authentic Japanese syllables. In this case it would seem that the application of the duration rules for Swedish quantity could have yielded a rather good rendering of the Japanese contrast. These results indicate that in the case of the realization of Japanese quantity, the transfer of at least some of the aspects the Swedish quantity contrast pattern is part of the Swedes' strategy in learning Japanese quantity. The durational aspects of the English tense-lax contrast present a somewhat less clear picture of the transfer phenomenon. It looks like the Swedish natives are attempting to render the contrast but could be unsuccessful because of their tendency to continue to apply the L1 pattern in their L2 use.

Further work on this material can give us a clearer idea of what residue from the L1 there might be in the phonetic realization of an L2 contrast.

References

Figure 2 shows the calculated V/C duration ratios for the short-long-vowels in Japanese averaged over both isolated words and words which occurred in a sentence.

Figure 2

Further work on this material can give us a clearer idea of what residue from the L1 there might be in the phonetic realization of an L2 contrast.

References
than female speakers. It would be interesting to know if the three male speakers instead use other phonetic cues more actively than female speakers.

Table 1. Cross-speaker variation in FO range for attitude (FO maxima minus (final) FO minima in semitones).

<table>
<thead>
<tr>
<th>Attitude/speaker</th>
<th>Female speakers</th>
<th>Male speakers</th>
<th>All the speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U</td>
<td>V</td>
<td>W</td>
</tr>
<tr>
<td>Q</td>
<td>11.4</td>
<td>14.7</td>
<td>14.5</td>
</tr>
<tr>
<td>SUS</td>
<td>12.0</td>
<td>19.2</td>
<td>17.7</td>
</tr>
<tr>
<td>JOY</td>
<td>15.5</td>
<td>16.3</td>
<td>20.1</td>
</tr>
<tr>
<td>DIS</td>
<td>10.8</td>
<td>10.8</td>
<td>10.4</td>
</tr>
<tr>
<td>NEU</td>
<td>12.5</td>
<td>13.1</td>
<td>12.4</td>
</tr>
</tbody>
</table>

2.2 Pitch range for initial rise and final fall

The pitch range was calculated for the initial FO rise and final FO fall (cf. Figure 1 below). A typical manifestation of the pitch range of initial FO rise is in ascending order DIS<NEU<Q<JOY. The cross-speaker pitch range variation for the initial FO rise is far more consistent than that of the final fall in relation to attitude type. Note that there is a considerable cross-speaker variation in the manifestation of pitch range for the FO fall for attitudes SUS and JOY, but not for the initial FO rise.

Figure 1. Pitch range in semitones for the initial FO rise (left) and for the FO fall (right).

2.3 Pitch range for final rise (Q and SUS)

Question utterances in Japanese are typically accompanied by a terminal rising contour. In the present data, even SUS utterances had a regular terminal rise. However, final FO rise for Q and SUS were consistently differentiated in the magnitude (cf. Figure 2). The average FO rise for Q was 9.1 semitones (SD=3.04) while that for SUS was 12.9 semitones (SD=4.72). The magnitude clusters around 2-4 semitones for most speakers, but speakers U and W had more extreme values.

Figure 2 (left). Magnitude of utterance final FO rise in semitone for Q and SUS.

Figure 3 (right). FO peak values for the six speakers. U, V, W are female speakers.

2.4 FO peak value

Figure 3 shows the FO peak values for different attitude types. Four out of six speakers had the same order in the FO peak value, which in ascending order is DIS<NEU<Q<JOY. Note that this order is similar to that of the pitch range of initial FO rise except that of SUS and Q. Two speakers, both male, had their highest FO peak for Q rather than JOY.

2.5 FO peak delay

The relevance of the FO peak delay, i.e. the FO peak is not on the expected syllable to which the phonological accent is affiliated to, has been discussed for some time in relation to pragmatics. In the present data, the FO peak delays were common even for NEU (cf. Figure 4). All except of one case (speaker X for NEU) had a peak delay of varying from one mora-delay to six morae-delay. All the speakers had the least peak delay for NEU while the peak delay in relation to other attitude types varied considerably across speakers with SUS showing most agreement in delay. Since the diversity among the speakers is great, it seems that the FO peak delay per se is not a reliable correlate for attitude types.

Figure 4 (left). The timing of FO peak with the mora. When it is 0, the FO peak is in the syllable (mora) to which the accent is phonologically affiliated and there is no delay.

Figure 5 (right). Intensity peak measurements in dB.

3 Intensity peak

There was a good cross-speaker agreement among the speakers in the intensity peak value with attitude type. The highest intensity peak value (average 75dB) was found for JOY while the highest (average 68dB) for DIS (cf. Figure 5 above). Intensity peaks in relation to the attitude types varied less across speakers. In contrast to the difference between JOY and DIS, variations in the intensity peak value for other types of attitudes is small (71-2 dB on average). However, speakers differ considerably in the magnitude of intensity peaks. Some speakers vary the intensity greatly for attitude types (speaker W and Z) while speaker U hardly varied...
it. Intensity peaks and F0 peaks correlate to some extent, yet it is clear that the two parameters should be treated separately. Note that speakers W and U have very similar F0 peak values but different intensity peak values.

4 Duration (speaking rate) and pause
Average total utterance duration for the three utterances for each speaker is presented in Figure 6. Of the three utterances, the utterance /a-soo-desuka/ permits the insertion of a pause after the initial interjection /a/. When pause duration is included, it shows the same durational pattern as the other two utterances without a pause in reflecting the attitude types. Therefore, we interpreted pause as part of durational manifestation and included it in total utterance duration. The smallest cross-speaker variation was found for NEU for which all except one speaker used the shortest duration, clustering around 600-800ms. In the absolute duration value, speakers were also uniform for SUS which falls in the range between 1000 to 1200ms. The greatest cross-speaker variation was found for DIS for which the duration of the utterance varied from 800ms to 1250ms.

5 Vowel quality
Auditory impressions suggested considerable intra- and cross-speaker variation in the use of voice quality in general as well as in the specifically tested attitude types. Since the acoustic cues for voice quality are less straightforward than other acoustic cues, we only present the differences in vowel quality in this paper. Figure 7 above shows the manifestation of vowel quality by speaker Z. This speaker differentiated the vowel quality of /a/ in such a way that SUS and JOY had a more front quality than NEU, Q, and DIS. The figure also shows the formants values of /a/ in nonsense words /mamamama/ spoken neutrally by the same speaker.

6 Summary and discussion
Together with our earlier report on F0 shape and phrasing (Nagano-Madsen & Ayusawa 2005), both agreement and discrepancies were observable among the six speakers in their manifestation of attitudes. It seems that pragmatic information can be expressed in at least a few alternative ways in Japanese and that this line of research needs more attention.

References