Respiratory Control, Pauses, and Tonal Control in L1’s and L2’s Text Reading –
A Pilot Study on Swedish and Japanese –

Toshiko Isei-Jaakkola¹, Yasuko Nagano-Madsen², Keiko Ochi³

¹Department of English Language and Culture (Graduate School), Chubu University, Japan
²Department of Languages and Literatures, University of Gothenburg, Sweden
³School of Media Science, Tokyo University of Technology

tisei@isc.chubu.ac.jp, yasuko.nagano-madsen@sprak.gu.se, ochik@stf.teu.ac.jp

Abstract

This paper reports the results of a pilot study, which examines the respiratory control exerted by chest and abdominal-muscles during the reading of a long text in the mother tongue (L1) and a targeted foreign language that is being learned (L2), with reference to syntax and prosody in Japanese and Swedish. Three datasets of read speech were obtained from Swedish speakers (SwL1), Swedish learners of Japanese (SwL2), and Japanese speakers (JL1). The results showed that the subjects used respiratory control differently while reading L1 texts and L2 texts, respectively. Both SwL1 and JL1 used chest and abdominal-muscles almost simultaneously, and the peaks and troughs of their muscular movements co-occurred at the onset of major syntactic units such as sentences and clauses. SwL2 used more chest muscles than abdominal muscles, with muscular movements being more frequent, irregular, and small. There was no significant difference between JL1 and Swedish L1 and L2 in terms of the tonal control (pitch range). Some pitch peaks and pauses that appeared at the major syntactic boundaries coincided with the peaks of the muscular movements, but other pitch peaks and pauses did not. These results led to the hypothesis that the acquisition of intonation precedes that of respiratory control in L2 learning.

Index Terms: respiratory muscular control, Swedish L1 and L2, Japanese L1, pauses, syntactic units, tonal (pitch) control

1. Introduction

Second language acquisition has been studied from various perspectives. It is well known that the neural basis of first language (L1) processing differs from that of second language (L2) processing ([1]). This is true even for the neural processes underlying a native speaker’s perception and production (L1) and that of a foreign language learner (L2) ([2]).

The present study continues to examine the acquisition of second language (L2) speech, with a focus on prosody. In the previous studies, the characteristics of Swedish learners’ speech read out in Japanese at the beginner level were reported in [5]. Moreover, we have been examining how SwL2 acquire Japanese and/or Chinese ([3], [4], [5], and [6]) in terms of production and perception. These studies have focused on phonetic and prosodic dimensions, based on acoustic experiments.

In this study, we add a physiological dimension to acoustic-phonetic and acoustic dimensions, focusing on the relationships between respiratory control, pauses, and intonation. We will examine how chest and abdominal-muscles are regulated during the L1 and L2 speakers’ read speech by utilizing the respiratory strain-gauge transducer (hereafter, RST). RST is used in the diagnosis of the sleepless syndrome, and its purpose is to know if the upper body is inflamed by examining the signals of a patient’s respiration. The two RST have been used previously in examining the differences in speech rhythm ([7], [8]), the correlation between pitch and intensity ([9]), gender differences ([10]), the F0 declination ([11]), and turn-taking in dialogue ([12]), among other uses.

Hitherto, two different methods have been known in acquiring the data of respiratory muscular movements: electromyography (EMG) and RST. The former, however, would require a large number of electrical needles since the muscles for respiration are spread widely across the upper body. Even the use of surface EMG can be unreliable because it can pick up irrelevant movements such as heart beats. Therefore, the present method of using the thin-wired RST (See Figure 1), which is not invasive, is preferred. Through our pilot test, we confirmed that a thin-wired RST picks finer movements than a belt-type RST. Furthermore, during respiration, not only chest and abdominal-muscles but also the back muscles are used, particularly the ones in the upper body; this also speaks of the suitability of the present methodology.

In this study, we deal with Swedish and Japanese. Swedish and Japanese share both a similarity and a difference in terms of prosodic typology. Swedish is a stress-timed language like other Germanic languages, including English and German, while Japanese is a non-stress language with a mora-timed rhythm. Both Swedish and Japanese, however, distinguish between two kinds of lexical pitch accent.

Our study’s ultimate goal is to clarify how respiratory muscular movements (chest and abdomen in this case) are related to syntactic units, pauses, and intonational units for Swedish L1 and L2 speakers, whose target language is Japanese, by having them read a story and comparing their output and processes to those of JL1 reading the same story.

Our research questions are as follows:

(1) What are the differences and similarities between different language speakers (JL1 and SwL1) in terms of respiratory muscular movements?
(2) What are the differences and similarities between L1 (JL1, SwL1) and L2 (Swedish) in terms of respiratory muscular movements? Here, we define pauses as both a silence in the speech waveforms and as a perceived pause.
(3) What are the correlations between respiratory muscular movements, pauses, and intonational cues in the above-mentioned L1 and L2 speakers?
(4) What can we predict about the linguistic rhythms of Japanese and Swedish respectively?
2. Experiments

2.1. Material
The material was a fable, “The North Wind and the Sun,” which
was provided in Swedish and Japanese versions; it was adopted
from [13]. When the number of simple sentences is calculated,
both the Swedish version and Japanese version were found to
contain nine sentences.

2.2. Subjects
The subjects were two SwL1 and two JL1. The two Swedish
speakers (SwL2) were intermediate level learners of Japanese.
All the subjects were male university students in their twenties.
The Swedish subjects read the story both in their mother tongue
and in Japanese. The Japanese subjects, serving as referential
informants, read the story only in their mother tongue. For the
purpose of each data type, all subjects read the story five times.
Before the recording, the subjects were given enough time to
learn to read the material fluently.

2.3. Experimental procedures
The experiments were conducted in Japan. The thin-wired RST
was placed on the subjects‘ chests and abdomens to record their
respective muscular movements during utterance (as shown in
Figure 1) along with a headset microphone to simultaneously
record a subject‘s voice. The distance between the mouth and
the microphone was secured while recording so that the
intensity would not be affected during the recording (Sampling
rate: 44.1 kHz). The signals from the chest and the abdominal-
muscular movements were sent directly to the PC through
LabChart, and the speech signals were sent to LabChart through
EDIROL. All these flows are illustrated in Figure 2.

For this pilot test, two JL1, two SwL1, and two SwL2 (who
were the same speakers as SwL1) uttered the story five times.
Thus, we acquired 30 signal data for the muscular movements
and 30 speech data from the recording.

3. Results

3.1. Observation of signals
Figure 3 shows the movements of the chest (top) and
abdominal-muscles (middle). It also depicts the speech
waveforms (bottom) for the first five sentences (out of a total of
eight sentences) in the Japanese version of “the North Wind and
the Sun,” which were produced by JL1. The figures show that
the number of sentences marked by a period (.) in the two
versions was five (Swedish) and eight (Japanese), respectively.
However, some of the sentences consisted of two or three
shorter sentences, which were connected by conjunctions.
In addition, in the figure, ‘S’ represents a sentence, ‘(,)’ represents
a comma, ‘(and)’ represents a conjunction, ‘((P))’ represents a
long pause, and ‘(p)’ represents a short pause. In Figure 3,
regular peaks of the chest and abdominal-muscular movements
appear at the onset of every sentence. An additional peak(s) also
occurred at the onset of the second part of a long sentence,
which was read in two parts. The peaks of the chest and
abdominal-muscular movements tended to co-occur regularly,
although the abdominal muscular movement started slightly
earlier than the chest muscle movement. Pauses were inserted
regularly after each sentence. With regard to SwL1, tendencies
from JL1’s speech were also found in SwL1’s speech. The
regular peaks of the chest and abdominal-muscular movements
appeared at the onset of every sentence and clause. Even here,
the peaks of the chest and abdominal-muscular movements
tended to co-occur regularly. Pauses were inserted regularly
after each sentence. In terms of SwL2, the difference in the
muscular movements was obvious in comparison to SwL1. As
Figure 4 shows, there were many small peaks for both chest
and abdominal-muscles. The peaks of these muscular movements
did not co-occur unlike they did for his speech in L1 text, but
more peaks were found for the chest muscle. Many irregular
pauses were inserted not only at the sentential and clausal
boundaries but also at the phrasal boundaries.

![Figure 1: RST with a wire on the chest and abdominal-muscles.](image1)

![Figure 2: The flow chart for acquiring data.](image2)

![Figure 3: JL1’s chest and abdominal muscular movements with speech waveforms for “The North Wind and the Sun” when read in Japanese, with regard to the first five sentences out of eight.](image3)

![Figure 4: SwL2’s chest and abdominal-muscular movements with speech waveforms for “The North Wind and the Sun” when read in Japanese, with regard to the entire eight sentences.](image4)
typological difference, the results were surprisingly similar with regard to the use of chest and abdominal muscles in L1 speakers’ text reading. In both languages, the peaks of the chest and abdominal muscular movements co-occurred at the onset of syntactic units such as sentences and clauses. The second JL1 speaker inserted many pauses. However, these pauses did not affect the number of muscular peaks. This may indicate that, although syntactic units are usually marked by pauses, pauses as such do not induce the muscular movements.

In contrast to the L1 speakers’ text reading, the occurrences of the muscular movement peaks showed very different patterns in the L2 speakers’ text reading. Muscular movements were more frequent and irregular, and the amplitude of the muscular movements was small. Unlike both Swedish and JL1’s speech, in which chest muscular movement and abdominal-muscular movement co-occurred, more chest muscular movements were observed in SwL2’s speech. Both SwL2’s inserted pauses in similar numbers to those of JL1. However, in JL1’s speech, the number of the peaks of muscular movements was not affected by the pauses.

3.2. Observation of manipulated signals
Figures 3 and 4 were acquired directly from speech signals utilizing LabChart. It has been found from the previous experiments that respiratory muscular movements during speech depend on language, individuals, physical shape, etc. Therefore, we need to normalize the signals in order to acquire the statistical data. For that purpose we changed these signals: L1 and L2 speakers’ chest and abdominal muscular movements, and their speech, into Matlab data. And thus we could acquire the peaks of the signals automatically as shown in Figure 5. In Figure 5, the first and third rows show the chest and abdominal signals respectively. The second and fourth rows have two kinds of lines (in blue and black) : the blue lines represent the $\Delta$ (the first order derivative) of the signals and the black lines represent the smoothed signals. The red circle on the black line of the second and fourth rows shows each location of peaks.

3.3. Cross-correlation between the signals
First, we confirmed whether the peaks occurred simultaneously between the chest and abdomen. We then acquired the cross-correlation between the chest and abdominal signals to measure the similarity of the two signals.

Figure 6 shows the cross-correlation at the lag 0 of the three groups. The result of a one-way analysis of variance (ANOVA) shows the significant main effect ($p < 0.05$), and the post hoc test of Bonferroni indicates the significant difference (shown by an asterisk) between JL1 and SwL1 and between JL1 and SwL2 ($p < 0.05$).

Figures 7 and 8 show the numbers of the peaks of chest and abdominal signals. The results of the ANOVA show the significant main effect only on the peaks of chest muscles (see Figure 7), not abdominal muscles (see Figure 8). The post hoc test shows the significant difference between SwL1 and SwL2.

3.4. Prosodic units and pauses
We compared the respiratory control exerted by the chest and abdominal-muscles with the intonational patterns in JL1’s speech. All the peaks of the muscular movements coincided with pitch peaks, but not vice versa. This is similar to the distribution of pauses, that is, the way pitch peaks appear in relation to the muscular movements is similar to the way pauses appear. All the muscular movements appear with or after a pause, but not vice versa. We therefore interpret these results to mean that the respiratory control, intonation, and pauses are all
regulated separately. Pitch peaks or pauses as such do not automatically induce the peaks of the muscular movements. Rather, all the three are set at the onset of a major syntactic unit.

The present study adapts the AM theory as a conceptual background and method [14] in analyzing L1 and L2 Japanese intonation. The F0 contour was segmented according to the hierarchical intonation units, such as accentual phrase (AP), intermediate phrase (iP), and utterance (U), by measuring a pitch reset. The results are presented graphically in Figure 9 below. We used the median value of the number of AP, iP, and pauses of five utterances for each subject. JL1 showed similar tendencies, and one SwL2 (SwL2_B) had a pattern that was similar to that of JL1, with regard to the balance between APs and iP. This may prove that SwL2_B had acquired a high degree of L2 Japanese intonation that was comparable to an advanced level, even though his respiratory control pattern was still typical of L2. With regard to the number of APs, the other SwL2 (SwL2_A) had a patterning that was different from that of JL1 and SwL2_B. Figure 10 shows the speech energy, phrasal units and meaning, F0 contours, and annotated prosodic units by SwL2 (SwL2_A). The figure demonstrates that there were hardly any hierarchical organizations of intonation in this speaker’s utterance. Instead, each utterance (U) consisted of a short and flat F0 fragment that was hard to define within the AM theory. This was also a typical pattern found at the early stage of L2 intonation acquisition [5].

Figure 9: The number of pauses, intermediate phrases (iP) or utterances (U), and accentual phrases (AP) of JL1 and SwL2. X-axis is the median value of five repetitions.

Figure 10: Phrases, intonational contours, and prosodic units in SwL2_A’s Japanese. U = utterance.

3.5. Pitch range

Since intonation differs from utterance to utterance and depends on different languages and speakers, we acquired inter-quartile range (IQR), which is the quantity related to pitch range in an utterance. (See Figure 11) As a result, we found that JL1 tended to exhibit a slightly larger range than SwL1, and SwL2 tended to exhibit a smaller range. It seems that JL1 had a more stable pitch range than SwL1 and SwL2. However, we calculated the median of logF0 of the pitch range in each sentence uttered by JL1, SwL1, and SwL2. At this stage, it is difficult to determine the differences between the pitch ranges of these three groups.

Figure 11: IQR of the pitch in each sentence by speakers and uttered language for JL1, SwL1, and SwL2, respectively.

4. Discussion and conclusion

The present study confirmed that the subjects who participated in this experiment control their respiratory muscles differently while reading L1 and L2 texts. The findings were as follows: (1) In L1 speakers’ text reading, chest and abdominal muscles were used almost simultaneously, and the peaks of the muscular movements co-occurred nearly at the onset of major syntactic units such as a sentence and a clause. In L2 speakers’ text reading, chest muscles were used more than abdominal muscles, and the occurrence of peaks of the muscular movements was more frequent, more irregular, and smaller. (2) Both SwL1 and JL1 showed similar results despite the fact that the two languages differ in terms of prosodic typology. All the peaks of the muscular movements coincided with pitch peaks, but not vice versa. This is similar to the distribution of pauses, as all the peaks of the muscular movements appeared immediately after the pause, but not vice versa. (3) The cross-correlation of the signals from JL1’s chest and abdominal muscular movements was the smallest of all the three groups. This implies that the time lag between the peaks of the chest muscular movements and abdominal muscular movements was longer for JL1 than SwL1 and SwL2.

As far as the present experiment is concerned, we therefore conclude that the respiratory muscular control, pauses, and intonation are all regulated separately. Pitch peaks and pauses as such do not automatically induce peaks of the muscular movements. Rather, all three are set at the onset of a major syntactic unit. The results also led to a hypothesis that the acquisition of intonation precedes that of respiratory muscular control in L2 learning. With regard to the pitch range (tonal control) and the time lag in (3), further investigation is necessary.

Further studies with more speakers and different language pairs are needed to confirm the present results and a hypothesis regarding the study of L1 and L2, and to distinguish the function of pauses in relation to syntactic units.

5. Acknowledgements

This study is partly supported by Grant-in-Aid KAKENHI for Scientific Research C (NO. 17K02698) of Japan and by the Royal Society of Arts and Sciences in Gothenburg (KVVS).

6. References


