

THE SPEECH RHYTHM OF VIETNAMESE SPEAKERS OF ENGLISH

B. Byrne, A. Butcher, & P. McCormack

Department of Speech Pathology
Flinders University of South Australia

ABSTRACT - The durational characteristics of the speech rhythm of Vietnamese speakers of English was compared to that of native speakers of Australian English. The 2 aspects of English speech rhythm examined were compensatory shortening of stressed syllables when unstressed syllables are added to an interstress interval, and rhythmic stress shifts. Results showed several areas where the Vietnamese subjects' performance differed to that of the native subjects. They evidenced a lesser degree of shortening, or compression, of unstressed syllables in interstress intervals. In stress shift contexts, although marking shift, they exhibited a lesser degree of durational adjustment to the second stressed syllable. The results indicate an overall difficulty in marking durational contrasts with native-like proficiency.

INTRODUCTION

This study was concerned with the acquisition of English speech rhythm in people whose first language is Vietnamese. In the last decade, migrants from Vietnam have formed a small but not insignificant proportion of people arriving in Australia - approximately 7.5% of the total migrant intake between 1984 and 1995, over 86,000 in all. Vietnamese learners can have particular difficulty with this aspect of English, because its rhythmic characteristics differ widely to those of Vietnamese (Nguyen, 1970). Although speech rhythm is generally related to the timing features of a language, it is also closely related to stress patterns in English.

Vietnamese is usually described as a "syllable-timed" language (Nguyen, 1970; Honey, 1987). The syllable-timed nature of Vietnamese means that native Vietnamese speakers can be expected to have difficulty acquiring appropriate English stress and rhythm patterns. Nguyen (1970) suggested that prosody is probably the most difficult feature of spoken English for Vietnamese learners to master. Differences between the two languages' prosodic systems described by Nguyen include the fact that Vietnamese, as a tonal language, has no system of word stress. Further, sentence stress is more frequent in Vietnamese and is realised through marked differences in loudness. Weak stress is accompanied by very short syllables, with the overall effect that conversational Vietnamese is "syncopated". Although not directly involved in this study, it is pertinent to note that Vietnamese consists almost entirely of monosyllables and the consonant system, in particular, is very different to that of English (Honey, 1987). Hence, articulatory difficulties associated with unfamiliar consonants, consonant combinations and polysyllabic words, as well as the tendency to omit particular final consonants, may also contribute to Vietnamese speakers' difficulty with the rhythmic features of English. There has generally been little documented research on the acquisition of this aspect of English prosody by Vietnamese, or any other non-native speakers. Vietnamese learners of English have been described as tending to "transfer their entire speech habits in the process" (Nguyen, 1970, p.25). Thus the speech stress patterns and rhythm of the first language will, along with other features, tend to be transposed into English. One study which focused specifically on Vietnamese speakers of English is that of Pittam & Ingram (1992). Vietnamese Australians' production and perception of English compound and phrasal stress contrasts were considered. The results showed that subjects had difficulty with some contrasts and that the degree of ease or difficulty experienced was affected by phonological complexity. Length of residence in the country was found to have a high positive correlation with both production and perception tasks, indicating that contact with the native language plays an important role in the acquisition of such contrasts.

Despite the controversy surrounding the concepts of stress and syllable-timing, these terms continue to be used widely in the current literature and they will thus also continue to be referred to here. Although studies have failed to find evidence of absolute isochrony or measurable differences between language rhythms (Dauer, 1983; Roach, 1982), it is clear that languages do nevertheless differ in their rhythmic qualities, which reflect language-specific characteristics. Stress-timing is said to

be characterised by isochronous interstress intervals and variable syllable duration. Unstressed syllable reduction and compression, as well as stress shift, are identified features of English rhythm. The rhythm of syllable-timed languages, in contrast, is based on numbers of syllables in a rhythmic group, leading to more equal syllabic duration and more variable interstress intervals than in stress-timing. Research has, however, failed to find evidence of measurable durational differences between languages, or of isochrony. Some researchers continue to support a tendency towards isochrony in English, while the classification of languages based on timing distinctions remains widespread. Research has supported the proposal that perceived rhythmic differences reflect differences in linguistic structure between languages. Language-specific rhythmic features have been found to operate in speech processing, while studies have shown the important role of stress and rhythm to overall comprehension and intelligibility (Cutler & Mehler, 1993).

The aim of this study was to investigate the acquisition by native Vietnamese speakers of the durational characteristics of two aspects of English speech rhythm. Firstly, the widely reported adjustments to stressed syllable duration as unstressed syllables are added to an inter-stress interval (Fowler, 1977). This process is sometimes referred to as compensatory shortening. Secondly, the durational adjustments made to the second stressed syllable in words that undergo "stress shift" to avoid rhythmically unacceptable "stress clashes" (McCormack & Ingram, 1995).

METHOD

Subjects

In order to control English language competence as a variable, this study focused on Vietnamese people who had English skills adequate for tertiary level education in Australia. Investigation of these speakers would help clarify the degree to which the influence of native rhythm persists in people with a good command of English. Anecdotal evidence suggests that well-developed language skills do not necessarily lead to good speech intelligibility in non-native speakers. Two groups of subjects were recruited: adult native Vietnamese speakers and adult native English speakers. All subjects were enrolled in undergraduate or postgraduate courses at universities in Adelaide, or had completed tertiary studies in Australia prior to the study. Fifteen native Vietnamese speakers and twenty native English speakers volunteered to participate. The Vietnamese speakers comprised eight females and seven males and were aged from 17 to 50 years. Their age of learning English ranged from 7 years to 29 years and their length of residence in Australia at the time of the study varied between 1 month and 14 years. Native speakers consisted of twelve females and three males, aged between 17 and 47 years. All subjects reported having no difficulty speaking their native language.

Procedure

Hearing was assessed using standard audiometry. Samples of the subjects' speech were recorded for analysis, and were made in a quiet room using a Sony analogue tape recorder (model TCM 5000EV) and a lapel microphone. Thirty five sentences incorporating different patterns of syllable duration were presented in random order for the subjects to read. Subjects were instructed to read each sentence silently first and then to read it aloud three times in their normal speaking manner. Three repetitions were requested because a fluent production was required for analysis and it was assumed that by the third repetition the subject would be speaking more fluently and naturally.

Sentence type

A) Increasing interstress intervals. This was similar to the paradigm used by Mochizuki-Sudo & Kiritani (1991). Three sets of sentences (a total of fifteen) were constructed. Each consisted of sentences to which an extra unstressed syllable was progressively added. Interstress intervals ranged from one to five; that is, intervening unstressed syllables ranged from zero to four, in each set. For example, "We have ham for dinner" has no unstressed syllable between the target stressed syllables "have" and "ham"; the interstress interval is one. "We're having some ham for dinner" has two unstressed syllables between the target stresses, with an interstress interval of three. Native productions would be expected to show shortening of the initial stressed vowel and reduction, or compression, of unstressed syllables with increasing numbers of unstressed syllable in the interstress interval.