

## THE STRUCTURE OF THE CHINESE TEXT: PROSODY AND GRAMMAR\*

VADIM KASSEVITCH – NIKOLAI SPESHNEV\*\*  
(St. Petersburg)

This paper presents an overview of the authors' recent findings concerned with the rhythmic structures typical of the Chinese text. The proportion of disyllabic and trisyllabic words is shown, disyllabic words or *binomials* being predominant. The intervals between two strong (full-toned) syllables are found to reproduce the same two-syllable pattern. Chinese is proposed to be classified as a *binomial*-timed language. With accelerating speech rate, certain syllables tend to lose their inherent full-tone characteristics, thus becoming "weak" or even light (neutral). Since full-toned "strong" syllables are often interpreted as stressed, their "disappearance" under certain conditions seems to be another evidence of superfluity of word-level phonological stress in Chinese.

*Key words:* Chinese, tone, stress, rhythmic patterns, word, grammatical structure.

0. At the very outset, it seems appropriate to define, at least loosely, the central terms we are going to use. We mean, first of all, such terms as *stress*, *accent* and *syllable tone*. Unfortunately, the widespread belief in the rigorous nature of phonetic/phonological terms is something of an illusion. The above mentioned terms are no exception, which makes it necessary to show our conceptualisation of the terms.

Generally speaking, lexical *stress* is a functionally relevant relative saliency of a syllable within the word. At least in the context of this paper, we would not differentiate between stress and accent. For the sake of simplicity, we put aside the problem of secondary stresses; as a result, the above preliminary definition implies that there is just one stressed syllable in each (phonological) word. This means, in turn, that the stress produces a *syntagmatic* contrast within the word. Given the number of syllables, the position of stress makes it possible to *predict* the accentual (stress) pattern of the word.

\* This is a slightly enlarged and updated version of the paper submitted at the International workshop "Tone, stress and rhythm in Spoken Chinese" (Prague, 1999). For the sake of brevity, bibliographic references are reduced to the absolute minimum.

\*\*# Vadim Kassevitch, University of St. Petersburg, P.O. Box 14, St. Petersburg, 191025, Russia, e-mail: kasevich@vbk.usr.pu.ru

It follows from the above that the number of stressed syllables and the number of words within a text are in one-to-one relationship, which makes it reasonable to argue that the main function of stress is the segmentation of speech into word-sized chunks rather than contributing to the word's distinctive potential as believed by many.

Unlike stress, *tone* is an inherent prosodic feature of *each* syllable, also phonologically relevant; in tone languages, syllables are opposed to one another in terms of their tones precisely as they are with respect to their consonants and vowels (or, rather, to their initials and finals, but this is a special problem in its own right, which cannot be discussed at the moment, see Kassevitch 1983).

Tones thus produce a *paradigmatic* contrast. The information about an individual syllable's tone within the word cannot and does not predict the latter's prosodic pattern, since, generally, individual tones are not constrained in their linear distribution.

**1.** The point that makes stress and tone functionally commensurable phenomena is as follows: Both are associated with the lexical units which are central to the system. There is no word lacking stress in such languages as Russian or English and, likewise, there is no syllable (or, rather, *syllabomorpheme*, but this, again, is a special problem) in Chinese which would lack a tone.<sup>1</sup>

If we agree to the above points, we will be prepared to accept that tone and stress are hardly mutually tolerant. Given the functional similarities and dissimilarities of the prosodic phenomena in question, they could be expected to be mutually exclusive rather than co-existent.

Without going into many details, we could refer to the perceptual experiments run by Michail Rumjantsev and his associates, where it has been shown that the putative minimal pairs like /pau<sup>4</sup> 'kau<sup>4</sup>/ '(a) report' vs. /'pau<sup>4</sup> kau<sup>4</sup>/ '(to) report' cannot be disambiguated by native speakers of Chinese (Rumjantsev 1974).

It could be added, just to complete the picture, that, alongside tonemicity, there is still another source that provokes the lack of stress or accent: it is vowel *harmony* which makes such languages as Mongolian equally stressless.

**2.** Yet, at this stage of our reasoning, it could be recalled that in many languages like English, there are two ways of classifying syllables prosodically. One is based on the [±stressed] distinctive feature, while the other on the [±strong] feature. The two classifications obviously overlap, but cannot be reduced to just one. The above claimed lack of lexical phonological stress in Chinese in no way rules out a possibility of having the same kind of opposition, that between strong vs. weak syllables.

It is important to emphasise that strong syllables in Chinese are *tonally* strong, and the weak syllables equally display their weakness from a tonal point of view. In other words, tones can be fully realised, their set of features, both distinctive and in-

<sup>1</sup> We would refrain from discussing both problems of so called "stressless" words in Russian or toneless words in some of Sino-Tibetan languages and dialects (for a discussion see Kassevitch 1983).

herent, approaching that associated with the citation form, in which case the respective syllables are strong; or their realisation can be partial, in which case the weak syllables are present. The light-tone syllables, although marked with a phonologically distinct zero tone, also enter the same opposition; as a result the opposition acquire its three-way shape, viz. strong-tone syllables vs. weak-tone syllables vs. zero-tone syllables.

Of course, we cannot discuss here the acoustic features responsible for identifying syllables as strong, weak or zero-toned. We will give, instead, a few examples that would show which typical syllable collocations are construed as incorporating the above postulated tonal categories (in what follows S stands for “strong”, W for “weak”, and Z for “zero (light) tone”). E.g. *la<sup>3</sup>ba<sup>0</sup>* ‘trumpet’ (SZ), *he<sup>1</sup>shui<sup>3</sup>* ‘drink (drink water)’ (WS), *mu<sup>4</sup>jiang<sup>4</sup>* ‘carpenter’ (SW), *fu<sup>4</sup>mu<sup>3</sup>* ‘parents’ (SS or WW). The latter type is remarkable for its equally “weighted” component syllables which can be realised as both strong or both weak depending on the position and/or speech rate (cf. below).

Even this very small set of examples given above seems to hint at a possible role of grammar and semantics in assigning prosodic weights to syllables within two-syllable words. Weak syllables are often associated with (partial) grammaticalisation, where the second syllable typically has agentive or similar meaning. Light tone is very often used with grammatical markers.

The case is more controversial when we come to three-syllable words. In such words, the second syllable tends to exhibit certain lenition, but its phonological interpretation is problematic. The trouble is that the aforesaid lenition is realised as a tone leveling, e.g. *xian<sup>1</sup>ren<sup>2</sup>zhang<sup>3</sup>* ‘cactus’ → *xian<sup>1</sup>ren<sup>1</sup>zhang<sup>3</sup>*; *han<sup>2</sup>shu<sup>3</sup>biao<sup>3</sup>* ‘thermometer’ → *han<sup>2</sup>shu<sup>2</sup>biao<sup>3</sup>* → *han<sup>2</sup>shu<sup>1</sup>biao<sup>3</sup>*; *hai<sup>2</sup>mei<sup>2</sup>lai<sup>2</sup>* ‘did not come yet’ → *hai<sup>2</sup>mei<sup>1</sup>lai<sup>2</sup>*; *er<sup>4</sup>nian<sup>2</sup>ji<sup>2</sup>* ‘second grade’ → *er<sup>4</sup>nian<sup>1</sup>ji<sup>2</sup>*. In our transcription, we use the 1st tone index to mark the reduced, weakened middle-position syllables. In other words, the transcription reflects an analysis, where the “weakness” of the syllables is understood as their *unmarkedness*, since the 1st (level) tone is the least marked in the tone system of Chinese. Yet, it has to be verified, if these middle-position syllables are phonologically identified with non-derived 1st tone syllables.

Not all three-syllable words are liable for the lenition (weakening) of their central syllables, cf. *lian<sup>2</sup>he<sup>2</sup>guo<sup>2</sup>* ‘United Nations’, *bei<sup>4</sup>wang<sup>4</sup>lu<sup>4</sup>* ‘memorandum’, *xi<sup>1</sup>hong<sup>2</sup>shi<sup>4</sup>* ‘tomato’, etc. We fail to offer an explanation to such a different behaviour of the apparently similar structures. The only tendency we can observe is a relatively more pronounced easiness of 2nd tone syllables to undergo levelling.

**3.** Practically, in the Chinese running text, one can find 5 different types of two-syllable *rhythmic feet*, namely, strong syllable + strong syllable; weak syllable + weak syllable; strong syllable + weak syllable; weak syllable + strong syllable; strong syllable + zero syllable. As a matter of fact, just the latter 3 are real feet, since the first is either distributed between adjacent words or, anyway, has a short pause between the 2 syllables, while the ‘weak+weak’ pattern is speech-rate dependent, it is not attested in (prosodically) strong positions, nor is it possible in isolation.

As can be seen, the strong-comes-first pattern predominates, which, however, does not mean that, rhythmically, Chinese is a predominantly *trochaic* language (cf. below).<sup>2</sup>

**3.1.** Without attempting to present any formal rules for the assignment of S/W/Z weights to the syllables, we will just mention a few specific regularities that govern the distribution of the markers in the text.

As is well known, light-toned particles and most function words rhythmically go with their left neighbours; it is also the case with the copula verb *shi*, cf. *wo shi*, where the *shi* may be realised as either weak or zero syllable, cf. *wosh*. As distinct from that, conjunctions, negation markers may join either right-hand or left-hand syllables (syllabomorphemes).

In conformity with the predominantly *binomial* structure of the Chinese linguistic units (which will be discussed in some more detail later), one can find cases of a re-distribution of S/W positions, depending on which binomial the syllable belongs with, cf. *zhi dao* ‘(I) know’, where the first syllable is strong, but *ni zhi dao bu zhi dao* ‘whether you know’, lit. ‘you know, (you) don’t know’, where the final *dao* becomes strong, while the first *zhi* acquires a near-to-zero shape losing its inherent 1st tone. It can be seen from the above example that, in some cases, the zero-tone syllable may not count as a binomial-foot member, which is in agreement with the prosodic rules of the Chinese poetry, where light-tone syllables may fall outside of the syllable count that underlies the poetic rhythm.

**3.2.** With these and some other regularities in mind, we have marked “manually” as S, W or Z all the syllables in a modern prosaic text consisting of more than 5,000 syllables. The text was read out twice by a native speaker of Chinese, the rate of speaking being normal and accelerated. Our aim was to find out how many syllables make the interval between two successive strong syllables.

**3.2.1.** To explain the purpose of the said statistical survey, we will have to recall the old problem of classifying languages into stress-timed vs. syllable-timed types originally introduced by Kenneth Pike. The syllable-timed languages are supposed to make the speech pattern regular (from a rhythmic point of view) due to their syllables being isochronous. As distinct from that, a fairly regular rhythm of the stress-timed languages is provided by levelling the intervals between successive stressed syllables. According to Pike, Chinese is likely to fall within the second class (with such diverse languages as Spanish or Japanese), while English, among others, will be considered as a typical stress-timed language.

As shown by Cutler (1980) and some other authors, the Pikean paradigm is borne out by English experimental data which seem to nicely agree with the pre-

<sup>2</sup> It is interesting to add that the distribution of tones within two-syllable words shows a clear tendency to mark the 2nd syllable with the 4th tone. The 4th tone is *falling*, which produces an effect close to that of *declination* within the sentence. In other words, basic lexical structures display a prosodic behaviour somehow similar to that of basic syntactic structures.

dictions made by the theory. In Cutler's experiments, the subjects went so far as to completely suppress syllables whose presence could violate too drastically the trend to maintain the inter-stress intervals roughly equal.

On the other hand, numerous experiments, both earlier and recent, have made it fairly evident that the clear-cut distinction suggested by Pike is an oversimplification (cf. Lehiste 1977; Dauer 1983; Borzone de Manrique – Signorini 1983; Bertinetto 1989; Carvalho 1989; Crystal – House 1990; Nootboom 1997, etc.). At least the languages, expected to display a regular syllable duration, have been found to exhibit widely varying strategies of timing, far from the simplistic trend to make all the syllables isochronous. As regards Chinese, the very presence of the zero-tone syllables with their extra-short duration makes the claim for the isochronous-syllable pattern clearly unrealistic.

**3.2.2.** At the same time, the three-way opposition of the syllables referred to above leads us to suspect that there may be certain regularities concerning the dimensions of the “inter-strong”, so to speak, intervals. Our statistical survey has shown that no *single* interval-pattern can be found. Yet there are good reasons to speak of *statistically predominant* trends, viz.: 30 per cent of all the intervals are two-syllabic in size; if the statistics for the three-syllable intervals are added, the two types account for more than 40 per cent of the total, whereas percentage of all the other intervals shows no comparable figures. It can be suggested, as a result, that the predominant rhythm pattern of Chinese with its typically two-syllable interval between successive strong syllables is *binomial*; in other words, Chinese is a *binomial-timed language*.

**3.2.3.** It is of utmost importance that, with accelerating the speaking rate, some *strong* syllables of the Chinese text (“stressed”, from a traditional point of view) tend to become *weak*, while the latter may be further reduced to the zero (light) type. This seems to testify once more to the *lack* of any lexical stress in Chinese, since the word-level phonological stress can shift rather than “disappear” at all.<sup>3</sup>

**4.** The “fore-grounded” two-syllable rhythmic pattern makes us recall that the two-syllable linguistic units are of special importance to Chinese lexicon as well as to the lexical and grammatical structure of the Chinese text. The point is that Chinese words and word-like units are typically *disyllabic*. Our recent results have shown quite convincingly that such disyllabic units may be stored in the speaker's mental lexicon even irrespective of the grammatical role of the units, notably irrespective of their being words or word-combinations. This claim is substantiated by our finding that Chinese subjects analyse a running text into fairly big chunks, which would definitely overload their short-term memory, if counted in terms of syllables. If, however, counted in terms of binomials, the size of the chunks is easily accounted for.

<sup>3</sup> In Russian, under the same experimental conditions, no significant changes in distribution of stressed and unstressed syllables have been attested, which is, of course, quite predictable.

The above feature of the Chinese language – the disyllabic size of its operative units – adds a new dimension to our theory. If the word is practically equal to the rhythmic foot and, moreover, the interval between two strong syllables tends to be also composed of two syllables, this means, in effect, that, both rhythmically and lexically (grammatically), the Chinese text is structured in terms of binomials. The rhythmic pattern and the functional (lexical and grammatical) structures are found closely correlated. In fact, such a correlation is an argument in itself for the proposed type of rhythm, as the latter is rightly believed to be *functional*.

From a purely rhythmical point of view, the combination of the two basic foot-types (S+W/Z and W+S) and the binomial intervals leads us to conclude that Chinese is neither trochaic nor iambic language, since, typically, trochaic and iambic feet must alternate more or less regularly.

**4.1.** In Western linguistics, this tendency of the Chinese language to operate with disyllabic units was underscored as early as 19th century in the classic works by Georg von der Gabelentz (1881/1953) who attached much importance to *Wortpaare* in Chinese. In the German author's opinion, the said tendency should be explained by a necessity to compensate for the predominantly monosyllabic *Stammwörter*. It does not escape his attention that this predisposition towards disyllabic units results, among other things, in the typically four-syllable structure of the Chinese clause where both NP and VP (to use modern terminology) mostly consist of two syllables. It could be added that in cases where leftmost NP is the topic of the utterance, it is always separated from what follows by a pause. Such a pause functions as the major boundary, both syntactic and rhythmic, within the utterance, it is most resistant to the rate-of-speaking changes and thus contributes greatly to the overall rhythmic pattern of the utterance.

**4.2.** Speaking more generally, Gabelentz is inclined to interpret the “bias for even-numbers”, implicitly referred to above, in terms of Chinese culture. He emphasises a general inclination of the Chinese to operate with “parallel and antithetic structures” (*frühentwickelter feiner Sinn des Volkes für Symmetrie und Antithese...*). What is considered to be the parallelism in the Chinese text, says Gabelentz, is the combined effect of its specific rhythm and the antithetic semantic structure.

In fact, Gabelentz crucially anticipates modern theories of such writers as Barthes who finds it an important feature of the traditional-type discourse to operate with parallel structures: the reproduction of certain constituents subsumed by the parallelism makes the text more coherent, since constituents become more predictable, their distribution acquires a more regular character.

**5.** At this point we come back to the general problem of the Chinese rhythmic structure. Our approach, roughly outlined above, seems to agree with the general theory of rhythm. Loosely defined, rhythm is a reproduction, usually in time, of certain *hierarchical structures* (cf. Fonagy 1982). Any hierarchy presupposes foregrounded head constituents against backgrounded ones. Yet, if the syllable-timed pattern is

construed as a mere reproduction of isochronous syllables, no hierarchy seems to come into play (at least, if one doesn't take into consideration the inner structure of the syllable as such). As contrary to that, the reproduction of disyllabic units (=binomials), where one finds the strong syllable as the prosodic head, definitely restores the needed hierarchy.

6. Of course, it is only reasonable to admit that language speakers generate a number of different rhythm patterns. As a result, a running text is also characterisable in terms of different rhythm patterns, each of them serving its special function. The binomial rhythm of the Chinese text appears to provide a "base-line"; the other rhythm patterns, including those associated with syntactic and topic/comment structures, are presumably superimposed on that prosodic base.

### References

- Bertinetto, P. M. (1989): Reflections on the Dichotomy 'Stress vs. Syllable Timing'. *Revue de Phonétique Appliquée* 91/93.
- Borzone de Manrique, A. M. – Signorini, A. (1983): Segmental Duration and Rhythm in Spanish. *Journal of Phonetics* Vol. 11, No. 2.
- Carvalho, J. B. de (1989): Phonological Condition on Portuguese Clitic Placement: On Syntactic Evidence for the Stress and Rhythmical Patterns. *Linguistics* No. 27.
- Crystal, T. H. – House, A. S. (1990): Articulation Rate and the Duration of Syllables and Stress Groups in Connected Speech. *JASA* No. 88.
- Cutler, A. (1980): Syllable Omissions Errors and Isochrony. In: Dechert, H. W. – Raupach, M. (eds): *Temporal Variables in Speech: Studies in Honor of Frieda Goldman-Eisler*. The Hague etc.
- Cutler, A. (1996): Prosody and the Word Boundary Problem. In: Morgan, J. L. – Demuth, K. (eds): *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale, N.J.
- Dauer, R. M. (1983): Stress Timing and Syllable Timing Reanalyzed. *Journal of Phonetics* Vol. 11, No. 1.
- Fonagy, I. (1982): *La ripetizione creativa: Ridondanze espressivo nella opera poetica*. Bari.
- Gabelentz, G. von der (1881/1953): *Chinesische Grammatik*. Berlin.
- [Kassevitch] Касевич, В. Б. (1983): *Фонологические проблемы общего и восточного языкознания* [Kassevitch, V. B.: Phonological problems in general and oriental linguistics]. Moscow.
- Lehiste, I. (1977): Isochrony Reconsidered. *Journal of Phonetics* Vol. 5, No. 1.
- Nooteboom, S. (1997): The Prosody of Speech: Melody and Rhythm. In: Hardscastle, W. J. – Laver, J. (eds): *Handbook of Phonetic Sciences*. Oxford.
- [Rumjantsev] Румянцев, М. К. (1974): К проблеме ударения в современном китайском языке путунхуа [Rumjantsev, M. K.: The problem of stress in modern Chinese]. In: *Вопросы китайской филологии* [Issues in Chinese philology]. Moscow.