Development of Japanese length contrast:
A longitudinal study of L2 vowels produced by Australian learners of Japanese

Akiko Kato and Felicity Cox
SHLRC, Department of Linguistics
Macquarie University, Australia
Akiko.Kato@ling.mq.edu.au  Felicity.Cox@ling.mq.edu.au

Abstract

In this study, we investigate the acquisition of Japanese vowel duration and phonemic vowel length contrast by inexperienced late onset learners whose native language is Australian English (L1). The learners’ Japanese vowels were obtained by longitudinal data collection over a 1.5 year period beginning at the initial stage of L2 learning. In addition, recordings of Australian English vowels (produced by the learners) and Japanese vowels (produced by native Japanese speakers) were also collected. The results show that Australian English learners were able to adjust vowel duration contrast in L2 Japanese by lengthening long vowels to produce ratios of long to short vowels approaching that of native Japanese speakers. However, the change in duration across vowel type was non-uniform with greater difference between the native language and the target language generating relatively greater change in L2 production. Adjustment of duration in L2 vowels typically appeared at the very early stage of learning but there was some variability between individuals in the patterns of achievement.

1. Introduction

1.1. L2 contrastive cue learning

It is generally accepted that both phonetic and phonological differences between native language (L1) and non-native/second language (L2) have an influence on L2 learning. One significant model that discusses this influence is Best’s Perceptual Assimilation Model (Best, 1995). Proponents of this model have examined how and when phonetic and phonological changes develop from native/L1 to non-native/L2 (e.g., Best, McRoberts & Goodell, 2001). Best and Tyler (in press) suggest that much phonetic perceptual learning may occur at the beginning of the late onset L2 experience. However, the nature of the relationship between perception and production at the early stage of late onset learning requires further examination. Diachronic analyses have been conducted on L2 vowel production from L2 late onset in immersion contexts (Baptista, 1992; O’Brien, 2003). However, the immersion context onset does not always correspond to that of L2 experience since often late learners receive L2 instruction before immersion commences. How L1-L2 production changes at the initial contact with L2 remains unclear. In immersion contexts, L2 is learnt through natural communication and plays institutional and social roles in the L2 spoken community. In classroom contexts, L2 is learnt more structurally through formal instruction and generally without any social role outside of classroom (see Ellis, 1994).

The current paper examines the diachronic changes in vowel production from L1 to L2 for late onset learners at the initial stage of L2 contact in the classroom context. The question of whether phonetic changes occur in L2 vowel production at the onset of L2 learning will be addressed by examining the change in vowel duration patterns over a 1.5 year period. Production studies focussing on the influence of L1 on achievement of the L2 durational feature (e.g., McAllister, Flege & Piske, 2002; Ingram & Park, 1997) have found that the presence of the durational cue in the native phonology assists in L2 durational achievement. Ingram and Park (1997) showed that the durational cue present in L1 (Japanese) phonology led to a tendency toward categorical contrast in the production of L2 (Australian English) vowel duration. McAllister et al. (2002) examined the length distinction of Swedish vowels produced by L2 learners with L1 backgrounds that use differently weighted durational cues. The results showed a gradient of influence on L2 production based on the weighting of the durational cue in the L1. Languages that used duration prominently as a contrastive feature favoured achievement of Swedish durational contrast whereas those that did not utilize
duration contrast, inhibited the development of this contrast in the learners. They concluded that achievement of L2 duration contrasts reflects the importance of durational cues in the L1 vowel system and they proposed the “feature prominence hypothesis” to explain this finding. In the present study, we investigate the “feature prominence hypothesis” by examining how the variable weighting (or feature prominence) of the duration cue within a single language affects L2 vowel production.

In AusE, monophthongal vowels exhibit contrast in spectral features and duration however the functional load of duration varies across spectral categories (Bernard, 1967; Cox, 2006; Harrington & Cassidy, 1994). The vowels that are primarily differentiated by duration are [e] and [æ] which share the same vowel quality. For most other AusE vowels, spectral quality is the primary differentiating feature and duration has secondary importance.

Standard Japanese has five vowel pairs, [i]-[e], [a]-[u], [o]-[ö] and [u]-[u:], which have minimal spectral differentiation and contrast primarily by duration (e.g., Han, 1962; Hirata 2004). Japanese durational contrasts display stable ratios of short to long vowels across speakers and speech rates (Hirata, 2004).

As a preliminary analysis, Japanese vowels produced by native Japanese speakers (henceforth NJSs) are compared with AusE vowels produced by Australian learners of Japanese. This comparison aims to determine the differences in the durational feature between the two languages across vowel types and to ascertain the starting point and goal of learning. Changes that occur in L2 vowel durations over a 1.5 year period of learning Japanese are examined.

We hypothesise that at the initial stage of L2 learning, duration will be an easy cue for Australian learners to use contrastively as it is employed as a secondary but important cue in the L1 and according to Ingram and Park (1997) should therefore be readily available for L2 differentiation. However, the achievement of the L2 durational contrast may vary due to the differential role of duration across vowel types in AusE. Feature Prominence Hypothesis (McAllister et al., 2002) would predict that vowel types which use duration as their primary cue would lead to relatively greater success in the production of phonetically similar vowels in L2.

2. Method

2.1. Speakers and corpus

Four female Australian learners of Japanese, aged from 19–22 years participated in the study. All started learning Japanese at Macquarie University in 2005 without any previous Japanese experience. Their AusE vowels were obtained in a single recording session and their Japanese (L2) vowels were obtained in three recording sessions, four, eight and sixteen months after they started learning L2. In addition, three native female Japanese speakers (henceforth NJSs) from Tokyo or its surrounding area, aged 26–29 years, were recorded in a single session.

Recordings were carried out in a sound treated recording studio at Macquarie University. Participants were asked to produce single disyllabic words presented in random order on a computer screen (Table 1). English words were presented in English orthography and Japanese words were presented in Japanese hiragana orthography. The elicitation task was restricted to isolated words because it was difficult for the learners to produce Japanese phrases at the time of recording. Both real words and non-words were used to control phonetic context and the investigator controlled the timing between each utterance. Five tokens of each word were obtained in each of the first three recording sessions and ten tokens were obtained from AusE learners in the final L2 (Japanese) recording session.

2.2. Analyses

Speech data was digitized at 44.1 kHz with 16 bit resolution. The data were segmented and labelled using Praat (Version 4.3.15). Vowel duration was established from the onset of vowel phonation to vowel offset defined as cessation of periodicity or marked reduction in F2 amplitude. The ratio of long to short vowels was also calculated by dividing a long vowel’s duration by each short vowel in the pair. Ratio is examined to control for possible rate effects.

Mixed model analysis was used to examine significant effects as this procedure accounts for multiple comparisons and repeated measures (Max & Onghena, 1999).

3. Results

3.1. AusE (L1) vs. Japanese (L1) comparison

3.1.1. Duration

Table 2 presents the mean vowel durations for AusE produced by the four learners and Japanese produced by three NJSs. Japanese long vowels were significantly longer than their AusE counterparts while the only short vowel to show significant effects was Japanese [a] which was longer than its AusE counterpart [n].

Durational differences between long and short paired vowels were greater in Japanese than in AusE as this is the primary contrastive cue. The phonemic contrast of
AusE [i]-[e] was realised by the largest duration difference (90.4 msec). However, the difference was still smaller than the Japanese contrast at 136.9 msec. AusE long vowels, [i], [o] and [u], had average durations that were closer to their Japanese short vowel counterparts, [i], [o] and [u] than long vowel counterparts, [i], [o] and [u]. The average duration of AusE [e] was midway between its long and short counterparts in Japanese.

Table 2. Duration means (msec): L1 (AusE) and L1 (Japanese produced by NJSs).

<table>
<thead>
<tr>
<th></th>
<th>AusE</th>
<th>Japanese (NJSs)</th>
<th>Difference</th>
<th>AusE (stdv)</th>
<th>Japanese (stdv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AusE [i] &amp; J [i] (S)hort</td>
<td>62.4 (10.0)</td>
<td>73.7 (13.6)</td>
<td>11.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [i] &amp; J [i] (L)ong</td>
<td>118.0 (25.7)</td>
<td>219.3 (30.8)</td>
<td>101.4 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference L-S</td>
<td>55.6</td>
<td>145.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [u] &amp; J [a]</td>
<td>87.3 (11.0)</td>
<td>109.2 (16.2)</td>
<td>21.9 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [u] &amp; J [a]</td>
<td>177.6 (32.6)</td>
<td>246.1 (26.3)</td>
<td>68.5 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference L-S</td>
<td>90.4</td>
<td>136.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [o] &amp; J [o]</td>
<td>78.0 (12.2)</td>
<td>95.7 (14.3)</td>
<td>17.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [o] &amp; J [o]</td>
<td>138.4 (31.4)</td>
<td>243.9 (34.1)</td>
<td>105.5 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference L-S</td>
<td>60.5</td>
<td>148.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [a] &amp; J [u]</td>
<td>64.9 (11.2)</td>
<td>84.1 (18.6)</td>
<td>19.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AusE [a] &amp; J [u]</td>
<td>103.4 (21.2)</td>
<td>213.7 (34.0)</td>
<td>110.3 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference L-S</td>
<td>38.5</td>
<td>129.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 presents individual speakers’ vowel duration. The asterisk indicates whether individual speakers differ significantly from the average of the NJSs (p<.01). Individual data for most AusE speakers show that all AusE long vowels were significantly shorter than Japanese counterparts. Short vowel comparisons between the two languages showed few significant differences.

3.1.2. Durational ratio

Table 3 presents the durational ratios for Japanese of NJSs and AusE.

Figure 1. Vowel duration for AusE and Japanese native speakers.
the short vowel comparison only Japanese [a] and AusE [ə] were significantly different from each other (see Table 2). AusE long vowels tended to be closer in duration to Japanese short vowels than to Japanese long vowels. AusE [r]-[ə] showed greater duration difference compared to other AusE vowel pairs and were more similar to Japanese. The ratios of Japanese vowel pairs were greater than those of AusE counterparts, but significant difference only occurred for high vowels.

3.2. L2 (Japanese) development

3.2.1. Duration

The individual learner’s duration changes from AusE to Japanese (L2) are shown in Figure 3. In general, the learners’ L2 (Japanese) vowels were longer than their L1 (AusE) counterparts. Greater lengthening occurred for long vowels but the degree of lengthening differed across vowel type. For example, the change from AusE [r]- to J [ə] was smaller than that from AusE [u] to J [u]. Larger L1 (AusE)-L1 (Japanese of NJS) difference observed in 3.1. resulted in larger change from L1 to L2. L1-L2 lengthening showed a tendency to compensate for the L1-L1 difference.

Three out of the four learners showed significant long vowel duration enlargement from L1 to L2 by the third recording session (16 month L2 learning experience). Over-lengthening of long vowels was often observed, especially in NW’s data.

Learners lengthened long vowels during the 1.5 year data collection period. Short vowel lengthening was less prominent. A tendency towards linear lengthening was observed in the long vowels, [ı], [a] and [o]. The pattern of long vowel lengthening was learner specific. For example, KC and NS showed gradual lengthening and did not show any statistical difference from the NJS’s average at the third recording session. AP showed no significant L1-L1 enlargement and her Japanese long vowels are always significantly different from the NJSs’ average throughout the three recording sessions. On the other hand, NW showed the greatest change in the initial recording and over-lengthening for all vowel types.

3.2.2. Durational Ratio

The individual learner’s ratio changes from AusE to Japanese (L2) are shown in Figure 4. The learners typically enlarged ratios from L1 to L2 within the 1.5 year period. Three learners, KC, NS and NW, showed significant change in the first/second recording session for [ı]-[ı], [a]-[a] and [u]-[u]. Over-enlargement occurred relative to NJS but this was non-significant.

Clear linear increase was observed in [ı]-[ı] and [a]-[a], and long vowel lengthening was a main factor contributing to ratio enlargement. The patterns of ratio change and the difference from NJSs’ average varied across individuals. NW showed more over-enlargement than others. AP showed very small ratio L1-L2 change.

Vowel type did not appear to affect L2 ratio success as all vowels were equally achieved.

---

### Table 3. Durational ratio means: L1 (AusE) and L1 (Japanese produced by NJSs).

<table>
<thead>
<tr>
<th>AusE</th>
<th>Japanese (NJS)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (stdv)</td>
<td>Mean (stdv)</td>
<td>L1-L2</td>
</tr>
<tr>
<td>AusE [ı]-[ı] &amp; J [ı]-[ı]</td>
<td>1.90 (0.34)</td>
<td>3.06 (0.63)</td>
</tr>
<tr>
<td>AusE [n]-[n] &amp; J [a]-[a]</td>
<td>2.05 (0.35)</td>
<td>2.29 (0.34)</td>
</tr>
<tr>
<td>AusE [s]-[oσ] &amp; J [o]-[œ]</td>
<td>1.81 (0.49)</td>
<td>2.57 (0.32)</td>
</tr>
<tr>
<td>AusE [o]-[u] &amp; J [u]-[u]</td>
<td>1.62 (0.36)</td>
<td>2.64 (0.62)</td>
</tr>
</tbody>
</table>

The phonemic durational contrast in Japanese was acoustically realized by large durational ratios. The ratios of Japanese vowel pairs, [ı]-[ı] and [u]-[u], were significantly larger than those of AusE counterparts for [ı]-[ı] and [u]-[u], and the ratio differences were the smallest between J [a]-[a] and AusE [r]-[r] (0.24). The largest AusE durational ratio is seen in [r]-[r].

**Figure 2.** Durational ratio for AusE and Japanese.

**Figure 2** shows the long/short durational ratios for AusE and Japanese of NJSs. Individual data showed that the AusE ratios were significantly shorter than NJSs’ averages. Exceptions were observed for AusE [r]-[r]. Two speakers display ratios that are comparable with NJS for [r]-[r]. An interesting difference between the two languages can also be seen in the ratio pattern across vowel types. Japanese speakers showed the smallest ratio for J [a]-[a] while AusE [n]-[n] ratios was the largest compared to other vowel pairs reflecting the importance of this cue in phonemic contrast.

In summary, Japanese long vowels were significantly longer than their AusE counterparts while in...
**Figure 3.** Individual mean durations across recording sessions: AusE (L1) and 3 Japanese (L2) recording sessions.

**Note:** The results of four recording session are presented for each learner. The dark bar indicates AusE and the three grey bars from left indicate Japanese recording 1, 2 and 3, respectively. Horizontal lines indicated NJSs’ Japanese vowel average duration.

**Figure 4.** Individual durational ratios across recording sessions: AusE (L1) and 3 Japanese (L2) recording sessions.

In summary, the learners lengthened long vowel durations and enlarged ratios in L2 at the initial stage of learning. Larger L1-L2 difference resulted in more change from L1 to L2 as learners attempted to develop duration and ratio contrasts approaching the NJSs’ norm. There were, however, learner dependent vowel type effects on L1-L2 change. By the second recording session, eight months after they first began learning Japanese, three of the four learners demonstrated significant increase in the durational ratio from L1 to L2 for the non-low vowels. For the low vowel contrast, which had similar ratios in Japanese and AusE, less increase from L1 was necessary to achieve the required L2 ratio.

**4. Discussion and conclusion**

In the current data, the vowel lengthening and ratio enlargement from L1 to L2 appeared at the initial stage of learning. The early phase perceptual learning suggested by Best and Tyler (in press:14) may therefore also apply to production of duration as a contrastive feature.

The results show that the degree of difference between the two languages predicts L1-L2 change. Larger L1-L1 differences resulted in greater change from L1 to L2. The learners adjusted L2 duration and ratios to compensate for the L1-L1 differences.

In general, development of the ratio towards NJSs’ norm was achieved by lengthening long vowels. Learners were also found to exaggerate the long vowel durations and
therefore durational ratio. Similar results were reported in previous studies which examined the durational ratio development of English learners of Japanese (Mah & Archibald, 2003:210; Ueyama, 2001:97). Exaggeration indicates some inaccuracy in the use of the L2 duration feature.

Long vowel lengthening appeared in the learners’ production of Japanese. Such lengthening was required because the average duration of the AusE long vowel counterparts was relatively close to the Japanese short vowel durations for [i], [o] and [u], or midway between Japanese long and short vowels for [a].

The “categorical representation” of L2 duration was found in the current data in accordance with Ingram and Park (1997). However, duration was the central contrastive feature in the L1 of Ingram and Park’s subjects whereas in the current research the durational feature was not the primary contrastive cue for most vowel types. According to “feature prominence hypothesis” (McAllister et al., 2002) we should expect a difference in the achievement of durational contrast across vowel types depending on the prominence of the duration feature in the L1. In the current study, differences between vowel types did not clearly occur. We expected [a]-[aː] to be more accurately produced than other vowel types because the AusE counterparts reply primarily on this feature for differentiation. However, statistical test did not reveal more accurate production for this vowel contrast in the L2 as would be predicted by feature prominence hypothesis. The results suggest that within a language, if the duration cue is used phonemically, regardless of its prominence, this cue can then be recruited in the L2 contrast.

Comparison with the other acoustic features, which are dominant in L1 or which are not used in L1 vowel system, will be required in future research to provide a more comprehensive picture of the relationship between L1-L1 difference and L2 approximation at the onset of L2 learning.

5. Reference


---

1 Asterisks, ** indicates significant difference with p<.01 and * indicates significant difference with p<.05 throughout all tables in this paper.