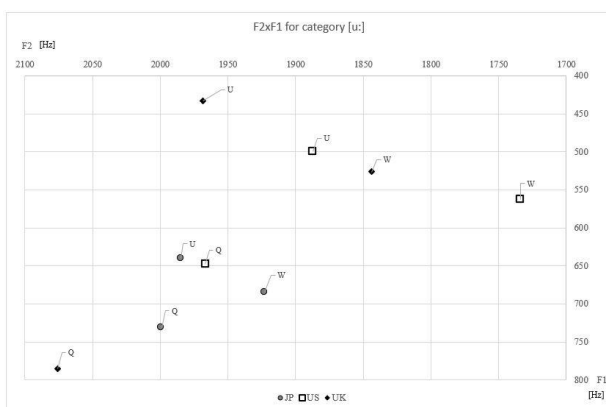


Fig. 10 Graph of the formants for category [u:] (letters q, u and w.) ●: Japanese students. □: native speakers of American

English. ◆: native speakers of RP.

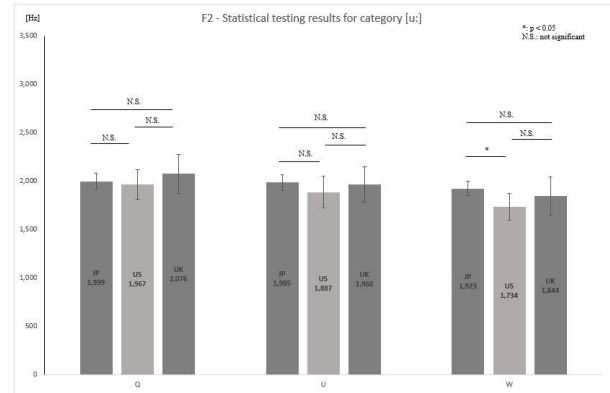


Fig. 11 Statistical testing of the formants F2s for category [u:] (letters q, u and w.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

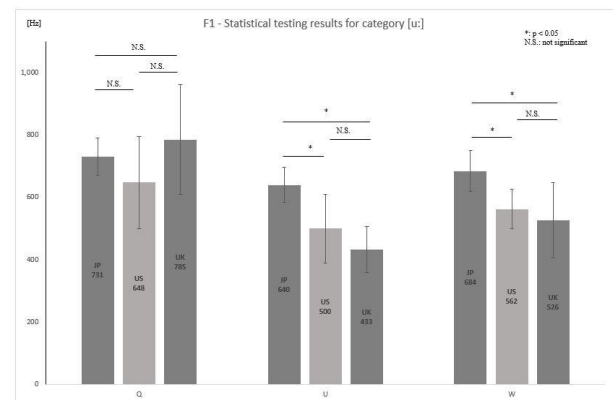


Fig. 12 Statistical testing of the formants F1s for category [u:] (letters q, u and w.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

Fig 13 describes three different categories. The letter ‘i’ were located in the lower left region of the graph with the points having Formant values very close to each other – F2 around 1850 Hz and F1 850 Hz. On the other hand, the points related to the letter ‘y’ were distributed in a region between 1600 and 1850 Hz for F2, and between 700 and 850 Hz for F1. Scattered were also the formants of ‘o’, which fitted in the range from 1300 to 1750 Hz along the horizontal axis and from 600 to 800 Hz for the vertical axis. As for ‘r’, they were clustered in a middle region in the lower part of the graph. Now, comparing F2s, Fig.14 exhibits that JP (i: 1875 ± 68 Hz, y: 1826 ± 69 Hz, o: 1557 ± 103 Hz, r: 1681 ± 98 Hz) and UK (i: 1805 ± 141 Hz, y: 1621 ± 120 Hz, o: 1739 ± 253 Hz, r: 1425 ± 185 Hz) were statistically related for ‘i’ and ‘o’ whereas JP and US (i: 1835 ± 187 Hz, y: 1697 ± 167 Hz, o: 1313 ± 212 Hz, r: 1542 ± 153 Hz) were only for ‘i’. For F1s as presented in Fig. 15, JP (i: 880 ± 50 Hz, y: 838 ± 60 Hz, o: 795 ± 49 Hz, r: 874 ± 56 Hz) and UK (i: 863 ± 185 Hz, y: 715 ± 94 Hz, o: 635 ± 141 Hz, r: 848 ± 117 Hz) were similar for ‘i’ and ‘r’ whereas JP and US (i: 836 ± 103 Hz, y: 730 ± 92 Hz,

o: 633 ± 104 Hz, r: 787 ± 99 Hz) were for only the letter 'i'.

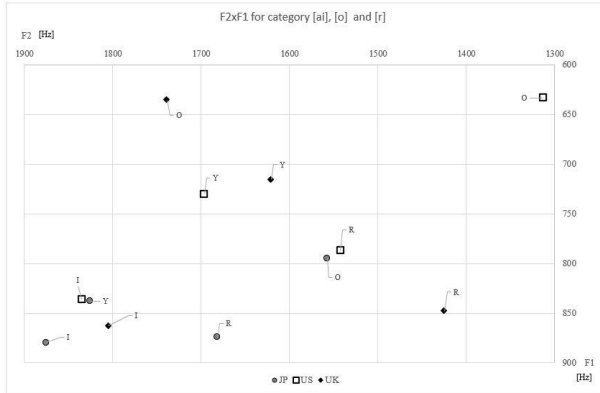


Fig. 13 Graph of the formants for categories [ai], [o] and [r] (letters i, y, o and r.) ●: Japanese students. □: native speakers of American English. ◆: native speakers of RP.

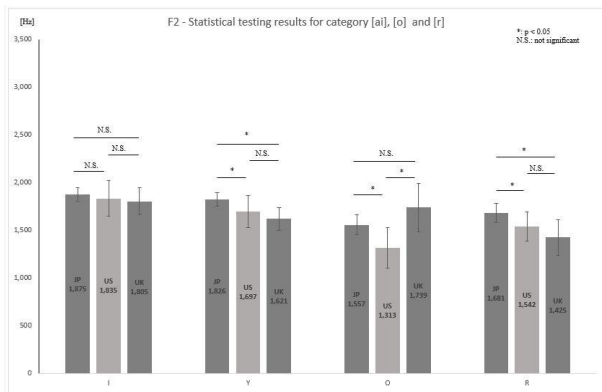


Fig. 14 Statistical testing of the formants F1s for categories [ai], [o] and [r] (letters i, y, o and r.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

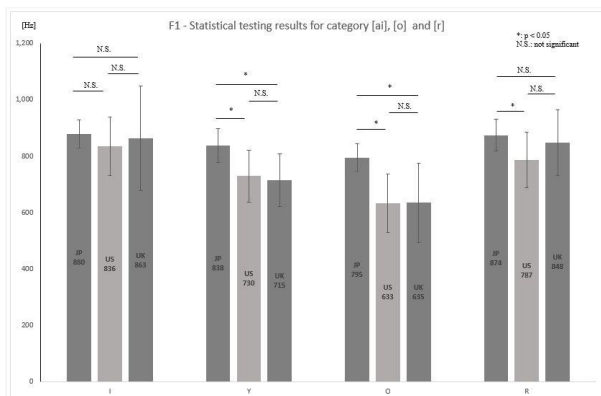


Fig. 15 Statistical testing of the formants F1s for categories [ai], [o] and [r] (letters i, y, o and r.) Left bar: Japanese students. Middle bar: native speakers of American English. Right bar: native speakers of RP. *: $p < 0.05$, two tails. N.S.: not significant.

IV. DISCUSSION AND CONCLUSION

The results lead to the following table which shows how the sounds made by Japanese students tested for the formants and to which English pronunciation they were more likely related.

TABLE I
AFFINITY OF THE SOUNDS MADE BY THE STUDENTS

	Both F2 and F1	Only F2	Only F1	neither F2 nor F1
UK-like	f, n, s, q, i	x, u, w, o	c, d, e, g, p, t, v, j, k, r	b, t, l, m, y
US-like	s, x, q, i	c, d, e, g, p, v, a, u	f, h, j	

It suggests that 'f', 'n', 's', 'q' and 'i' are positively related to UK for both F2 and F1 whereas 's', 'x', 'q' and 'i' are similar to US. Moreover, the sounds of 's', 'q' and 'i' correlates with both UK and US. The majority of the letters in the category [i:/i] fitted in US-<Only F2> and UK-<Only F1>; and since F2 is related to the <forward/backward> positioning of the tongue whereas F1 to the openness of the mouth, it is likely that these utterances were made with US-like tongue positioning while the openness tended to be UK-like. Interestingly, contrary to the view among the Japanese students that 'r' has a very peculiar sound in the sense that it is difficult to pronounce it, the formants were not in the classification <neither F2 nor F1>, which had 'b', 't', 'l', 'm' and 'y' in it. Thus, considering that the Japanese educational system includes a program which assigns native speakers of English language as "assistant language teachers (ALTs)" to all junior and high schools throughout the country, and that the great majority of the ALTs are from the USA, these results call on further investigations to understand the absence of a dominant "English" in the speaking of Japanese female students.

REFERENCES

- [1] T. Riney and J. Anderson-Hsieh, "Japanese Pronunciation of English," *JALT Journal*, vol. 15, no. 1, May 1993, pp. 21-36.
- [2] B. Smith, "Pronunciation patterns of Japanese learners and their implications or teaching," *Polyglossia*, vol. 23, October 2012, pp. 199-206.
- [3] K. Ohata, "Phonological Differences between Japanese and English: Several Potentially Problematic Areas of Pronunciation for Japanese ESL/EFL learners," *Asian EFL Journal*, vol. 6, issue. 4, paper 5, December 2004, pp. 1-19.
- [4] J. Noguchi, "Contrastive Analysis between Japanese and American English Sound Systems: From an Articulatory Setting Perspective," *The Journal of Kanda University of International Studies*, vol. 26, 2014, pp.293-309.
- [5] S. W. Carruthers, "Pronunciation Difficulties of Japanese Speakers of English: Predictions Based on a Contrastive Analysis," *Tesol Working Paper Series*, vol. 4, issue 2, fall 2006, pp. 17-23.
- [6] T. Nagamine, "Effects of Hyper-Pronunciation Training Method on Japanese University Students' Pronunciation," *Asian EFL Journal*, vol. 53, July 2011, pp.35-50.
- [7] W. Shudong, M. Higgins and Y. Shima, "Training English Pronunciation for Japanese Learners of English Online," *The JALT CALL Journal*, vol. 1, no. 1, 2005, pp. 39-47.
- [8] G. Izuta, "An assessment of the utterances of some English vowels spoken by Japanese Students," *Bulletin of Yonezawa Women's J. Coll.*, no. 49, 2013, pp. 63-82.