

FOREWORD
Hana Tříšková

The present volume results from *the International Workshop Tone, Stress and Rhythm in Spoken Chinese* held in Prague in May 1999. The workshop was jointly organized by the CCK International Sinological Center at the Charles University, and the Oriental Institute of the Academy of Sciences of the Czech Republic.

In comparison to studies on written languages, research on spoken languages does not have such a long history. In recent years we can observe a growing interest especially in suprasegmental features of languages (one of the reasons being the needs of rapidly developing speech technologies). The above holds good for Chinese linguistics too. The aim of the Prague workshop was to bring together specialists working in this field. The meeting proved that substantial progress has been made in the past years, although the approaches to this subject are diverse. Besides the importance of the topic itself, there were also 'historical' reasons for organizing this event in Prague. The tradition of phonological studies carried out by the Prague Linguistic School reaches back to the 1930s. Furthermore, research on Chinese phonology and phonetics was conducted here in the course of several decades by Prof. Oldřich Švarný, who turned eighty last year. This volume is dedicated to him.

The workshop offered an international context for Švarný's work, which is pioneering in many aspects. His research on Mandarin prosody, launched in early 1950s¹ got a major impetus during his stay at the University of California at Berkeley in 1969/1970. Švarný carried out an instrumental analysis of fluent Chinese speech in the Phonology Laboratory of Prof. William S. Y. Wang. He experimentally verified several levels of stress in Pekinese and acoustic cues for segmentation. In subsequent research Švarný studied accentuation of compounds. Relying on broad statistics, he outlined seven 'accentuation types' of disyllabic words and described major factors conditioning their variability. Švarný's studies

on Madarin prosody resulted in a design of prosodic transcription, based on *pinyin*. The system has a strong theoretical base and was successfully tested in the teaching process. It should be noted that Švarný's scholarly erudition was always inseparable of his willingness to take up educational responsibilities. Thanks to him, Czech students of Mandarin have teaching materials at their disposal, which stand up to theoretical standards in their description of prosody.

A unique feature of all of Švarný's language teaching works is voluminous exemplificative material available both on tapes and in prosodic transcription. Numerous attempts to mark prosodic features of Mandarin speech for pedagogical purposes were made in the past (e.g. N. A. Speshnev: "Fonetika kitajského jazyka", Leningrad 1970; "Practical Chinese Reader", Beijing 1988; Wu Jiemin: "Xinbian putonghua jiaocheng", Hangzhou 1988). However, Švarný is undoubtedly the first one to implement a prosodic transcription on such a large scale and in such a systematic way. The ability to employ theoretical findings in pedagogical materials compiled for practical use is one of the major Švarný's merits.

At the end of 1990s, Švarný published an extensive dictionary "Učební slovník jazyka čínského" (Learning Dictionary of Modern Chinese²) in four volumes³. This work has two unique features distinguishing it from a standard dictionary. First, entries (i.e. characters in a certain reading) are analyzed into semantic fields – *yusus*⁴. Every *yusu* is equipped with numerous examples of both free and/or bound usage. The second major objective of the dictionary is to describe the prosody of Mandarin utterances. Prosodic transcripts of 16 000 exemplificative sentences⁵ make up an essential part of the dictionary. Prosody is viewed as a mean for expressing numerous linguistic functions beyond lexical tones, including accentuation of compounds, sentence stress, focus, sentence intonations etc. This voluminous work has to be considered as the outcome of Švarný's lifelong research on Chinese grammar and prosody.

The papers presented at the workshop (altogether sixteen) touched upon the subject of Mandarin prosody from various angles – they dealt with tonal variations in connected speech, speech rhythm and nature of stress, comparison of accent phenomenon across Chinese dialects, rhythm as a stylistic device, intonation, relationship between prosody and grammar, or prosodic annotation of

a speech database. Some contributions offered a historical perspective, or a language teaching perspective of the topic. To make the present volume coherent, the editors decided to choose out of all papers mainly those dealing with the experimental phonetics. However, it has to be pointed out that the papers not included here brought many new ideas and substantially contributed to the overall success of the workshop.

Human speech is materialized in sound waves. However, the communicative information encoded in acoustic waveforms is extremely complex. To reveal the contribution of particular factors influencing the prosodic shape of Mandarin utterances and to find proper tools for its description are among the major research objectives of the studies on Chinese prosody. While encouraging results were achieved in many aspects (e.g. the effects of downstep and declination, or the interplay of adjacent tones) are rather well documented, other effects are not profoundly explained yet (e.g. stress assignment rules, the interplay between prosody and grammar, pragmatic and emotional functions). The authors of the following pages concentrate on various aspects of prosody of Mandarin (i.e. of Standard Chinese, only in case of Chang Yueh-chin Taiwan Mandarin): sources of F_0 variations of lexical tones (Xu, Shih), rhythm (Cao), links between prosody and grammar (Třísková and Sehnal, Chang, Feng), and historical development of stress rendering (Endo). The speech materials on which the experimental studies are based are either read speech recorded in laboratory conditions (Shih, Xu, Chang, Třísková and Sehnal), or TV news and broadcasting (Cao).

In all languages, prosodic features are carried by three major acoustic parameters: fundamental frequency, intensity, and duration. However, the specific ways of their utilization for expressing particular linguistic functions vary. To give an example - while in some languages, for instance in Czech, duration has a distinctive function at the segmental level, in Mandarin the increased duration of a syllable typically signals stress. Yet another example: while in non-tone languages we are accustomed to attribute the F_0 modulations primarily to the factors rooted at sentence level (such as sentence intonation), in Mandarin, pitch is functionally used also on the lowest prosodic level - level of syllables - to distinguish meanings of various *yusus*. Sometimes superficial observers wrongly assume there is no room for sentence intonation in Mandarin.

as both tones and intonation are manifested by pitch changes. However, as Xu and Shih point out, sentence intonation, focus and tones are realized by different aspects of F_0 contours (tones are shaped by local F_0 contours, while focus and intonation are expressed by pitch range variations). The term 'intonation' is commonly used as a general term covering pitch variations of speech. Xu suggests that there is in fact nothing left for an independent entity called intonation, once various F_0 shaping factors are identified.

Tone in Mandarin is characterized by a set of acoustic features, distinctive F_0 curve being the most striking one. In connected speech, canonical forms of tones undergo more or less dramatic changes in all acoustic parameters. Thus, one of the research objectives is to find out how lexical tones are realized in utterances, and to disclose the sources of their behavior. Xu and Shih make a substantial contribution to this issue in their papers. Both of them are focused on F_0 variations. Xu sheds light on the complexity of factors affecting F_0 curve, identifying and categorizing them. Lexical tones, prosodic structure, syntax, pragmatics and emotions are listed among major *voluntary factors*. On the other hand, *involuntary factors* he defines as the limitations of the articulators. Xu demonstrates how some of the voluntary and involuntary factors interact with one another in producing F_0 contours. His experiments deal with three types of effects, related to different linguistic levels: (1) pitch contour variations due to adjacent tones, (2) interplay of tone and focus, (3) mechanism of downstep and declination. Xu concludes that to obtain a clear picture of F_0 variations in Mandarin, the distinction between communicative intent (reflected in voluntary factors), and involuntary articulatory constraints always needs to be made.

Shih attempts to isolate effects of individual factors for intonation analysis and data normalization, and to combine them for intonation generation. She draws a hierarchical prosodic structure, where particular layers of intonational effects are rooted in different linguistic levels. Similarly to Xu, various effects are treated separately as additive components contributing to the surface F_0 contour. Shih analyzes the *segmental effects*, rooted at the segment level, and the *declination effect*, rooted at the sentence level. The results of experiments encouragingly show that segmental effects are quite predictable. Experiments on declination effect observed its interaction with sentence length and focus. Concluding experiment on F_0 generation was done by summing

various effects. The clear advantage of Shih's model of F_0 normalization and generation is its modularity, which allows exploring the effect of particular factors separately and to utilize results obtained from other studies.

Speech rhythm is related to both speech production and perception. Perceived rhythmic organization of speech usually corresponds to certain acoustic-phonetic correlates. However, there is no straight correspondence between the measured values and the perceived qualities. To paraphrase Xu, we can suggest that there is no independent entity of 'rhythm'. It is just a cover term for all relevant factors contributing to the overall rhythmic percept. Speech rhythm is often defined and treated in various ways. We still lack a generally accepted notion (Švarný's works⁶ offer one of the scarce systematical concepts of rhythm in Mandarin). It is commonly recognized that speech rhythm forms a hierarchy. However, there are differences between the number of hierarchical levels that particular authors recognize. Švarný marks two rhythmical levels in his prosodic transcription: 'rhythmical segment' (composed of disyllables and/or odd syllables), and 'colon'. Cao recognizes three hierarchical levels above the syllabic level: 'minor rhythmic unit', 'intermediate rhythmic chunk' and 'major rhythmic group' (corresponding to prosodic word, prosodic phrase and intonation phrase of metric phonology). She attempts to find acoustic cues for the boundary markers of these rhythmical units, and the coherence features bonding together their components. Cao's hierarchy of junctures is supported by pitch and duration measurements and perception tests as well. As a material she uses TV news and broadcasted speech. Mandarin Chinese is traditionally viewed as a stress-timed language with a strong tendency towards isochrony. However, the theory of isochrony as such has its critics. Cao claims that she found no evidence for so called isochrony in Mandarin (unlike Švarný, who strongly advocates plausibility of the concept of isochrony for Mandarin). Further on, Cao questions the relationship between prosody and syntax. Similarly to other authors, she concludes that the correspondence between prosody and syntax is not direct.

Třísková and Sehnal approach the issue of rhythm and its relationship with grammar from the angle of corpus annotation and statistical processing. They introduce the PALM software, designed to grasp and analyze the basic rhythmic structure of Mandarin utterances. A small corpus was prosodically transcribed and annotated for various prosodic and grammatical features.

Tříšková explains theoretical basis of her prosodic transcription, which was partly inspired by Švarný's system (a simplified version is proposed for pedagogical purposes). Statistical analysis of the annotated database is carried out, observing various combinations of prosodic and/or grammatical features of either syllables, or words. Several examples of utilization are offered. Tříšková's examples deal with stress and tone features of the syllables depending on speech tempo, Sehnal is focused on mutual dependence between the grammatical features of words, and their stress features. The PALM project is one of a few labeling systems devised for Mandarin which includes prosodically labeled data. The software can be applied to a larger database to study the links between rhythmical structure of the Mandarin utterance, its grammatical structure and speech tempo.

Feng's paper deals with the historical development of *ba* sentences, explaining synchronic phenomena with diachronic studies. Prosody is viewed here as an important factor contributing to syntactic changes. Besides the links of stress assignment rules to the syntactic structure of the sentence, Feng also discusses the relationship of these rules to the semantic structure. He suggests that *ba* sentences spread out to natural speech from poetry while changing their structure, semantics and consequently the stress rules in the course of this process. Feng argues that the *ba* construction first appeared in early Tang poetry. *Ba* sentences of this type [*ba* NP V] had the main stress falling on NP. With further development the structure and consequently the semantics of *ba* constructions changed - the predicate became more complex, expressing a delimitative event. However, delimitation requires the object to be specific. In natural speech, the inevitable result was the loss of stress of the NP. *Ba* became out of focus and was reduced to an empty verb. In natural speech this was grammaticalized as a new pattern with an unstressed NP and stressed predicate.

Chang studies prosodic cues for disambiguation. It is well known that Mandarin Chinese is highly homonymous. This phenomenon has several sources - in particular a restricted choice of syllabic structures, the rarity of polysyllabic words (according to "Xiandai hanyu pinyin cidian" 1986, in colloquial speech about 75% of word occurrences fall to monosyllabic words), lack of inflection etc. Consequently, sometimes the sole phonetic information is not sufficient to distinguish between unambiguously structured words/phrases. On the other hand,

there can be pairs of phrases or words, which are structurally ambiguous, and prosodic features of speech can help to disambiguate them. Chang is testing both lexically ambiguous phrases, and structurally ambiguous phrases. The experimental data showed no significant acoustic difference for lexically ambiguous phrases. For structurally ambiguous phrases, though, she found differences in duration in some types of syntactic structures (while no consistent differences in F_0 were discovered). The perception tests showed that if there was an acoustic difference, the subjects could perceive this difference well and use it to disambiguate sentences. Duration proved to be a more robust acoustic cue than F_0 . If there was no significant acoustic difference, the subjects tended to rely on sentence frequency or word frequency to disambiguate.

Endo offers a historical perspective to the reflection of stress in Mandarin. He shows historical evidence of the existence of stress phenomenon. The evidence of stress can be found in old poetry (i.e. the rhyming features), or transcription materials between Chinese and some other language (e.g. Tibetan, Sanskrit, Khotan, Persian, Korean, and Russian). The data show the existence of stressed and unstressed versions of pronunciation of the same word. The stress-conditioned phonetic change in many cases led to a phonological change, where doublet readings were codified and eventually written by two different characters. Introducing various transcription sources, Endo shows that stress-related phenomena were not only conveyed in the transcriptions, but were also actively recognized and described by the authors (the earliest description dating back to the Ming dynasty). Other interesting sources are dictionaries and language teaching materials. Endo compares transcription systems as used in several textbooks and other materials (Seidel 1901, Arendt 1918, and Chinese Linguaphone 1928). He points out that the modern dialects also provide promising source for the reconstruction of the history of stress in Chinese.

Last but not least: the fact that many of the prosody-related issues do not have a satisfactory solution in linguistic research reflects upon the state of the art of dictionaries, textbooks and methodology of teaching Chinese as a second language. For instance, one of the issues of Mandarin prosody frequently glossed over by lexicographic works is the variation of stress in compound words. Number of exceptions which attempt to reflect the stress features of compounds can be found, though. Perhaps the earliest example of such dictionary is the

"Russko-kitajskij slovar" (Russian-Chinese dictionary, Isaia 1867) quoted by Endo. One of the more recent works is "Kitajsko-russkij slovar" edited by I. M. Oshanin (Chinese-Russian Dictionary, Moscow 1955), or "Chugoku jiten" by Kuraishi Takeshiro (Chinese-Japanese dictionary, Tokyo 1966). Švarný's dictionary mentioned above is the most recent case coping with the problem.

If we take a look at the language teaching materials, we note that even the phenomena which were already successfully described by the linguists often do not find proper treatment in these practical areas. E.g. a third tone is traditionally brought out in the textbooks in its canonical form as high-low-high, instead of being primarily described as a low tone. Insufficient rendering of changes of citation forms of tones in connected speech regularly causes puzzlement to the elementary students of Chinese. I recall a *liuxuesheng* complaining that she had to spend arduous time at school to learn the lexical tones, yet as soon as she walked out of the classroom, she got impression the Chinese did not actually speak in tones at all! This little story indicates that there must be something wrong with our teaching methods. Modern methodology of teaching Mandarin phonetics requires more frequent contact between those working in theoretical research and the language teaching community.

The advantage of workshops and seminars on a small scale is undoubtedly a chance to become very intense and focused. The organizers trust that the Prague event, hosted by the ancient walls of the Charles University, was such an example. It undoubtedly helped to establish the contacts among the foremost researchers engaged in the discipline and provided a distinct perspective on the field. The participants came up with a broad variety of views and new linguistic data. The future task is to integrate them in a systematic framework. The following pages offer an insight into the field from different angles - be it experimental phonetics, studies on grammar, language teaching or historical development. At the same time, hitherto unresolved problems are pointed out. We hope this volume can serve as a stimulation for future research.

NOTES

1. Švarný's thesis on prosodic qualities of the Chinese syllable and its modification in connected speech dates back to 1952.

2. Švarný 1998-2000. *Učební slovník jazyka čínského* (Learning Dictionary of Modern Chinese). Olomouc: Publishing House of the Palacký University.
3. The supplementary fifth volume facilitating the usage of the dictionary for English speakers is being compiled.
4. Švarný outlined the concept of the monosyllabic semantic unit without functional specifications in his PhD dissertation in 1963, that is earlier than the Chinese linguists came up with their concept of *yusu*.
5. Exemplificative sentences were recorded by a native speaker from Peking - Mrs. Tang Yunling Rusková in considerably natural, spontaneous manner. The recordings are available on 39 tapes at present, on four CD-ROMs in the future.
6. Švarný 1991. The functioning of the prosodic features in Chinese (Pekinese). *Archív Orientální* 59, 2: 208-216; Švarný, and Tang Yunling Rusková. 1991. Prosodic features in Chinese (Pekinese) - prosodic transcription and statistical tables. *Archív Orientální* 59, 3: 234-254.

ACKNOWLEDGEMENTS

We would like to acknowledge the numerous inputs which made this volume possible. Special thanks should be offered to Prof. William S. Y. Wang from the City University of Hong Kong, who supported both the idea of organizing such a workshop in Prague and the idea of publishing collected papers in the JCL monograph series. Prof. Mathew Chen from the City University of Hong Kong helped with his valuable suggestions in the early stages of the preparatory process of the monograph. Of course, all our efforts would be in vain, if not for the generous financial support for both the workshop and the monograph by the Chiang Ching-kuo International Sinological Center in Prague represented by Prof. Oldřich Král of the Charles University. Thanks are also due to the head of the Institute of East Asian Studies of the Charles University, Prof. Olga Lomová, who came with the original idea of organizing this workshop and did a large part of the preparatory work.

Lastly, we express our gratitude to Prof. Oldřich Švarný, whose lifelong research on Chinese prosody was the initial motivation for holding such an event in Prague. We believe this undertaking will not remain the last one of its kind.