Temporal Variables in Read and Spontaneous Speech[1]

I-li Yang
Graduate Program of Linguistics

（收稿日期：1999年6月5日；修訂日期：1999年8月25日；接受刊登日期：1999年9月27日）

Abstract

This study explored the temporal patterns in oral reading and spontaneous speech of Mandarin Chinese. Read speech was collected from eight adult speakers of Mandarin by asking them to read three texts (a classical poem, a modern poem, and a modern essay) and spontaneous speech was collected from the same subjects in an informal interview. Eight temporal variables were analyzed—syllable duration, pause duration, pause proportion, utterance length, articulation and speech time, and articulation and speech rates. We hypothesized that different types of texts would trigger different temporal patterns and that oral reading and spontaneous speech are very different on the basis of the temporal variables investigated. Our findings are summarized as follows.

1. Text type had a clear effect on syllable duration; it was 335 ms in classical poetry, 277 ms in modern poetry, 252 ms in modern poetry, and 208 ms in spontaneous speech.
2. Text differences also affected the other temporal variables to a significant degree. Mean pause time was longer in both classical and modern poetry (555 ms and 558 ms) than in modern prose (479 ms) and spontaneous speech (491 ms). Also,

[1] This paper is one of the papers resulted from a project funded by NSC (NSC 87-2411-H-004-006).

articulation time was longest in classical poetry (2187 ms), and yet shortest in spontaneous speech (1145 ms). These two variables directly affected the results of articulation and speech rates, both of which were slower in reading than in spontaneous speech.

3. Finally, oral reading and spontaneous speech demand rather different processes in production, as evidenced from the temporal patterns observed. Similarly, poetry reading and prose reading also demand different processes.

Key Word：temporal variables, pause, read speech, spontaneous speech

I. Introduction

Speech is by nature a temporal event; that is, it must be produced within the real time course. As it is constrained by the human speech mechanism physiologically and psychologically, its emergence must alternate with silence, resulting in some kind of rhythm.[2] These two aspects of speech--its temporality and its alternation with silence--manifest themselves not only in spontaneous speech but also in reading. However, oral reading differs from spontaneous speech in that the "what" of speech production has already been pre-determined, leaving only the "how" to the reader. Typically, one reads and pauses only when necessary, without bothering about what to say next. But when the "what" is of a special kind, e.g., a poem or a sermon, then the "how" changes too. When one is reading a poem, he might pause, often a long time, to give a special color to the emotion of a line. Also, it is not unusual for a preacher to pause at the end of a sermon to allow some time for the audience to reflect on themselves. Different speech tasks would certainly trigger different temporal patterns.

By temporal patterns, we refer to such durational and temporal variables as syllable duration, pauses, articulation rate, etc. Researchers interested in such variables not only enjoy the discovery of the facts *per se* but also strive for a theory to account for language production after cumulating enough evidence. While a number of studies have been done on temporal patterns in Western languages, relatively little is known about them in the Chinese language, be it read or spontaneous speech.

Thus, this paper aims to present a comprehensive picture of the temporal variables of read and spontaneous speech of Mandarin Chinese. As we all know, oral reading takes many

---

[2] For the Chinese people who see everything in perfect harmony with everything else, there is rhythm in everything, including language (Zhu 1976, Zeng 1976). Recently, the study of pauses in the West has also come up with the idea that there is cognitive rhythm in language production (Henderson et al. 1965).

different forms. There is the reading of stories to children; there is also the reading of texts in the classroom. Still, there is the reading of poetry to entertain self and others. We assume that different types of reading materials would trigger different temporal patterns. We also assume that read speech would have a different temporal pattern from spontaneous speech. Specifically, we want to know if there are different temporal patterns in reading Chinese poetry, classical as well as modern, Chinese modern prose, and spontaneous speech. We want to know if the traditional analysis of classical poetry can find evidence in the present study. We also want to know if our results can be compared with those obtained by other studies for other languages. Therefore, the following section will briefly introduce the traditional view of the form and structure of Chinese poetry, followed by a section of a review on related literature.

II. A Brief Description of the Form and Structure of Chinese Classical Poetry

Speaking of Chinese classical poetry, one is apt to think of Tang poetry which is a strictly structured form of verse, often in four (絕句*jue-ju*) or eight lines (律詩*lyu-shi*), each having either five or seven syllables arranged according to a predetermined scheme of two different tones—the level tone (平*ping*) and the non-level tone (仄*ze*). The principal rule is that if the first line is 11221 (1 for level, 2 for non-level), then the next line should be 22112. This is called a "pair (*du*對)." If this rule can not be observed, then the two lines are said to "attach" (*nian*黏) to each other (Zeng 1976). Such alignment thus creates a rhythm that can not be found elsewhere. The following eight-line *ly-shi* by Wei Ying-wu (a Tang poet) illustrates this point.

(1) qu nian hua li feng jun bie	去年花裏逢君別，	2112112
(2) jin ri hua kai you yi nian	今日花開又一年。	1211221
(3) shi shi mang mang nan zi liao	世事茫茫難自料，	2211122
(4) chun cou an an du cheng mian	春愁黯黯獨成眠。	1122211
(5) shen duo ji bing si tian li	身多疾病思田里，	1122112
(6) yi you liuwang kui feng qian	邑有流亡愧俸錢。	2211221
(7) wen dao yu lai xiang wen xun	聞道欲來相問訊，	1221122
(8) xi lou wang yue ji hui yuan	西樓望月幾回圓。	1122211

Here, Lines 1 and 2, and Lines 7 and 8 are "attached" to each other, and Lines 3 and 4 and Lines 5 and 6 make two "pairs." This pattern is most often followed by practitioners of poetry writing. However, great poets often create their own variations, thus bringing new delights to the reader.

What are the level and the non-level tones like? The level tone is neither rising nor falling. As to the non-level tone, actually there are three tones that qualify as the non-level tone. The first is the falling-rising tone (*shang*), the second the falling tone (*qu*), and the third the glottal tone (*ru*). It is said that the level tone is the longest, and the falling-rising the next longest, and the glottal the shortest (Gu Yan-wu, quoted from Zhu 1976). However, due to dialectal differences, the lengths of the four tones remain unsettled (Zhu 1976).

The reading or reciting of a poem must follow some rules too. Thus, for the five-syllable line, it must be read with two stops or *duns* (頓)[3], one after the second syllable and the other after the last; and for the seven-syllable line, there are three stops, one after the second, another after the fourth, and the third after the last syllable (Zeng 1976). However, these stops may not be real silence; they are often realized by a lengthening of the preceding

[3] A *dun* is more or less like the English foot or the French caesura. It is a prosodic unit, often coinciding with the meaning unit.

syllable (Zhu 1976). The *dun* is a necessary maneuver for creating the rhythm of classical poetry.

Most Chinese people have some knowledge about the form and structure of classical poetry from their schooling, as the saying goes, "Even if you can not write poetry, you will be good at reciting after you read the 300 Tang poems well." To read or recite a poem in front of an audience or by oneself is a great enjoyment to most Chinese people, especially among the educated, even when the dialect used is Mandarin.

Mandarin differs from classical Chinese in that the glottal tone has disappeared and there is a rising tone. So, the four tones now are the first (level), the second (rising), the third (falling-rising), and the fourth (falling). There is also a fifth tone—the neuter tone which occurs only with certain function words. As the second tone begins in the middle-pitch and ends in a pitch as high as the first tone, the two are considered as belonging to the *ping* tone, and the third and the fourth to the *ze* tone. To do without the *ru* tone can create some problems in poetry reading. However, with the help of some dialects where the glottal tone is preserved, e.g., Southern Min or Hakka, one can easily restore the tone pattern of any classical poem (The example given is analyzed in this way).

As just mentioned, the arrangement of varying tones and the *dun* phenomenon of recitation have been fixed, but little is known about the exact nature of syllable duration and other temporal variables when a poem is read or recited by means of objective measurements. Are *ping* syllables truely longer than *ze* syllables? Are syllables preceding *duns* really longer than those not preceding them? Are syllables at the line-final position longer than those at the medial position? That is: is there prepausal lengthening? What do *duns* look like? Is there a genuine pause when a *dun* appears? Are *duns* at line-final position longer than those at line-medial position? Finally, is the temporal pattern observed in poetry reading the same as that found in prose reading and in spontaneous speech? Are the findings comparable to those found for other languages? These are the questions we want to

address in the present study. But we must first look at some of the findings in recent research on pausing and syllable duration.

III. Related Literature

In this section, we shall report research findings on the concerned temporal variables in reading for the Western languages first, then those dealing with related matters in the Chinese language.

3.1 Studies on Temporal Variables in Reading for Western Languages

First, the speech of reading is mostly fluent; pauses at minor constituent boundaries (such as phrases and words) disappear and those at major constituent boundaries (clauses or sentences) are concentrated about a mean value and relatively invariant, ranging from 300 to 500 ms (Barik 1977, Kowal et al. 1975, O'Connell & Kowal 1972, Sabin et al. 1979, Hardin et al. 1989, Kien & Kemp 1994). Extremely long pauses are seldom found in read speech (Goldman-Eisler 1972, Grosjean 1980). In other words, most reading is fluent, as distinct from spontaneous speech where disfluency alternates with fluency in a rhythmic way (Henderson et al. 1965). In prose reading, 18% to 20% of the time is spent in pausing (Huggins 1964, Goldman-Eisler 1968, Butcher 1981, O'Connell 1989). But in spontaneous speech, pausing might take as much as 50% of the total speaking time (Klatt 1975).

Pauses are found to be directly related to the rate of speaking (Goldman-Eisler 1956). The normal speech rate for reading ranges from 150 to 250 words per minute or 4-7 syllables per second (Goldman-Eisler 1968, Barik 1972, Sabin et al. 1979). When speech rate is accelerated, not only is the duration of segments (vowels, consonants, and thus syllables) shortened, but also pauses decrease in number as well as in duration (Goldman-Eisler 1968). In accelerated speech, pauses at minor constituent boundaries, e.g., within phrases,

disappear (Grosjean 1980).

There is evidence that pauses between sentences are generally longer than those between phrases (Brubaker 1972). In fact, pauses often correspond to punctuation—periods result in longer pauses than commas (O'Connell, et al. 1989). Clearly, syntactic structures play a role in the length and frequency of pauses.

Yet, syntax is not the only factor that affects speech rate and pauses. Sentences near the end of a paragraph are often read faster than those occurring earlier (Brubaker 1972). A semantically bizarre sentence will often induce longer pauses (Kowal et al. 1975). Besides, reader characteristics such as sex , age, and proficiency level also affect pause patterns. Kowal et al. (1973) found that boys produced longer pause time than girls of the same age, and that speech rate and length of utterance per pause correlate with age growth. It was also found that less competent readers often produce more longer pauses (Kowal et al. 1975).

However, longer pauses do not necessarily mean lower proficiency. In poetic and dramatic reading, pauses are often two or three times longer than those found in normal prose reading (Clemmer et al. 1979). Political speech and dramatic or poetic reading often have slower speech rate (3-4 syllables per second) and more and longer pauses (ca. 700-800 ms) (Clemmer et al. 1979, Kien & Kemp 1994). Clearly, pauses in such readings serve other functions, rhetorical or persuasive (Duez 1982, O'Connell et al. 1989).

The average syllable duration is 200 ms. But many factors have been found to affect syllable duration—linguistic as well as non-linguistic ones (Klatt 1975). First, a syllable at phrase final or clause final or even word final positions (Lehiste 1975, Klatt 1975, Oller 1973) is often lengthened—the so-called prepausal lengthening. Second, when a syllable is stressed, its duration increases 10-20% or even more (Coker, et al. 1973, Oller 1973). Also, an unusual word appearing the first time in a connected discourse is always longer (Umeda 1975). It has also been found that the metrical structure of a sequence directly affects syllable duration: a stressed syllable is shortened when the following foot is an anapest

rather than an iamb (Fourakis & Monahan 1988).

Some language specific factors seem to play a role too. For instance, Han (1962) observed that the average duration ratio of a Japanese heavy syllable to a light syllable is 2:1, later corrected to 1.7 : 1 by Hoequist (1983). In both Spanish and Japanese, it is common to lengthen syllables towards the end of a word, with initial syllables shortest, medial a little longer, and final longest. But in English, medial syllables appear to be the shortest (Hoequist 1983).

3.2 Studies on Temporal Variables in the Chinese Language

For the Chinese language, not many studies have been done on syllable duration and pauses. As we have said, of the traditional four tones (the level, the falling-rising, the falling, and the glottal), the level tone is said to be the longest, and the glottal shortest, but the length of the four tones remains a debatable issue due to dialectal differences (Zhu 1976). For the Mandarin four tones, in the citation form the third tone is the longest and the neutral tone the shortest (Cheng, C-C. 1973), with Tone 2, Tone 1, and Tone 4 in decreasing order (Kuo 1993; Cheng, C-Y 1994). However, Tseng (1990) found that in spontaneous speech any tone can be longer than any other depending on its position as well as on its semantic content. The third tone often lacks the rising portion, resulting in shorter duration (Tseng 1990). The duration of the third tone has been found to be playing a facilitating effect on the perception of this tone (Blicher et al. 1980). As the third tone often undergoes sandhi, it is often confused with the second tone perceptually (Chang 1992; Kuo 1993)

Recently, Lin (1997) found that in spontaneous speech the neutral tone is the shortest and that the third tone is not longer than any other lexical tone. In his corpus of spontaneous speech by four people (two over the age of 50 and two around their 20's), syllable duration ranges from159 ms to 310 ms. Lin also discovered a prepausal lengthening effect in the speech, 289 ms as opposed to the total average of 210 ms or to the non-

prepausal of 203 ms. Lin also found that most lengthened syllables are content words (90%).

A couple of studies have also been done on pausing and speech rate in the Chinese language. First, by asking their subjects to read as fast as they could, Cheung and Kemper (1993) found an articulation rate of 4.06 syllables in reading Chinese (Cantonese) one-syllable words.

Second, Yang (1997a) observed that mean pause time in spontaneous speech can vary from 400 to 1500 ms, depending on syntactic location as well as speaker characteristics. In another study (1997b), the mean pause time in spontaneous speech is 287 ms, with a speech rate of 4.96 syllables per second, and an articulation rate of 5.93 syllables per second. Yang's 1997a study used a total of 720 paused-defined units from four adult speakers, all over 45 years, and her 1997b used a total of 5524 intonation units from 24 speakers all round the age of 20.

Recently, Hardin et al. (1998) compared the reading (a fable) of German and Chinese by native speakers of the respective language and found that while both readings had the same mean pause time (490 ms), the reading of Chinese exhibited a larger proportion of pausing (26.1% to 14.7%), a smaller articulation time (1370 ms to 2590 ms), a shorter mean utterance length (6.5 syllables to 14.8 syllables), and a smaller articulation rate (4.74 syllables/second to 5.74 syllables/second). The authors suggested that the main cause for such a different pattern was the writing system; Chinese was a logographic language while German was an alphabetic language.

IV.Methodology

From the review, we know that little has been done on the temporal variable in the oral reading of Chinese, let alone that of Chinese poetry. This prompts us to take up the present study, with a view not only to describing the temporal patterns in oral reading but also to

understanding the processes underlying read speech as well as spontaneous speech.

The specific research questions we wanted to address were

(1)Does syllable duration remain constant across different types of speech, read as well as spontaneous?

(2)Is syllable duration affected by tones, one set being classical distinction between level and non-level tones and the other being the Mandarin four/five tones?

(3)Does syllable duration differ because of the position of the syllable--before *duns* in classical poetry, and before major constituent breaks in modern poetry and prose?

(4)What do *duns* look like? Are they a mere artifact or are they real silence? If they are real pauses, at what positions are they more likely to appear? What about their durations?

(5)Where are pauses likely to appear in reading modern poetry and prose?

(6)What temporal patterns can be found for different types of speech?[4] Why?

So we selected three pieces of reading materials. The first piece is a *lyu-shi* by Wei Ying-wu (hereafter RS1); the second a modern poem (by a contemporary writer called Dai Wang-shu; hereafter RS2); and the third a modern essay (from a magazine, author unknown; hereafter RS3). The first piece has 8 lines with 7 syllables in each line and lines 2, 4, 6, and 8 rimed[5]. The modern poetry, the second piece, is full of emotion but with little syntactic constraints of the classical verse and is more colloquial (*bai-hua*); it has 18 lines, each having

---

[4] As mentioned earlier, temporal patterns refer to PT, AT, PP,etc. A short definition of each of these variables is given below.
Pause time (PT) refers to the period of time in speech (in milliseconds) when absolute silence occurs.
Pause proportion (PP) refers to the proportion of pause time to the total speaking time (in %).
Articulation time (AT) refers to the amount of time (in milliseconds) used in actual articulation.
Speech time (ST) refers to the time of speaking or the summation of pause time with articulation time (in milliseconds).
Utterance length (Length) refers to the number of syllables within an utterance which is defined by pauses.
Articulation rate (AR) is calculated by dividing Length with AT (syllables per second).
Speech rate (SR) is calculated by dividing Length with ST (syllables per second).

[5] The line here is defined poetically; that is each line does not necessarily correspond to a complete syntactic unit such as the clause or the phrase.

the length between 3 and 10 syllables, with a loose riming scheme. The third piece is a short essay describing a natural scene and the feelings aroused by the scene.[6] Besides, we also collected spontaneous speech from the subjects by interviewing them both before and after the reading task (hereafter SS).

Our first hypothesis was that classical poetry would be rendered more slowly, hence having longer syllable duration as well as more long pauses and slower speech rate, while spontaneous speech would be rendered much faster, with shorter syllable duration and fewer long pauses and faster speech rate, and the modern poem and prose would bring out results in-between the two extremes.[7]

We also hypothesized that the rhythmic structure—the arrangement of *ping* and *ze* tones as well as the occurrence of *duns*—would affect syllable duration in the reading of the classical poem. As to the modern poem and prose, we hypothesized that the four tones and the position of the syllables in a syntactically complete unit[8]—initial, medial, or final—would make a difference. Of course, we also predicted that the three pieces of reading would result in different pause patterns from that of the spontaneous speech.

We asked eight people, four males and four females in two age groups (the young around the age of 20 and the old around the age of 50)[9] to read the three pieces of reading materials as well as to introduce themselves briefly before the reading and to comment on their own reading afterwards. These two interludes made up the spontaneous speech. All the speech was first recorded in the regular tape recorder, and then it was analyzed by means of SoundEdit Pro on Power McIntosh. Syllable duration as well as pauses were measured and

---

[6] Only a portion of the piece was adopted for the reading task.
[7] We may imagine the four types of speech as forming a continuum, but in actuality each type has rather different characteristics. We shall deal with it in Section 6.
[8] This unit is often marked by overt punctuation marks such as the comma, the semicolon, or the period.
[9] Our projert initially recruited 20 people, but here we used only data collected from 8. Due to space limit, we do not want to explore the effect of sex and age in the present paper although we know they do play a part on temporal variables.

recorded for statistical analysis.

Thus, for the dependent variable Syllable Duration, the independent variables were Speech Type (Classical Poetry, Modern Poetry, Modern Prose, and Spontaneous Speech), Rhythm (*ping* and *ze*) or/and Tone (1, 2, 3, 4, 5), and Position.[10] But for the other temporal variables (PT, AT, Length, ST, AR, SR, PP), the independent variable was Speech Type only.

V. Results

The read corpus used for durational analysis by each subject had a total of 238 syllables (56 syllables for Type 1 (*lyu-shih*), 89 for modern poetry, 93 for expressive prose), and the total syllables analyzed were 1903 (one subject (No.6) missed one syllable during reading).[11] The data base of read speech for pausal analysis was slightly larger—with a total of 4304 syllables in 654 pause-defined units. As to the spontaneous speech, we had obtained a total of 1113 syllables for durational analysis (each subject producing a number ranging from 109 to 194 syllables) and 499 pause-defined units (ranging from 23 to 114 units) for pausal analysis. So, there were a total of 3016 syllables for durational analysis and 1153 units for pausal analysis. We present the results of the six research questions in the order mentioned earlier.

1. Does syllable duration remain constant across different types of speech ?

The following table gives the answer to this question in terms of the individual speakers as well of the type of speech. First of all, from the figures in the rightmost column we see

[10] There are only two third tone syllables that underwent tone sandhi; i.e., they were realized as the second tone. And in this paper, they are treated as Tone 2 syllables.

[11] Subjects No. 1-4 are older people, and 5-8 are younger people; 1, 2, 5, 6 are females, and 3, 4, 7, 8 are males.

that individual differences in MSD were very great (F=35.03, P=0.000). Furthermore, these differences were also significant in different types of speech, read as well as spontaneous (for RS1, F=23.75, P=0.000; for RS2, F=21.66, P=0.000; for RS3, F=17.52, P=0.000; for SS, F=17.05, P=0.000). Finally, with the exception of Subject 4's RS 2 and SS, most variation coefficients (=SD/Mean) were far below 50%, indicating that each subject was rather constant in syllable duration even while producing different types of speech.

Table 1. Mean Syllable Duration (MSD, further abbreviated as M) and Its Standard Deviation (SD, in milliseconds) for the Three Types of Read Speech (RS) andSpontaneous Speech (SS) by Each Subject

	RS1(N=56)		RS2 (N=89)		RS3 (N=93)		SS (N=*)		Total	
Subject	M	SD	M	SD	M	SD	M	SD	M	SD
1	361	123	329	143	253	103	164	70	258	133
2	287	100	270	118	238	116	231	126	248	121
3	423	117	378	117	349	99	226	90	329	128
4	442	173	250	140	247	96	253	158	276	159
5	345	74	295	99	260	90	163	71	252	107
6	274	75	245	99	244	114	236	118	245	108
7	307	108	221	67	209	76	172	72	215	90
8	242	60	228	63	219	78	168	77	208	77
Total	335 (N=448)	127	277 (N=712)	120	252 (N=743)	105	208 (N=1113)	115	254 (N=3016)	124

*The numbers of syllables for each subject in spontaneous speech are: 124 (S1), 188 (S2), 112 (S3), 194 (S4), 109 (S5), 153 (S6), 116 (S7), 117 (S8).

As this was a rather big sample size, we would not brood over individual differences (although they did exist) any further in the following. Let us now turn to the effect of different types of speech on syllable duration by looking at the totals.

When we examined the bottom line of the table, again we found that MSD in one type differed greatly from that in the other: for RS1, it was 335 ms; for RS2, 277ms; for RS3, 252ms; for SS, 208 ms ($F=139.85$, $P=0.000$). The Tukey-B test revealed that all four of them differed from one another. Clearly, Type of Speech (i.e., the type of reading material) did make a great difference in MSD.

2. Is syllable duration affected by the variable of Tone/Rhythm ?

We used the term Rhythm to refer to the classical distinction of *ping* and *ze*. The results are presented in the following table. From the table, we can see that the *ping* words were longer than the *ze* words in RS1. In the total column (the rightmost), we see that Tone 1 was longest, Tone 2 second, Tone 3 third, Tone 5 fourth, and Tone 3 the shortest. But within each type, the picture was more complex; in RS1, Tone 2 was longest (356 ms), but in RS2, Tone 4 (305 ms), and in RS3 and SS, Tone 1 (264ms and 241 ms). The most striking MSD was found for Tone 5 in RS3 speech, the shortest of all (180 ms). Clearly, Tone 3 was never longer than any other tone except in RS3 where Tone 5 was the shortest.

Table 2. MSD's Calculated in Terms of Rhythm and Tone in the Three Types of RS and SS

		RS1 (N)	RS2 (N)	RS3 (N)	SS (N)	Total (N)
Rhythm	*ping*	355 (232)	~	~	~	~
	ze	314 (216)	~	~	~	~
Tone	1	337 (88)	288 (168)	277 (96)	241 (157)	280 (509)
	2	356 (160)	251 (168)	264 (152)	204 (206)	265 (702)
	3	311 (40)	250 (72)	246 (120)	184 (205)	221 (421)
	4	319 (160)	305 (208)	257 (304)	203 (388)	256 (1060)
	5	~	262 (96)	180 (71)	226 (157)	226 (324)
Total		335 (448)	277 (712)	252 (743)	208 (1113)	254 (3016)

One-way Anova tests showed that for the total speech sample, durational differences caused by tone or rhythm differences were significant (F=19.07, p=0.000); that is, Tone 1, Tone 2, an Tone 4 were significantly longer than the other two tones. Similarly, the difference between *ping* and *ze* in RS1 speech was significant (F=11.69, p=0.0007). And the differences between the different tones were also significant for all the other types of speech respectively (for RS1, F=2,78, p=0.04; for RS2, F=6.51, p=0.000; for RS3, F=11.11, p=0.000; for SS, F=6.75, P=0.000). In RS1, Tone 2 was significantly longer than Tone 3 (356 ms to 311 ms). In RS2, Tone 4 was significantly longer than Tones 2, 3, and 5 (305 ms to 251, 250, and 262 ms), and Tone 1 significantly longer than Tone 2 (288 ms to 251 ms). In RS3 speech, Tone 5 was significantly shorter than the other four tones among which the differences were not significant. And in SS, Tone 1 was significantly longer than Tones 2, 3, and 4, and Tone 5 significantly longer than Tone 3.[12] In other words, Tone 1, Tone 2, and Tone 4 were usually the longest, depending on the type of speech, and yet Tone 3 and Tone 5 were never longer than the other three tones in any situation.

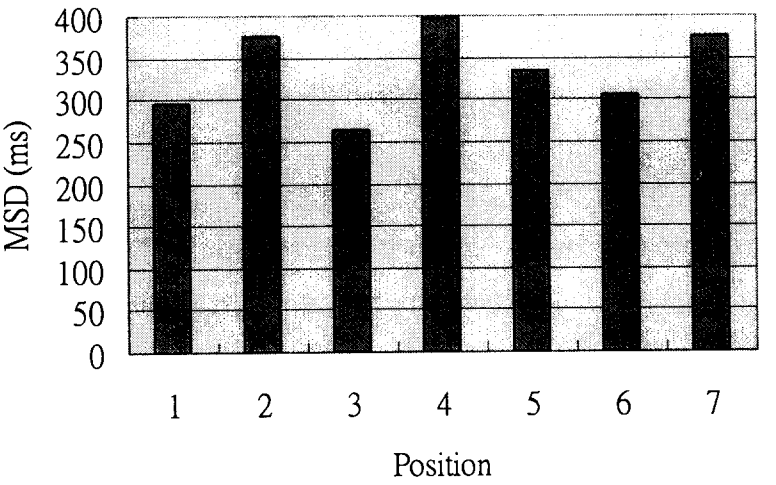
Another way to look at the table is to compare the durational differences within the same tone across the different types of speech. Thus, we find great differences between RS1 and SS for Tone 1, Tone 2 and Tone 3; between the first two types and the last two types for Tone 4; between RS3 and RS2, SS for Tone 5. By filtering out minor inconsistencies, we might say that as speech becomes more modern and more spontaneous, all the four/five tones become shorter. This is another way to understand the effect of speech type on the tones.

[12] That Tone 5 could be longer than Tone 2, Tone 3, and Tone 4 in this sample can be explained on two grounds. First, the subjects might not have properly said the tone. Second, many of the Tone 5 syllables might have appeared at hesitation points. But more evidence is needed for this explanation.

3. Does syllable duration differ because of the position of the syllable?

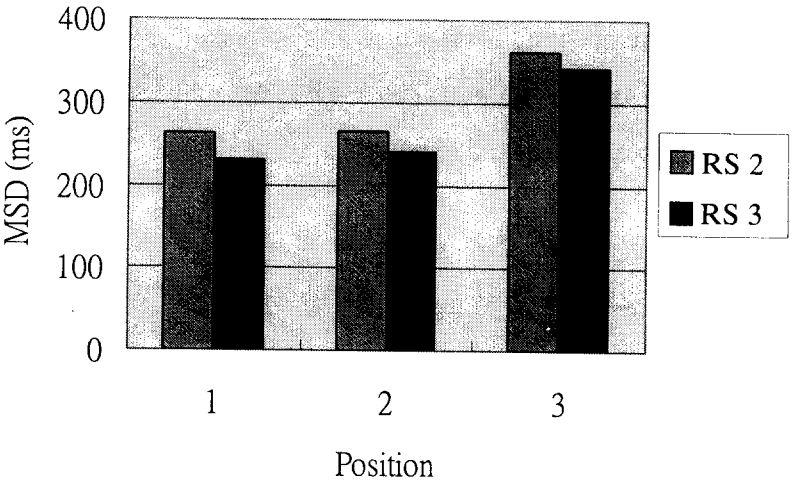
For the classical verse where syllable position is fixed, we had 7 positions. But for the modern poem and prose, we used a simplistic schema—unit-initial, medial and final on the basis of punctuation. As to the spontaneous speech, no analysis about position was done.[13] The following two figures present the results.

Figure 1. MSD at 7 Positions in RS1 Speech



[13] There were two reasons. First, spontaneous speech was not the focal interest in the present study. Second, as the speech was pausally defined, it was rather difficult to assign positional factor.

Figure 2. MSD at 3 Positions inRS2 and RS3 Speech



As the figures show, there were significant differences in MSD at different positions in each of the three types of read speech (for RS1, F=10.82, p=0.000; for RS2, F=26.06, p=0.000; for RS3, F=42.87, p=0.000). In RS1, syllables were longest at Position 4 (397 ms), and then at Positions 2 and 7 (both 375 ms), shortest at Position 3. In both RS2 and RS3, syllables were longest at final position (361 ms and 340 ms respectively), and small differences were found at initial and medial positions.[14]

4. What do *duns* looks like?

There were a total of 60 pauses in RS1, three times at Position 4 (mean=356 ms), one time at Position 5 (463 ms), and 56 times at Position 7 (564 ms). In other words, our subjects

[14] The position variable is actually a syntactic factor. Syllable lengthening at the points mentioned is a reflection of constituent boundaries—mainly phrase and clause boundaries. For the classical poem, it is a manifestation of the *dun*, but for the modern poetry and prose, it is often succeeded by real pauses at constituent boundaries.

paused at the end of each line (Position 7), and Position 4 had a slightly better chance than Position 2 in inducing pauses. However, the frequency was too small to result in any significant differences. As to the duration of pauses, line-final pauses were much longer (no significance tests were performed because of the paucity of cases at other positions). Clearly the so-called *dun* in the middle of the line was a perceptual artifact; it was not realized as real pauses.

5. Where are pauses likely to appear in reading modern poetry and prose?[15]

In both situations, pauses occurred only once at initial position (mean duration being 92 ms and 418 ms respectively). This happened because the subject was hesitant. Pauses at medial position were again relatively few (f=29, 27.4%, and f=12, 12.4%), and their duration was not as long as those at final position (524 ms to 569 and 342 ms to 523 ms). However, the differences were not significant due to too few cases at the initial position. The following table summarizes the results.

Table 3. Frequencies, Percentages, and Mean Pause Time (ms) in Type 2 and 3 Speeches at the Three Positions

	Initial			Medial			Final			Total N
	F	%	mean	f	%	mean	f	%	mean	
RS2	1	0.9	92	29	27.4	524	76	71.7	569	106
RS3	1	1	418	12	12.4	342	84	86.6	523	97

Clearly, pauses at other positions than the final were rather rare. We shall have more to say about pauses in the following section.

6. What temporal patterns can be found for different types of speech? Why?[16]

---

[15] These pauses were those found in the speech data selected for durational analysis only. The database used for pausal analysis was much larger.

Of the seven temporal variables, Pause Time (PT), Articulation Time (AT), Length of Utterance (Length) were the three basic measures because Speech Time (ST), Speech Rate (SR), Articulation Rate (AR), and Proportion of Pause Time to Speech Time (PP) were all derived from them. (Refer to Note 6 for more information.)

There were a total of 1153 pause-defined units, with a mean pause time of 498 ms, mean articulation time of 1496 ms, and a mean utterance length of 6.08 syllables. The total speech time is 2299926 ms or 38 minutes and 19 seconds and 926 milliseconds. The average speech rate is 3.06 syllables per second, and the mean articulation rate is 4.21 syllables per second. The following table presents the statistics for each individual speaker as well as for the total.

Table 4. Statistics of the Seven Temporal Variables in the 8 Individual Samples

Subject	PT(ms)	AT(ms)	ST(ms)	Length(sy)	AR(sy/s)	SR(sy/s)	PP(%)	N
1	620	1498	2202	5.95	4.03	2.67	33	137
2	481	1461	1896	5.92	3.95	2.98	25	164
3	438	1716	2252	5.15	3.25	2.30	28	153
4	495	1512	1983	6.03	3.99	2.98	25	191
5	513	1816	2268	7.71	4.40	3.43	21	121
6	278	1703	2028	7.03	4.09	3.40	17	98
7	481	1195	1673	5.96	5.17	3.62	29	150
8	536	1169	1722	5.53	4.90	3.29	32	139
Total	498	1496	1994	6.08	4.21	3.06	27	1153

Individual differences were quite great for all the temporal variables but PT (refer to the P-level in the following table), probably due to a sample size that was rather big. Of much more importance was the question whether different speech types resulted in different temporal patterns. The following table presents the statistics of PT, etc. based on speech type.

[16] We shall answer the why question in Section 6.

Table 5. Statistics of PT, etc. in Different Types of Speech

	RS1 (N=62)	RS2 (N=142)	RS3 (N=450)	SS (N=499)	Total (N=1153)	P-level
PT (ms)	555 (236)*	558 (376)	479 (252)	491 (442)	498 (361)	0.07
AT (ms)	2187 (623)	1682 (960)	1731 (905)	1145 (765)	1496 (900)	0.000
ST (ms)	2742 (667)	2240 (1056)	2210 (957)	1637 (848)	1994 (969)	0.000
Length (sy)	6.54 (1.22)	5.92 (3.38)	6.79 (3.42)	5.43 (4.17)	6.08 (3.74)	0.000
AR (sy/sec)	3.12 (0.68)	3.64 (0.92)	3.99 (0.84)	4.69 (1.52)	4.21 (1.27)	0.000
SR (sy/sec)	2.46 (0.57)	2.67 (0.83)	3.05 (0.83)	3.25 (1.48)	3.06 (1.17)	0.000
PP (%)	20 (9)	26 (15)	23 (12)	31 (20)	27 (17)	0.000

*The figures in the parentheses are standard deviations.

By examining the mean and the standard deviation within each cell of the three basic measures (PT, AT, and Length), we found that all these measures were most constant in RS1, and least so in SS. The variation coefficients (=SD/Mean) of the three measures in RS2 and RS3 were somewhere between 50-60%. The Anova tests showed that the differences found for all the temporal variables, except PT, caused by type of speech were significant. For AT and ST, RS1 and SS contributed most to the differences; for Length, significant differences were observed between RS3 and SS only; for AR, all differed from each other; for SR, big gaps were found between the first two and the last two types of speech; and for PP, SS differed most from the rest.

We wanted to know if the same thing would occur if the calculation was done within each individual speaker. The following table summarizes the results.

Table 6. Summary of Tests of Significance for the Effects of Type on the 7 Temporal Variables for Individual Speakers (* Significant; -- Not significant)

Speaker	AR	AT	Length	PP	PT	SR	ST
1	*	*	--	--	*	*	--
2	--	*	*	*	--	*	*
3	*	*	*	*	*	*	*
4	*	*	*	*	--	--	*
5	*	*	--	*	--	*	*
6	--	--	--	--	--	--	--
7	*	*	*	*	--	*	*
8	*	*	*	--	--	*	*

The results show that AT was most vulnerable to a change in Speech Type—all but Subject No 6 were affected, and then AR, SR, and ST, and PT was the least affected. The table also shows that some people were more sensitive to Type change than others, e.g., Subject No 3 compared with No 6. Most people were affected to a certain degree in many of the temporal variables. Note also that Length was affected only 5 out of 8 times, slightly over the chance level.[17] The effect of Speech Type was too great to be pure chances (37/56*100=66%), and it was AT that was consistently affected. This finding has important bearings on the issue of syllable duration (to be discussed later).

The following table presents the frequencies and percentages of PT of various lengths in the four speech types.

[17] We are pretty sure that the factors of Sex and Age played a role here. For the majority of subjects who showed more effect of Speech Type in the seven variables were mostly older and male subjects. However, in the present paper we did not want to pursue this matter further.

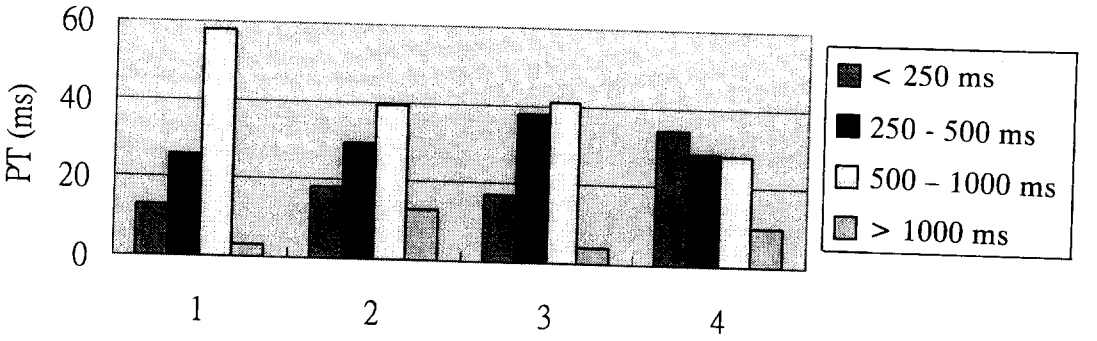
Table 7. Frequencies and Percentages of PT of Various Lengths in the Four Speech Types

	<250 ms	250-500 ms	500-1000 ms	>1000 ms
RS1	8 (12.9)	16 (25.8)	36 (58.1)	2 (3.2)
RS2	26 (18.3)	42 (29.6)	56 (39.4)	18 (12.7)
RS3	77 (17.1)	169 (37.6)	185 (41.1)	19 (4.2)
SS	170 (34.1)	143 (28.7)	138 (27.7)	48 (9.6)
Total	281 (24.4)	370 (32.1)	415 (36)	87 (7.5)

One thing that is common with all four speech types is that pauses over 1000 ms were the minority. However, there were great differences between SS and the three RS's. First, in ss pauses shorter than 250 ms had the greatest percentage (34.1%), while those at 250-500 ms and 500-1000 ms had an equal chance of occurrence (28.7% and 27.7%). But with the three RS's, most pauses clustered between 250 and 1000 ms, with the ones lasting 500-1000 being of the majority (58.1%, 39.4%, 41.1%). Secondly, there were a greater percentage of pauses longer than 1000 ms in SS than in RS1 and RS3 (9.6% to 3.2% and 4.2%). In fact, there were one between 2000-2500 ms, five between 2500 and 3000 ms, and one over 3000 ms. Although the percentage of pauses over 1000 ms in RS2 was greater than that in SS, there was only one instance of pause over 2000 ms.

What all this amounts to is that the distribution of pauses of different lengths was most diverse in SS, less so in RS2, still less in RS3, and least in RS1. In other words, with read speech, especially with classical poetry, most pauses concentrated around the mean (VC=42%, 67%, and 52% for RS1, RS2, and RS3 respectively), but with spontaneous speech, great variance was found (VC=90%). The following figure presents the same story in a more vivid way.

Figure 3. PT in Four Durations--(<250, 250-500, 500-1000, >1000 ms)



Type of Speech (1=RS1/classical poetry, 2=RS2/modern poetry, 3=RS3/modern prose, 4=SS/interview

VI. Discussion

Let's repeat some of our findings on syllable duration.

First, syllable duration was longest in the classical poem (335 ms), second longest in the modern poem (277 ms), third longest in the modern prose (252 ms), and shortest in spontaneous speech (208 ms). Besides, all the four differed from one another significantly although the difference between RS2 and RS3 was rather small.

Our speech data, read or spontaneous, were all connected speech. When compared to data collected through word list reading, our MSD's of different tones were not as long, as shown in the following table.

Table 8. A Comparison of Our MDS (ms) with Those of Lin (1997) and Cheng (1994)

	Lin (1997)	SS	Cheng (1994)	RS1	RS2	RS3
Tone 1	232	241	344	337	288	277
Tone 2	211	204	397	356	251	264
Tone 3	211	184	516	311	250	246
Tone 4	219	203	311	319	305	257
Tone 5	144	226	--	--	262	180
Total	210	208	--	335	277	252

*Lin's data was spontaneous (lectures), and Cheng's was the reading of word lists.

Presumably, the reading of word lists is different from the reading of a connected piece of discourse, poetry or prose. In word list reading, each syllable is pronounced in the most complete way, but in connected speech there is bound to be some linking, compression, and lengthening, or even deletion, from syllable to syllable.

The table also presents Lin's results, and it stands to reason that our SS had MSD's that bore some resemblance to Lin's, for both sets of data were spontaneous speech in casual situations.

As shown in the table above, Cheng's third tone was the longest of all. But our third tone was never longer than other tones except the neuter tone. Tone 2 was longest in classical poetry, Tone 4 in modern poetry, Tone 1 in modern prose and spontaneous speech. This was one big difference found for the four tones in citation form and in connected speech.

As we examined the position factor, we found that Tone 1 was always the longest, Tone 2 the second longest, and Tone 4 the third longest at final position, as shown in the following figures.

Figure 4. MSD in terms of Tone by Position in RS1

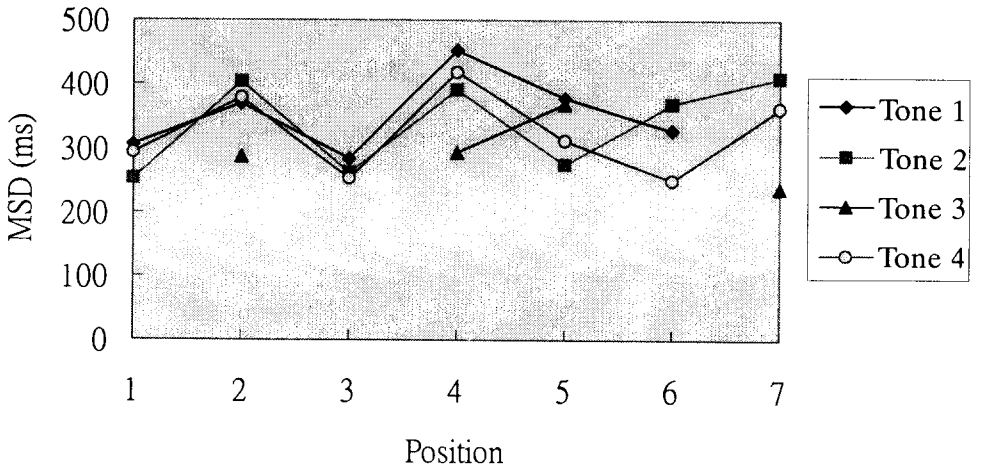


Figure 5. MSD in terms of Tone by Position in RS2

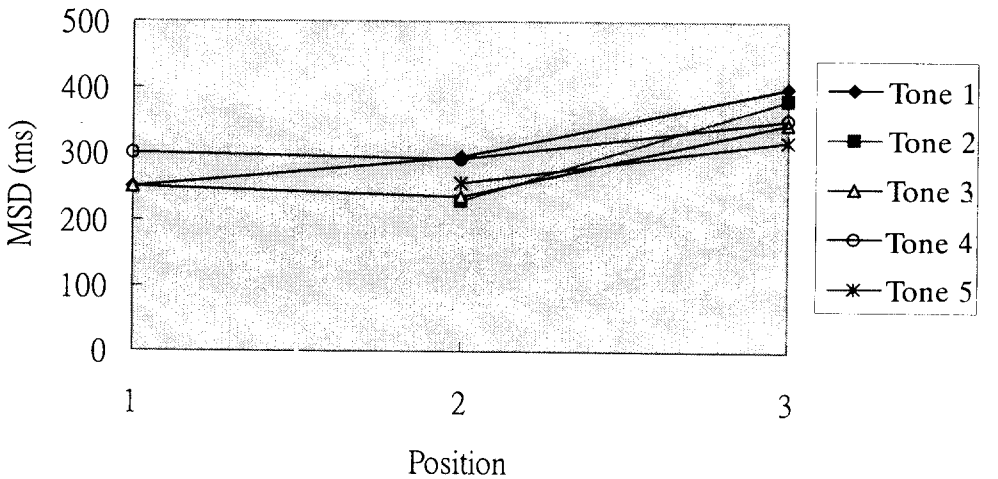
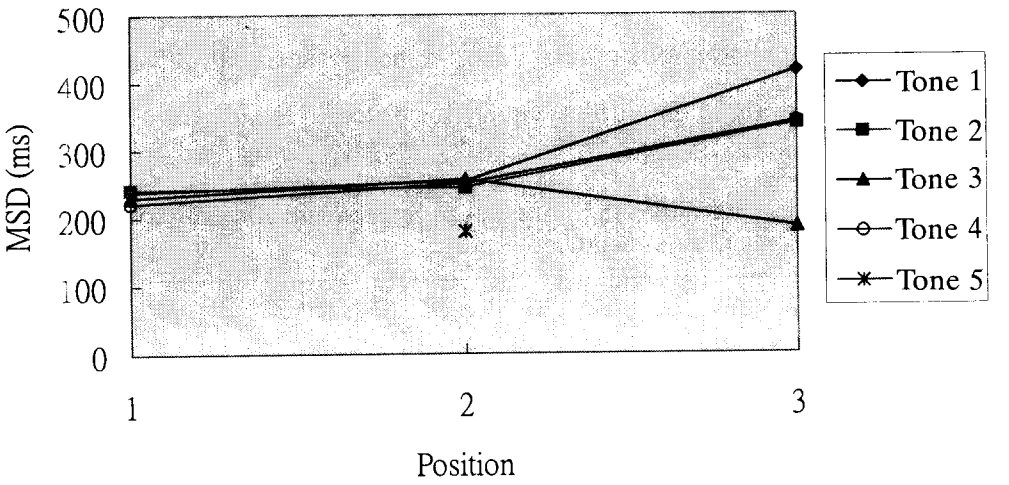


Figure 6. MSD in terms of Tone by Position in RS3



Although Tone 1 did not appear at Position 7 in Figure 4, we can claim quite strongly that it would be the longest if it had (because it was the longest at four positions and the second longest at two positions). Also note that Tone 3 in this speech was further shortened at the final position.

Figure 5 shows that at both initial and medial positions, MSD's for the four/five tones were much shorter than those at the final position. In Figure 6, the four/five tones almost collided at the same points at both initial and medial positions. It was only at the final position that significant differences were observed. Mark again that the third tone was further reduced in RS3 at final position in this speech sample.

Now, let's turn to other temporal variables. As we have mentioned earlier, of all the temporal variables, PT and AT are the most determinant ones. However, PT has remained quite constant across different types of speech (from 555 ms in RS1 and 558 ms in RS2 to 479 ms in RS3 and 491 ms in SS, $p=0.07$), but AT changed drastically, from 2187 ms in RS1

to 1145 ms in SS, a fact that resulted in variance in MSD.

As we compared our results with those found by other researchers, we found that our PT behaved the same as found in other studies—it centered around the mean in read speech, but had more variance in spontaneous speech. But in terms of PP, our figures were often much larger than those found by others. Also, we had rather small AR and SR. Let us put our results along with those obtained by others for different languages.

Table 9. Temporal Variables in Spontaneous Speech and Reading for Three Languages in Four Studies

		PT	AT	ST	Length	AR	SR	PP
SS	This study	491	1145	1637	5.43	4.69	3.25	31
	Duez (F)*	711	2063	--	--	5.2	--	22
Prose reading	This study	479	1731	2210	6.79	3.99	3.05	23
	Hardin et al.(M)	490	1370	--	6.5	4.74	--	26.1
	Hardin at al.(G)	490	2590	--	14.8	5.74	--	14.7
Poetry reading	This study	555	2187	2742	6.54	3.12	2.46	20
	K & K (G)**	540	1700	--	--	--	--	--
	K & K (K)	590	2240	--	--	--	--	--

*Duez (1982) had his subjects, six political figures, talk about their family life, etc.
**Kien & Kemp (1994) did not give other measures than PT and AT. Also, their figures are medians rather than means. [18]

Note that our PP was larger than that for French in spontaneous speech, and both ours and Hardin et al's were larger than that for German in prose reading. But Duez had quite long PT and AT for spontaneous speech. The fact that Hardin et al. had rather long AT in prose reading is probably due to the long utterances produced in that sample.

In poetry reading, again this study did not differ from Kien & Kemp in terms of PT. However, in terms of AT, Mandarin seems to bear more resemblance to Korean than to

[18] Kien & Kemp's study was not reviewed in section 3.2 because they were concerned with a motor theory of language production. However, their statistics on poetry reading in German and Korean were included for comparison.

German, both having longer AT. [19]

As differences observed across languages are not so straightforward, it is hard to arrive at any conclusion cross-linguistically. But we felt that Chinese is a language that is more likely to result in syllable lengthening (having longer articulatiom time), with a larger proportion of pausing time (having more pauses and shorter utterances). When reading is involved, the mean pause time does not differ very much across different types of speech.

We have treated the four types of speech as if they were of the same variable. Actually, there were three factors involved, one being Read vs. Spontaneous Speech (RS's vs. SS), another being Poetry vs. Prose (RS1 and RS2 vs. RS3 and SS), and the third one being Classical vs. Modern (RS1 vs. RS2, RS3, and SS). Thus, each type of speech had rather different features, as shown in the following table.

Table 10. The Four Types of Speech Defined by Three Factors

	RS1	RS2	RS3	SS
Read (+) vs Spontaneous (-)	+	+	+	-
Poetry (+) vs. Prose (-)	+	+	-	-
Classical (+) vs. Modern (-)	+	-	-	-

These three factors seemed to affect the temporal variables in different ways and to different degrees.

First, the mean pause time (PT) seemed to be a direct reflection of the factor of Poetry vs. Prose. As we have just presented, the differences among the four PT's across the four types of speech were not significant enough. However, there were obvious differences between PT's in the first two types (RS1 and RS2) and those in the last two (RS3 and SS).

---

[19] This is probably due to the form of poetry in the languages. Classical poems seem to be more fixed in form in both Chinese and Korean than in German. Still, another reason could be that Kien & Kemp included in their poetry reading a piece of modern poem. This could make a difference for modern poetry is more loosely structured than classical poetry in German too.

But, the differences observed in Figure 3 in terms of PT in various durations seemed to delineate the distinction between Read vs. Spontaneous speech. With spontaneous speech, PT of longer durations were found (resulting in greater variance), but in read speech PT was concentrated around the mean. This is espeically acute in RS1. Such a phenomenon seems to be quite universal (Goldman-Eisler 1968).

On the other hand, the issue of PP seems to reflect to a certain degree the Classical vs. Modern factor, for in RS1, PP was the smallest (20%), compared to 26%, 23%, and 31% for the other three.

Of special interest is the variable AT in which we can see the interplay of the three factors. First, poetry demands to be read more slowly than prose. Second, classical poetry must be read more slowly than modern poetry because of its highly concise form and condensed meaning. Finally, spontaneous speech can not permit too much lengthening unless hesitation occurs. The three factors working together resulted in very long AT for RS1 and very short AT for SS, and very close AT's for RS2 and RS3.

Our discussion has finally led us to say a few words about the psychological process underlying speech production of the Chinese language.

Clearly, classical poetry was read in a very special way. As we all know, a Chinese classical poem has a very concise form, i.e., its utterance length is fixed at 8 syllables a line. Besides, it has very condensed content—each of the eight lines is a complete sentence/proposition. These two features decide the way it is read. First, each line has to end in a pause that properly reflected the structural status of the line. While *duns* does not actually appear, syllables, especially those at *dun* positions, must be lengthened to convey every possible meaning, literal and metaphorical, and verbal as well as emotional, thus resulting in frequent syllable lengthening. Here, syllable lengthening, not pausing, is used to create poetic effects.

But in reading the modern poem, that is a slightly different story. Not only is its form

much looser--with lines of different lengths in run-on forms, but also its meaning is obvious—being more wordy and more repetitious. The 18-line modern poem makes only four complete sentences—more than twice the length of the classical poem but half the content. Given such a different structure and content, the modern poem was read differently—with more pauses (not longer mean pause time) but less syllable lengthening. Here, pauses, not syllable lengthening, were used to create the poetic flavor or emotional meaning, hence having the shortest utterance length. Presumably, syllables were lengthened just enough to express emotional meaning. This accounts for the small difference observed between its MSD and that of modern prose reading.

Modern prose is even more loosely structured, more colloquial, and more redundant. In reading the modern prose, AT was set to more or less the same parameter of tempo as in reading the modern poetry (1731 ms and 1682 ms). Besides, PT as well as PP was reduced such that it properly signaled constituent boundaries, with little concern for other meanings—emotional or rhetorical. The result was longest Utterance Length (6.79 syllables). The pause pattern observed for this type speech was that of the ideal delivery, resulting in very fluent speech.

The great PP found in the spontaneous speech was a totally different story. Here, pauses were not used to create a certain poetic or emotional meaning, but rather a result of disfluency or hesitation, for in this type of speech, there was the extra demand for verbal planning, global as well as local. Thus, pauses of much longer duration, say over 2000 ms, were bound to occur. But in reading, few pauses lasted longer than 2000 ms. This is best understood as we compare the variation coefficients of PT in RS2 (where it was the greatest of the three types of RS) and in SS—67% to 90%.

VII. Summary and Conclusion

The results of the present study suggest that there are different processes involved in the production of different types of speech. The act of reading aloud is basically a process in that the reader has to be careful in making out the sense of the read text and in getting the message across to the listener (including himself) with as much interpretation as required. On such occasions, speech must be delivered more slowly. But, on the other hand, as the reading text has already been decided, there is no need to plan for the content as well as the form (lexical items and sentence structures, etc.) of the speech. The very task of reading aloud has programmed the reader to read in a careful way such that the meaning of the text can be expressed. Careful reading results in prolonged articulation but not in more pauses. Pausing occurs only when it is necessary, that is, at major constituent boundaries as realized in the form of punctuation marks. This is why we have obtained for the three types of read speech longer syllable duration (because of longer articulation time), fewer pauses, moderate pausing time and utterance length. This is the typical form of ideal delivery. The process is mainly automatic. But, when the purpose of reading is to arouse emotions or feelings, articulation is further slowed down to express richer meanings, as in poetry reading (contrary to prose reading).

But in the spontaneous speech delivered by our subjects in the casual interview, the process is different. Here, speech is found to alternate between fluency and hesitation. As the speaker has to plan on the content as well as on the form of the speech, more pauses of a longer duration are often demanded. He often has to pause to search for the right word to say or the right way to say it, thus resulting in pauses of longer duration or in more pauses. As he is engaged in an interview, he must keep his speech in good tempo—he can't drawl. The results are faster speech and articulation rates, with little syllable lengthening (shorter AT), a greater range in pause duration, a larger PP and shorter utterance length.

The psychological processes underlying oral reading are very complex. For one thing, readers often have to have some degrees of understanding the text before the reading. Secondly, several linguistic levels (e.g., lexicon, syntax, phonology, semantics, etc.) and non-linguistic factors (e.g., background knowledge, cognitive functioning, emotional states, etc.) are involved. Finally, after everything has been understood and the reader begins to read, he is faced with the time pressure as well as the purpose of reading.

Traditionally, research on reading (aloud or silent) have mainly focused on the decoding of the linguistic aspects of the text—e.g., lexical access and retrieval or phonological encoding, paying little attention to the on-line processes themselves. Thus, most reading models fail to take into consideration the temporal variables in oral reading. By looking into the temporal variables in oral reading (or in speech production), we provide another perspective in understanding the real-time on-line processes of language production.

References

Barik, H. C. 1977. Cross-linguistic study of temporal characteristics of different types of speech materials. *Language and Speech*, 20:2, 116-127.

Blicher, D.L., R. L. Diehl & L. B. Cohen. 1990. Effects of syllable duration on the perception of the Mandarin tone 2/tone 3 distinction: evidence of auditory enhancement, *Journal of Phonetics*, 18:1, 37-49.

Brubaker, R. S. 1972. Rate and pause characteristics of oral reading. *Journal of Psycholinguistic Research*, 1:2, 141-147.

Chang Hsin-wu (張欣戊). 1992. The Acquisition of Mandarin by Chinese Children: Progress report (中國兒童學習語言的研究現況。中國語文心理學研究第一年度結案報告). Chia-yi: National Chung-cheng University.

Cheng, Chin-chiuan. 1973. *A Synchronic Phonology of Mandarin Chinese.* The Hague: Mouton.

Cheng, Ching-yi(鄭靜宜). 1994. The effect of tones, rymes, and syllabicity on the perception of Mandarin phonemes (聲調、韻母及音節性對漢語音素偵測的影響), MA thesis, the Graduate Program of Psychology, National Chung-cheng University.

Cheung, H. & S. Kemper. 1993. Recall and articulation of English and Chinese words by Chinese-English bilinguals. *Memory and Cognition*, 21:5, 666-670.

Clemmer, E. J., D. C. O'Connell, & W. Loui. 1979. Rhetorical pauses in oral reading, *Language and Speech*, 22:4, 397-405.

Duez, D. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech*, 25, 11-28.

Fourakis, M. & C. B. Monahan. 1988. Effects of metrical foot structure on syllable timing, *Language and Speech*, 31:3, 283-306.

Goldman-Eisler, F. 1968. *Psycholinguistics: Experiments in Spontaneous Speech.* London: Academic Press.

Goldman-Eisler, F. 1972. Pauses, clauses, sentences. *Language and Speech*, 15, 103-113.

Grosjean, F. 1980. Linguistic structures and performance structures: studies in pause distribution. *Temporal Variables in Speech: Studies in Honor of Frieda Goldman-Eisler,* ed. by Dechert, H. W. & M. Raupach (The Hague: Mouton).

Hardin, E. E. , D. C. O'Connell, & S. Kowal. 1998. Reading aloud from logographic and alphabetic texts: comparisons between Chinese and German. *Journal of Psycholinguistic Research*, 27:3, 413-420.

Han, M. S. 1962. The feature of duration in Japanese. *Onsei No Kenkyu*, 10: 65-80.

Hoequist, C., Jr. 1983. Syllable duration in stress-, syllable- and mora-timed languages, *Phonetica*, 40, 203-237.

Huggins, A. W. F. 1964. Distortion of the temporal pattern of speech: interruption and alternation, *Journal of Acoustic Society of America*, 36, 1055-1064.

Kien, J. & A. Kemp. 1994. Is speech temporally segmented? Comparison with temporal segmentation in behavior, *Brain and Language*, 46, 662-682.

Klatt, D. H. 1975. Vowel lengthening is syntactically determined in a connected discourse, *Journal of Phonetics*, 3, 129-140.

Klatt, D. H. 1976. Linguistic uses of segmental duration in English: acoustic and perceptual evidence, *Journal of Acoustic Society of America*, 59:5, 1208-1221.

Kowal, S., D. c. O'Connell, & E. J. Sabin. 1975. Development of temporal patterning and vocal hesitations in spontaneous narratives, *Journal of Psycholinguistic Research*, 4:3, 195-207.

Kuo Chin-fu (郭錦桴). 1993. *An Introduction and exploration of Mandarin tones* (漢語聲調語調闡要與探索). Beijing: Beijing Language Institute.

Lehiste, I. 1975. The phonetic structure of paragraphs, *Structure and Process in Speech*

Perception, ed. by A. Cohen and S. Nooteboom (Springer-Verlag, Heidelberg).

Lin, C. C.(林千哲) 1997. Syllable duration in spontaneous speech of Mandarin, unpublished paper, Graduate Program in Linguistics, NCCU.

O'Connell, D. C. & S. Kowal. 1972. Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research*, 1, 155-164.

O'Connell, D. C. , S. Kowal, U. Bartels, H. Mundt, & D. A. van de Water. 1989. Allocation of time in reading aloud: being fluent is not the same as being rhetorical. *Bulletin of the Psychonomic Society*, 27: 3, 223-226.

Oller, D. K. 1973. The effects of position in utterance on speech segment duration in English, *Journal of the Acoustical Society of America*, 54: 1235-1247.

Sabin, E. J., E. J. Clemmer, D. C. O'Connell, & S. Kowal. 1979. A pausological approach to speech development. *Of Speech and Time*, ed. by A.W. Siegman & S. feldstein (Hissldale, NJ).

Zeng, Yung-yi (曾永義) 1976. Elements of Poetry Rhythm (影響詩詞曲節奏的要素). *Zhung-wai-wen-xue* (*中外文學*), 4:8, 4-29.

Tseng, Chiu-yu (鄭秋豫). 1990. An Acoustic Phonetic Study on Tones in Mandarin Chinese. Nankang: Institute of History and Philology, Academia Sinica.

Umeda, N. 1975. Vowel duration in American English, *Journal of the Acoustical Society of America*, 58: 434-445.

Yang, I-li. 1997a. The pattern of pauses in monologues: a case in Mandarin Chinese, *Journal of NCCU*, 75, 207-256.

Yang, I-li. 1997b. Project report; A Study of Pauses in Mandarin Speech, NSC86-2411-H-004-015.

Zhu, Kuang-chien (朱光潛). 1976. The rhythm and prosody of Chinese poetry (中國詩的節奏與聲韻). *Zhung-kuo-shi-ji-kan* (*中國詩季刊*), 7:1, 15-67.

Appendix[20]

1.寄李儋元錫—唐韋應物

1.去年花裏逢君別，

2.今日花開又一年。

3.世事茫茫難自料，

4.春愁黯黯獨成眠。

5.身多疾病思田里，

6.邑有流亡愧俸錢。

7.聞道欲來相問訊，

8.西樓望月幾回圓。

2.雨巷—戴望舒

1.撐著油紙傘，獨自

2.彷徨在悠長、悠長

3.又寂寥的雨巷，

4.我希望逢著

5.一個丁香一樣地

6.結著愁怨的姑娘。

7.她是有

8.丁香一樣的顏色，

9.丁香一樣的芬芳，

10.丁香一樣的憂愁，

11.在雨中哀怨，

12.哀怨又徬徨；

13.她徬徨在這寂寥的雨巷，

---

[20] In the modern poem and the modern essay, only the sentences in double parentheses and in bold face are used for durational analysis.

14.撐著油紙傘

15.像我一樣，

16.像我一樣地

17.默默彳亍著，

18.冷漠，淒清，又惆悵。

3.抒情文

山裡有風景，但是難以形容。走過清香細細的松林，渡過獨木的窄橋，沿一條流聲若隱若現的山溪，上到岩崗，下到谷地，是絕不同於走過鬧市的體驗。平時上街，走過就走過了，並不希求發現什麼。而山徑迂迴，陌生，奇妙，轉折處可能有一池清潭、一湖苔藻，或一隻鹿、一對鴨，或山谷驟然開朗，或奇峰巍巍聳立。**((景致不斷變化，如一曲複雜的樂章。這樣一路走去，彷彿走在一闕長長的詩裡，自己也變成一句婉轉的詩。**

有時景緻令人厭倦:單調的樹林，雜沓的亂石，寸草不生的巨岩。或是只是重複出現的景色，見慣不奇，失去了當初的驚奇和興奮。)) 到洛磯山國家公園許多次以後，對一些景緻逐漸因爲熟悉而漠然了。即使壯偉的連綿巒峰，深鑒的翠綠河谷，最後徒然感覺不過爾爾:山就是這樣了!雖然這一切，其實無需我的肯定。

事實上，有時出發勝於到達。彷彿寫作，有一個意象，一個原始的構想，腦中充滿了光，像擰亮了一盞燈，整個人爲了即將開始的作品非常燦爛。等到完成以後，卻難免不滿意。那心目中的境界永遠懸在鼻尖，咫尺而不可及，如腑臟中搔不到的殺人的癢。

Temporal Variables in Read and Spontaneous Speech

朗讀與自然口語中的時間變項

楊　懿　麗

政大語言所教授

（收稿日期：1999年6月5日；修訂日期：1999年8月25日；接受刊登日期：1999年9月27日）

摘　要

本論文探討朗讀與自然口語中時間變項之出現情形。我們請八位成人用國語朗讀三段文字，之後並且和他們自由交談。然後分析這四種語料中的音節長度、停頓等時間變項。我們假設朗讀和自然口語需要不同的心理歷程，而且不同文字的朗讀也會產生不同的時間形態。結果如原先所預料，摘要如下：

1.不同的文字確實影響音節的長度。朗讀唐詩時的音節為335毫秒，現代詩為277毫秒，現代散文為252毫秒，而自然口語則為208毫秒。

2.不同的文字也影響其他的時間變項。朗讀唐詩及現代詩時所產生的平均停頓時間最長（555和558毫秒），朗讀現代散文及自然口語時所產生的停頓平均卻只有479和491毫秒。同時發聲時間在詩的朗讀也最長，而在自然口語則最短。因是之故，說話速度有相當明顯的不同。

3.從這些時間變項的不同，我們覺得朗讀和自然口語有不同的心理歷程。同時，詩詞朗讀與散文朗讀亦有不同的歷程。