

A Preliminary Investigation of the Effects of Study Tours on L2 Speech Rhythm

Donald White, Peggy Mok

Department of Linguistics and Modern Languages, Chinese University of Hong Kong, Hong Kong
donalddtimothywhite@gmail.com, peggymok@cuhk.edu.hk

Abstract

This study is a preliminary investigation of L2 speech rhythm in a study abroad context. Five students from an English-medium secondary school in Hong Kong were recorded twice: once before a 16-day study tour to New Zealand, and once after they returned. Both spontaneous speech and read speech were recorded. Several durational speech rhythm metrics were employed to measure utterances from the two sets of recordings (%V, PVI and Varco of vocalic, intervocalic, and syllabic intervals). The results indicate wide individual differences among the students. Two students showed differences in vocalic and syllabic variability while the other three students had no differences between recordings. As for vowel reduction, four out of five students had significant differences in F2 values in read speech after the study tour. One of these four also had significantly different F1 values. The results, as stated, are preliminary and call for further longitudinal study in this area.

Index Terms: speech rhythm, L2 acquisition, Cantonese, English

1. Introduction

This study investigates the speech rhythm patterns of Cantonese-speaking students learning English as a second language (L2) in a study abroad (SA) context. Students acquiring a second language will often go to live in a different country where the target language is spoken. Generally, the goal is to leave the comfort zone of their first language (L1), and immerse themselves in the culture of the second language in order to facilitate acquisition. While the prosodic effects of SA remain largely uninvestigated, studies of L2 speakers of English who immigrate to English-speaking countries suggest research in this area would be fruitful: the concomitant changes in L2 input seem to lessen the degree of perceived foreign accent in new immigrants, especially within the first six months of arriving in a new linguistic environment [1]. When compared to the permanence of immigration to a new country, however, the study abroad context seems relatively ephemeral. If there were prosodic changes in students' spoken English after a short time studying abroad, these changes would likely be subtle and difficult to perceive. With this in mind, this study uses some acoustic measurements (rhythmic metrics and formant frequencies) to quantify the effects that learning English in an SA context may have on students' speech rhythm.

1.1. L2 in a Study Abroad Context

The efficacy of studying abroad is a largely unsettled question among researchers of second language acquisition (SLA). In fact, there is a considerable body of literature to date in which SLA in an SA context is investigated. A recent summary of this research presents the progress of linguists working in this area [2]. There is optimism with regard to the direction that this sub-field of SLA is taking, but [2] also expresses surprise at the dearth of studies that investigate the development of

phonological accuracy: while a handful of studies have examined this aspect of SLA in an SA context, the results have been largely inconsistent. Additionally, studies of L2 fluency often measure students according to the subjective judgments of the researchers. To the best of our knowledge, there have been no studies to date in which the acoustical phonetic correlates of students studying a second language in an SA context have been examined. Fluency is an elusive and highly subjective concept; an examination of speech rhythm using acoustic measurements may provide objective data in order to track the development of fluency in L2 speakers.

1.2. Speech Rhythm

From a rhythmic perspective, the L1 and L2 in the present study are quite different: Cantonese is a syllable-timed language and English is a stress-timed language [3,4]. The hallmarks of stress-timing are stressed syllables, vowel reduction, and greater complexity of syllable structure; syllable-timed languages tend to have lower amounts of these three aspects [3]. Rhythmic patterns are not easily mastered by L2 learners, especially those whose L1 has little to no prosodic similarity with the target language.

In recent years, there have been a number of attempts to quantify speech rhythm by measuring the durational components of the speech signal, such as the overall percentage of vowels in an utterance, and the durational variability of vocalic, intervocalic, and syllable intervals [5,6,7,8].

While the effectiveness of the speech rhythm metrics has been the subject of some debate [9], it is clear that the metrics are more reliable when used in certain areas of linguistic research. One such area is longitudinal studies: a recent survey of the speech rhythm metrics concluded that one of their best applications is in tracking the phonological development of individual speakers over a period of time [9]. In studies of new immigrants, the changes in L2 phonology are most dramatic in the very early stages of residency in a new linguistic environment [1, 10, 11]. The present study hypothesizes that the L2 phonology of students in a study abroad context will also be affected, and will be reflected in their speech rhythm scores.

2. Method

Five boys were recruited from an English-medium secondary school in Hong Kong. Three of the boys aged 15 (Form 3), one was 14 (Form 2), and one was 13 (Form 1). Each boy was interviewed twice: the first time was about three weeks before they departed on a 16-day study tour to New Zealand; the second recording was made within 48 hours of their return to Hong Kong, 41 days after the first recording. (Student 4's study tour was cut short because of a family emergency. He had to return to Hong Kong after only nine days in New Zealand.) A few days before the first interview, the students received a copy of the short story called *The North Wind and the Sun*. They practised reading the story so that they would be prepared to read it aloud in the upcoming interview. In both

interviews, the students were engaged in casual speech with the first author, who is a native speaker of Canadian English, discussing their studies and the tour. After this discussion, the students were asked to read *The North Wind and the Sun*. The students were recorded with a solid state recorder in a quiet classroom at their school.

2.1. Procedures

Two types of utterances were examined: the spontaneous speech of the casual interviews, and utterances from *The North Wind and the Sun*, which was read aloud. To be deemed acceptable for the use of rhythmic metrics, an utterance had to be greater than six syllables in length, and articulated within the same breath group. In the read speech, the utterances were acceptable only if matching pairs could be found in the first and second interviews that conformed to the criteria above. A total of thirty utterances were selected for each student: 20 spontaneous utterances (10 before and 10 after); and ten read utterances (5 before and 5 after). The selected utterances were segmented into vocalic, intervocalic and syllabic intervals [in milliseconds] using the Praat software. The criteria for segmentation followed those in [6], i.e. mainly acoustic. For syllables, the segmentation principles in [8] were adopted. The segmentation was cross-checked by both authors.

A total of seven rhythmic metrics were used in this study: Percent V (%V), normalized Pairwise Variability Indices for intervocalic, vocalic and syllabic intervals (nPVI_C, nPVI_V, nPVI_S) and the normalised standard deviations of the durations of each type of interval within an utterance (Varco C, Varco V and Varco S) [5,6,7,8]. Percent V measures the overall percentage of an utterance that is composed of vowels. A higher score in this metric indicates greater syllable timing. The PVI and Varco metrics both measure durational variability, which is higher in stress-timed languages. PVI compares the durations of like intervals adjacent to one another. The Varco metrics measure global variability by measuring the standard deviation of like intervals within an utterance. All metrics measuring durational variability are normalised for speech rate. Paired t-tests were used to compare the scores of each student before and after the study tour.

Additionally, in order to further investigate the vowel reduction patterns of the students, the formant frequency values of several schwa vowels occurring in the read passage were measured. In the read passage, there were 16 tokens of the word “the” occurring before words with onsets. The vocalic portions of these 32 tokens of schwa (16 before; 16 after) were segmented, and the F1 and F2 values were measured and cross-checked by both authors. Paired t-tests were used to compare the scores of each student before and after the study tour.

3. Results

Out of the 35 possible comparisons of the rhythmic metrics, there were only 6 significant differences. All of these differences were found in students 1 and 2. For student 1, the differences were mostly in the utterances that were read aloud: %V ($p = .005$); Varco S ($p = .04$); nPVI_S ($p = .02$). For student 2, the vocalic variability of the spontaneous speech was significantly different both in Varco V [$p = .001$] and nPVI_V [$p = .002$]. Additionally, student 2 displayed greater consonantal variability in the utterances that were read aloud [$p = .03$].

The differences in student 1 were somewhat perplexing because the three metrics that contained differences actually changed in the opposite direction to what would be expected (see Figure 1). That is, if L2 English were becoming more stress-timed, an increase in syllable variability would be expected. For student 1, however, the syllabic variability decreases significantly in both the Varco S and the nPVI_S metrics. Similarly, %V would be expected to decrease if the student’s L2 English were becoming more stress-timed, but the %V scores for student 1 actually increase significantly in the second recording.

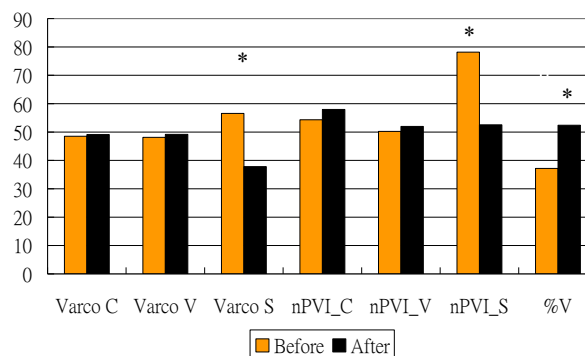


Figure 1: Rhythmic metrics for Student 1 Read Speech ($*=p<0.05$).

In contrast, Student 2’s results were significantly different in the expected direction (see Figure 2). The vocalic variability of the spontaneous utterances increases significantly in the second recording. The consonantal variability of the utterances read aloud also increase significantly in the second recording. These significant differences in student 2’s results corresponded with the impression of the first author, who was also the interviewer for both recordings. The first author felt that, of all the students, student 2’s was the only one whose English sounded more stress-timed in the second recording.

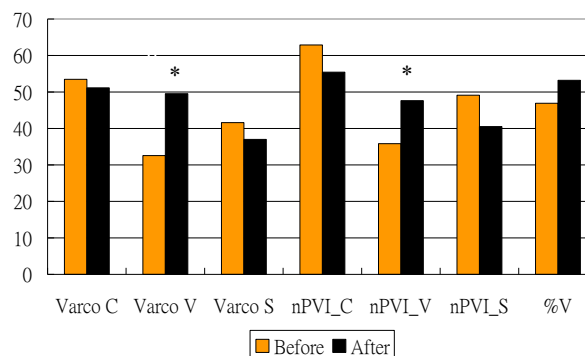


Figure 2: Rhythmic metrics for Student 2 Spontaneous Speech ($*=p<0.05$).

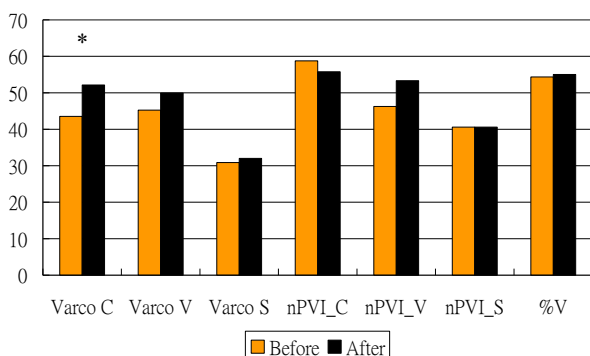


Figure 3: *Rhythmic metrics for Student 2 Read Speech* (*= $p < 0.05$)

Among the other students, there were no significant differences. There was one noteworthy pair of scores, however, that just missed significance: the Varco S scores of student 5 ($p = .053$). The data indicate that the student's syllabic variability seemed to be moving in the direction of greater stress-timing, although the difference is not significant.

In the measurements of the vowel formants, all five students had significant differences in their F2 values. These changes, however, were quite inconsistent. Three of the students' F2 values were significantly higher while two of them were significantly lower. Only one out of five students had a significantly different F1 value for the schwa tokens after the study tour. Student 2's F1 value was significantly lower.

Table 1. *F1 and F2 values of the schwa in the word 'the'.*

Student	Formant	Before	After	p
1	F1	476	467	0.636
	F2	1488	1431	0.034
2	F1	590	554	0.036
	F2	1639	1769	< 0.001
3	F1	471	461	0.262
	F2	1687	1632	0.042
4	F1	466	457	0.450
	F2	1596	1701	0.023
5	F1	508	502	0.479
	F2	1645	1698	0.016

4. Discussion

This study was a preliminary examination of how a short immersion in an SA context can change L2 speech rhythm. While there seems to be no consistent pattern of changes among the five students in the present study, the results suggest that even after a very short time in a new linguistic context, their speech patterns were affected. After only 16 days in New Zealand, one student out of five had speech rhythm scores that were significantly more variable than before the study tour. It seems likely that these are the incipient effects of immersion in an L2 culture: a longer time in an SA context would probably yield greater differences in the rhythmic scores.

The fact that the other four students did not have any measurable difference in their rhythm may be attributed to individual differences. It is well known that students acquiring a second language will have different levels of success. The

school system in Hong Kong is in many ways a hierarchy that stratifies students based on their varying capacities to speak English. At the top of this hierarchy are the Band One schools, which attract the best English students because they use English as a medium of instruction in all subjects except for Chinese. The students in the present study are all Band One students, and, therefore, skilled enough in English that they can use their L2 to study multiple subjects at the secondary level. That being said, it is also quite clear when speaking to these students that English is not their first language. First, their spoken English contains a variety of fairly frequent grammatical errors that are uncommon among native speakers. Second, and more relevant to the present study, there are a number of phonological and phonetic errors that are made repeatedly and consistently. Some of these errors are easy for a native speaker to detect, e.g. errors in word and sentence stress. Two of the students had greater mastery over English stress than the other three, but it was evident that none of them had obtained native-like proficiency in this area. This is just one example, but it serves to illustrate the considerable amount of individual differences in the L2 phonologies of the students involved in this study.

These differences also extend to the attitudes of the students. When they were interviewed, it was clear that some students were more receptive to the idea of speaking to a native English speaker. Some of them were more enthusiastic about the interviews than others. These varying levels of enthusiasm are not terribly surprising, but they may provide an explanation for the range of results in the present study. In other words, the likelihood that study tour will improve a student's spoken English will depend largely on that student's willingness to interact with his English-speaking hosts. If the interviews in the present study are any indication, it is clear that some students were more comfortable with the idea of interacting with native speakers. In fact, the student that seemed most enthusiastic in the interviews was also the student that had the most significant differences between his first and second interview: student 2. This student, who happened to be the youngest of the five in the present study, was also remarkable because his English actually sounded different to the interviewer after the student had returned from New Zealand. That is to say, before the speech rhythm scores were known in the present study, there were indications that student 2's English had undergone the greatest change during the study tour. This bodes well for the use of acoustic metrics in tracking changes in speech rhythm. While an impressionistic appraisal of foreign accent does not have quantitative power, it is not meaningless. In other words, because native-speakers are fine-tuned to the intelligibility of their interlocutors, it is likely more than coincidence that student who sounded more intelligible after the study tour was also the student whose speech rhythm scores reflected greater stress-timing.

The impression of greater stress timing in student 2 was further supported by the changes in his F1 and F2 tokens for his schwa vowels in the read passage. Two of the rhythmic metrics that scored significant differences in student 2 were measurements of vocalic variability. If this represents a movement toward more stress-timed rhythmic patterns, we would expect to find greater vowel reduction in syllables that are usually reduced in English. The "the" token in the present study are paragons of this phenomenon. The vowel in pre-consonantal "the" is largely realized as schwa by native English speakers. The formant values indicate that the quality

of student 2's schwa vowel changed significantly after the study tour.

While it is true that student 2 only had significant vocalic differences in his spontaneous speech, the vocalic variability in the read speech was also higher, if not significantly so, in the same two metrics: Varco V and nPVI_V. It seems that the study tour had a quantifiable effect on student 2's L2 stress-timing, and the quality of the vowels that are an integral part of speech rhythm.

The changes in student 2's schwa vowel seem to reinforce the findings of some previous similar studies. In [1], for example, which investigated L2 English among new immigrants to Canada, the intelligibility of the subjects' English was found to change soon after arrival in the new linguistic context. Several English vowels were investigated in [1]. A group of 44 newly arrived L2 speakers of English (20 native Mandarin Chinese speakers; 18 Russian; 5 Ukrainian; 1 Croatian) were interviewed at two months intervals over the course of one year. Among these L2 speakers, the largest gains in intelligibility were observed after the first four months; in later interviews, while there was still improvement, it was not as extreme. The vowels investigated in [1] did not include schwa. However, the early changes in vowel quality soon after immigration seem to be observable in the present study. Significant changes in formant frequency values were seen after only 16 days in a new SA context. This may represent the early stages of what was observed in [1]. Moreover, the significant differences in the present study suggest that it is possible to quantify L2 development using tests that do not depend upon the subjective intelligibility judgments of the researchers.

In order to solidify this claim, more studies should be carried out, especially those that focus on SA context for a longer period of time. Sixteen days is a very brief time. A longer study tour would provide an opportunity to look at the effects of longer-term immersion, and how it influences the speech rhythm scores of L2 learners. Student participants could also be screened more carefully in order to ensure that the students are eager to speak with foreigners and not overly shy. Certainly, the result of future studies of this kind would be more fruitful if the students interact more with their host using their L2 and less with their classmates using L1. This could be further ensured if the students were monitored during their study time abroad. Ideally, phonetic researchers could accompany students on tours such as these in order to maximize the amount of interaction in L2. Furthermore, data from a greater number of participants is needed to further investigate the effects of immersion on L2 speech prosody, as the present results show much individual variation. Nonetheless, the present study does indicate that further studies into this area will yield interesting results.

5. Acknowledgements

Thanks to Kwun Tong Mary Knoll school, the students that participated in the study, and especially to Mr. Kevin Ng, their teacher, for organizing the interviews. Thank you also to Robert Xu and Donghui Zui for cross-checking the utterances and formant values, and for conducting the statistical analyses.

6. References

- [1] Munro, M. and Derwig, T., "Segmental Acquisition in Adult ESL Learners: a Longitudinal Study of Vowel Production", *Language Learning*, 58(3): 479-502, 2008.
- [2] Freed, B., "Second Language Learning in a Study Abroad Context", in N. Van Deusen-Scholl and N.H. Hornberger [Eds], *Encyclopedia of Language and Education 2nd edition Volume 4, Second and Foreign Language Education*, 113-125, Springer Science+Business Media, 2008.
- [3] Dauer, R. M., "Stress-timing and Syllable-timing Reanalyzed", *Journal of Phonetics*, 11(1): 51-62, 1983.
- [4] Mok, P., "On the Syllable-timing of Cantonese and Beijing Mandarin", *Chinese Journal of Phonetics*, 2: 148-154, 2009.
- [5] Ramus, F., Nespor, M. and Mehler, J., "Correlates of Linguistic Rhythm in the Speech Signal", *Cognition*, 73(3): 265-292, 1999.
- [6] Grabe, E. and Low, E., "Durational Variability in Speech and the Rhythm Class Hypothesis", *Papers in Laboratory Phonology*, 7: 515-546, 2002.
- [7] Dellwo, V., "Rhythm and Speech Rate: A Variation Coefficient for ΔC ", in P. Karnowski, and I. Szigeti [Eds], *Language and Language Processing*, 231-241, Peter Lang, 2006.
- [8] Deterding, D., "The Measurement of Rhythm: A Comparison of Singapore and British English", *Journal of Phonetics*, 29(2): 217-230, 2001.
- [9] Wiget, L., White, L., Schuppler, B. and Grenon, I., Rauch, O. and Mattys, S., "How Stable are Acoustic Metrics of Contrastive Speech Rhythm?", *Journal of the Acoustical Society of America*, 127(3): 1559-1569, 2010.
- [10] Flege, J. E., "Factors Affecting Degree of Perceived Foreign Accent in English Sentences", *Journal of the Acoustical Society of America*, 84: 70-79, 1988.
- [11] Flege, J. E., Munro, M. J. and Skelton, L., "Production of the Word-Final /t/-/d/ Contrast by Native Speakers of English, Mandarin, and Spanish", *Journal of the Acoustical Society of America*, 92: 128-143, 1992.