

Some Stress Related Acoustic Features of Disyllabic Words in Mandarin Chinese

Chiu-yu Tseng

The aim of the present study is to investigate acoustic features related to stress patterns of disyllabic structure in Mandarin Chinese (MC). The stress patterns under investigation include the stress-weak (SW) pattern and the normal-stress (NS) pattern in both non-emphatic and emphatic speech produced in citation forms. (The SW patterns include both morphemically determined weak stress as well as lexically determined weak stress.) Test tokens also include disyllabic words that could be produced with either stress pattern. The acoustic features being examined include fundamental frequency (F0) patterns, syllable duration and amplitude. The subject (S) was a female native speaker of Peking MC who had resided in Taiwan for nearly 40 years. Overall results of the S's performance regarding non-emphatic speech with no knowledge of the nature of the experiments yield a preference towards the NS pattern, i. e., 75% of the morphemically determined SW pattern was produced with the NS pattern, and 54% of the lexically determined SW pattern was produced with the NS pattern. 72% of the tokens that can be pronounced with either the SW or NS pattern was produced with the NS pattern. Nevertheless, S's performance changed towards the SW pattern after learning about the aim of the study. For both morphemically and lexically determined weak stress, 100% of the tokens were produced with the SW pattern, whereas 96% of the test tokens that can be pronounced with either stress pattern was produced with the SW pattern. However, test tokens in all conditions were again produced with the preference of the NS pattern in emphatic speech, namely, 93% of the emphatic data. Analyses of the F0 patterns include overall F0 shape, and the syllable beginning, middle and final frequency in all conditions. Mean syllable duration of the test tokens in each condition and the ratio of the two syllable in each condition were derived. The amplitude of syllables across all test tokens was also derived. It is found that the major difference in terms of acoustic features between non-emphatic and emphatic speech lies in the domain of syllable duration. A general tendency for MC spoken in Taiwan to adopt the NS pattern over the SW pattern is also discussed.

I. Introduction

The objectives of this study are mainly to examine the stress patterns of disyllabic words in Mandarin Chinese (hereafter MC) from an acoustic

phonetic perspective. The scope of study includes the so-called normal vs. weak stressed distinction as well as the neutral tone (Chao, 1968:35); while the acoustic features under investigation include the fundamental frequency (hereafter F0) pattern, syllable duration and the corresponding amplitude.

Existing literature on stress related phenomena and the neutral tone in MC generally falls into the following five types of studies. The first type aims at describing the phenomena from an articulatory phonetic and somewhat phonological perspective. The most detailed can be found in Chao (1968:35-39; 1980). The norm of study is MC by native speakers from the Peking area. Similar studies, though mainly more of presentation of examples rather than discussion of the phenomena can be found in Jiang (1956) and Shiu(1956). The second type of study is of more theoretical concern, emphasizing the phonemic nature of lexical tones, and the acoustic properties of tones. This type of study can be represented by Wang & Li(1967). The third type of study consists of more recent acoustic measurements of weak-stress related phenomena. This type of work can be found in Howie (1976); Lin, Lin, Shia & Tsau (1980); Lin and Yan (1980); Yin (1982), Wu (1982b) and Lin (1983). The fourth type of study proposes a so-called mid-level-tone, which was derived from phonetic measurements from production data, to be the general phonetic output of weak-stress and neutral tone (Hu, 1988), but does not address to weak-stress and/or neutral tone related problems in linguistic or extra-linguistic context. The fifth type of work differs from the above-mentioned studies in the following two aspects, namely: (1.) This type of study either points out the fact that MC has been going through considerable phonological and phonetic change due to contact with Southern Chinese dialects, and is thus of more sociolinguistic interests (Kubler, 1979, 1985b; Chen, 1984; 1986); (2.) The foci of these studies emphasize the rather confusing norm of stress patterns and stress related phenomena in MC, thus proposing a new way to look at MC. Chen (1984) argues that there exists considerable

difference of opinion regarding weak-stress/neutral tone in both various written sources as well as the speech of native speakers via a paper-and-pencil test, and is thus worth more elaborate re-examination.

The present study differs from the above-summarized five types of studies in the following aspects: (1.) The reported study includes only disyllabic words produced in isolated, read form and is thus more limited in scope. (2.) Although the results reported here are from one native speaker born and raised in the Peking area till adulthood, the fact that the subject has resided for nearly four decades in Taiwan was taken into consideration. Whether the speech patterns of such speakers have changed due to long-term language/dialect contact is of particular interest to the investigator. In other words, the subject here undoubtedly is able to speak the so-called Peking MC, and producing all weak-stress/neutral tone by the rules of Peking phonology, but just how much of it is still retained in actual speech has been sampled and measured. The motivation of such a study stems from rather independent considerations: The first being to further examine some acoustic properties related to stress patterns in MC. The second being to apply the results to the synthesis of Mandarin speech, especially in the area of improving the speech output. The scope of the study will be limited at the level of the actual phonetic output, and subsequent discussion will also not involve other linguistic levels that might also trigger a syllable to become weak stressed and thus neutralized on the surface.

II. The problem

Following Chao's definition (1968:35)¹, it is quite clear that stress, in

1 Chao states (1968:35): "...Stress in Chinese is primarily enlargement in pitch range and time duration and only secondarily in loudness. Thus when a 3rd Tone is stressed it is dipped lower, and when a 4th Tone is stressed it starts higher and falls lower. Stress, therefore, can be pictured by stretching the tone graph on an elastic background, as described above [Sec. 1. 3. 4(1)]. There are physically many perceptible degrees of stress, but phonemically we have found it best to set up no more than three degrees of stress: normal, contrastive, and weak..."

the strict sense, is an acoustic phenomenon that carries phonetic and phonemic function. The three degrees of phonemic stress he sets up are normal, contrastive and weak stresses. However, what makes the stress phenomena rather confusing on the surface is the fact that the weak stress appears to be indistinguishable from the neutral tone in strict phonetic sense. That is, by definition, in weak stress the lexical tone is reduced to a neutral tone, thus sharing the characteristics of neutral tone. In other words, weak stressed lexical tones become, theoretically, subject to the same sandhi rules that neutral tones are, which specifies that the preceding stressed lexical tone conditions the following neutral tone. By this definition, weak stressed lexical tone, as the neutral tone, should not be in any position to trigger sandhi rule that causes modification of tone of the preceding syllable. Nevertheless, as Chen (1984) pointed out, such is not the case at all. There exists neutral tone in certain forms that does not trigger a sandhi change, e. g., *jiějie* 'elder sister', *yǐz* 'chair', etc.; whereas the neutral tone in other forms does, e. g., *shǐǎujie* 'unmarried woman', *tzǒutzou* 'to take a walk', etc. (Chen, 1984). Moreover, there are stressed neutral-toned syllables as well as weak stressed toned syllables (Chen, 1984). Chen argues that neutral tone and weak stress are two separate entities, and do not interact in terms of cause and effect. In MC, there are a handful of grammatical morphemes which lack tone at both the underlying and the surface level, and are usually weak-stressed. Even when they are normal-stressed, they do not trigger sandhi changes. Such morphemes include suffixes, particles and reduplicated kinship reference terms. On the other hand, when a toned syllable is weak-stressed, the tonal characteristics may become less distinct, but such syllables are still viable in triggering sandhi changes. She stressed that toned syllables may fluctuate between normal and weak stress in accordance with various factors such as sentence intonation, familiarity of the term, emotion, etc, and attributed the chaotic surface instability also to dictionary listings as well. For the present study, the focus is not on

further elaboration of the chaotic surface level, nor the forward and/or background sandhi changes that occur, but rather on the actual phonetic output. The subject under investigation is a controlled situation, being a native speaker of Peking MC. That is, the tendency for even neutral toned morphemes to take up lexical tones is the point of interest here. Consider, if, by definition, lexical tones are phonemic whereas stresses are not, could it be that stress is thus freer to vary, resulting the coexistence of the many surface forms of phonetic output of weak stress and neutral tones. Or, could it also be that stress is undergoing changes due to language and dialect contact and is thus predictable, as Kubler (1979, 1985a) has observed with respect to grammatical structures in MC? Cheng (1985) presented a detailed comparison of stress patterns between what he termed Peking dialect and Taiwanese in which he classified stress patterns into four types of neutralization, namely lexically determined neutral tone, morphemically determined neutral tone, syntactically determined neutral tone, and finally semantically focused neutral tone. He states that the main difference regarding stress patterns between these two dialects are in the domain of lexically determined neutral tone, which occur frequently in Peking MC (about 10% to 20% of the most commonly used multisyllabic words) but almost do not exist in Southern Min. The problem addressed in this study is precisely this particularly aspect of MC, which I believe has been going through some considerable changes, moving from MC to Southern Min in Taiwan. Because the interest lies in the actual speech output, the focus will be on the acoustic side of the problem, rather than discussion involving other linguistic levels that might also interplay with the phenomena.

III. Methodology

The experiment here was to assess three acoustic properties related to stress patterns in MC, namely the F0 pattern, syllable duration and the corresponding amplitude. Citation forms were the only forms used for the

study. Because one of the purposes of the study was to derive stress rules for synthetic speech output, one speaker was acceptable. The subject (hereafter S) was a 60-year-old female speaker of MC born and raised in Peking with college education, currently working as a teacher of MC to foreign students. S has lived in Taiwan since approximately 1950, and is a monolingual MC speaker.

Three groups of disyllabic MC words were selected for the experiment. These tokens were extracted from *Guóyǔrèbaà Tsǎdiǎn* 'Mandarin Daily Dictionary' (1974) and consulted by a native speaker of MC from Peking who writes teaching materials for foreign students of the Chinese language. The consultant did not serve as a subject of the study. The first group (Group A) of tokens consists of 20 disyllabic morphemes that could be analyzed as combination of two monosyllabic morphemes, as a large portion of disyllabic words are. According to the two sources consulted, these tokens should be read only with the stress-weak (hereafter SW) pattern, with the first syllable being the stressed syllable, and the second syllable receiving the weak stress. To borrow Cheng's (1985) term, these are morphemically determined neutral tone. For instance, *pútāu* 'grape', *pípa* 'loquat' or 'a string instrument', etc. The second group (Group BI) consists of 100 tokens of MC disyllabic words which could have both the normal stress reading, i.e., both syllables have neither weak nor contrastive stress; and the weak stress reading, i.e., the second syllable receive the weak stress. In other words, such tokens could be produced phonetically either with the next-strong-then-strong stress pattern or the strong-weak stress pattern. For instance, *jàngren* or *jàngren* 'father-in-law', *dūngshì* or *dūngshì* 'thing, object', etc. The third group (Group BII) consists of 100 MC disyllabic words that possess the SW pattern as those in Group A, but differs in the sense they are not always disyllabic morphemes, but compound words that may be analyzed as two morphemes combined, and the stress pattern can be determined at the level of the lexical item itself.

To borrow Cheng's (1985) term again, these would be lexically determined neutral tone. For instance, *jǎnliang* 'food capacity, appetite', *tóufa* 'head hair, hair on the head', etc. Except for Group A which consists only 20 tokens, the test tokens from the other two groups were further broken into two group of 50 tokens and presented on two different sheets of paper. S was presented with test tokens in 5 groups, with one group of 20 tokens and four groups of 50 tokens, before the first recording session and was given enough time to study the items before recording. Three recording sessions were administered, and the orders of the five groups of test tokens were varied for the first two recording sessions. For the first recording session, S was simply asked to read each token from five sheets of paper in citation forms, that is, with a pause between each token. One reading of 220 tokens were obtained. For the second recording session, S was presented with the test tokens in the order of Groups A, BI and BII. And S was now told that the purpose of the project was to study stress patterns, and examples of both the normal stress and the weak stress were provided by the experimenter. S was then asked to read the tokens in the stress patterns again accordingly, if she was able to. Special interest was devoted to test tokens in Group BI, which consists of disyllabic-words tokens that can be produced either with the SW pattern or the normal-stress(hereafter NS) pattern. Another reading of 220 tokens were obtained. The third recording session was scheduled during which S was asked to read the 100 tokens in emphatic form, but again using her preferred stress pattern. One reading of 100 tokens were obtained. Altogether, 520 tokens were the data collected for the study. The recording was done at S's office on three different days approximately a week apart, using a UHER 4400 reel-to-reel tape recorder, and a Beyernamic M420 N(C) microphone. The data were then analyzed using a Digital VAX 11/730 computer and the Interactive Laboratory System (ILS) software. F0 pattern, syllable duration, and amplitude information were obtained.

IV Results and analysis

The results of the S's production in terms of stress patterns is summarized in Table 1.

Table 1. S's production regarding stress patterns of disyllabic words in MC.

designated stress patterns S's production		Group A (stress-weak)	Group BI (stress-weak or normal)	Group BII (stress-weak)	Group BIF (stress-weak or normal)
stress-weak (SW)	before instruction	75%	28%	46%	/
	after instruction	100%	96%	100%	7%
normal-stress (NS)	before instruction	25%	72%	54%	/
	after instruction	/	4%	/	93%

Note: / in table 1 denotes that no tokens of the particular stress pattern could be found. S's production before instruction, that is, S's voluntary choice of preferred reading yields that for Group A, morphemically determined SW pattern, 75% of the tokens were produced with the normal stress pattern. For Group BI, disyllabic words that can be produced either with the SW pattern or the NS pattern, 28% of the tokens were produced with the SW pattern, whereas 72% of the tokens were produced with the NS pattern. For Group BII, disyllabic words that involve lexically determined neutralization on the second syllable and consequently should be produced with the SW pattern, only 46% of the tokens were produced with the SW pattern, whereas 54% of the tokens were produced with the NS pattern. However, after being told that the aim of the study was to examine stress of disyllabic words in MC and given production examples of what is

meant by SW pattern and the NS pattern, the S's production yields the following results: For Group A, disyllabic morphemically determined neutral tone that results the SW pattern, all the tokens, i.e., 100% of the tokens, were produced with the SW pattern. For Group BI, disyllabic words that can be produced with either pattern, 96% of the tokens were produced with the SW pattern, whereas 4% were produced with the NS pattern. For Group BII, disyllabic words that are lexically determined SW pattern, 100% were produced with the SW pattern. As for emphatic reading regarding test tokens in Group BI (BIF in Table 1), i.e., disyllabic words that can be produced in either stress pattern, 7% of the tokens were produced with the SW pattern, whereas 93% were produced with the NS pattern.

Three acoustic properties that bear direct relationship with stress patterns were examined, namely, F0 pattern, duration and amplitude. To investigate the F0 patterns, the mean F0 patterns for the four lexical tones in MC across the two stress patterns were examined and derived, i.e., the SW vs. the NS patterns, is collapsed and summarized in Tables 2 and 3. The range of the F0 pattern in terms of their respective syllable beginning F0, syllable middle F0, and syllable ending F0 are reported, together with the mean value of these points.

IV 1. Results and analyses regarding the F0 Patterns

Results and analyses regarding the F0 patterns are presented with respect to the three stress patterns under investigation.

IV 1.1. The SW pattern

The tokens that were actually produced with the SW pattern were analyzed and summarized in Table 2. This means that the phonetic output was the only variable considered, regardless of their respective reasons of neutralization. In the table both the stressed syllable and the weak syllable

from the SW pattern were included.

Table 2. Analysis of disyllabic words produced with the stress-weak patterns. The range and mean frequency of each lexical tone at the syllable beginning, syllable middle and syllable final position are derived.

		syllable beginning F0		syllable middle F0		syllable ending F0	
		range	m	range	m	range	m
stressed syllable	Tone 1	96-286 Hz	194 Hz	137-213 Hz	188 Hz	115-200 Hz	182 Hz
	Tone 2	76-204 Hz	160 Hz	141-178 Hz	156 Hz	145-217 Hz	175 Hz
	Tone 3	100-189 Hz	158 Hz	111-146 Hz	133 Hz	96-145 Hz	121 Hz
	Tone 4	87-244 Hz	199 Hz	119-202 Hz	164 Hz	72-189 Hz	133 Hz
weak syllable		74-217 Hz	146 Hz	72-212 Hz	135 Hz	63-217 Hz	128 Hz

Figure 1. a. and Figure 1. b. are the graphic plot of the results from Table 2. For Tone 1 that occurred as stressed syllables from the SW pattern, the beginning frequency ranges from 96Hz to 286Hz, with the mean value of 194Hz across all such tokens; the syllable middle frequency ranges from 137Hz to 213Hz, with the mean value of 188Hz; and the syllable ending frequency ranges from 115Hz to 200Hz; with the mean value of 182Hz. Thus the overall shape of the F0 pattern is a high level one. For Tone 2 that occurred as stressed syllables from the SW pattern, the beginning frequency ranges from 76Hz to 204Hz, with the mean value of 160Hz; the syllable middle frequency ranges from 141Hz to 178Hz, with the mean value of 156Hz; and the syllable ending frequency ranges from 145Hz to

217Hz, with the mean value of 175Hz. Thus the overall shape of the F0 pattern displays a mid beginning and a rise of approximately 21Hz (between 156Hz to 175Hz) from the middle of the syllable, namely, the second half of the syllable. For Tone 3 that occurred as stressed syllable from the SW pattern, the beginning frequency ranges from 100Hz to 189Hz, with the mean value of 158Hz; the syllable middle frequency ranges from 111Hz to 146Hz, with the mean value of 133Hz; and the syllable ending frequency ranges from 96Hz to 145Hz, with the mean value of 121Hz. As a result, the overall shape of the F0 pattern, with the beginning at approximately the same frequency height of Tone 2, displays a gradual fall from the beginning of the syllable, rather than the traditionally termed dipping shape. For Tone 4 that occurred as stressed syllable from the SW pattern, the beginning frequency ranges from 87Hz to 244Hz, with the mean value of 119Hz; the syllable middle frequency ranges from 199Hz to 202Hz, with the mean value of 164Hz; and the syllable ending frequency ranges from 72Hz to 189Hz, with the mean value of 133Hz. Thus Tone 4 begins at approximately the same frequency height as Tone 1, and displays a rather sharp fall of approximately 66Hz (between 199Hz to 133Hz). The last column of Table 2 reports the results of the range as well as the mean values of all the weak-stressed syllables across all samples collected, disregarding the possible changes of F0 shape due to the preceding lexical tone. The reason was because the data were collected under natural room setting when the recording sessions were performed, making it impossible to rid all background noises. This resulted in the weak-stressed syllables to become almost indistinguishable from the background noises. The ILS software was thus incapable to produce reliable F0 information from such tokens. Nevertheless, mean value regarding available data was summarized in Table 2. Figure 1. b. also plots the F0 pattern of the available data regarding weak stressed syllables across the test conditions.

The beginning frequency of all weak syllable from the SW pattern

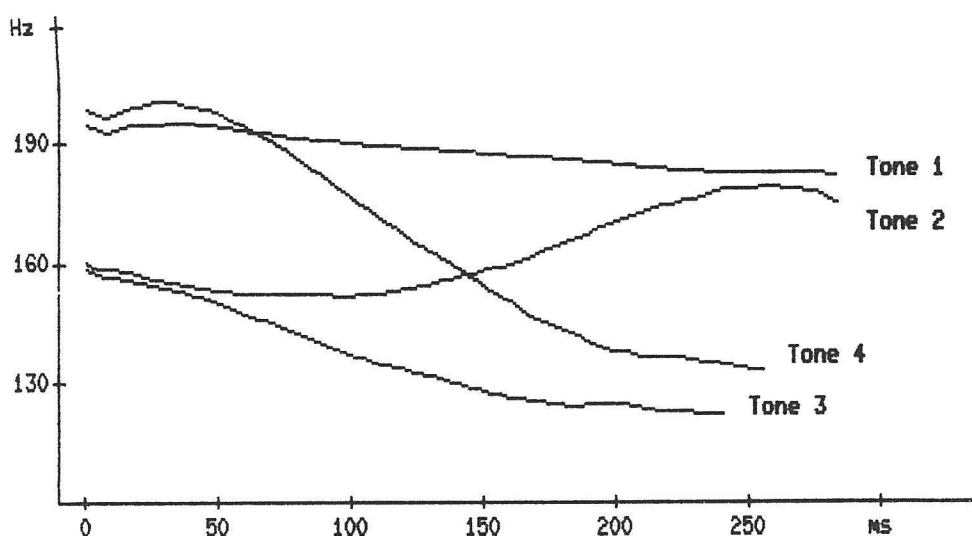


Figure 1. a. Mean F0 patterns (vowel portions only) derived from the stressed syllables of disyllabic words produced with the stress-weak (SW) pattern with respect to the four lexical tones in MC.

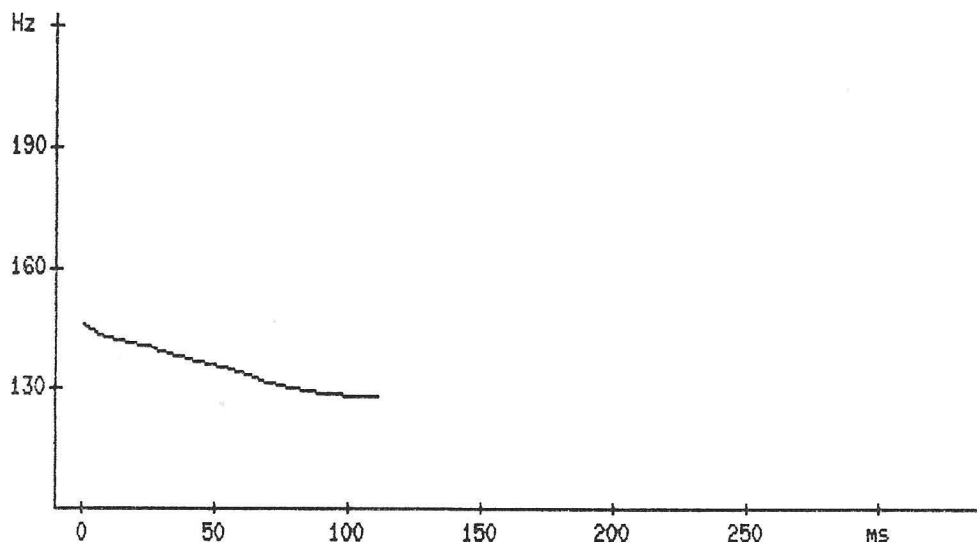


Figure 1. b. Mean F0 pattern (vowel portions only) derived from the weak syllable of disyllabic words produced with the stress-weak (SW) pattern.

ranges between 74Hz and 217Hz, with the mean value of 146Hz; the mid syllable frequency ranges between 72Hz and 212Hz, with the mean value of 135Hz; and the syllable ending frequency ranges between 63Hz to 217Hz,

with the mean value of 128Hz. The overall F0 shape of the weak syllable is a relative low beginning frequency height followed by a gradual fall of approximately 18Hz(between 146Hz to 128Hz). Note that such a syllable also possesses a much shorter duration (See following discussion on corresponding syllable duration.).

IV 1.2. The NS pattern

Table 3 summarized the F0 patterns of the four lexical tones in MC in NS conditions across the data collected.

Table 3. Analysis of disyllabic words produced with the normal-stress pattern. The range and mean frequency of each lexical tone at word initial (1st syllable) and word final (2nd syllable) positions are derived.

		syllable beginning F0		syllable middle F0		syllable ending F0	
		range	m	range	m	range	m
Tone 1	1st syllable	81-238 Hz	198 Hz	149-200 Hz	184 Hz	164-196 Hz	182 Hz
	2nd syllable	112-196 Hz	172 Hz	147-182 Hz	166 Hz	114-189 Hz	155 Hz
Tone 2	1st syllable	76-179 Hz	156 Hz	136-170 Hz	151 Hz	147-189 Hz	169 Hz
	2nd syllable	82-182 Hz	140 Hz	115-136 Hz	124 Hz	93-149 Hz	128 Hz
Tone 3	1st syllable	145-179 Hz	162 Hz	110-149 Hz	132 Hz	115-159 Hz	127 Hz
	2nd syllable	84-185 Hz	137 Hz	80-147 Hz	116 Hz	63-169 Hz	118 Hz
Tone 4	1st syllable	98-256 Hz	204 Hz	143-185 Hz	163 Hz	115-159 Hz	140 Hz
	2nd syllable	132-208 Hz	176 Hz	88-184 Hz	148 Hz	75-156 Hz	120 Hz

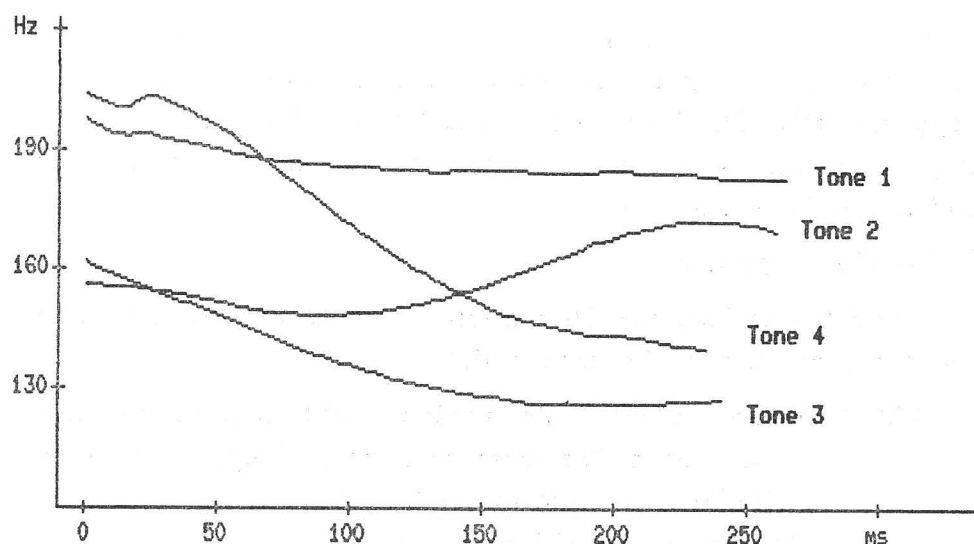


Figure 2. a. Mean F0 patterns (vowel portions only) derived from the stressed syllables (1st syllables) of disyllabic words produced with the normal-stress (NS) pattern with respect to the four lexical tones in MC.

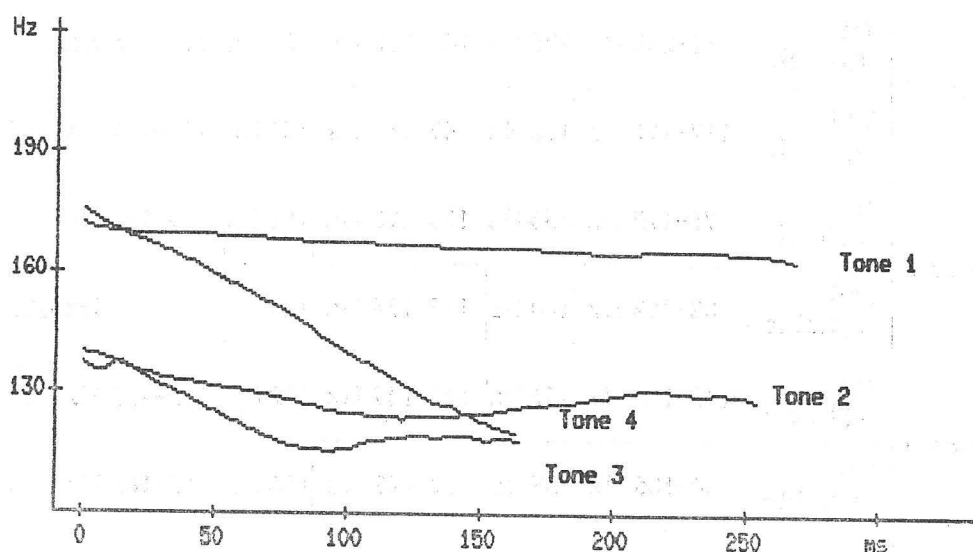


Figure 2. b. Mean F0 patterns (vowel portions only) derived from the next-to-stressed (2nd syllables) of disyllabic words produced with the normal-stress (NS) patterns with respect to the four lexical tones in MC.

Table 3 is the results of the F0 range and mean values of the four lexical tones when occurred in the first syllable of the NS pattern, as well as the same tones when occurred in the second syllable of the NS disyllabic words. Figures 2. a. and 2. b. plots the results from the collected data regarding the NS pattern.

For Tone 1 that occurred in the first syllable of a disyllabic normal-stressed word, the beginning frequency ranges from 81Hz to 238Hz, with the mean value of 198Hz; the syllable middle frequency ranges from 149Hz to 200Hz, with the mean value of 184Hz; the syllable ending frequency ranges from 164Hz to 196Hz, with the mean value of 182Hz. The overall shape of Tone 1 is again high level at approximately the same height as Tone 1 that occurred as the stressed syllable of a SW disyllabic word. For Tone 2 that occurred in the first syllable of a disyllabic NS word, the beginning frequency ranges from 76Hz to 179Hz, with the mean value of 156Hz; the syllable middle frequency ranges from 136Hz to 170Hz, with the mean value of 151Hz; the syllable ending frequency ranges from 147Hz to 189Hz, with the mean value of 169Hz. The overall shape of Tone 2 exhibited a somewhat greater dip than that of stressed Tone 2 in a SW word (See Figure 1. a.) before the rising pattern started. For Tone 3 that occurred in the first syllable of a disyllabic NS word, the beginning frequency ranges from 145Hz to 179Hz, with the mean value of 162Hz; the syllable middle frequency ranges from 110Hz to 149Hz, with the mean value of 132Hz; and the syllable ending frequency ranges from 115Hz to 159Hz, with the mean value of 127Hz. The overall shape of Tone 3, like its counterpart in the SW condition, is a gradual fall that begins at approximately the same starting frequency as Tone 2. Once again, the traditionally termed mid-falling-rising tone is a clear “half Tone 3 (benching in MC)”, that is, without the rising portion of Tone 3. For Tone 4 as the first syllable of a NS disyllabic word, the beginning frequency ranges from 98Hz to 256Hz, with the mean value of 204Hz; the syllable middle frequency ranges from

143Hz to 185Hz, with the mean value of 163Hz; and the syllable ending frequency ranges from 115Hz to 159Hz, with the mean value of 140Hz. The overall shape of Tone 4 also displays a relatively sharper fall of approximately 64Hz (between 204Hz and 140Hz), as its stressed counterpart in a SW word.

Table 3 also summarizes the results of the second syllable in NS disyllabic words with respect to tones. For Tone 1, the syllable beginning frequency ranges from 112Hz to 196Hz, with the mean value of 172Hz; the syllable frequency ranges from 147Hz to 182Hz, with the mean value of 166Hz; the syllable ending frequency ranges from 114Hz to 189Hz; with the mean value of 155Hz. The overall shape of Tone 1 as the second syllable in a NS disyllabic word is also a level tone, but at a lower frequency height of approximately 20Hz. For Tone 2, the syllable beginning frequency ranges from 82Hz to 182Hz, with the mean value of 140Hz; the syllable middle frequency ranges from 115Hz to 136Hz, with the mean value of 124Hz; and the syllable ending frequency ranges from 93Hz to 149Hz, with the mean value of 128Hz. The overall shape of Tone 2 as the second syllable in a NS disyllabic word displays no distinct rising. Note that the mean of syllable ending frequency (128Hz) is lower than the syllable beginning frequency (140Hz) by 12Hz. Note also the mean of syllable middle frequency is lower than the syllable ending frequency (124Hz vs. 128Hz), though the difference might not be a significant one. Nonetheless, it appears that the rising contour of Tone 2 is not maintained. For Tone 3, the syllable beginning frequency ranges from 84Hz to 185Hz, with the mean value of 137Hz; the syllable middle frequency ranges from 80Hz to 147Hz, with the mean value of 116Hz; and the syllable ending frequency ranges from 63Hz to 169Hz, with the mean value of 118Hz. The overall shape of Tone 3 as the second syllable in a NS disyllabic word displays a fall of 21Hz (between 137Hz to 116Hz) at approximately 90msec into the syllable, followed by slight rise (See Figure 2.b.). Note that the duration

of Tone 3 is considerable shorter. For Tone 4, the syllable beginning frequency ranges from 132Hz to 208Hz, with the mean value of 176Hz; the syllable middle frequency ranges from 88Hz to 184Hz, with the mean value of 148Hz; and the syllable ending frequency ranges from 75Hz to 156Hz, with the mean value of 120Hz. The overall shape of Tone 4 as the second syllable in a NS disyllabic word displays a distinct falling pattern for approximately 56Hz (between 176Hz to 120Hz). Again, when Tone 4 occurred in the second syllable of a NS disyllabic word, its frequency height is about 28Hz (between 204Hz to 176Hz) lower than Tone 4 in the first syllable of the same stress pattern.

IV 1. 3. The emphatic reading of the SW pattern

Table 4 summarizes the F0 patterns of the four lexical tones from the emphatic reading of tokens that were the ambiguous cases, i.e., disyllabic strings that could be produced with either stress pattern, that is, tokens from Group BI. Note that 93% of the tokens were produced with the NS pattern. Thus Table 4 reports rather the F0 information of the majority of the data.

Table 4. Analysis of 93% of the disyllabic words from Group BI, i.e., disyllabic tokens of the stress-weak (SW) pattern. The range and mean frequency of each lexical tone at word initial (1st syllable) and word final (2nd syllable) positions are derived.

		syllable beginning F0		syllable middle F0		syllable ending F0	
		range	m	range	m	range	m
Tone 1	1 st syllable from normal-stress	192-233 Hz	207 Hz	175-208 Hz	190 Hz	164-196 Hz	170 Hz
	2 nd syllable from normal-stress	156-196 Hz	175 Hz	154-189 Hz	171 Hz	139-189 Hz	183 Hz

Tone 2	1 st syllable from normal-stress	143-167 Hz	156 Hz	130-169 Hz	152 Hz	137-200 Hz	174 Hz
	2 nd syllable from normal-stress	114-182 Hz	141 Hz	112-156 Hz	128 Hz	90-156 Hz	136 Hz
Tone 3	1 st syllable from normal-stress	152-182 Hz	166 Hz	128-145 Hz	136 Hz	109-143 Hz	123 Hz
	2 nd syllable from normal-stress	111-196 Hz	146 Hz	85-126 Hz	110 Hz	82-139 Hz	120 Hz
Tone 4	1 st syllable from normal-stress	104-250 Hz	202 Hz	148-213 Hz	168 Hz	116-189 Hz	139 Hz
	2 nd syllable from normal-stress	125-208 Hz	180 Hz	130-187 Hz	158 Hz	95-161 Hz	126 Hz

Figures 3. a. and 3. b. is the corresponding graphic plot of the results of analysis of the F0 patterns for the tokens of the emphatic readings, i. e., the results summarized in Table 4.

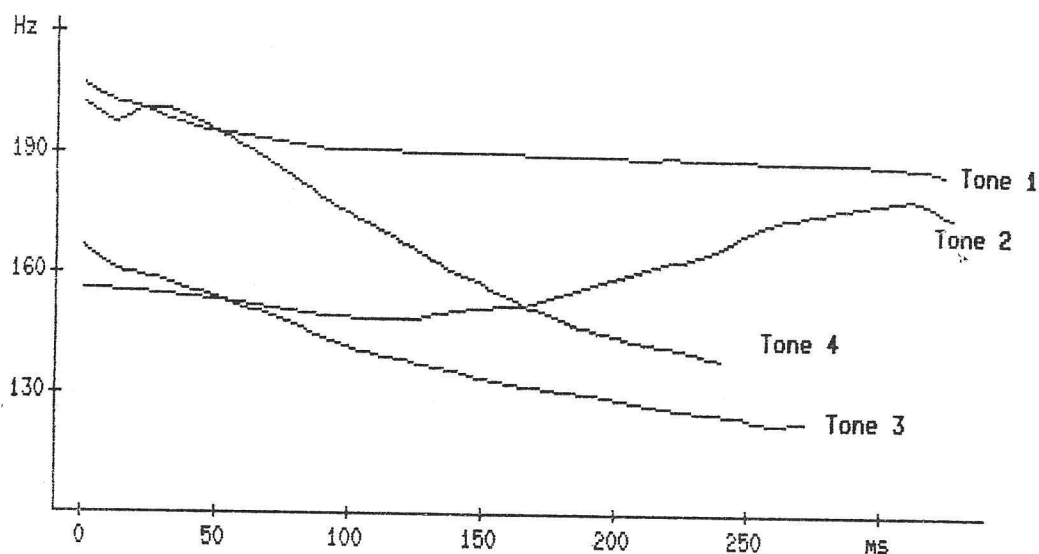


Figure 3. a. Mean F0 patterns (vowel portions only) derived from the stressed syllables (1st syllables) of emphatic reading of disyllabic words in MC produced with the normal-stress (NS) pattern.

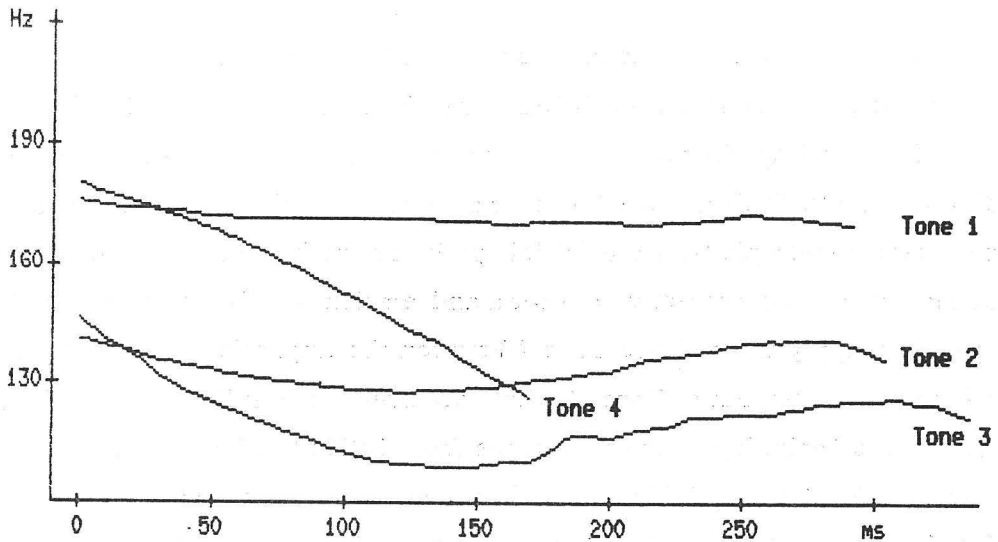


Figure 3. b. Mean F0 patterns (vowel portions only) derived from the next-to-stressed syllables (2nd syllables) of emphatic reading of disyllabic words in MC produced with the normal-stress (NS) pattern.

For emphatic Tone 1 that occurred as the first syllable of disyllabic words, the syllable beginning frequency ranges from 192Hz to 233Hz, with the mean value of 207Hz; the syllable middle frequency ranges from 175Hz to 208Hz, with the mean value of 190Hz; and the syllable ending frequency ranges from 164Hz to 196Hz, with the mean value of 170Hz. The overall shape of emphatic Tone 1 at the word initial position of a disyllabic word is nearly identical with word initial Tone 1 in both SW as well as the NS conditions. Thus, the difference may lie in the domain of duration and amplitude. For emphatic Tone 2 that occurred as the first syllable of disyllabic words, the syllable beginning frequency ranges from 143Hz to 167Hz, with the mean value of 156Hz; the syllable middle frequency ranges from 130Hz to 169Hz, with the mean value of 152Hz; and the syllable ending frequency ranges from 137Hz to 200Hz, with the mean value of 174Hz. The overall shape of emphatic, word initial Tone 2 is a more gradual rise from approximately syllable middle that stretches over a longer duration. For emphatic Tone 3

at word initial position, the syllable beginning frequency ranges from 152Hz to 182Hz, with the mean value of 166Hz; the syllable middle frequency ranges from 128Hz to 145Hz, with the mean value of 136Hz; and the syllable ending frequency ranges from 109Hz to 143Hz, with the mean value of 123Hz. Such a Tone 3 does not differ much from Tone 3 in the above two non-emphatic word initial positions with respect to F0 shape, but maybe in their respective duration and amplitude. The same relationship appears to apply to emphatic Tone 4 between the emphatic vs. non-emphatic positions. For emphatic Tone 4 that occurred in disyllabic word initial position, the beginning frequency ranges from 104Hz to 250Hz, with the mean value of 202Hz; the syllable middle frequency ranges from 148Hz to 213Hz, with the mean value of 168Hz; and the syllable ending frequency ranges from 116Hz to 189Hz, with the mean value of 139Hz, although the emphatic Tone 4 appears to be slightly shorter than its non-emphatic counterpart. Table 4 also summarizes analysis of the F0 information regarding the four MC lexical tones that occurred in the second syllable of disyllabic words. For emphatic second-syllable Tone 1, the syllable beginning F0 ranges from 156Hz to 196Hz, with the mean value of 175Hz; the mid-syllabic frequency ranges from 154Hz to 189Hz, with the mean value of 171Hz; and the syllable ending frequency ranges from 139Hz to 189Hz, with the mean value of 183Hz. For emphatic second-syllable Tone 2, the syllable beginning frequency ranges from 114Hz to 182Hz, with the mean value of 141Hz; the syllable middle frequency ranges from 112Hz to 156Hz, with the mean value of 128Hz, and the syllable ending frequency ranges from 90Hz to 156Hz, with the mean value of 136Hz. For emphatic second syllable Tone 3, the syllable beginning frequency ranges from 111Hz to 196Hz, with the mean value of 146Hz; the syllable middle frequency ranges from 85Hz to 126Hz, with the mean value of 110Hz; and the syllable ending frequency ranges from 82Hz to 139Hz, with the mean value of 139Hz. For emphatic second syllable Tone 4, the beginning frequency ranges from 128Hz to 208Hz, with

the mean value of 180Hz; the syllable middle frequency ranges from 130Hz to 187Hz, with the mean value of 158Hz; and the syllable ending frequency ranges from 95Hz to 161Hz; with the mean value of 126Hz.

IV 2. Results and analyses regarding syllable duration

Table 5 summarizes results of both the SW pattern and the NS pattern from the production data regarding syllable duration.

Table 5. Mean syllable duration of disyllabic words produced with the stress-weak(SW) pattern and the normal-stress (NS) pattern.

	sample #	m of syllable duration
stressed syllable from stress-weak	285	380 msec
weak syllable from stress-weak	199	266 msec
1st syllable from normal stress	132	339 msec
2nd syllable from normal stress	120	381 msec

For tokens that were produced with the SW pattern, the mean syllable duration of stressed syllable is 380 msec; whereas the mean syllable duration of weak syllable is 266 msec, yielding a ratio of 1:0.7. For tokens that were produced with the NS pattern, the mean syllable duration of the first syllable is 339 msec; where as the mean syllable duration of the second syllable is 381 msec, yielding a ratio of 0.89:1.

Table 6 summarizes results of the emphatic reading from the production data. The test tokens were the same tokens from Group BI, i.e., MC disyllabic words that possess the SW pattern.

Table 6. Mean syllable duration of disyllabic words produced in emphatic reading with the stress-weak (SW) pattern and the normal-stress (NS) pattern.

	sample #	m of syllable duration.
stressed syllable from stress-weak	7	429 msec
weak syllable from stress-weak	7	344 msec
1st syllable from normal-stress	93	425 msec
2nd syllable from normal-stress	93	440 msec

For token that were produced with the SW pattern, the mean duration of the stressed syllable is 429 msec; whereas the mean duration of the weak syllable is 344 msec, yielding a ratio of 1:0.8. For tokens that were produced with the NS pattern, the mean duration of the first syllable is 425msec; whereas the mean duration of the second syllable is 440 msec, yielding a ratio of 0.97:1.

II. 3. Results and analyses regarding amplitude.

By using the rms energy function provided by the ILS software, the strength of a digitized speech signal can be represented graphically. Figures 4. a. and 4. b., to be viewed as the amplitude plot of Figures 1. a. and 1. b., plots the mean of energy measurement of stressed syllables with respect to tones from the SW pattern, as well as the mean of energy of measurement of weak syllables. Note that all the energy plot appears to have longer duration than their F0 equivalent. This is due to the fact that the F0 measurements are derived from the vowel portion of the syllables only, while the energy measurements are derived from the entire syllable.

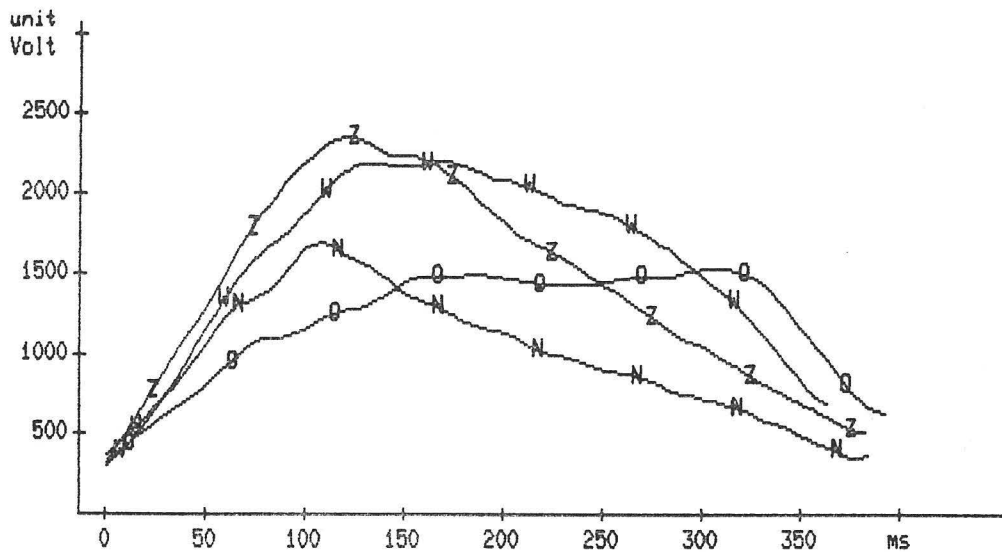


Figure 4. a. rms energy measurement of stressed syllables from the stress-weak (SW) pattern with respect to tones. W, O, N, and Z represents lexical tones 1, 2, 3 and 4 respectively.

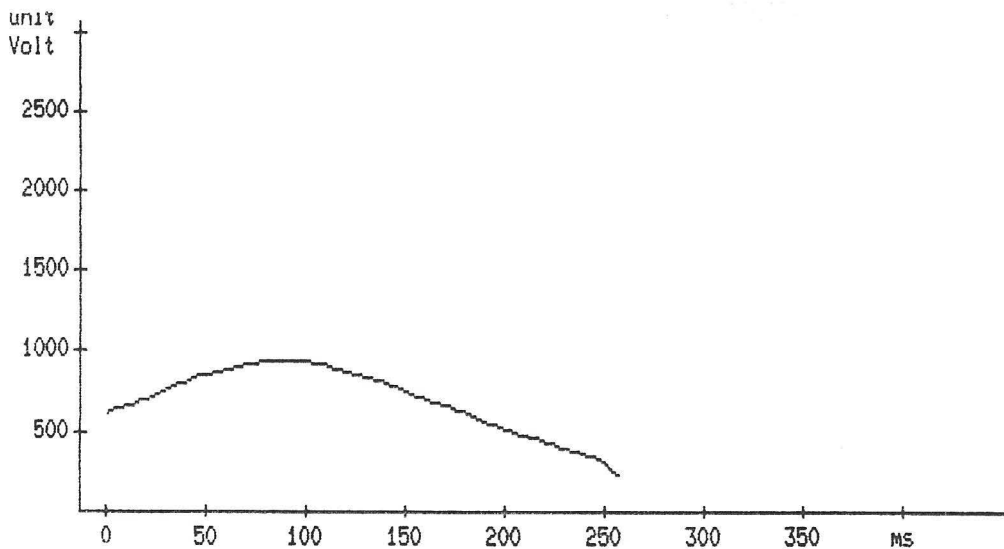


Figure 4. b. rms energy measurement of weak syllables from the stress-weak (SW) pattern.

Figures 5. a. and 5. b., to be viewed as the amplitude plot of Figures 2. a. and 2. b., is the mean of energy measurement of the first vs. the second syllable from the NS pattern.

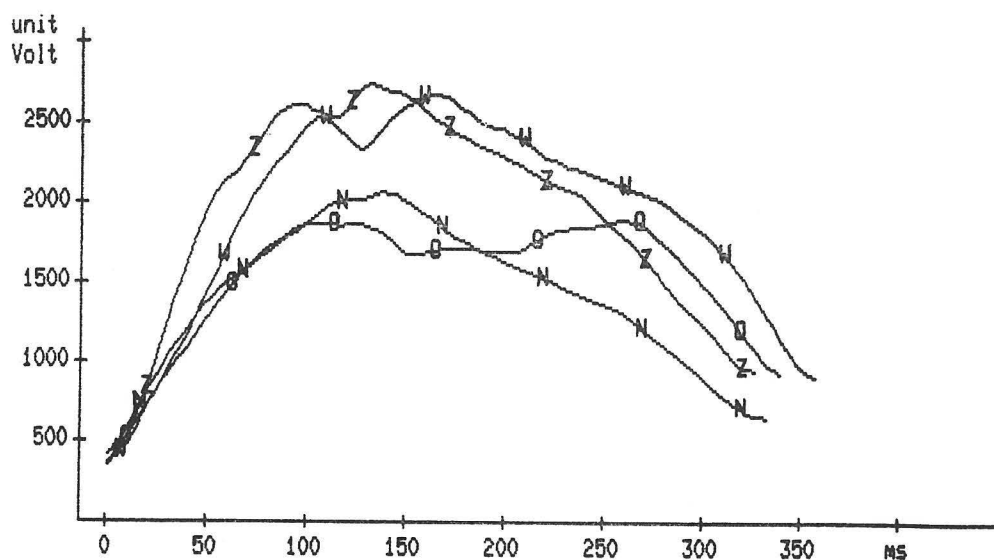


Figure 5. a. rms energy measurement of stressed syllables(1st syllables) from the normal-stress (NS) pattern with respect to tones. W, O, N and Z represent lexical Tones 1, 2, 3 and 4 respectively.

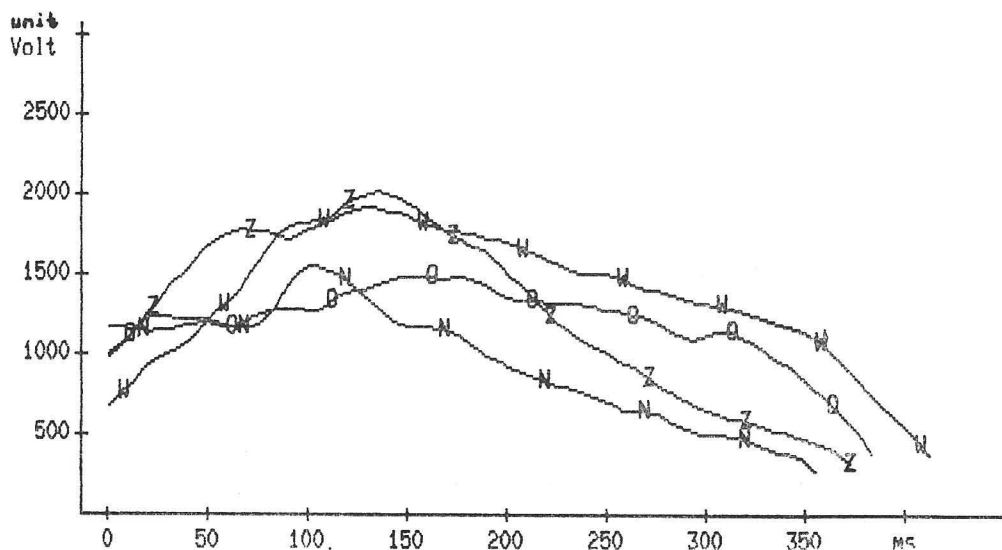


Figure 5. b. rms energy measurement of next-to-stressed syllables (2nd syllables) from the normal-stress (NS) pattern with respect to tones. W, O, N and Z represent lexical Tones 1, 2, 3 and 4 respectively.

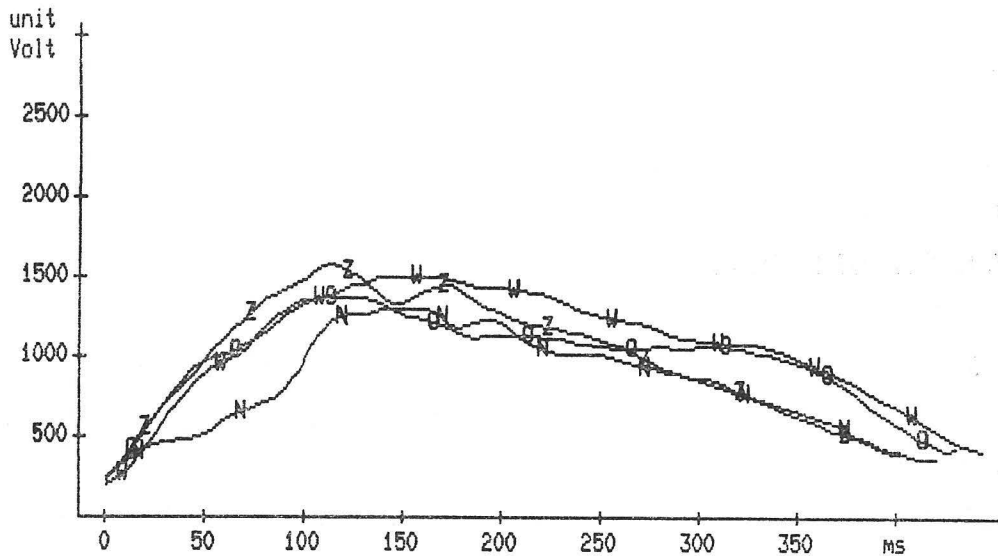


Figure 6. a. rms energy measurement of emphatic reading of stressed syllables (1st syllables) from the normal-stress (NS) pattern with respect to tones. W, O, N and Z represent lexical Tones 1, 2, 3 and 4 respectively.

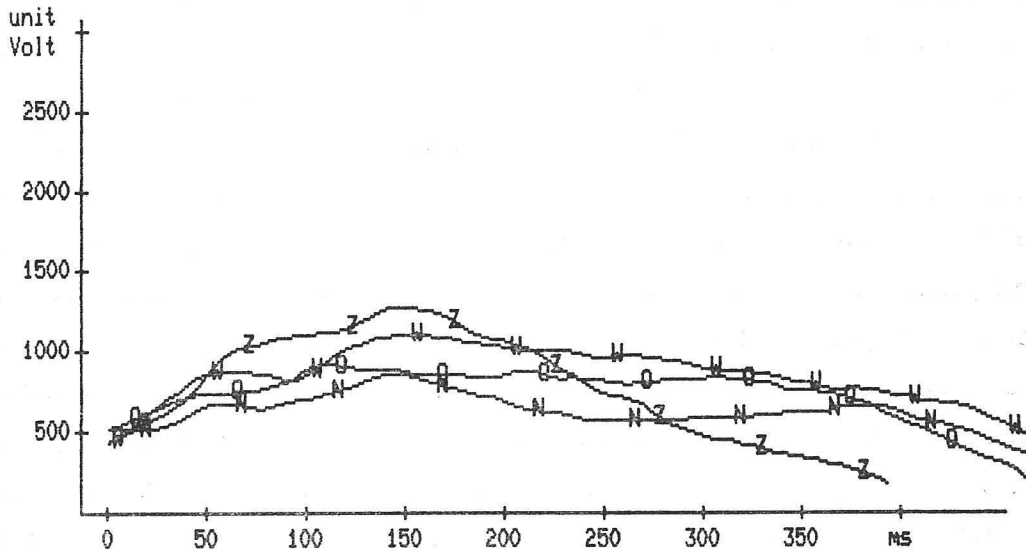


Figure 6. b. rms energy measurement of emphatic reading of next-to-stressed (2nd syllables) from the normal-stress (NS) pattern with respect to tones. W, O, N and Z represent lexical Tones 1, 2, 3 and 4 respectively.

Figures 6.a. and 6.b., to be viewed as the amplitude plot of Figures 3.a. and 3.b., is the mean of energy measurement of the first vs. the second syllable from the emphatic reading that was produced with the emphatic NS pattern rather than the emphatic SW pattern.

V Discussion

V.1. General findings

The acoustic properties examined in this study are lexical tone shapes (F0 patterns), syllable duration and energy measurements of the stress patterns of MC disyllabic words. The data demonstrate that for MC disyllabic word initial syllable, regardless of the stress pattern, the tonal patterns are very similar in terms of F0 contour, syllable beginning frequency and syllable ending frequency (See Tables 2 & 3). The overall tonal shapes are also maintained, with Tone 3 being *bānshǎng* 'half shang or half Tone 3', that is, without the rising portion of the dip. The possible difference between these two types of word initial syllable, if there is any, is then in the domain of duration and amplitude. In this case, the data show that the difference of mean syllable duration of the word initial syllable for the SW pattern is 380 msec, while the mean syllable duration for the first syllable of the NS pattern is 339 msec. A discrepancy of 41 msec in mean syllable duration is found. Note however, the first syllable of the SW pattern is the stressed syllable, whereas the first syllable of the NS pattern is not. If comparison is made between the stressed syllables for both stress patterns, the data demonstrate difference in all three acoustic features being examined. In terms of F0 patterns, there is a distinct difference in frequency height as well as overall tone shape (See Figures 2.a. and 2.b.). Duration and amplitude also vary. For F0 patterns, the difference can be found in overall frequency height and shape of the contour. For frequency height, the second syllable of the NS word is around 21Hz lower than its word initial counterpart. Such first syllables include the stressed syllable in the SW pattern as well as the next-to-stressed stress

in the NS pattern. For overall tone shape, Tone 1 maintained the general level shape. The shape of Tone 2 does not display a slight dip instead of a distinct rise. The beginning frequency and ending frequency differ in 12Hz, i.e., between 140Hz and 128Hz, with the ending frequency being the lower end. The lowest point is also not the beginning frequency, but rather somewhere in the middle. This is largely due to the fact that all the data collected were produced in citation form, making the second syllable of each disyllabic word in the study a breath-group (Lieberman, 1967) final syllable. As a result, the frequency height is lowered, duration expended, and amplitude tapered off. For Tone 3, a somewhat sharper dip than the above described Tone 2 is found, again with the ending frequency lower than the beginning frequency, that is, 137Hz VS. 118Hz, and a dip of approximately 21Hz, that is, from 137Hz to 116Hz. For Tone 4, a distinct fall is displayed, namely, from 176Hz to 120Hz. Note that the main difference for normal-stressed second syllables is in the domain of duration (See Tables 3) and amplitude (See Figures 5). Tables 2 & 3 further break down the data by tones and present the mean syllable duration by tones as the stressed syllable in the SW pattern and the two syllables in the NS pattern. Note the inconsistent relationship regarding duration to tones in breath-group final position. The second syllable Tone 1 in normal-stressed pattern is considerably longer than Tone 1 in breath-group initial position, be it a stressed one from a SW pattern or the next-to-stressed one from a NS pattern. But Tones 2, 3 and 4 display an opposite relationship in that they are shorter in breath-group final position. The case is further displayed in Figures 2. a. and 2. b. where only the vowel portion of the syllable is plotted for pitch extraction. It is clear that both Tones 3 & 4 are distinctly shorter in vowel duration. To further investigate the situation, it would be necessary to control the consonants for the disyllabic words examined, and will not be included in the current study. What is clear is that within each stress pattern, the SW pattern goes without discussion, the

ratio between the stressed and the weak or the stressed and next-to-stressed is consistent across conditions, namely, the stressed syllables being the longer one. For amplitude, it is evident that the stressed syllable is not necessarily the one produced with more energy. The physiological effect of breathing clearly overrides stress pattern. For the SW pattern, since it coincides with the breath-group, the stressed syllable is the syllable produced with more energy. But for the NS pattern, since it is contrary to the breath-group, the stressed syllable, i.e., the second syllable, is produced with less energy coupled with longer duration. This can be at least one explanation that although the vowel portion in such syllables might be shorter, the entire syllable is not necessarily shorter as well. More detailed analyses regarding consonant duration of MC disyllabic words will be dealt with in another study, which may enable the author to better explain duration as an acoustic feature in MC. However, for the weak stress in the SW pattern, the mean syllable duration of 266 msec is derived, and a ratio of 1:0.7 is found between the stressed and weak syllable. This ratio is approximately the same in Lin (1983) for synthetic speech, but disagree with the results found by Lin and Yan (1980) since their data yielded an approximately 1:0.5 ratio. Since Lin and Yan's 2 subjects and the subject for the present study speak essentially the same dialect, and Lin's (1983) norm of speech is also Peking MC, a tentative reason for the discrepancy in duration can best possibly be attributed to individual differences since neither study has a large subject pool. Nevertheless, it is fairly certain that both duration ratios are acceptable for Peking MC, which can be seen as evidence that duration ratio is somewhat flexible and thus is not a primary acoustic feature for stress patterns. In other words, it may not be acceptable for a weak stressed syllable to be longer in duration than a stressed syllable, and the same may also apply to the NS pattern in the sense that the next-to-stressed syllable may not be longer in duration than the stressed syllable, but the ratio could be

somewhat fuzzy. If this is indeed the case, syllable duration may also be a feature that is susceptible to change.

The emphatic data of the SW pattern yield two kinds of stress patterns in actual production. The NS pattern is the preferred choice, since 93% of the data were produced with this pattern, despite the fact that all tokens included for emphatic reading are of the SW pattern. The tonal information for such tokens were clearly demonstrated in their respective F0 shape (See Figures 3. a. and 3. b.). The breath-group effect is again demonstrated in the emphatic data, that is, lowering the overall frequency height of the second syllable, this time by 22Hz, showing that the range of frequency lowering to breath-group final position is approximately constant within a speaker regardless of modes of speech. Since only 7% of the tokens were produced with the SW pattern, in this case 7 tokens altogether and not covering every one of the four tones, no quantitative analysis was performed for F0 patterns. It is very clear that in emphatic speech, the NS pattern is the mode of speech, although it is possible to emphasize the SW pattern. In terms of syllable beginning frequency height, by comparing the emphatic NS pattern with the non-emphatic counterpart, no consistent pattern was found. The mean beginning frequency of emphatic Tone 1 (207Hz) is the only frequency height that is higher than both the non-emphatic stressed Tone 1 of SW pattern (194Hz) and the first syllable Tone 1 of NS pattern (198Hz). However, no distinct frequency height difference is found for the other three tones (See Tables 2, 3, & 4). No consistent frequency height difference is found throughout the entire tones, thus making frequency height a likely concomitant cue for emphatic reading. As for duration information, the emphatic reading is consistently longer in duration than the non-emphatic counterpart, and the ratio within each stress pattern is also different from the non-emphatic reading. For the emphatic SW pattern, the mean duration of stressed syllable (429 msec) is 49 msec longer than the non-emphatic stressed syllable (380 msec). The mean

duration of emphatic weak syllable (344 msec) is 78 msec longer than the mean duration of the non-emphatic weak syllable (266 msec). Similar results can be found in the comparison between the mean duration of emphatic first syllable of the NS pattern (425 msec) and the non-emphatic first syllable (339 msec); and the mean duration of emphatic second syllable (440 msec) and the non-emphatic second syllable (381 msec). A difference of 86 msec and 59 msec is found respectively. Within each stress pattern, the ratio of syllable duration between the emphatic stressed syllable ($m=429$ msec) and the emphatic weak syllable ($m=344$ msec) is 1:0.8; whereas the ratio between the emphatic first syllable of the normal-stressed tokens ($m=425$ msec) and the emphatic second syllable ($m=440$ msec) is 0.97:1 (See III.2). The ratios obtained from the emphatic data is not significantly different from the ratios obtained from the non-emphatic data, which shows that the ratio between syllables within each stress pattern is generally maintained regardless of modes of speech. But the main characteristic of emphatic apparently lies in the domain of syllable duration, because the amplitude of the emphatic syllables, for this matter the normal-stressed ones, is seen to be stretched longer due to the effect of duration (See Figures 6.a. and 6.b.). This is particularly salient in the weak syllable that were produced in emphatic SW pattern and the first syllable of the emphatic NS pattern, that is, 78 msec and 86 msec longer than their non-emphatic counterpart. In other words, in emphatic speech, it is the weak and next-to-stressed syllable that receive longer syllable expansion than their respective stressed mates. In consequence, the discrepancy of syllable duration within each stress pattern is reduced, giving the effect that both syllables are not very much different in terms of duration. Thus, the rhythm of speech might be quite different effect. As a result, it appears that syllable duration is the primary cue to achieve the effect of emphasis in speech for disyllabic words produced in citation forms.

V.2 Modification of stress pattern in Taiwan

Results from the S's overall production suggest that S has a tendency to choose the NS pattern as the preferred stress pattern. Aside from the acoustic phonetic aspect, the data also demonstrate that S's production performance before and after instruction was rather inconsistent. Before instruction, for Group A, 25% of the test tokens were produced with the NS pattern; and for Group BI, 28%, and Group BII, 54%. That is, for both morphemically and lexically determined weak stress, approximately one fourth of the tokens were produced with the NS pattern. It is especially interesting to observe tokens in Group BII, tokens whose SW pattern involve lexically determined neutral tone, because the decision between choosing the SW pattern and the NS pattern appears to be split in half. In other words, the tokens bear a near half chance to be produced with either stress pattern. S's preference of the NS over the SW pattern is even more obvious for tokens in Group BI, i.e., tokens that may be produced with either stress pattern. A high 72% was produced with the NS pattern, displaying the S's preference of the NS pattern over the SW pattern once again. However, S's production after instruction appears to bear some self-consciousness to overproduce the SW pattern. That is, after S was given production examples of both stress patterns, S's production was then more of the SW pattern. For both Groups A and BII, S produced all the tokens, i.e., 100%, with stress-pattern. Whereas for Group BI, 96% were produced with the SW pattern. This suggests that the NS pattern has become rather the unmarked stress pattern. The same phenomenon is further evidenced in emphatic reading of words that involved lexically determined weak stress. A mere 7% was produced with the SW pattern, showing that the SW pattern can be produced even for emphatic reading. But the fact that an overwhelming 93% were produced with the NS pattern in emphatic reading clearly suggests that the NS pattern is the preferred stress pattern for emphatic reading. Consider the fact that S is a native speaker of Peking MC in which both stress patterns exist and the reduction of a weak

stressed lexical tone to a neutral tone applies to a significant portion of the most commonly used words (Cheng, 1985), and the fact that the NS pattern is the stress pattern for Southern Chinese dialects, especially Southern Min which is the dominant dialect in Taiwan, S's preference of the NS pattern reveals some insights regarding the prosodic features of speech patterns in MC spoken in Taiwan. As Cheng (1985) observed, the main difference regarding stress patterns between MC and Southern Min lies in weak stress in the sense that the SW pattern is a much more active pattern in MC rather than in Southern Min. He pointed out specifically that there are approximately 800 multisyllabic words in MC that involves lexically determined neutral tone, such as tokens in Group BII, which constitutes around 10% to 20% of the mostly commonly used vocabulary. But such lexically determined neutral tones are very few in Southern Min. In fact there exists so few of such words that Cheng claims that there is almost no lexically determined neutral tones in Southern Min, Taiwanese among them naturally. This feature is one of the features carried over to spoken MC by speakers in Taiwan who may or may not be native speakers of Taiwanese. The inconsistency of S's production performance thus suggests that S has developed a code switching strategy that is rather situation conditioned. It is quite clear that S's mode of speech, when it is Peking Mandarin, is somewhat different, if not drastically, from that of the society she belongs to, and she has since taken up some of the traces of the speech pattern that the majority of speakers possess. Kubler (1985a, 1985b) has observed some similar conditions in terms of grammatical structures as well as the lexicon. Chen's studies (1984, 1986) noted the discrepancies between the phonology and the speaker's intuition reflected by her test. The author would argue that there are by now a significant population of native speakers of MC who may or not speak a second dialect, and who may or may not even speak sufficient Southern Min, but the MC they speak is certainly not the same as Peking MC. The stress patterns of MC

spoken in Taiwan have gone through considerable modifications, mostly likely from the influence of stress patterns of Southern Min, and the norm of MC in Taiwan is, in my opinion, of some distance from MC spoken around the Peking area (cf. Cheng, 1985). For one thing, textbook MC taught at schools does not seem to emphasize stress patterns at all. Except the few neutral tone morphemes, morphemically determined weak stress is not emphasized, nor is lexically determined weak stress. What is emphasized, on the contrary, is the lexical tone of each individual morpheme. Thus for those speakers whose native tongue is not Peking MC, but MC spoken by people who use stress patterns from Southern Min, it is quite logical that very little attention was ever given to stress patterns of Peking MC. As a result, there exists a significant population who may not be native speakers of Southern Min, but whose stress patterns are distinctly from that of Southern Min. Such may also be a salient feature of MC spoken in the Taiwan area. Such an outcome is, and should be expected, given the fact that the so-called guoyu 'national language' is by its nature an artificial product by now, if not in the beginning stage of becoming one. To have chosen one dialect as the national language means that a large portion of the population whose native language is anything other than MC have to learn MC as a second dialect. This in turn removed MC from its original geographical location into direct contact with many dialects that it normally would not. The MC spoken in Taiwan thus has most constant and long term contact with Southern Min. Modification of stress patterns is but one small facet of the changes that MC has been going through in Taiwan.

VI Conclusion

In this paper, production analysis of MC disyllabic words produced in citation forms in both emphatic and non-emphatic forms was reported. Three acoustic features that are directly related to MC stress patterns were examined, namely, F0 patterns, syllable duration and amplitude. For stressed

syllables in both stress patterns (SW and NS), lexical tone information was fairly straight forward for all three acoustic properties examined, and discussion with respect to the two stress types was given. Unfortunately, no detailed analysis was available regarding the weak stress and neutral tone due to the limitation of the implementation techniques used. The main difference between emphatic and non-emphatic appears to be in the domain of syllable duration rather than a wider F0 range or more energy information. However, it is found that in current MC spoken in Taiwan, the NS pattern is the preferred stress pattern over the SW pattern. Such is the case in both non-emphatic and emphatic speech. As a result, weak-stressed lexical tones that are supposed to be reduced to neutral tone and subsequently become subject to sandhi rule specification regarding neutral tone tend to take up the stressed form, maintaining more tonal information. The ratios between stressed and weak syllables or the two syllables of the NS pattern are also getting closer, giving the effect of a different pattern of rhythm. In short, the prosodic aspect of current MC appears to be somewhat different than the way prosody was described in earlier works. It seems fair to claim that MC spoken outside the Peking area has been undergoing changes due to constant contact with other dialects; stress pattern being but one aspect of many possible ones. This is evidenced from a native speaker of Peking MC who has resided in Taiwan for the longest part of her life, and is expected to be more so for younger generations who were born and raised on the island. In turn, it may not be unlikely for the two stress patterns to merge, leaving the lexically determined neutral tones intact and the morphemically determined neutral tone to become exceptions in the phonology. Moreover, even the morphemically determined neutral tone may also take up lexical tones in the future. Subsequent studies will be conducted among younger age groups to further test the prediction, as well as to discover phonetic characteristics of current MC.

Appendix

I. (Group A)

玫瑰	戒指	哈蟆	胳膊	掃帚	葡萄
枇杷	蹄膀	狐狸	玻璃	葫蘆	琵琶
螃蟹	馬虎	哆嗦	簸箕	螞蚱	里肌
啞巴	眼睛				

II. (Group B I)

分量	燒餅	燒賣	雞眼	招呼	受用
照顧	男人	女人	丈人	故事	過年
煎餅	下水	下子	小月	準頭	炒麪
生意	生氣	造化	裁縫	爲難	委屈
運氣	告訴	造就	隨和	開化	見證
口頭	江湖	自然	自由	神氣	勾當
打點	大夫	大方	地上	地下	地方
地道	姑娘	王道	東西	東家	便當
四海	規矩	老實	裹腳	綁腿	寬綽
冷戰	火燒	頑皮	提防	提拔	土地
土氣	單薄	官司	內人	家務	過去
生性	財主	霸道	鋪陳	買賣	門道
方丈	扶手	打手	挑剔	拿手	來往
利害	龍頭	幹事	跟前	姑爺	工夫
公道	管子	開發	考究	口舌	交代
老子	下場	秀氣	香烟	兄弟	出息
出身	勢利	人性	算計		

III. (Group B II)

膽量	口袋	腦袋	服侍	玉米	蝦米
寡婦	見識	毛病	舒服	笑話	佩服
俐落	閨女	香甜	進去	錯處	鋪蓋
鑰匙	鈴鐺	飯量	翻騰	妥當	拿開

教訓	兆頭	念頭	行頭	後頭	行家
腩腴	囉唆	囑咐	噤咕	嘮叨	打開
解開	情分	情面	使喚	便宜	頭髮
抖擻	俏皮	包袱	打聽	迷糊	潑辣
福分	脾氣	和氣	老氣	腥氣	福氣
名堂	外頭	風頭	留下	留難	踮躑
巴掌	巴結	架式	搗肢	將就	妯娌
嫁妝	妖精	姊夫	妹夫	喜好	薄荷
荷包	朦混	芝麻	花消	花稍	芥末
鼓搗	苔帶	蓼花	莊稼	含糊	念叨
街坊	街上	街道	腳錢	臉面	胳膊
周到	房錢	皮匠	皮膚	力量	力氣
力錢	女家	嬌貴	作坊		

References

- Chao, Yuen-Ren 1933. "Tone and intonation in Chinese". *Bulletin of the institute of history and philology*, VI, 121-134.
- Chao, Yuen-Ren 1956. "Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese". *For R. Jakobson*, 52-59, The Hague: Mouton.
- Chao, Yuen-Ren 1968. *A Grammar of Spoken Chinese*, Berkeley: University of California Press.
- Chao, Yuen-Ren 1980. "Chinese tone and English stress". in L.R. Waugh and C. H. van Schooneveld. *The Melody of Language*, Baltimore: University Park Press.
- Chen, Chung-yu 1984. "Neutral tone in Mandarin: Phonotactic description and the issue of the norm". *Journal of Chinese Linguistics*, Vol. 12, No. 2, 299-333.
- Chen, Chung-yu 1986. "Salient segmental features in Singapore Mandarin". *Journal of Chinese Linguistics*, Vol. 14, No. 1, 114-151.

- Cheng, Robert L. 1985. "Tsung shehueiyuyanshiue he yuyanjiayude guan-dian kan Beiping kouyu he Taiwanhua jrjian chingjungyinde yitung. (A comparison of the stress patterns in Peking dialect and Taiwanese from the sociolinguistic and language educational points of view.) *Proceedings of the First International Conference on the Teaching of Chinese as A Second Language*, 369-382 World Chinese Language Association, Taipei, Taiwan, ROC.
- Howie, John M. 1976. *Acoustical Studies of Mandarin Vowels and Tones*, New York: Cambridge University Press.
- Hu, Baihua 1988. "Huayu de jungpingdiao (Mid-level tone in Mandarin Chinese)" *Huawen Shrie* Vol. 47, 30-37
- Jiang, Cheng 1956. "Fuyintsz de chingjungyin jengli wenti". *Jungguo Yuwen*, Vol. 5, 28-29.
- Kubler, Cornelius C. 1979. "Some differences between Taiwan Mandarin and textbook Mandarin". *Journal of Chinese Language Teachers Association*. Volume 14, Number 3. 27-39
- Kubler, Cornelius C. 1985a. *A Study of Europeanized Grammar in Modern Written Chinese*, Student Book Co. Taipei.
- Kubler, Cornelius C. 1985b. *The development of Mandarin in Taipei: A Case Study of Language Contact*, Student Book Co. Taipei.
- Li, Chen-Ching 1981. "The sociolinguistic context of Mandarin in Taiwan: Trends and developments". *Proceedings of the 14th International Sino-Tibetan Conference*, 257-278.
- Lieberman, Philip 1967. *Intonation, Perception and Language*, Cambridge (MA): MIT Press.
- Lin, Mau-tsan; Lin, Lian-he; Shia Guang-rung; Tsau, Yu-sheng, 1980. "Putunghua Ertztsz biandiaude Shryan yanjiou". (An experimental study on tone sundhi of disyllabic words in Putunghua) *Jungguo Yuwen*, Vol. 1, 74-79.
- Lin, Mau-tsan, Yan, Ching-ju 1980. "Beijinhua chingshengde shengshiue

- shingjr". (The acoustic characteristics of neutral tone in Beijing dialect) *Fangyan*, Vol. 3, 166-178.
- Lin, Tao 1983. "Tantau Beijinghua chingyin shingjr de chubu shryan". (A preliminary experiment to investigate the quality of the weak stress in Peking dialect) *Yuyangshiue Luentsung*, Vol. 10, 16-37.
- Shen, Susan X. N. (forthcoming) *Tone, Stress and Intonation of Mandarin Chinese*. Unpublished monograph.
- Shih, Chi-lin 1986. *The Prosodic Domain of Tone Sandhi in Chinese*, Ph. D. dissertation, University of California, San Diego.
- Shiu, Shr Rung 1956. "Shuangyin jueitsz de jungyin gueiliu". (Stress rules for disyllabic words) *Jungguo Yuwen*, Vol. 1, No. 2, 35-37.
- Wang, William S. Y. and Kung-Pu Li, 1967. "Tone 3 in Pekinese". *Journal of Speech and Hearing Research*, 1967, Vol. 10, 629-636.
- Wu, Tsung-ji 1982b. "Putonghua yujiu jung de shengdiao bianhua". (Tone variation of Mandarin sentences) *Jungguo Yuwen*, Vol. 6, 439-449.
- Yin, Zuo-yan 1982. "Guanyu Putunghua shuanyin changyungtsz chingjung yin de chubu kaucha". (Preliminary examination on the stress patterns of disyllabic items in Putunghua) *Jungguo Yuwen*, Vol. 168, 168-173.

國語中雙音節詞輕重音的聲學特徵

鄭 秋 豫

本研究的目的是在於自聲學特徵的觀點探討國語中雙音節詞的輕重音問題。研究範圍在重音形式方面包括重輕型(stress-weak, (SW) 及一般型(normal-stress, (NS)) 二種, 其中 SW 又包括了由詞素本身的輕聲以及由詞彙結合所引起的弱化及輕聲二種現象; 以及兩種重音型都存在的詞組。在聲學語音特徵方面則包括了基頻軌跡(含調型與調值)、音節長度及強度三方面。發音的方式是朗讀, 包括正常語調及加重語氣兩種。受試者為一以北平話為母語, 但在臺灣地區居住了四十年的女性, 測試結果顯示出受試者在不知道測驗性質時的表現, 以一般型重音使用最多, 就正常語調而言, 在全部受試詞組中, 含有詞素輕聲的部分, 有 75% 是以一般型重音發出; 而含有詞彙組合弱化輕聲的部分, 有 54% 以一般型重音發出。而兩種重音型皆可的詞組中, 有 72% 是以一般型重音發出。不過, 當受試者在得知實驗性質以後, 她的表現有很大的轉變: 凡含有詞素輕音及詞彙組合弱化輕聲的詞組, 全部(100%) 以重輕型發出, 而兩種重音型皆可的詞組中, 則有 96% 是以重輕型發出。就加強語調而言, 一般型仍為使用最多的重音型, 即在所有受試的加重語調詞組中, 有 93% 是以一般型發出。本研究並分析了所有受試狀況的基頻軌跡, 包括調型、音節開始、中段及結尾的頻率。而各項受試詞組中的平均音節長度以及音強也加以分析。本研究發現, 正常語調與加強語調最大的差異是在音節長度這部分。此外, 本文並討論了目前在臺灣所通行的國語的重音型趨勢, 似以一般型為最常用的形式。