

# Training Effect on Speech Rhythm of Nepalese Non-Native Speakers of English

*Ananda Aryal*

## Abstract

This project aimed to ascertain the effect of training on speech rhythm of Nepalese non-native speakers of English. Six Nepalese and one native speaker of English participated in this research. A six-week training program was administered and pre- and post-training guided speech samples were acoustically analysed. Acoustic measurement of the pre- and post-training speech samples showed that training played an important role in improving the rhythm of Nepalese non-native speakers of English.

**Key words:** English Rhythm, Nepalese speakers of English, pronunciation training, acoustic measurement.

## Introduction

Pronunciation teaching plays a vital role in English as a second language (ESL) and English as a foreign language (EFL) teaching (Morley, 1991; Wong, 1987). However, rhythm in speech has been basically ignored despite it being the indispensable component which builds fluency, intelligibility and comprehensibility in communication (Hood & Dixon, 1969; Munro, 2008; Kennedy & Trofimovich, 2008).

In order to acquire native-like speech, rhythm should be emphasised because it is an essential component in having a natural and free-flowing speech (Cai, 2008). Non-native speakers (NNSs) often fail to communicate well with native speakers (NSs) due to their lack of correct speech rhythm (Adam, 1979; Setter, 2006). Nepalese learners of English are no

exception. My personal experience as an EFL teacher in Nepal has been that aspects of pronunciation teaching are basically ignored. Incompetent teachers, lack of teacher training, exam oriented teaching, oversized classes and teacher attitudes are some of the reasons for this lack of pronunciation teaching (Shrestha, 2008). Similarly, Nepalese EFL classes are often criticised for teaching about the language and not the language itself; thus focusing on subject matter rather than communicative language skills (Sharma, 2003). It has been evident that training plays a significant role in improving pronunciation. The rhythm of English is particular and its mastery is considered one of the greatest difficulties faced by EFL learners (Gong, 2002). Hence, this research attempted to determine if training has positive impact on enhancing the rhythm of Nepalese NNS of English.

## Literature review

An accurate rhythm is one of the most essential phonetic aspects for the auditory comprehension and intelligible oral production of English (Prator, 1971; Morley 1979; Savignon, 1983; Faber, 1986; Anderson, 1993). Speech is a continuous stream of sounds, without clear-cut borderlines between them where stress and rhythm are the key to gaining a natural, smooth-flowing style of speech (Cai, 2008). Therefore, the rhythm of English should be emphasised in language teaching. Lin, Fan & Chen (1995) believed that pronunciation could be taught in four parts: intonation, stress and rhythm, consonants and vowels. They strongly emphasised the value of rhythm in pronunciation teaching.

### Defining the terms: Rhythm, stress-timed and syllable-timed languages

Rhythm can be defined as the patterned, recurring alternations of contrasting elements of speech. Abercrombie (1967) and Pike (1945) explain rhythm as the pattern which we can perceive in speech sound as a result of the repetition at regular intervals of prominent elements, such as stressed syllables. In the same way, Dauer (1983) described rhythm as:

When we speak naturally, words are parts of phrases and longer sentences. What we hear is a sequence of syllables in time, like notes in music. The time relationships among syllables make up the rhythm of language.

Pike (1945) was the first person who claimed the existence of stress-timed and syllable-timed languages whose view has been supported by many other scholars. Syllable-timed languages are those where all syllables are of approximately equal length, whereas stress-timed languages have approximately equal intervals of time

between stressed syllables (Adams, 1979; Marks, 1999). Setter (2006) explains that, in stress-timed languages (e.g. English), the start of each stressed syllable is said to be equidistant in time from the start of the next stressed syllable. This kind of rhythm is in contrast to syllable-timed languages (e.g. French, Spanish, Cantonese, and Nepalese), in which the start of each syllable is said to be equidistant in time from the start of the next.

Despite some controversies, many researchers believe English is a stress-timed language (Dauer, 1983; Deterding, 2001; Ling, Grabe and Nolan, 2000; Setter, 2006). Couper-Kuhlen (1993) studied the acoustic and perceptual isochrony of English, and Auer, Couper-Kuhlen, & Muller (1999) investigated on acoustic characteristics of rhythmic classes across languages. Their research resulted in cross-varietal differences in rhythmic patterning of the languages. Additionally, the basis for attributing stress, or syllable, timing may involve vowels rather than syllables. Syllables in a foot are compressed to approach foot isochrony, and vowel reduction is exploited to achieve this compression in English. Syllable-timed languages like French do not need to achieve foot isochrony and do not exploit vowel reduction for rhythmic purposes (Brakel, 1985). In the light of these studies,

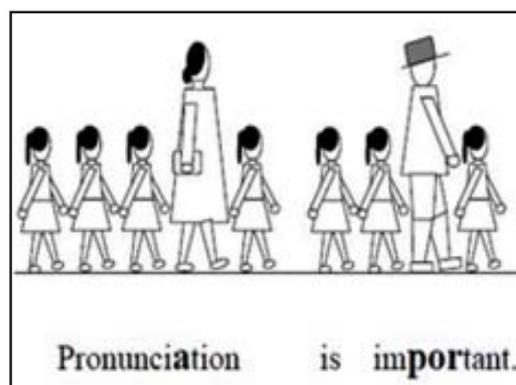


Figure 1: Stress-timed language

even though there are contrasting ideas related to the syllable-timed and stressed-timed languages, we cannot deny the fact that English has a certain rhythmic pattern which causes problem of intelligibility for the NNSs of English. Stress-timing in English is affected by the grammatical category of words. English words mainly belong to one of the two types of grammatical categories: content words and function words. Content words (e.g. 'ship') carry meaning in a sentence and are normally stressed, while function words (e.g. 'and') have no meaning in isolation and are de-stressed (Barrera-Prado, 2008). Moreover, McQueen and Cutler (1997) state that the English language contains strong and weak vowels which indicate language rhythm. They claim that rhythmic speech can be achieved if we emphasise strong vowels and give less prominence to the weak ones in an utterance. Praetor and Robinett (1985) demonstrated the stress-timed and syllable-timed languages in the following figures. They also described a stress-timed language as a language that has stressed syllables in an equal interval and a syllable-timed language as a language having all syllables in equal distance.

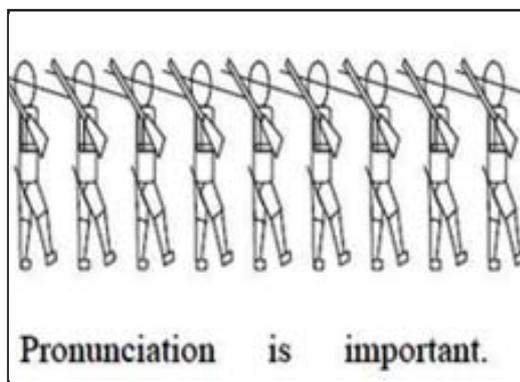


Figure 2: Syllable-timed language

### Problems of NNSs

Rhythm of English has created problems in fluency, intelligibility and

comprehensibility for NNS (Chen, Fan & Lin, 1996; Setter, 2006). Problems with mastering the rhythm of English are common to all NNSs even though they are from diverse language backgrounds (Taylor, 1981; Gong, 2002). Chela-Flores (1994) claimed that the variation in length between the vowels in stressed and unstressed syllables appears to be the main difficulty for NNSs of English. Similarly, non-rhythmic speech is much more difficult for the NSs to comprehend as compared to more naturally flowing rhythmic speech (Buxton, 1983; Kenworthy, 1987).

Mother tongue influence is another vital cause of difficulty for learners of English which frequently causes them to misgroup phrasing, fail to link sounds, have irregular pauses, and/or to have a tendency to stress incorrect words. This makes their English pronunciation sound strange and unnatural (Gong, 2002) and is the reason why NNSs often fail to establish effective communication with NSs of English. To overcome these problems, rhythm teaching should be emphasised in EFL and ESL classes with implementation of relevant teaching methodology (Wong, 1987; Orton, 1994). Barrera-Pardo (2008) believed that most EFL learners have a problem with decoding unstressed, weak function words contained in naturally spoken English. So a listening strategy should be adopted and rehearsed.

### Teaching methodologies and materials

Rhythm teaching is emphasised in the syllabus of different non-English speaking nations (Anderson-Hsieh, 1990; Chela-Flores, 2004). However, it has been found that appropriate teaching methodology is not specified for teachers, while non-native teachers themselves often lack confidence in teaching rhythm of English (Gong, 2002). Researchers (Anderson-Hsieh &

Venkatagiri, 1994; Tajima, Port, & Dalby, 1997) have provided evidence that rhythm of English can be improved by using certain feasible training methods.

Celce-Murcia, Brinton, and Goodwin (1996) designed a framework for teaching pronunciation using five stages. Their teaching framework was presented as (1) description and analysis, (2) listening discrimination, (3) controlled practice and feedback, (4) guided practice and feedback and (5) communicative practice and feedback. These methods seem practical and useful for the teaching of rhythm in English. Likewise, Brooks (1964) promoted the audio-lingual method for pronunciation teaching. He stated that the audio-lingual and gestural-visual methods should be emphasised with re-organisation and discrimination of structure followed by imitation, repetition and memorisation of the structure. Although the use of the audio-lingual method is thought to be the old and traditional method of pronunciation teaching, it is still used worldwide in EFL and ESL classes (Jones, 1997).

Other linguists have illustrated more procedures on teaching rhythm of English. Orton (1994) stated that rhythm can be taught through stepping on every beat while walking. Jones (1972) and Wong (1987) claimed that rhythm can be taught tapping on every stressed syllable in a sentence. Similarly, Chela-Flores (1994) suggested using graphic representation, whereas Gilbert (1984) recommended using rubber bands. Although linguists have suggested different procedures on teaching rhythm, all have a similar view that the regular occurrence of stressed syllables should be emphasised. In the same way, rhymes, poetry, jazz chants and limericks that have strong rhythmic patterns can be also used as teaching materials (Adams, 1979; Bray, 1995).

Murphy (2003) discussed further classroom techniques and tasks for teaching pronunciation which also appear useful in teaching rhythm. According to him, use of contextualised minimal pairs, gadgets and props, slow motion speaking, tracking the natural speech of native speakers and the use of techniques from drama and theatre arts can better serve in pronunciation teaching. He has suggested listening and repetition, teaching sounds in isolation as well as communicative and task-based pronunciation teaching methodologies. This later method of focussing on communicative and task-based teaching is considered important and the most advanced methodology for pronunciation teaching. Different linguists (Brown & Yule, 1983; Elliott, 1997; Hinkel, 2006) believe that the communicative approach is more practical in ESL and EFL classes. Proponents of communicative language teaching assert that language should be learnt from real life situations; not through the setting of rules (Widdowson, 1989).

### **Assessment criteria**

Ramus, Nespor, and Mehler (1999) stated that acoustic measurement which focuses on variability in vocalic and intervocalic intervals to analyse rhythm is one of the best methods to assess rhythm. Several researchers have used conversations or read speech methods to measure the rhythm in speech while others have used both. Low and Grabe (1995) used conversational speech as well as read speech in an acoustic measurement. Although it is evident that there is significant intersyllabic variation, their research demonstrated that the differences in rhythm along a stress syllable-timed scale can be instrumentally measured which confirmed two varieties of rhythm.

Low, Grabe, and Nolan (2000) and Deterding (2001) used conversational

recordings. Their intention was to measure whole syllables rather than analysing the vowel production only. They eliminated articulation of the final syllable and used a normalization of speaking technique based on the whole utterance. Normalization of speaking rate in acoustic analysis is essential because syllable duration of a slow speaker will be very high even if the speaker has rhythmic speech. They also tried to investigate speaking rate and the frequency of reduced syllables which proved that there are two varieties of rhythm in speech.

Low (1998, cited in Deterding 2001) and Deterding (2001) both used normalization of speaking rate in their analysis but in different ways. Low used pair-wise normalisation because she used read speech. She normalised comparisons of vowel durations by dividing the difference between each pair of successive vowels by the average duration of the two. In contrast, Deterding divided the duration of each syllable by the average duration of all the syllables (except the last) of the utterance as he was using conversational speech. According to Deterding, pair-wise normalisation cannot be used in conversational speech as it would be extremely difficult to exactly locate a boundary between two short syllables.

Derwing, Munro and Wiebe (1997, 1998) also carried out research on ESL students in Canada. Participants of their research were recorded before and after 12 weeks of training. Recordings were transcribed to evaluate intelligibility and rated on a 9-point likert scale (1=no accent to 9=very strong accent). Results of both studies demonstrated an improvement in the participants after the training. Similarly, Setter (2006) used auditory/perceptual analysis in his research. He scaled recordings into four categories: weakened, unstressed, stressed, and tonic. These

procedures of analysing the researched participants had limitations, one being the possibility of bias in rating the level of accent, intelligibility, and comprehensibility.

## **EFL context in Nepal**

Several studies concluded that English language teaching (ELT) in Nepal is far from satisfactory (Shrestha, 2008; Giri, 2010; Bista, 2011). Traditional teaching methods are still practiced in EFL classes with limited use of pair and group activities which are central to communicative language teaching (Shrestha, 2008). Most EFL classes lack student centred activities. Additionally, there is excessive use of the mother tongue (Khatri, 2011). Almost all EFL teachers in Nepal are non-native speakers and their use of English is largely confined to the classroom. They are not conscious of worldwide changes in ELT and most are unaware of practical methodologies for pronunciation teaching (Bhattarai & Gautam, 2005).

Bhattarai and Gautam (2005) also claim that Nepalese EFL teachers lack professional development and knowledge of networking with national and international experts. EFL learners in Nepal can barely communicate in English in a real life scenario (Shrestha, 2008; Adhikari, 2010) with pronunciation remaining as the predominant drawback. For this reason, this research study sought to test the hypothesis that training can improve the rhythmic English of Nepalese EFL learners. Subsequently, it answered the following research question:

Does training improve Nepalese non-native English speakers' rhythm of English ?

## Methodology

### Participants

Six Nepalese and one native speaker of English participated in this research at the University of Aberdeen, Scotland in 2010. All Nepalese participants were living in Aberdeen with a length of stay ranging from two to seven months. Most expressed their readiness to join the research believing that their rhythmic speech would improve after training. The only selection criterion was that they must be Nepalese speakers of English.

Four males and three females with an average age of 27.16 were divided into three groups: five non-native experimental (NNE), one non-native control (NNC) and one native control (NC). The NNC participant was included for control purposes, while the NC was used for sample recording only. A pre-training questionnaire was designed to obtain all the participants' personal information.

### Procedure

Pre and post-tests were administered. The post-tests were conducted a week after the training was completed. NNE participants were recorded before and after the training, while pre and post-test recordings of the NNC participant were made without any training. Only one sample recording was made of the NC participant. Participants were recorded in a sound proof phonetic laboratory at the University using a high-quality microphone and recorder. Web Surfer was used to generate the wide-band spectrogram for the acoustic measurement.

### Pre-test and post-test recordings

Read speech material was used in this research. The participants were requested to read out the short story 'The North Wind

and the Sun' during the pre- and post-test. Read speech was considered to have several advantages, including the avoidance of sequences that are hard to syllabify or segment. Moreover, using the same read speech for different groups allowed for paired comparisons to eliminate one source of variation from the statistics.

Only three sentences (See Appendix-1) from the middle of the read script were edited and used for analysis. All these sentences consisted of stressed and unstressed syllables. The researcher's prime intention for editing from the middle of the recordings was to provide the participants with some initial phonation in order to adjust and relax their voices, and to become accustomed to the presence of the researcher during their reading.

### Training

A six-week training was provided to the experimental participants. The training was conducted by a linguistics lecturer, a phonetician at the University of Aberdeen. Materials from a book 'English Pronunciation in Use Advance (Hewing, 2007), compact disc (CD) from the same book and BBC learning English website (<http://www.bbc.co.uk/worldservice/learningenglish/index.shtml>) were used in the training. The training was basically focussed on raising consciousness about the rhythm of the English language by exemplifying function words, content words, prominent words, and weak and strong vowels using power-point presentations and hand-outs from the above mentioned book. The participants were asked to stress and de-stress the words in a conversation and in isolation after they became able to identify these words. They were encouraged to speak aloud in the training session as repetition drill and group conversation were emphasised.

## Acoustic measurement

Acoustic measurement methodology, validated by Deterding (2001), was employed in this research. Variability index (VI), which was obtained from the differences in the duration of neighbouring syllables, was used in the analysis period. The three mid script sentences of the read speech were further divided into six stretch utterances wherein there were 38 syllables in total.

## Number of syllables

The number of syllables in a given stretch of speech can create a problem when comparing syllable duration due to the merging of words (Deterding, 2001). Equally, there are variations in pronunciation of lexical items which also depend on the speaking rate of a speaker. For example, the word 'comfortable' might consist of three or four syllable (Jones, Roach & Hartman, 1997, cited in Deterding, 2001).

This study was based upon the comparison of pre- and post-test recordings of guided speech. The problem of participants not being consistent when reading the same stimuli twice was acknowledged and therefore, pre and post-test syllables were made equal. The final syllable of every utterance was deleted in the analysis because final syllables tend to be lengthening, and it is often very difficult to trace-out the exact end of a syllable in the final position.

## Normalization of speaking rate and the Variability Index (VI)

Normalization of the speaking rate was essential to obtain VI because the effect of speaking rate gives much longer syllable duration for the participant who is speaking at a very slow speed. In order to

combat this, and to achieve normalization of speaking rate, whole utterances were firstly divided into smaller utterances for each participant. After that, all syllable durations in the utterances were added to achieve the whole utterance duration. Then, the whole utterance duration was divided by the total number of syllables in the utterance (except for the final syllable). Finally, the syllable duration was divided by the average syllable duration to achieve normalization of speaking rate. The same procedure was followed to obtain normalization of speaking rate in pre- and post-test utterances for all the participants. The methodology and formula followed to acquire variability index (VI) in this research is exactly the same as used by Deterding (2001). Moreover, the normalized syllable duration of all syllables (except the final syllable) was subtracted from that of the following syllable because the analysis of this research was also focussed on comparing the duration of neighbouring syllables. Average syllable variability was then calculated for each syllable using the formula given below where the normalized duration of the  $k$ th syllable is  $d_k$ , and the number of syllables in the focused utterance is  $n$ :

$$VI = \frac{1}{n-2} \sum_{k=1}^{n-2} |d_{k+1} - d_k|$$

The distance between the syllables is the main characteristic of stress-timed languages and VI indicates stress-timing or the rhythm of languages. NNSs of English are considered to have an English rhythmic pattern in speech if their VI is very close to the VI of native speakers. For this reason, specified syllables of all utterances made by the participants were measured, and then used further to generate their individual VI.

After that, all the VI of the participants were added and divided by the number of the utterances, in order to achieve the average variability index (AVI) of the individual participant. Then pre- and post-test AVI of individual participants were compared to ascertain their improvement. Finally, individual AVI was further compared with the VI of the native speaker to measure the native-like rhythm of non-native participants.

This table will assist in the understanding of procedures used to calculate normalization of speaking rate.

Table 1: *The VI calculation of participant no. 111 is illustrated in the following Table*

Syllable order	Syllable	Preduration (ms)	Normalized syllable duration	Difference from next syllable	Whole utterance duration (ms)	Average syllable duration (ms)	VI
1.	we	196	0.749	0.202	1307	216.4	0.050
2.	will	143	0.547	0.401			0.100
3.	see	248	0.948	0.818			0.204
4.	A(bout)	34	0.130	1.189			0.287
5.	a(bout)	345	1.319	-			
6.	that	341	-	-			
<b>VI for the utterance 1=0.653</b>							

The same procedure was followed to calculate normalization of speaking rate and VI of pre and post-test recordings for all the participants.

**Result**  
**Pre and post-test AVI of individual participants**

Participant 301 was the NC participant in this study and her AVI was 0.6665.

Therefore if, after training, the AVI of non-native speaking participants, either experimental or control, became closer to the AVI of this native speaking participant, it can then be asserted that they have native-like rhythm.

The following table provides individual results demonstrating overall performance after training. NNE participants 111 and 118 had an AVI very close to the AVI of the NP indicating native-like rhythm. Similarly, NNE participants 108 and 113 also demonstrated significant improvement. Despite this, participant 108 was considered to have non-native rhythm due to his pre and post-test AVI being very low compared to the others although he significantly improved after training. In

contrast, NNE participant 110 only had minimal improvement and NNC participant 204 had a significant decrease in his rhythm of English.

Table 2: Pre and Post-test AVI results of individual participants

Participant No.	Group of the participants	Number of utterances	AVI	
			Pre-test	Post-test
108	Non-native experimental	6	0.4145	0.4986
110	Non-native experimental	6	0.5489	0.5504
111	Non-native experimental	6	0.5633	0.6137
113	Non-native experimental	6	0.5190	0.5906
118	Non-native experimental	6	0.6718	0.7125
204	Non-native Control	6	0.5474	0.4940
301	Native control	6	0.6665	-

### AVI of all groups

NNE participants improved after their training as demonstrated in Figure 3 which reflects the increment. The AVI of the experimental group in the pre-test was 0.5435 which increased to 0.5891 in the post-test. Therefore, the experimental group reduced their AVI gap with the native participant by 0.0456 which confirms that they improved their rhythm of English.

On the contrary, the NNC participant had a considerable decrease in post-test culminating in a reduced AVI of 0.0534. His VI pre-test was 0.5474 which reduced to 0.4940 indicating a reduction in English rhythm.

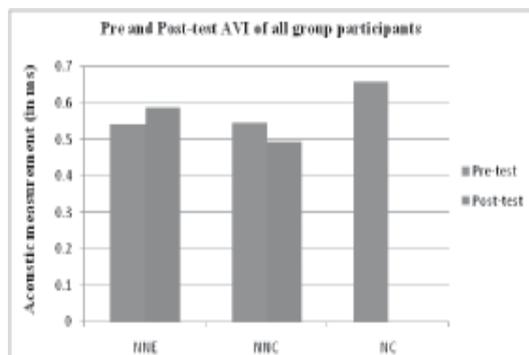


Figure 3: Pre and Post-test AVI of all group participants

### Class observation

The researcher wrote class observation journal at the end of each training session. Those class observation journals were analysed after the training was over. The class observation reflected that participants 110, 111, 113 and 118 seemed to have good English language knowledge and their spoken English was considered to be reasonably fluent. Additionally, participants 111, 113 and 118 were self-motivated and eager to learn in class; actively participating in each and every activity of their training. Hence, their post-test AVI was closer to native rhythm in acoustic measurement. However, participant 110 showed less motivation to learn from the training sessions. He hardly participated in group conversation as well as in other activities of the training and this was reflected in his marginal post-test improvement.

Participant 108 was considerably improved in the post-test; however, he had commenced from a low base with only average English knowledge and was also absent from one of the training sessions. Post-test acoustic measurement of his speech confirmed that he had non-native rhythm.

## Discussion and conclusion

Rhythm is primarily a perceptual property. Rhythmic distinctions suggested by measures of vocalic interval variability were supported by perceptual discrimination tasks (Ramus, Dupoux & Mehlers, 2003). White and Mattys (2007) also believed that the possible link between rhythmic classification and the realisation of prosodic timing processes was worthy of further cross-linguistic investigation.

This research is, therefore, considered to be an additional demonstration that differences in rhythm along a stress syllable-timed scale can be measured instrumentally. Further, a comparison of results for this study with previous research on the impact of training on the segmental or supra-segmental aspects of the English language indicates an obvious improvement after training. It concurs with the findings of Saito (2007), Burleson (2007) and Derwing, Munro & Wiebe (1997, 1998).

Aside from the training, two other reasons may have contributed towards the improvement of participants' rhythm in this study when compared with the above mentioned research. The prime reason for improvement in acoustic measurement could be the use of stimuli (Guided Speech). It is widely believed that acquired pronunciation knowledge can be used more effectively in guided speech because there is some extra time to think, and then apply knowledge, when reading script; however, less time is available when speaking spontaneously. Therefore, the participants of this study may have concentrated on pronouncing prominent words and stressed syllables with more conviction after they became aware of this from the training. Secondly, use of a computer-based assessment instrument in this study may have also contributed towards

improvement due to its objectivity. Any type of subjective assessment method may have generated different results.

Motivation of the learner has a vital role in second language learning (Dörnyei, 1998). From observing the participants in a classroom situation, it was evident that motivation, English language proficiency and individual attitudes may have played a major role in their performance. Regular participants in the training sessions achieved higher AVI when compared to the participant who missed a training session. Similarly, participants who actively participated in training activities outscored others who seemed less motivated. Moreover, English language proficiency was not considered in the selection of participants in this research; therefore participants who reported that they were students from the English faculty in pre-training questionnaire achieved AVI closer to the NS participant. These participants were of varying language proficiency, so a better, more equal selection of participants may have generated a more valid result. Additionally, it would have been easier to design the training lessons for participants of a similar language proficiency level and this again may have yielded a different outcome. Regardless, the decrease in AVI of the NNC participant is inexplicable. To my knowledge, there has not been any research which has examined the post-test decrement of control group members. This is an issue for future research.

The findings of this research have theoretical implications in relation to teaching rhythm in ESL and EFL classes. Acoustic measurement carried out in this study confirmed that training can play a major role in improving rhythm of English in the speech of the NNS. After examining the results of acoustic measurement from NNE participants, all had improved their rhythm of English after the training.

Regardless of the considerable deterioration by the control group participant, this research has provided evidence that rhythm of English can be improved through training.

The focus of this research was on teaching rhythm of English language; however, it is not maintained that teaching rhythm is the only pronunciation teaching aspect. Rather, this study should draw the attention of EFL and ESL teachers to the importance of teaching rhythm of English in their classes in conjunction with current pronunciation teaching methodologies. Thus, this study favours strong support for teaching rhythm in classes, especially where pronunciation teaching is still being basically ignored. In addition, it has also specified several methods and materials for teaching rhythm of English as well as evaluating strategies for the EFL and ESL teachers. Teachers are encouraged to teach rhythm of English in their classes by using controlled and communicative practices, as it will significantly enhance the pronunciation of their students. Moreover, this study has provided other language researchers a direction for further research in the future. Finally, EFL and ESL language policy makers from various nations, including Nepal, are encouraged to include teaching of rhythm in their curriculum and to prepare teacher training programs which includes this with the other aspects of pronunciation teaching.

There were some uncontrollable factors during this research which resulted in a limitation of this study. The length of time accorded to training sessions was considered one of its major limitations. Short training sessions were designed in consideration of the participants' availability of time. Longer training sessions may have resulted in a better outcome. Derwing, Munro and Wiebe (1998) designed 12 weeks training sessions

to examine accentedness and comprehensibility of their participants. Their participants demonstrated significant improvement after the training. Moreover, the sample size was too small in this study and a larger selection may have produced different results. A future study could be expanded by increasing both the training period as well as the number of participants. In the same way, there was no sufficient time gap between last training session and post-test recording in this study. The post-test recordings of this study were made a week after the last training session. As a result, it is believed that participants may have transferred their recently gained knowledge in the post-test. A delayed post-test may have generated more authentic result. Nonetheless, the researcher had to complete this study in limited time. Hence, it also provides a new direction for the future research.

More research on rhythm teaching should be conducted to establish the most effective ways of assessing methods and results as well as identifying the factors which contribute to same. Finally, objective assessment measures should be included where possible in order to achieve a more reliable outcome.

### **Acknowledgment**

I wish to acknowledge the continuous support and constructive feedback provided by my supervisor Dr. Haisheng Jiang during the process of writing this paper.

### **About the author**

*Ananda Aryal is a PhD candidate in Faculty of Education in the University of Tasmania, Australia. He is researching 'Self-efficacy and Motivation of Nepalese English as a Foreign Language Teachers.' He has research interest in pronunciation teaching. He is also particularly interested in researching issues of*

*English as a second language and English as a foreign language teaching.*

## References

- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Adams, C. (1979). *English Speech Rhythm and the Foreign Learner*. The Hague: Mouton.
- Adhikari, B. R. (2010). Teaching speaking in the Nepalese context: Problems and ways of overcoming them. *Journal of NELTA*, 15(1-2), 1-9.
- Anderson, P. (1993). Defining intelligibility parameters: The secret of sounding native. In *XXVII TESOL Annual Convention, Atlanta*.
- Anderson-Hsieh, J. (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes*, 9(3), 195-214.
- Anderson-Hsieh, J., & Venkatagiri, H. (1994). Syllable duration and pausing in the speech of Chinese ESL speakers. *TESOL Quarterly*, 28(4), 807-812.
- Auer, P., Couper-Kuhlen, E., & Muller, F. (1999). *Language in time: The rhythm and tempo of spoken interaction*. New York: Oxford University Press.
- Barrera-Prado, D. (2008). The reality of stress-timing. *ELT Journal*, 62(1), 11-17.
- Bhattarai, G. R., & Gautam, G. R. (2005). *English Language Teachers at the Crossroads*. *Journal of NELTA*, 10(1-2).
- Bista, K. (2011). Teaching English as a Foreign/Second Language in Nepal: Past and Present. *English for Specific Purposes World*, 32 (11).
- Brakel, A. (1985). Towards a morphophonological approach to the study of linguistic rhythm. *Chicago Linguistic Society*, 21, 15-25.
- Bray, E. (1995). *Using Limericks to help Japanese students improve their Pronunciation*. ERIC Document Reproduction Service (ED404862).
- Brooks, N. (1964). *Language and Language Learning: Theory and Practice* (2<sup>nd</sup> Ed.). New York: Harcourt Brace.
- Brown, G., & Yule, G. (1983). *Teaching the spoken language* (Vol. 2). Cambridge University Press.
- Burleson D. F. (2007). Improving intelligibility of non-native speech with computer assisted phonological training. *IULC Working Papers Online Volume*, 7.
- Buxton, H. (1983). Temporal predictability in the perception of English speech. In *Prosody: Models and Measurements* (pp. 111-121). Springer Berlin Heidelberg.
- Celce-Murcia, M., Brinton, D.M., Goodwin, J.M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.
- Chela-Flores, B. (1994). On the Acquisition of English Rhythm -Theoretical and Practical Issues. *Iral: International Review of Applied Linguistics in Language Teaching*, 32(3), 232-242.
- Chela-Flores, B. (2004). Optimizing the Teaching of English suprasegmentals. *Bells: Barcelona English language and literature studies*, (12), 5.
- Chen, C. F., Fan, C., & Lin, H. (1996). *A New Perspective on Teaching English Pronunciation: Rhythm*. ERIC Document Reproduction Service (ED397638).
- Couper-Kuhlen, E. (1993). *English Speech Rhythm: Form and Function in Everyday Verbal Interaction*. Amsterdam; Philadelphia: J. Benjamins.

- Cai, C.Y. (2008). The features and training of English stress and rhythm. *US-China Education Review*, 5 (11), 61-65.
- Dauer, R. (1983). Stress-timing and syllable timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Derwing, T. M., Munro, M. J. & Wiebe, G. (1997). Pronunciation Instruction for "Fossilized" learners: Can It Help? *Applied Language Learning*, 8 (2), 217-235.
- Derwing, T.M., Munro, M. J. & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48 (3), 393-410.
- Deterding, D. (2001). Letter to the editor The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics*, 29, 217-230.
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31(3), 117-135.
- Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80 (1), 95-108.
- Faber, D. (1986). Teaching the rhythms of English: a new theoretical base. *International Review of Applied Linguistics*, 24, 205-216.
- Gilbert, J. B. (1984). *Clear Speech: Pronunciation and listening comprehension in American English*. Cambridge, England: Cambridge University Press.
- Giri, R.A. (2010). English language teachers' resource centre: A model for developing contexts. *Journal of NELTA*, 15 (1-2), 64-76.
- Gong, J. (2002). Introducing English Rhythm in Chinese EFL Class rooms: A Literature Review. *Post-Script*, 3,(1).
- Hewings, M. (2007). *English Pronunciation in Use Advanced*. Cambridge: Cambridge University Press.
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109-131.
- Hood, R. B., & Dixon, R. F. (1969). Physical characteristics of speech rhythm of deaf and normal-hearing speakers. *Journal of Communication Disorders*, 2(1), 20-28.
- Jones, D. (1972). *An Outline of English Phonetics* (9th Ed.). Cambridge: Heffer.
- Jones, R. H. (1997). Beyond "Listen and Repeat": Pronunciation teaching materials and theories of second language acquisition. *Elsevier Science Ltd*, 25 (1), 103-112.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64(3), 459-489.
- Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.
- Khati, A. R. (2011). When and why of mother tongue use in English classrooms. *Journal of NELTA*, 16 (1-2), 42-51.
- Lin, H., Fan, C., & Chen, C. (1995). *Teaching Pronunciation in the Learner-Centered Classroom*. (ERIC Document Reproduction Service No. ED393292).
- Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: syllable-timing in Singapore English. *Language and Speech*, 43(4), 377-401.
- Low, E. L. & Grabe, E. (1995). Prosodic patterns in Singapore English. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm (Vol. 3, 636-639).

- Marks, J. (1999). Is stress-timing real? *ELT Journal*, 53 (3), 191-199.
- McQueen, J. & Cutler, A. (1997). 'Cognitive processes in speech perception'. In W.J. Hardcastle and J. Laver (Eds.), *The handbook of phonetic sciences*. Oxford: Blackwell.
- Morley, J. (1979). *Improving Spoken English*. Ann Arbor: University of Michigan Press.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-520.
- Munro, M. J. (2008). Foreign accent and speech intelligibility. In J. G. H. Edwards & M. L. Zampin (Eds.), *Phonology and second language acquisition* (pp. 193-218). John Benjamins Publishing.
- Murphy, J. (2003). Pronunciation. In D. Nunan (Ed.), *Practical English Language Teaching* (pp. 111-128). MacGraw-Hill.
- Nakatani, L. H., O'Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38(1-3), 84-106.
- Orton, J. (1994). The Demands of Intensity: Reflections on the Design of a 60-hour Intensive Advanced Chinese Course. In J. Gassin & M. Smith (Eds.), *Innovations in Language Teaching. The Horwood language centre occasional papers* (pp. 107-125). Melbourne: The University of Melbourne.
- Prator, C. (1971). Phonetics vs. phonemics in the ESL classroom: when isallophonic accuracy important? *TESOL Quarterly*, 5 (1), 61-72.
- Prator, C., & Robinett, B. (1985). *Manual of American English Pronunciation* (4th Ed.). New York: Holt, Reinhart and Winston.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15<sup>th</sup> international congress of phonetic sciences* (pp. 337-342). Barcelona.
- Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 72, 1-28.
- Saito, K. (2007). The Influence of Explicit Phonetic Instruction on Pronunciation Teaching in EFL settings: The Case of English Vowels and Japanese Learners of English. *The Linguistics Journal* 3 (3), 16-40.
- Savignon, S. (1983). *Communicative competence: Theory and classroom practice*. Reading MA: Addison-Wesley.
- Scarcella, R. & Oxford, R. L. (1994). Second Language Pronunciation: State of the Art in Instruction. *System*, 22 (2), 221-230.
- Setter, J. (2006). Speech Rhythm in World Englishes: The Case of Hong Kong. *TESOL Quarterly*, 40(4), 763-782.
- Sharma, U.N. (2003). *A Textbook of Applied Linguistics*. Kathmandu: Stud.
- Shrestha, P. N. (2008). ELT, ESP & EAP in Nepal: Whose interest are served? In K. Mark (Ed.), *EAP and ESP in developing countries: State of play vs actual needs and wants* (pp. 191-210). Canterbury: IATEFL (ESP SIG).
- Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25(1), 1-24.
- Taylor, D. (1981). Non-native speakers and the rhythm of English. *International Review of Applied Linguistics*, 19, 219-226.
- White, L., & Mattys, S. L. (2007). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.J. Sole (Eds.), *Segmental and prosodic issues in romance phonology (Current issues in linguistic theory series)*.

Amsterdam, Philadelphia: John Benjamins.

Widdowson, H. G. (1989). Knowledge of language and ability for use. *Applied Linguistics*, 10, 128- 137.

Wong, R. (1987). *Teaching Pronunciation. Focus on English Rhythm and Intonation*. New Jersey: Prentice Hall.

## Appendix-1

*Only bold sentences below were used for generating AVI of individual participant from the story **The North Wind and the Sun**.*

"I am stronger than you," said the North Wind to the Sun. "That is not true;" said the Sun. "Everyone knows that I am the stronger."

"Show me that you are stronger than I," replied the North Wind. "You know very well that you are not." "Do you see that traveller coming? I can make him take off his coat. You cannot," said the Sun.

"We will see about that," answered the North Wind. "The one that makes the traveller take off his coat is the victor." "All right," said the Sun, "and you may have the first trial."

"Whew! How the North Wind blows," said the traveller. "Whew! Whew! Hold on there, North Wind; I would rather walk than fly. Whew! Whew!"

"How cold it is! I must button my coat uptight! Whew! Whew! Whew! I never felt such a wind before," said the traveller, as he folded his arms over his chest. "It seems determined to tear off my coat. I will turn my back to it. Whew! Whew! Whew! Whew!"

At last the North Wind said, "I will try no longer, but you, Sun, can do no better."

The Sun said nothing but came out from under a cloud and smiled down upon the traveller. "How good that feels," said the traveller. The Sun shone on. "It is getting warm," said the traveller, unbuttoning his coat. It was now past noon. "The Sun is too much for me," said the traveller, and he threw off his coat and hunted for a shady place.

The North Wind's harshness had failed. The Sun's gentleness had won.