

Rhythmical Features of Spontaneous Spoken Narrative in Mandarin Chinese

Tereza SLAMĚNÍKOVÁ

Palacký University Olomouc, Czech Republic
tereza.slamenikova@upol.cz

David UHER

Palacký University Olomouc, Czech Republic
david.uher@upol.cz

Abstract

Every oral communication is shaped by a series of closely linked factors and rhythm is one of them. The rhythmical flow in colloquial Mandarin Chinese was thoroughly described by the Czech phonetician Oldřich Švarný, who invented a unique notation system dedicated to facilitating the teaching of Chinese as a foreign language. His prosodic transcription parses speech into a two-level arrangement of prosodic units. It displays the speech rhythm as an alternation of rhythmically significant and insignificant syllables of unequal tone prominence. The paper contributes to Švarný's findings by shifting the attention from prefabricated teaching materials to an example of unrehearsed dynamic speech. Specifically, it attempts to outline the basic rhythmical features of personal narrative and confront them with the current understanding of speech rhythm in colloquial Chinese. It also demonstrates how Švarný's methodological framework can be exploited to investigate different types of spoken discourse.

Keywords: Mandarin Chinese, speech rhythm, Oldřich Švarný, spontaneous speech, personal narrative

Povzetek

Vsako ustno komunikacijo oblikuje serija tesno povezanih faktorjev in ritem je eden izmed njih. Ritmični tok pogovorne kitajščine je temeljito opisal že češki fonetik Oldřich Švarný, ki je izumil edinstven sistem notacije z namenom poučevanja kitajščine kot tujega jezika. Njegova prozodična transkripcija razčlenjuje govor v dvonivojsko urejenost prozodičnih enot. Prikazuje govorni ritem kot izmenjavo ritmično pomembnih in nepomembnih zlogov neenake tonske pomembnosti. Naš članek Švarnýjeve ugotovitve, zasnovane na vnaprej pripravljenih učnih materialih, preverja na primeru spontanega dinamičnega govora. V njem popišemo osnovne ritmične značilnosti osebne pripovedi in jih soočimo z obstoječim razumevanjem govornega ritma v pogovorni kitajščini. V članku tudi prikažemo, kako je moč Švarnýjev metodološki okvir uporabiti pri raziskovanju različnih vrst govornega diskurza.

Ključne besede: kitajščina, ritem govora, Oldřich Švarný, spontani govor, osebna pripoved



1 Introduction

Prosodic research indicates that colloquial¹ Mandarin Chinese is characterized by tonal assimilation and neutralization, a phenomenon already pointed out by the Chinese-American linguist Chao Yuen Ren (1892–1982) in his widely respected study *A Grammar of Spoken Chinese* (1968). Although other researchers have further developed his observations, one cannot fail to notice that their considerations are either limited to a word or, eventually, a phrasal level (e.g., Duanmu, 2007; Lin, 2007) or are primarily concerned with the issue of intonation in the tonal language (e.g., Shen, 1989; Tao, 1996). As such, they do not pay adequate attention to how tonal weakening affects the overall rhythmical segmentation of spoken texts. The findings of the Czech sinologist and phonetician Oldřich Švarný (Švarný, 1991a, 1991b, 1998, 2000; Švarný & Rusková Tang, 1991;² Švarný & Uher, 2014) and scholars adopting his line of research (e.g., Uher & Slaměňíková, 2015; Uher, 2016; Pospěchová, 2016a, 2016b, 2022; Slaměňíková, 2016, 2018; Třísková, 2017, 2019; Slaměňíková & Uher, 2021), however, clearly indicate the importance of this phenomenon. To provide a basis for teaching Chinese as a foreign language,³ Švarný developed a unique prosodic labeling system. It is accompanied by a complex theory on speech rhythm in Mandarin Chinese grounded on the empirical and instrumental examination of acoustic cues.

Simply speaking, the philosophy underlying Švarný's system is built on an observation that rhythmical features are organized in terms of categorically distinct entities (e.g., stressed and unstressed syllables) and relations (e.g., a relatively stronger or weaker tone prominence). Syllables with different properties are grouped and create different rhythmical patterns. Švarný et al. (1998, p. 8) identify two basic rhythmical sections of the so-called linear segmentation: cola (sg. colon), i.e., basic units characterized as compact

¹ There is a great discrepancy between the spoken (kǒuyǔ 口语) and written languages (shūmiànyǔ 书面语) in Mandarin Chinese. The term 'colloquial' refers to the first mentioned variant. Gajdoš (2011) outlines the differences between these two language variants considering choosing language data for synchronous linguistic research.

² To increase their accessibility, Švarný's articles written in foreign languages (including Švarný, 1991a, 1991b, 2000; Švarný & Rusková Tang, 1991) were gathered together and published as a collected reprint under the title *Oldřich Švarný: Prosodia linguae Sinensis* (Uher & Slaměňíková, 2019). The edition also includes the English chapter Rhythmical Features and Prosodic Transcription of Chinese included in *Hovorová čínština u příkladech III* [Colloquial Chinese in Sentence Examples] (Švarný et al., 1998, p. 7ff).

³ Although the situation is changing, prosodic aspects of speech are still largely ignored in the practice of L2 teaching. This seems to be a mistake, especially in cases when L1 and L2 are typologically different, which is often the case with native Slavic speakers learning East Asian languages. One researcher who pays attention to this phenomenon is Golob (2011, 2022), targeting the (dis)similarities between Slovene and Japanese.

semantic sections with intonation features separated by a pause, and segments, i.e., minimal units with a certain rhythmical flow. Their research on speech rhythm has shown that the rhythmical features are “variable according to the speed of speech, its emotional coloring, the overall length of the predicative and nominal complexes and according to the possible cases of logical and contrastive stresses” (Švarný et al., 1998, p. 7).

For understandable reasons, Švarný focused on the practical ends of the developed theory. He analyzed the prosodic interpretation of an extensive set of example sentences in learning materials for Chinese language students. The most influential is *Učební slovník jazyka čínského* [Learning Dictionary of Modern Chinese] (1998–2002, abbrev. LDMC). In light of this, his works systematically describe rhythmical features of read speech. The developed theory and methodology are universal, however, and can therefore be applied to any spoken Chinese discourse. Taking a personal narrative as an example, this paper shifts the attention to a speech performance that represents a continuous flow of clauses created spontaneously at the moment of recording.

2 Theoretical background and research approach

Spontaneous spoken language is often defined in opposition to written language. Using such a comparison, Miller and Weinert (1998, p. 22) provide a list of key properties describing this type of discourse as a language that is produced in real-time by people talking face-to-face, is accompanied by suprasegmental and paralinguistic information, and is constrained by the capacity of both the speaker’s and hearer’s short-term memory. Halliday (1989, p. 81) focuses on the way reality is represented and argues that spoken language shows phenomena as dynamic processes, whereas written language reproduces them as ready-made products. Similarly, Crystal (2000, p. 181) assigns qualities of time-boundedness, dynamism, and transience to the speech and sets them in contrast with the space-bounded, static, and permanent nature of writing. A different point of view distinguishes (written or spoken) spontaneous forms of communication that are “created on the fly” from the manufactured varieties that are “prepared, reworked, and turned into an artifact or performance” (Clark, 2014, p. 292). This article analyzes a sample of verbal communication that meets the criteria of ‘spontaneity’ in both of these approaches: it is a story delivered ad hoc by a speaker.

The various general properties of spontaneous spoken communication are reflected in certain linguistic features on both grammatical and lexical levels. The overall fragmented and unintegrated syntax is characterized by an

occurrence of relatively simple phrases and clausal constructions and a preference for paratactic over hypotactic patterns (Miller & Weinert, 1998, p. 22ff). The structure of clause complexes is intricate in a way that it is “able to represent not one process after another in isolation but whole configurations of processes related to each other in a number of different ways” (Halliday, 1989, p. 86). As such, the grammatical intricacy can compensate for the otherwise lower lexical density per clause (Halliday, 1989, p. 87). To help ensure messages are understandable, speakers “carefully stage” the sections with a higher amount of lexical information provided in the long strings of words (Miller & Weinert, 1998, p. 22ff). The prosodic effects are one of the main tools used for this purpose. They also have the ability to reflect elements of language that are not provided through the choice of vocabulary or encoded in the grammar (cf. e.g. Crystal, 2000, p. 179).

Despite the dominant role of conversational talk, spontaneous speech may also include numerous monologues. One of its forms is a spoken narrative, e.g., speeches in the shape of recounting an event, telling a joke, or narrating a story in dialogues (Clark, 2014, p. 303). Brazil (1993, p. 155) mentions two fundamental aspects that distinguish narratives from conversation. First, storytellers usually engage the other participants in active listenership through clearly marked openings and endings. Second, their linguistic choices are goal-oriented in that they carefully set the story’s endpoint. In contrast, Brazil also emphasizes that narrative is, just like a conversation, an interactive and contextualized phenomenon. Speakers adjust their choices to the moment-to-moment needs emerging during the course of communication.

From a broader perspective, the spoken narrative is an interdisciplinary subject of study that is in the interest of several major groups of scholars, including literary theorists, psychologists, folklorists, and anthropologists, as well as sociolinguists (cf. Juzwik, 2012, p. 331ff). According to Labov (1972, p. 359), the narrative is a “method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which (it is inferred) actually occurred.” The definition implies a certain interplay between several key elements. The framework developed by Bauman (1986) identifies the event (event recounted in the narrative, i.e., narrated event, be it an actually occurred or narratively constructed event), story (i.e., narrated text), and performance (the act of telling the story including the situated context, in which the narrative is told, i.e., narrative event). A similar threefold distinction was drawn by Blum-Kulka (1993), however, she leaves the narrated event from the model and separates the performer from the act of storytelling, who may or may not be the person accountable for the story. All in all, the so-called “dimensions of narrativity” distinguished by Blum-Kulka include the ‘tale,’ the

'teller,' and the 'telling.' This short introduction to several basic concepts briefly indicates the complexity of spoken narrative discourse and demonstrates why it can offer researchers from different disciplines a great deal of space to target different segments of its manifold organism.

The sociolinguistic approach has been influenced by the structural approach of Labov (1972, p. 354ff; Labov & Waletzky, 1967), who, among others, defined the formal units of personal narratives and established a six-category typology of clauses constituting the structure of the narrated text, i.e., a) abstract (opening clauses capturing the point of the narrative); b) orientation (descriptive clauses setting the scene); c) complicating action (sequentially ordered narrative clauses forming the plot); d) evaluation (any markers revealing the attitude of the narrator toward the narrative); e) result or resolution (the termination of the complicating action); f) coda (clauses signaling the end of the narrative). The conversation analytic scholars shifted the attention, from elicited, interview-style narratives collected from informants, to the stories occurring in everyday communication (Thornborrow, 2012). Their interest centers around the nature of narrative events, addressing issues such as the typology of social actions in which they occur, the role that the situated context plays, the functions that the narratives fulfill, the complexity of the participant roles structure, the conditions enabling participants to gain the storytelling rights or the mechanisms underpinning the transition from non-narrative talk to storytelling (cf. Norrick, 2007; De Fina, 2008; Thornborrow, 2012). Elements observed in the process of narrative organization analyses include formulaic sequences characteristic for specific parts of the narrative (i.e., openings, closings, transition or turning points), disfluencies (e.g., false starts, restarts, self-corrections, hesitation markers or pause-fillers), discourse markers, or attention signals. In contrast to the Labovian approach, the approach is also not limited to personal narratives but examines how these aspects play out in a wide range of storytelling types (cf. Norrick, 2000, p. 135ff).

The sociologically motivated interest in storytelling significantly contributed to analyzing their language structure from which the other branches of linguistics can benefit. This paper addresses the issue of speech rhythm in colloquial Mandarin. The narrative was chosen as one of the forms that the spontaneous spoken discourse can take. As mentioned earlier, Švarný's findings are grounded on an abundant but strikingly homogeneous corpus mainly composed of prefabricated example sentences read by a Pekinese native speaker. In contrast, the theoretical works on spontaneous spoken narrative mention two opposing tendencies relating to the chunking of messages: a) creation of long strings of words (with a lower amount of information); and b) sequencing of the speech into shorter units as a result of

disfluencies. The previous research also indicates that narratives are usually composed of several types of clauses differing in the function they fulfill. It is therefore reasonable to assume that the heterogeneous composition of the narrative impacts the overall rhythmical segmentation of speech. In this paper, we attempt to depict its rhythmical characteristics emerging from a comparison with findings based on an analysis of read sentences. Apart from Švarný's (1991a, 1998, 2000) theoretical works, we base our description on analyses of the sentences listed in the LDMC (i.e., Švarný, 1991b; Švarný & Rusková Tang, 1991; Slaměňíková, 2016, 2018). The term "LDMC sample" used in the analysis refers to the investigation of sentences included in entries 420–454 and 1926–1969 (Slaměňíková, 2018).

3 Research corpus and methodology

The analysis was undertaken with a speech by a Pekinese woman uploaded on the *Phonemica* website, an online platform for stories told in local vernaculars (van Dam et al., 2011–2021). The storyteller (Mrs. Qin, born in 1947) is a retired doctor (Zheng, 2013) sharing her emotionally intense experience during medical practice. The story is entitled *Bǎ tǔdòu gěi mǎile cái yǒu qián shūyè* 把土豆儿给卖了才有钱输液 *Selling Potatoes to Get Money for an Infusion* (Mrs. Qin, 2012). Considering the platform's objective, it is evident that the recipient elicited the narrative event in a way similar to how a host asks a participant in a TV talk show (cf. Thornborrow, 2012, p. 58ff). The storyteller takes the floor for an extended time without needing to set the stage for the story. While speaking, she is not interrupted by the hearer. His reactions are limited to a few minimal displays of reciprocity delivered in the background and one question produced partly in the overlap. What is different, however, from the talk show practices is that the narrative event does not show any signs suggesting that the hearer (unlike the talk show moderator) would be familiar with the tale before the recording. It is unclear, however, whether he elicited a talk on a specific topic from the speaker. The general conception of the *Phonemica* website, covering, e.g., nursery rhymes, humorous anecdotes, or local history (van Dam et al., 2011–2021), suggests that it might not be the case and that the speaker most likely decided to tell a story of personal experience and accomplishment. All in all, the delivery style is natural, without offering any evidence of being prepared in advance.

The story's plot comprises three events (E1, E2, E3) that are told chronologically. Each of them is introduced by a formulaic opener (E1 and E2 *yǒu yí cì* 'once'; E3 *zài yǒu* 'furthermore'). The E2 ended with a closing cliché *zhèi gùshi wǒ yí bèiz yě wàngbuliǎo* "I will not forget this story for the rest of

my life” (cf. [Norrick, 2000](#), p. 49ff). It is also one of the evaluative comments through which the storyteller enhances the tellability of her story. Another one is located at the end of the E3, i.e., *wǒ jiù tèbié gǎndòng* “I was very moved.” Apart from this, the analyzed narrative displays a set of other discourse-specific features, including the following formal verbal means:

- 1) In accordance with the observations made by [Li \(2019\)](#), the advancement of the story is highlighted by different perfective markers. One of them is the optional grammatical marker *-le* following a dynamic verb⁴ (e.g., E1.02 *lái* ‘come’ + *le*; E1.23 *bāo* ‘wrap’ + *le*; E1.26 *zǒu* ‘go’ + *le*). Apart from this, the devices of the so-called situation aspect also prove themselves productive. The two most frequent include: a) resultative and directional verb complements (e.g., E1.13 *pèishang* ‘join with’; *wèiwan* E1.15 ‘finish to feed’); and b) constructions with coverbs occurring after the main verb (e.g., E2.04 *zǒudao* ‘walk to’).
- 2) The discourse markers manage the flow of the narrative. Most of them may perform different functions that are not mutually exclusive and operate on several linguistic levels simultaneously ([Jucker & Ziv, 1988](#), p. 3). In this respect, Mandarin is no exception (e.g., [Liu, 2009](#)). Using the classification established by Y. Maschler (2009, p. 22ff), they can be divided into four prototypical groups. The textual structure of the analyzed narrative is dominated by two of them, i.e., the referential and cognitive discourse markers. The sequence markers *ránhòu* and *hòulái* ‘then,’ result marker *suǒyǐ* ‘therefore,’ causal marker *yīnwéi* ‘because’ and non-coordination marker *kěshì* ‘but’ are the most frequent examples of the first of them. As for the second mentioned, the speaker often uses the marker *jiùshì* ‘precisely be’ to rephrase her previous words. The expressions *nèi ge/ zhèi ge* ‘that is’ also signal her thought process. Another one is the interjection *wa*, indicating the processing of the information as a surprising fact.
- 3) As expected, the discourse markers belong to the most productive lexical items of the analyzed narrative (cf. [Jucker & Ziv, 1998](#), p. 3). As a natural consequence of telling a personal experience, the pronoun *wǒ* ‘I’ is the most frequent word in the analyzed narrative. The story’s plot involves more people whose appearance is associated with high use of the third-person deixis. The speaker also often utilizes dialogue as a dynamic device to advance the plot. This, in other words, means that the narrative contains

⁴ Similarly, M. Petrovčič (2009) argues that the so-called operator *le* “contributes to change in dynamics” (p. 4) and “tends to appear the most frequently in narrative sentences and further narrative passages” (p. 192). It should be noted that Petrovčič adopts the ‘one-*le* approach’ and uses the term operator *le* as a hypernym for both of the syntactic positions of *le*, i.e., a post-verbal and a sentence-final particle.

a relatively high proportion of clauses that indicate the upcoming act of speaking. Specifically, the speaker repeatedly uses the dialogue tag *shuō* ‘say,’ which represents the most frequent verb of the narrative.

The processing of the audio recording followed the procedure developed by Švarný. First, it was transcribed into Švarný’s prosodic labeling system. After parsing the clauses into rhythmically relevant units, we mark each syllable’s relative degree of tone prominence with respect to their significance in the rhythmical flow. We adopt the graphic marking introduced in *Učebnice čínské konverzace 2* [Textbook of Chinese Conversation 2] (Uher et al., 2016), which provides a more user-friendly version of the graphic labeling used in *LDMC*. Second, the transcript was analyzed according to the method established by Švarný. The following section provides basic guidelines for the understanding of both the prosodic transcription and the evaluation of rhythmical patterns. For more details, we advise the reader to consult the original texts written by Švarný listed in the previous section.

The prosodic transcript enclosed in Appendix 1 is, for clarity, divided into three parts, each represented by one event and supplemented by a brief plot summary.⁵ Apart from this, we used the segmentation into sections as published on the *Phonemica* website. This is because the timing of the *Phonemica* made it possible to provide each section with its speaking ratio, one of the variables monitored by O. Švarný. The sections of each event are numbered chronologically (e.g., E1.01, E1.02, etc.). Speech errors resulting in meaningless strings of phonemes or rhythmically hardly identifiable units were excluded from the analysis. These expressions are struck through in the transcript. Moreover, the transcript does not record the reactions signaling the hearer’s attention. Apart from the transcript, Appendix 1 also documents the rhythmical patterns of cola.

4 Analytical framework of speech rhythm in Chinese by Švarný

Švarný’s⁶ prosodic transcription depicts a clause complex⁷ as a combination of stressed and unstressed syllables with a certain tonal prominence. The first

⁵ This procedure was inspired by V. Štefek (2021).

⁶ This chapter is based on Švarný’s description of rhythmical features and prosodic transcription as provided in Švarný et al. (1998, p. 7ff); Švarný & Uher (2014, p. 9ff, 147ff).

⁷ Considering the nature of the analyzed corpus, Švarný uses the term ‘sentence.’ In our paper, we prefer Halliday’s (1989, p. 66) term ‘clause complex’ as an alternative to identify a sentence in the spoken language. Its component part is a clause, i.e., “the grammatical unit in which semantic constructs of different kinds are brought together and integrated into a whole.”

mentioned, called ictus-bearing syllables or arses, are more prominent (in terms of the expansion of pitch range and time duration) than the immediately adjacent syllables. The other syllables are called non-ictus-bearing syllables or theses. The three types of ictus-bearing and four types of non-ictus-bearing syllables establish in all a seven-degree scale of syllable prominence hierarchy. As mentioned above, the so-called segments are the minimal units carrying a rhythmical flow. Each segment contains at least one arsis of any type, and the syllable immediately preceding or following an arsis is always the thesis.⁸ The following table summarizes the types of syllables, including their graphic marking in the transcript. Appendix 2, listed at the end of the paper, provides example descriptions of syllable prominence and linear segmentation of clauses occurring in three selected sections.

Table 1: Types of syllables

Rhythmical significance	Syllable prominence	Marking
Ictus-bearing syllable, i.e., arsis (abbrev. A)	Stressed-prominence syllable*	mǎ
	Full-tone syllable**	mǎ
	Weakened-tone syllable	ma ³
Non-ictus-bearing syllable, i.e., thesis (abbrev. T)	Full-tone syllable	mǎ
	Weakened-tone syllable	ma ₃ / ma ^{***}
	Void-tone syllable****	Neutralized Toneless
		ma

Notes to the table:

* It either represents the most prominent syllable of a colon or indicates a logical or contrastive stress, putting certain aspects of the clause in the foreground. The last ictus-bearing full-tone syllable is the most prominent in cola without any stressed-prominence syllable.

** The traditional tone marks label both ictus-bearing and non-ictus-bearing full-tone syllables. Their distinction is based on their position within a frame of one segment. The full-tone syllables, surrounded by weakened-tone syllables or void-tone syllables, are arses. The so-called acronymic rule determines the rhythmical significance in an immediate sequence of full-tone syllables: a) in a two-syllable structure, the first of them is a thesis and the second arsis; b) in a three- or four-syllable structure, arses are located at the edges, i.e., at the last and the first syllable.

⁸This means, in other words, that two arses can never occur in immediate proximity to one other in the same segment.

*** Non-ictus-bearing weakened-tone syllables, located in the middle or at the end of a segment, are marked by a number in a lower index. Those at the beginning of a segment are transcribed without marking, e.g., ma. The fact that the first syllable of a segment is always a bearer of at least a weakened prominence eliminates the confusion with a void-tone syllable.

A rhythmical flow arises by joining syllables together. There are two basic types of connections: a) a descending sequence of two or more syllables with the arsis resting on the first of them; and b) an ascending sequence of two or more syllables with the arsis at the end. Some syllables remain single and join as proclitics or enclitics of the already existing descending or ascending sequences. Alternatively, they behave as the so-called odd ictus-bearing syllables. The two types of sequences and the odd ictus-bearing syllable represent the basic structural components of the rhythmical flow. Segments with a simple structure contain one, whereas segments with a more complex structure combine two or more of these items. In the transcript, hyphens connect syllables within a segment in cases where the syllables are not part of one word.

Cola, the highest units of the rhythmical flow, can be formed by one or more segments. The colon boundaries are labeled with punctuation marks. In the transcript provided in Appendix 1, we add a comma in cases when the original transcription does not depict the boundaries of cola. In other cases, we keep the original marks provided on the *Phonemica* website when they do not interfere with the prosodic segmentation of the audio recording.

Apart from the transcription, Švarný established the metalinguistic apparatus used for the quantitative analysis of rhythmical patterns of cola. These are as follows: / an ascending sequence; /° an ascending sequence followed by an enclitic syllable(s); \ a descending sequence; °\ a descending sequence preceded by a proclitic syllable; | an odd ictus-bearing syllable; V a trisyllabic two-arsis acronymic sequence of the so-called 'cretic' type (i.e., a combination of a descending sequence and an odd syllable; or less often, a combination of an odd syllable and an ascending sequence). In Appendix 1, we add these marks in a separate column on the right side of the transcribed text. The cola of each section are numbered, e.g., c1, c2, etc. The letters A–D preceding the description of rhythm indicate the type of colon as described below.

Švarný's classification of cola takes into account two aspects. The first of them is the number of rhythmical sequences. Considering this, the category of one-sequence cola is separated (hereinafter referred to as type A). More

than one sequence cola are then classified based on the position of arses into:

- cola without an immediate sequence of two arses (type B)
- cola with an immediate sequence of two arses (type C)⁹
- cola with an enclitic or proclitic syllable preventing the immediate sequence of two arses (type D)

As can be seen, the model accentuates whether two ictuses can be found in immediate contact next to each other. By its nature, this can happen strictly at the edges of two neighboring segments, represented by the type C cola in the model. The contact of two arses is indicated by the symbol +.

All types include a spectrum of rhythmical patterns, as will be shown in the analysis. At this point, the so-called acronymic rhythm principle must be explained. The simple acronymic structure of the so-called 'cretic' type comprises three syllables. Its arrangement has been described above. A slightly more complex structure combines a descending and an ascending sequence \ /, i.e., the so-called 'choriamb' type. Both of these patterns have two arses, the first located on the initial syllable and the second on the last syllable. Apart from these, three- and more arses acronymic structures are also quite common. They are created through the repetition of the descending sequences or ascending sequences. The core of these structures can be of both a cretic or a choriamb type (i.e., three-arses acronymic structures \V, \ \ /, V/, \ //; four-arses acronymic structures \ \ V, \ \ \ /, V //, \ \ \ /, \ V/, \ \ \ //, etc.).

5 Analysis and discussion of the findings

5.1 Speaking rate

The duration of the analyzed narrative is 336 sec. Compared to previous research samples, the speech of the storyteller differs in one crucial respect: the speaking rate varies widely during her speech. The up to now most extensive corpus of approximately 16,000 transcribed sentences collected in *LDMC* shows a certain acceleration of the speech tempo between the initial and the final part of the large corpus (Švarný, 1991b, p. 132). The complexes of sentences located in particular parts manifest, however, a high degree of homogeneity. In contrast, the speech analyzed in this paper is characterized

⁹ An immediate sequence of arses occurs when an ascending sequence of the first segment is immediately followed by a descending sequence of the second segment, including the ascending and descending sequences in any acronymic structure.

by significant tempo differences within a short amount of time that are associated with the speaker's emotional involvement as the story evolves. She manipulates the speech tempo to contextualize the clauses in a certain way. Specifically, her speaking rate varies between 1.9 to 5.7 syl/sec through the different sections.

Since Švarný et al. (1998, p. 7) identify the speech tempo as the variable influencing the rhythmical features, we pay attention to it while analyzing individual aspects and, once it is relevant, draw a comparison between slower and faster sections. In order to be able to determine this, we set the average speaking rate of 3.5 syl/sec as the central point and focus on the linear segmentation of the sections with a rate slower than 3.0 syl/sec and faster than 4.0 syl/sec. The first mentioned includes 53 cola composed of 238 syllables, and the second mentioned 58 cola composed of 436 syllables. In the transcript, the slower sections are presented with the background color blue and the faster with orange.

At this point, it is important to comment on the logic behind the speech segmentation into sections provided on the *Phonemica* website. They basically divide the speech into small information units. Their length ranges from one to six clauses. Having in mind the Labovians classification, multi-clauses sections are composed of both a series of one-type clauses or a combination of different types of clauses. A previous study on the relationship between speech rate and typology of clauses indicates that there is “no statistically significant effect” between these two variables on the global level (Oliviera, 2012, p. 640). The analyzed narrative seems to support this finding, as it can be determined from the inconsistent speech tempo of the sections containing one type clause. In light of this, the fact that a section might be composed of different types of clauses is not considered while observing the fluctuation in speaking rate.

5.2 Linear segmentation

As mentioned above, Švarný defines two rhythmical units of linear segmentation. Compact semantic sections with intonation features, i.e., cola, are composed of minimal rhythmical units, i.e., segments. When observing the different parts of the *LDMC* corpus, Švarný et al. (1988, p. 24) found out that while the length of segments is considerably dependent on the speech tempo, the length of cola does not seem to be influenced in this way. The syntactic complexity of the sentences can be, to a certain measure, significant in this way: the length of cola in longer sentences tends to be slightly higher.

Švarný also calculated the average length of both linear units: 2.5 to 4.5 syllables per segment and 6 to 7 syllables per colon.

The analyzed sample contains 197 cola, 367 segments, and 1,168 syllables. Considering the average length of segments, the number of 3.2 syllables places the analyzed narrative in the middle part of the spectrum determined by Švarný et al. (1998, p. 8). Just like it was reported in previous research, most segments contain two to four syllables (cf. Švarný et al., 1998). Specifically, two- and three-syllable segments account for almost 30% each, and four-syllable segments for almost 24% in the narrative. In addition, the average length of cola, i.e., 1.9 segments and 5.9 syllables, is essentially consistent with previous reports. What is different is the distribution of cola with a certain number of segments. About one-half of the *LDMC* sample (Slaměňíková, 2018, p. 157) comprises two-segment cola, whereas a similar proportion belongs to one-segment cola in the narrative. In addition, the lower occurrence of two-segment cola is compensated for by the higher occurrence of three- and more-segment cola.

This change in proportion arises from the characteristic nature of the spontaneous spoken language. Disfluencies such as false starts with abrupt cut-offs, pause-fillers, or discourse markers break the speech flow into shorter cola. Phrases and clauses advancing the plot of the story tend to, in contrast, rhythmically speaking, connect together and form longer cola. The impact of spontaneity is also reflected in the relation between the average length of cola and the speaking rate (see Table 2). The average length of 4.5 syllables per colon is significantly lower than the value of 6 to 7 syllables determined by Švarný et al. (1998, p. 24), and, more importantly, it is inconsistent with the assumption that the colon is a linear unit whose average length is not to be influenced by speech tempo. Against expectation, the average length of segments that was reported to be directly proportional to the speaking rate shows insignificant variations in average numbers. A certain difference, however, can be observed: the occurrence of two- to four-syllable segments in faster sections is reduced in favor of five-syllable segments.

Table 2: Linear segmentation

Sample	The average length of segments	The average length of cola
Whole narrative	3.2 syllables	5.9 syllables / 1.9 segments
Slower sections (< 3,0 syl/sec)	3.1 syllables	4.5 syllables / 1.5 segments
Faster sections (> 4,0 syl/sec)	3.3 syllables	7.5 syllables / 2.3 segments

5.3 Syllable prominence

Table 3 summarizes the occurrence of different types of syllables in the analyzed narrative. Compared to the *LDMC* sample (Slaměníková, 2018, p. 160), the narrative contains more weakened-tone and void-tone syllables. The most noteworthy difference is the amount of stressed-prominence syllables. As far as we can say, their occurrence in the different parts of the *LDMC* sample varies between 12.1 to 19.1% (cf. Slaměníková, 2016, p. 80; Slaměníková, 2018, p. 160). The reason for their lower occurrence in the narrative can be explained as follows. While recording the isolated sentences of *LDMC*, the speaker needed to emphasize a piece of certain information in each sentence or even a colon. In contrast, the speaker telling the continuing story places the stress with the aim to contrast entities or focus attention on entities within more complex textual units.

Table 3: Title of this table

Syllable prominence	Occurrence
Stressed-prominence syllable	4.3%
Full-tone syllable – arsis	28.9%
Full-tone syllable – thesis	5.3%
Weakened-tone syllable – arsis	12.8%
Weakened-tone syllable – thesis	8.8%
Void-tone syllable	39.9%

In agreement with Švarný's observations (Švarný & Rusková Tang, 1991, p. 241ff), the changing speaking rate of the narrative affects the proportion of syllables with different rhythmical significance. Table 4 demonstrates that the

occurrence of arses in the faster sections is about 4% lower than in the slower sections. As Švarný explains, clauses with faster speech tempo are rhythmically organized into larger groups of syllables within which some of the ictus-bearing units tend to be weakened or even neutralized and, thus, become theses.

Table 4: Occurrence of arses

Syllable prominence	Occurrence
Whole narrative	46.0%
Slower sections (< 3,0 syl/sec)	47.9%
Faster sections (> 4,0 syl/sec)	43.8%

5.4 Rhythmus of segments

The rhythmus of the segments relies on the number and position of the ictus-bearing syllables. In the analyzed speech, more than one-half of the segments contain one arsis. As can be seen in Table 5, the occurrence of the ascending sequence is almost twice as low as the occurrence of the descending sequence, which represents the most frequent rhythmical pattern overall. Descending sequences, or, to be specific, the combinations of the two of them, also play an essential role in the two-ictus-bearing segments. The acronymic structure prevails, however, in this type of segment and embodies the second most frequent pattern. Almost all tree-ictus segments are acronymic; only one segment comprises three descending sequences.

Table 5: Rhythmus of segments

Arses	Occurrence	Rhythmical pattern	Occurrence
One	57.6%	Odd syllable	3.8%
		Descending sequence	35.3%
		Ascending sequence	17.7%
		Others	0.8 %
Two	38.9%	Acronymic structure	24.5%
		Two descending sequences	13.0%
		Others	1.4%
Three	3.5%	Acronymic structure	3.3%
		Three descending sequences	0.2%

Generally speaking, the occurrence of the rhythmical patterns does not deviate from our current understanding of the matter. The numbers of patterns range more or less in the middle between their occurrence in the slower and faster speech in the *LDMC* sample (cf. [Slaměnková, 2018](#), p. 166). This does not, however, fully apply to one-arsis segments, in which case the occurrence of odd syllables and ascending sequences is slightly lower.

It should be further pointed out that both descending and ascending sequences are far less often composed of lexical compounds. The relatively lower lexical density of spontaneous speaking style is the reason for a higher occurrence of rhythmically joined groups of monosyllabic words. As was described by Švarný, a significant number of these monosyllabic words tie in with the rhythmical sequences as clitics. The most common proclitics, preceding a descending sequence, include singular personal pronouns and multifunctional adverbs *dōu* and *jiù*. The attributive marker *d* and sentence-final particles *le* and *ne* are the most frequent enclitics following the ascending sequences (cf. [Švarný et al., 1998](#), p. 39).

Finally, a note has to be added concerning the lexical structure of the segments. The use of discourse-specific elements can be, of course, observed all over the narrative. It is primarily, however, the group of two-syllable segments, in which case they occur with a significant concentration. Four categories of these elements can be distinguished. The first of them is related to the first person point of view and consists of two combinations: a) pronoun *wǒ* + verb *shuō* 'I said' (e.g., E2.21); b) pronoun *wǒ* + adverb *jiù* 'so I' (e.g., E1.11). The second category includes temporal nouns and conjunctions that indicate the sequence and causality, e.g., *hòulai* (e.g., E2.22). The third category is the connection of the demonstrative *nèi* with the general measure word *ge* carrying the meaning 'that,' used by the speaker both as a reference device (e.g., E2.02) or as a pause-filler for lexical or content search (e.g., E1.03). The last one is the discourse marker composed of the adverb *jiù* and the verb *shì*, together carrying the meaning 'precisely be' and used by the speaker as a rephrasing tool (e.g., E2.27).

5.5 Rhythmus of cola

The analysis has shown that most of the one-sequence cola (type A) are characterized by a descending rhythm. Consistent with previous findings, cola without an immediate sequence of two arses (type B) show a preference for two patterns, i.e., the repetition of descending sequences and different forms of acronymic structures. Generally speaking, type C and D cola share the same patterns. The only difference lies in that the first mentioned are

characterized by an occurrence of an immediate sequence of two arses, whereas the second mentioned contain a clitic preventing such a contact. From the variety of patterns these two types can be organized into, the analyzed narrative uses only one more often, i.e., a combination of two acronymic structures. The following table provides the relative occurrence of each pattern and lists some examples occurring in the narrative in the case of the most frequent poly-sequence patterns.

Table 6: Rhythmus of cola

Type	Rhythmical pattern	Occurrence
A	An odd syllable	5.1%
	A descending sequence \	14.7%
	An ascending sequence /	7.1%
	A three-syllable acronymic structure V	5.1%
B	An acronymic structure e.g., \ /, \ \ /, \ V, \ \ /, V \ /, \ \ \ V ^o etc.	23.9%
	Two or more descending sequences e.g., \ \, ^o \ \, \ \ \, \ \ \ \ etc.	17.8%
	Others	2.5%
C	A combination of two acronymic structures e.g., \ /+ \ /, V+ \ /, V+ \ \ /, \ \ V+V etc.	5.1%
	Others	7.6%
D	A combination of two acronymic structures e.g., V ^o \ /, \ / ^o /, V ^o \ V, \ \ / ^o \ \ /	3.6%
	Others	7.6%

Compared to the *LDMC* sample (Slaměňíková, 2018, p. 169), the narrative is characterized by an almost twice as high number of the rhythmically simple type A cola. Considering the above-discussed nature of the dependency between the average length of cola and the speaking rate, it is logical that this type is especially productive in the slower sections. It also has to be mentioned that, just as with the case of the segments, the rhythmically most simple cola are often formed by the referential and cognitive discourse markers, occasionally joined with a personal pronoun or sentence-final particle, e.g., E2.27 hòulai, E3.10 suǒyi, E1.22 ránhòu-ne, E2.07 yīnwei-tā, E1.26 jièguǒ-ne, E1.05 jiū⁴shi. Similarly, the most frequent pause-filler also tends to remain rhythmically isolated, e.g., E2.10 nèi-ge.

Naturally, all the types of more-sequences cola can be found more often in the faster sections, as seen in the table. Just as in the case of the *LDMC* sample (Slaměnková, 2018, p. 168), they are dominated by type B cola without an immediate sequence of two ictuses. As for the rhythmical patterns, it is worth mentioning that acronymic structures, consisting of at least one descending and one ascending sequence, represent the most frequent pattern in all the samples. The second most common pattern is of a descending nature in both rate-related samples, however, its form reflects the different complexity of slower and faster cola. The first mentioned bear more often one ictus, i.e., it is composed of one descending sequence (type A), while the second mentioned combines two descending sequences, each with its ictus (type B).

Table 7: Types of cola

Syllable prominence	Type A cola	Type B cola	Type C cola	Type D cola
Whole narrative	32.0%	44.2%	12.7%	11.2%
Slower sections ($< 3,0$ syl/sec)	45.3%	39.6%	11.3%	3.8%
Faster sections ($> 4,0$ syl/sec)	17.2%	51.7%	17.2%	13.8%

6 Conclusion

In 2000, O. Švarný (p. 149) argued that the prosodic features of today's Mandarin Chinese are "more and more encroaching on the lexically distinctive function of 'tones.'" He supported this statement with an extensive analysis of transcribed sentences through which he demonstrated how unequal syllable prominence generates specific rhythmical patterns. This paper builds upon the methodological and theoretical framework established by Švarný and extends his description of prosodic features of modern colloquial Chinese in a way that targets a different type of spoken discourse.

The analyzed narrative differs from Švarný's corpus in several main aspects. First, it is determined by the different circumstances of their origin. Švarný's corpus collects oral manifestations of isolated sentences. The fact that the speaker was encouraged to perform them as naturally as possible does not change the essential characteristics of a text being read out loud. In contrast, the available information about the analyzed storyteller's speech suggests that she was invited to choose a topic of her preference and use her

own words to deliver a speech. It is evident that each of the corpus is composed of a different content. Švarný's corpus collects illustrative sentence examples, each of which is tied with a specific agenda, i.e., explaining the grammatical unit in question. The narrative is a variable flow of syntactically not necessarily correct clauses that, as a whole, aim at achieving a particular communicative goal.

This article is a pilot study that attempted to depict general tendencies of speech rhythm in a spontaneous spoken narrative in colloquial Mandarin without a claim to exhaust the topic. A large amount of data is necessary to yield more conclusive findings. In any case, some preliminary conclusions can be drawn based on the comparison with the read speech. As expected, the narrative is composed of parts that differ from each other. The storyteller chunks the messages in constituents of variable length. We can also observe a significant fluctuation in speaking rate differences occurring within a short amount of time. In accordance with Švarný's et al. (1998, p. 7) observation that speech tempo is a variable influencing the rhythmical features, we have detected differences between slower and faster speech sections, especially on the level of linear segmentation, number of ictus-bearing units (*arses*) and typology of *cola*.

The analysis establishes two findings that seem germane to the issue of speech rhythm in spontaneous Mandarin Chinese communication. First, slower sections show a lower average length of *cola*, which are, according to Švarný, supposed to resist changes in speech tempo. This is not surprising, however, since the greater frequency of disfluencies results in shorter syntactic units and, consequently, shorter prosodic units. Moreover, discourse-specific markers, in particular, are often separated by longer pauses indicating the borderline between *cola*. Second, the narrative contains a considerably lower amount of stressed-prominence syllables. In this respect, it should be explained that their use in the *LDMC* sample seems to be a strong habit of routine: the speaker systematically emphasizes at least one syllable in each sentence example. In contrast, their use in the narrative is more selective: the storyteller highlights essential pieces of information in the context of the whole story plot, often occurring in longer strings of words.

On a global level, the rhythmical organization of the spontaneous narrative is similar to the read sentences of the *LDMC* sample. This feature is stronger in the case of the smaller linear units, i.e., segments. We can observe a high degree of regularity in the proportion of particular rhythmical patterns within the scope of the two samples. The issue of *cola* is more complicated. The ratio of the four basis rhythmical types in the narrative is greatly influenced by the heterogeneous nature of the syntactic structures

comprising both fragmented and highly elaborated clauses. Moreover, as has already been mentioned above, the occurrence of the first one is higher in the slower sections and of the latter in the faster section. The arrangement of rhythmical sequences within a certain type of cola shows, however, a similar proportional distribution. As far as we can conclude from the present data, the rhythmical features in general seem to be governed by universal principles, regardless of the speech mode in question. This is not to say, however, that there is no mutual connection between the spoken genre and the speech rhythm. Higher constituents of linear segmentation, i.e., cola, are the units that seem to exhibit certain discourse-specific rhythmical features that are worthy of attention in future research.

This paper opens a wide area of research in the field of speech rhythm in Mandarin Chinese. From the micro perspective, the present analysis, for example, points out the specific rhythmical function of discourse markers that is worthy of further investigation. Apart from this, examining the relationship between clause and colon might bring interesting findings regarding the position of their mutual boundaries. From the macro perspective, spoken communication offers numerous spoken genres. The scope of research can be expanded on finding the interrelationships between speech rhythm and various criteria used to classify them, including, but not limited to, the number of participants, degree of preparedness, communicative purpose, level of structural and context complexity, selection of lexical units, degree of elaboration of syntactic structures and occurrence of expressive elements.

Acknowledgments

Funding for the present research paper was provided by the Czech Ministry of Education, Youth and Sports for specific research (IGA_FF_2023_056).

References

- Bauman, R. (1986). *Story, Performance and Event: Contextual Studies of Oral Narrative*. Cambridge: Cambridge University Press.
- Blum-Kulka, S. (1993). "You gotta know how to tell a story": Telling, Tales, and Tellers in American and Israeli Narrative Events at Dinner. *Language in Society*, 22(3), 361–402.
- Brazil, D. (1993). Telling tales. In G. Fox, M. Hoey, & J. M. Sinclair (Eds.), *Techniques of Description* (pp. 154–169). London: Routledge.

- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Clark, H. H. (2014). Spontaneous Discourse. In M. Goldrick, V. Ferreira, & M. Miazzo (Eds.), *The Oxford Handbook of Language Production* (pp. 292–307). Oxford: Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199735471.013.023>
- Crystal, D. (2000). *The Cambridge Encyclopedia of Language*. Beijing: Foreign Language Teaching and Research Press.
- De Fina, A., & Georgopoulou, A. (2008). Introduction: Narrative Analysis in the Shift from Texts to Practices. *Text & Talk*, 28(3), 275–281.
<https://doi.org/10.1515/TEXT.2008.013>
- Duanmu, S. (2007). *The Phonology of Standard Chinese*. New York: Oxford University Press.
- Gajdoš, L. (2011). The Discrepancy between Spoken and Written Chinese – Methodological Notes on Linguistics. *Studia Orientalia Slovaca*, 10(1), 155–159.
- Golob, N. (2011). Acoustic Prosodic Parameters in Japanese and Slovene: Accent and Intonation. *Acta Linguistica Asiatica*, 1(3), 25–44.
<https://doi.org/10.4312/ala.1.3.25-44>
- Golob, N. (2022). Word-prosodic Typology and the Traps of False Similarity: Japanese and Slovene. *Acta Linguistica Asiatica*, 12(2), 115–123.
<https://doi.org/10.4312/ala.12.2.115-123>
- Halliday, M. A. K. (1989). *Spoken and Written Language*. Oxford: Oxford University Press.
- Jucker, A. H., & Ziv, Y. (1998). Discourse Markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Discourse Markers* (pp. 1–12). Amsterdam: John Benjamin Publishing.
<https://doi.org/10.1075/pbns.57>
- Juzwik, M. M. (2012). Spoken Narrative. In J. P. Gee, & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 326–341). London: Routledge.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov & Waletzky. (1967). Narrative Analysis: Oral Versions of Personal Experience. In J. Helm (Eds.), *Essays on the Verbal and Visual Arts* (pp. 12–44). Seattle: University of Washington Press.
- Li, W. (2019). Aspect in Chinese discourse. In Ch. Shei (Ed.), *The Routledge Handbook of Chinese Discourse Analysis* (pp. 160–174). London: Routledge.
- Lin, Y.-H. (2007). *The Sounds of Chinese*. Cambridge: Cambridge University Press.
- Liu, B. (2009). Chinese Discourse Markers in Oral Speech of Mainland Mandarin Speakers. In Y. Xiao (Eds.), *Proceedings of the 21st North American Conference on Chinese Linguistics* (Vol. 2, pp. 358–374). Smithfield, Rhode Island: Bryant University.

- Maschler, Y. (2009). *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam: John Benjamins Publishing Company.
- Miller, J., & Weinert, R. (1998). *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Clarendon Press.
- Norrick, N. R. (2000). *Conversational Narrative: Storytelling in Everyday Talk*. Amsterdam: John Benjamins Publishing Company.
<https://doi.org/10.1075/cilt.203>
- Norrick, N. R. (2007). Conversational Storytelling. In D. Herman (Eds.), *The Cambridge Companion to Narrative* (pp. 127–141). Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CCOL0521856965.009>
- Oliviera, M. (2012). A Study on Speech Rate as a Prosodic Feature in Spontaneous Narrative. *Alfa: Revista de Linguística*, 56(2), 623–651. DOI:
<https://doi.org/10.1590/S1981-57942012000200012>
- Petrovčič, M. (2009). *Operator LE in Chinese: Complexity within Simplicity and Simplicity within Complexity*. VDM Verlag. Retrieved July 17, 2023, from https://www.researchgate.net/publication/325828254_Operator_Le_in_Chinese
- Pospěchová, Z. (2016a). *Tónové změny v moderní čínštině ze sociolinguistické perspektivy*. Olomouc: Univerzita Palackého v Olomouci.
- Pospěchová, Z. (2016b). Prosodic Transcription of Standard Chinese and its Use in Teaching. *Acta Linguistica Asiatica*, 6(1), 35–45. <https://doi.org/10.4312/ala.6.1.35-45>
- Pospěchová, Z. (2022). Gendering Prosody in Communication: The Case of Standard Chinese. *Journal of Asian and Pacific Communication*, 32(1), 133–153.
<http://dx.doi.org/10.1075/japc.00087.pos>
- [Mrs.] Qin 秦女士 (2012). Ba tudou gei maile cai you qian shuye 把土豆儿给卖了才有钱输液 In K. P. van Dam, S. Hansen & J. Qi (Eds.), *Phonemica*. Retrieved October 21, 2017, from <http://phonemica.net/entry?mid=543748332bd553180882ab53&story=0>
- Shen, X. S. (1989). *The Prosody of Mandarin Chinese*. Berkeley, Los Angeles, Oxford: University of California Press.
- Slaměníková, T. (2016). Rytmičké členění příkladových vět Učebního slovníku jazyka čínského (hesla 1–50). *Dálný východ*, VI(2), 64–82.
- Slaměníková, T. (2018). An Analysis of the Correlation between Speech Tempo and Rhythmical Segmentation in Standard Chinese. *Studia Orientalia Slovaca*, 17(2), 151–169.
- Slaměníková, T., & Uher, D. (2021). Prozodická charakteristika textu: Připravenost a spontánnost mluveného projevu. *Dálný východ*, XI(1), pp. 104–122.
- Štefek, V. (2021). Prozodická charakteristika textu: Neformální promluva filmové postavy. *Dálný východ*, XI(1), 49–81.

- Švarný, O. (1991a). The Functioning of the Prosodic Features in Chinese (Pekinese). *Archiv orientální*, 59(2), 208–216.
- Švarný, O. (1991b). Rhythmical Features of Spoken Chinese: Quantitative and Grammatical Analysis. *Rocznik orientalistyczny*, XLVII(2), 132–137.
- Švarný, O., et al. (1998). *Hovorová čínština v příkladech III*. Olomouc: Vydavatelství Univerzity Palackého.
- Švarný, O. (1998–2002). *Učební slovník jazyka čínského I–IV*. Olomouc: Univerzita Palackého. (abbrev. LDMC)
- Švarný, O. (2000). Prosodical Transcription of Modern Chinese: Experimental Research and Teaching Practice. In Z. Palková & H.-W. Wodarz (Eds.), *Papers in Phonetics and Speech Processing* (pp. 149–159). Frankfurt am Main: Hector Verlag.
- Švarný, O., & Rusková Tang, Y. (1991). Prosodic Features in Chinese (Pekinese): Prosodic Transcription and Statistical Tables. *Archiv orientální*, 59(3), 234–254.
- Švarný, O., & Uher, D. (2014). *Prozodická gramatika čínštiny*. Olomouc: Univerzita Palackého v Olomouci.
- Tao, H. (1996). *Units in Mandarin Conversation: Prosody, Discourse, and Grammar*. Amsterdam: John Benjamins Publishing Company.
- Thornborrow, J. (2012). Narrative Analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (pp. 51–65). London: Routledge.
- Třísková, H. (2017). De-stressed in Mandarin: Clitics, Cliticoids and Phonetic Chunks. In Kecskes, I., & Sun, Ch. (Eds), *Key Issues in Chinese as a Second Language Research* (pp. 29–56). New York: Routledge.
- Třísková, H. (2019). Is the Glass Half-full, or Half-empty? The Alternative Concept of Stress in Mandarin Chinese. *Studies in Prosodic Grammar*, 2(4), 64–105.
- Uher, D. (2016). The Rhythmic Structure of Sentences in the Textbook of Chinese Conversation. *China & the World Cultural Exchange*, 5, 25–28.
- Uher, D., Jin, X., & Slaměňíková, T. (2016). *Učebnice čínské konverzace 2*. Voznice: Leda.
- Uher, D., & Slaměňíková, T. (Eds). (2019). *Oldřich Švarný: Prosodia linguae Sinensis*. Olomouc: Univerzita Palackého v Olomouci.
- Uher, D., & Slaměňíková, T. (2015). Prozodická analýza monologu. *Dálný východ*, V(2), pp. 104–115.
- van Dam, K. P., Hansen, S., & Qi, J. (Eds.). (2011–2021). *Phonemica*. <https://phonemica.net/about>, *Phonemica*. Retrieved July 19, 2023, from <https://phonemica.net/>
- Zheng, R. 郑汝可. (2013). Meigou xiaohou ban wanzhang shoulu Zhongguo fanyan 美国小伙办网站收录中国方言. *Hanwang – Changjiang ribao 汉网-长江日报*. Retrieved from July 17, 2023, http://hb.ifeng.com/news/fygc/detail_2013_07/12/986775_0.shtml

Appendix 1: Prosodic transcript of the analyzed narrative

List of abbreviations: c = colon; A = type A colon; B = type B colon; C = type C colon; D = type D colon

List of graphic symbols: / ascending sequence; /° ascending sequence followed by an enclitic syllable(s); \ descending sequence; °\ descending sequence preceded by a proclitic syllable; | odd ictus-bearing syllable; V trisyllabic two-arsis acronymic sequence; + contact of two arses in type C cola; ~~strikethrough text~~ indicates rhythmically hardly identifiable units that were excluded from the analysis

1. Event: Mrs. Qin tells about a farmer who came to her with a sick child suffering from severe diarrhea and dehydration. His father went to sell potatoes to earn money to pay for an infusion in the hospital. The child's health was already bad then, and the storyteller gave him some home-made medicine. Unfortunately, the father did not return and the child was not getting better. His mother decided to take it back home to the mountains because of the local tradition that people passing away outside cannot be buried on the village's ground. Before leaving, the storyteller gave the mother some more of the home-made medicine to take later; since there was no way to contact the family then, she did not know what happened with the child next.

Section	Syl/sec	Transcript	Rhythm
E1.01	2.6	you ² -yi-ci, zǎochen-qǐslai ² , wo-hǎi-mei qǐchuáng,	(c1) A: V (c2) B: \ / (c3) B: °\ /
E1.02	3.5	yi ² -ge-nóngmín bàozh-yi-xiǎoháir jiu-láile. ta ¹ -yao-kànbìng.	(c1) D: \ \ \ /°\ (c2) B: \ /
E1.03	3.3	wo ³ -yi-kàn nei-xiǎoháir jiushi, nèi-ge, lā-dùz, tuōshuǐ.	(c1) B: V // (c2) A: \ (c3) A: °\ (c4) A: /
E1.04	4.6	ta ¹ -d-yǎnjing yǐjing bibushàng-le, jiu ⁴ shi yǎnjing-bìshang-dou, zhōngjian-dou-you-fèngr.	(c1) B: \ \ \ V° (c2) B: \ \ \ (c3) B: \ V
E1.05	3.5	yǎnjing dou ¹ -shi-kōu-d, jiu ⁴ -shi, ránhou, ne, jiu-zhèi-yangr jiu-zhěnduàn nei-xiǎoháir dou-yǐjing-shi sǎn-du-tuōshuǐ-le.	(c1) B: V° / (c2) A: \ (c3) A: \ (c4) A: (c5) D: °\ /°\ /°
E1.06	2.9	tuōshui-yìhòu, ta-yǐnggai-shū-yè,	(c1) B: \ / (c2) B: °\ /
E1.07	2.9	kěshì-ne, yīyuan-lǐ ² tou ni-méi-you-qian ² mǎi-yào, ni ³ -gei-xiǎoháir shūbuliao-yè.	(c1) A: /° (c2) B: \ \ V° / (c3) C: \ /+ \ /
E1.08	2.8	ta-bà-ne, qù, dao ⁴ -nei-ge jiēshang-qu yao ⁴ -ba-tǔdòur gei-màile cai ² -you-qían.	(c1) A: °\ (c2) A: (c3) D: \ \ \ /° V
E1.09	4.6	nǎhui-qían-lai cai ² -neng-mǎi-yào gei-xiǎohair-zhìshang.	(c1) D: \ \ \ /°\
E1.10	3.6	hou ⁴ lai-wǒ kànzh-zhei-xiao ³ hair ² , wo-shuō zhei-zénme-bàn? ta ¹ -bu-tù, zhì ² -shi lā-d te ⁴ bie-lìhai.	(c1) C: V+ / (c2) D: /° V (c3) A: V (c4) B: \ \ \ \
E1.11	3.0	zhème-yangr ⁴ wo ³ -jiu gei ² -ta-yòng wǒ-d, e, táng, yán, he-nei ⁴ -ge sūdafěn.	(c1) C: V+ V+ \ (c2) A: (c3) A: (c4) A: (c5) B: °\ V
E1.12	4.2	zhei-sǎn-yangr-dōngxì, pèizai-yìqǐ, jiu-xiàng nei ⁴ -ge shūyè-d nei ⁴ -ge, wo-zhīdao-d nei ⁴ -ge-bìli.	(c1) B: °\ \ (c2) B: \ / (c3) C: /+ /°\ (c4) B: °\ \ /
E1.13	3.8	ránhou nei ⁴ -ge, wo ³ -jiu-gei-ta pèishang-shuǐ.	(c1) B: \ \ (c2) B: \ V
E1.14	5.1	wo ³ -gei-ta-yi ² -ge-wǎnr, gei-ta ¹ -yi-xiǎo-shǎor, wo-shuō ni-yidiānr-yidiānr manmānr-wèi, manmānr-wèi-ta.	(c1) B: \ V (c2) B: °\ / (c3) D: /°\ \ / (c4) B: \ /
E1.15	2.7	wèiwanle-yìhòu ⁴ -ne, ta-bà, hai ² -mei-mǎihui, hái-mei, hái-mei na ² hui-qían-lai.	(c1) B: \ \ (c2) A: \ (c3) B: \ \ (c4) A: \ (c5) B: \ \ \
E1.16	4.4	ta-mā-ne jiu-juéde zhei-xiǎoháir, shuō, dàifu, wo ² men-bu-zhìle, shuō zhei-xiaohair ² bù-xíng-le.	(c1) B: °\ °\ / (c2) A: (c3) A: \ (c4) B: \ \ (c5) B: //°
E1.17	2.9	hen-xiǎo, yi-liang sui? yi-liang-sui — liang — liang-sui-zuoyou ⁴ .	(c1) A: / (c2) A: V (c3) B: \ /
E1.18	5.5	ta-bàozh-xiaohair ² , ta ¹ men-dāngdi-d fēngsu-ne jiu ⁴ shi-shuo zhei-rén bù-neng sī ² zai-wàibianr yao ⁴ -sīzai jiā-libianr.	(c1) B: °\ / (c2) C: \ \ \ \ /+ \ \ \ \ \
E1.19	4.7	ta ¹ -xiang-neng-máizai tāmen-d dāngdi-d fēn-di-lǐ ³ bianr.	(c1) B: \ \ \ \ \ \
E1.20	4.3	yao ⁴ shi-sīzai wàibianr-d zhei ⁴ -xie, rén-ne, jiu ⁴ -bu-neng jiu-huí-cūn, jiu-zhì ² -neng zai-cūn-wàibianr-d dīshang-mǎi ² le.	(c1) B: \ \ \ \ (c2) A: \ (c3) B: \ / (c4) B: °\ / \ \ \
E1.21	3.5	suo ³ yi tāmen-d-ren ² , bu-zhì-dao zuihou ⁴ .	(c1) B: \ \ / (c2) B: °\ /
E1.22	3.3	ránhòu-ne, ta ¹ -jiu-bàozh zhè-xiaoháir jiu ⁴ -yao-zòu.	(c1) A: /° (c2) C: \ \ V+V
E1.23	3.2	ranhòu-ne, wo ³ -jiu gei-tā-bāole táng he-yán, zhei ⁴ -xie jì-yangr-dōngxì.	(c1) A: /° (c2) D: °\ /° / (c3) B: \ / \

E1.24	4.1	wo ³ -rang-tā, wo-shuo ¹ jia ³ -ru-huǐjiā xiāohair ² -hai-kěyi, wo —wo jiu— we— wo ³ -rang-ta-chī zhei-yào.	(c1) A: V (c2) C: /+ \ / /°\ (c3) B: \ / /
E1.25	4.9	ranhòu ba-wo ³ -zìjǐ-d shi-huángliansù-ma shi-shen ² me-dōngxì, wo ³ -zai- gēi-ta yì ² -piānr-liǎng-piānr ₄ , yin'wei-xiāoháir hen-xiǎo.	(c1) D: //°°V°°\ (c2) B: \ / / / (c3) B: \ / /
E1.26	2.9	jiēguō-ne, ta ¹ -jiu-zǒule.	(c1) A: /° (c2) B: \ /
E1.27	5.7	zou ¹ le-dào dà-shān-libianr. gen'ben jiu-méi-you-diànhuà, méi-you- shen ² me, gen'ben jiu-méifār-liánxi shì ⁴ -bu-shì?	(c1) B: \ / /°° (c2) B: \ / / (c3) B: \ / (c4) C: \ /° / + \

2. Event: After some time, Mrs. Qin came by coincidence together with a barefoot doctor to the village where the child was from. His mother dragged them to their home and told them the child had healed from the disease. Then, the mother served them some local dish that Mrs. Qin could not eat because of its extreme spiciness. Taking advice from the barefoot doctor, she asked the child's mother to let her take it away and eat later, so as not to show a lack of respect. At the end of this part, Mrs. Qin admires the sincerity of the farmer's gratitude.

Section	Syl/sec	Transcript	Rhythm
E2.01	2.8	you-yí-ci nei ⁴ -ge, tā-d-nei ⁴ -ge, chījiao yīshēng...	(c1) B: °\ (c2) B: \ / (c3) B: \ /
E2.02	4.8	nèi-shìhour jiu-guǎn-tamen jiao-chī jiu'shì guāngzh-jīāobuyār-d nei ⁴ -ge, chījiao-yīshēng.	(c1) C: /° \ / + \ / V°\ (c2) B: \ /
E2.03	4.3	ran ² hou-tā, wo ³ men-liǎng-ge, bu ⁴ -zhìdào yao ⁴ -qu-zuò-shenme, dao- nóngcun-lǐ ² bianr.	(c1) A: V (c2) B: \ / (c3) C: V+ \ / (c4) B: °\ (c5) A: \ /
E2.04	3.1	zhènghaor-zǒudào zhei ⁴ -ge-nǚd ta ¹ men-jīā.	(c1) B: \ / / / V
E2.05	3.4	wa! ta-fēi-yao ba ³ -wo-lādao ta ¹ men-jīā!	(c1) A: / (c2) B: ° \ / / V
E2.06	3.8	ran ² hou-wǒ, wo ³ -bu-rènshì-tā,	(c1) A: V (c2) B: \ /
E2.07	3.5	yin'wei-tā, —nong— nèi-xie-ren ² dou-zhǎng-d wo ³ -dou-juéde tèbie-yi- yàng.	(c1) A: V (c2) D: V° \ / / / / /
E2.08	3.4	suo ³ yi-hòulai, tā— ta ¹ -jiu-gēn-wo-shuō zhèi-jian-shìr, ta-shuō wo ³ -d- xiāoháir meì ² -you-sí.	(c1) B: \ / (c2) C: \ / V+V (c3) C: /+ \ / +V
E2.09	4.2	huǐjiā wo ³ -jiu-gei-tā-hē zhei-shuǐ, ta ¹ -jiu-shì-huó, jiu-hǎo, jiu-jiu-hao-le-nei bing-	(c1) C: /+ \ / / (c2) B: \ / (c3) A: /
E2.10	2.2	hòu ⁴ lai-nèi-ge, tā-gei-wo zuò-shenme-ne? nèi-ge	(c1) B: \ / (c2) B: \ / (c3) A: \ /
E2.11	2.7	qī qīshu, jiùshì, yóuqī-d-nei ⁴ -ge, shù-keyi chū-youqi ¹ wo ³ -bu-zhìdào.	(c1) A: \ (c2) A: \ (c3) D: /°\ (c4) C: \ / V+ \ /
E2.12	3.4	ta ¹ men-jiao-qīshu, jiu ³ shì yóuqi-d-qī.	(c1) B: \ / (c2) B: \ /
E2.13	2.7	qīshu-d, nei ⁴ -ge-shù-zǐr, zhàchulai-d-yóu	(c1) A: \ (c2) B: \ / (c3) B: \ /
E2.14	3.4	gei ³ -wo-zhā-yi-ge, jiu-jiu— fan ² zheng-wo-gūji-jiao, ta-shuō-shì, jiu-xiàng, tā-jiao wánz-ba, quèshì-jiùshì sū-mian-gēdar	(c1) B: \ / (c2) B: \ / (c3) A: °\ (c4) A: / (c5) B: \ / (c6) B: \ /
E2.15	3.8	nèi-ge-dong'xi, ke ³ shì-zhèi-ge	(c1) B: \ / (c2) B: \ /
E2.16	2.8	méi-you _s , jiu ⁴ -shì-miàn.	(c1) A: \ (c2) A: V
E2.17	2.9	ránhòu, nèi-ge. wa! chí ¹ -zhei-ge-yóu-ba, tèbie-là sang ² -z-yān-d-huang, là-d nei ⁴ -zhong-wèir.	(c1) A: \ (c2) A: \ (c3) A: (c4) B: \ /° (c5) C: V+ \ / (c6) B: \ /
E2.18	3.3	ni-gēnben yānbuliǎo.	(c1) B: ° \ /
E2.19	4.7	hòulai-wo ³ -jiu-gǎnkuai toutōu gen ¹ -nei chījiao-yīsheng-shuo ¹ , wo-shuo ¹ zhei-dōngxì-wo ² zhēn-d-chī ¹ buliǎo.	(c1) C: \ / / + \ / V (c2): C: /° V+V
E2.20	4.4	ke ³ shì-tā, jiu ⁴ shì-tāmen-jīā zuì-hao-d-dōngxì, ta ¹ -gei-ni-zuò yóu-zhā-d zhei-dōngxì.	(c1) A: V (c2) C: \ / V+ \ / (c3) D: \ / /°°
E2.21	3.6	hòulai-wo —we— wo-shuo ¹ wo-zhēnd-chībuliǎo. ta-shuō, wo-shuo ¹ ni- gǎnkuai xiāng-bànfā.	(c1) A: \ (c2) D: /° \ / (c3) A: / (c4) D: /°°
E2.22	4.4	hòulai ta ¹ -jiu gen ¹ -zhei-nǚd-shuō (+ a question of the man)	(c1) B: \ / / / V
E2.23	2.6	dāngdi-d, cūn-libianr-d	(c1) A: \ (c2) A: \
E2.24	3.1	ranhòu, ta ¹ -jiu-gǎnkuai gen ¹ -ta zěnme-shuō.	(c1) A: / (c2) B: \ / / / V
E2.25	3.9	shuo ¹ le-yìhòu ke-hái-gei-wo-dàishang. dàishangle, wo ³ -jiu gǎnkuai-zǒu-le.	(c1) D: \ /° \ / (c2) A: \ / (c3) B: \ /
E2.26	3.3	jiu-shuō hái-you bie ² d-shìr	(c1) C: /+ \ / V
E2.27	2.2	hòulai, suǒyì jiùshì, ta-dāngdi-d nóngmin-ba, jiùshì, tā...	(c1) A: \ (c2) B: \ / (c3) B: °\ (c4) A: / (c5) A:
E2.28	2.8	yin'wei-gēn wàibianr jiēchu-bu-duō, suo ³ yi-tā xīn tèbie-zhēncheng.	(c1) C: V+ \ / / (c2) C: V+ + \ /
E2.29	1.9	zhēnd-gǎnxiè-ni ³ !	(c1) B: \ /
E2.30	2.0	zhei-gùshì wo-yì-beiz ye-wàngbuliǎo	(c1) B: ° \ /° V

3. Event: Mrs. Qin elaborates her experience with farmers' genuine and sincere hospitality. She describes how she met an old farmer who offered her a dish that, as she assumed, consisted of ingredients valuable for the farmers, and she expressed her concerns that perhaps there might not be enough for the farmers themselves.

Section	Syl/sec	Transcript	Rhythm
E3.01	3.2	zai-yǒu, jiǎru-dao ⁴ tiáojiàn shāowei-hǎo-yidian ³ -d-a, nèi-ge, nei-nóngmín jiā-lìbianr	(c1) A: / (c2) C: V/+ \ V ^{oo} (c3) A: \ (c4) B: ° \ \
E3.02	4.0	ránhòu, wó ³ men you ³ -d-shíhòu yě-ke.neng gōngshe-d gen ¹ -wo yikuàir-qu	(c1) A: / (c2) B: \ \ \ \ \ \ / ° (c3) D: // ° \ \ //
E3.03	3.4	ránhòu jiùshì ⁴ làotàitai jiù ⁴ -gei-wo-zhǔ yi-wǎnr-miàn	(c1) C: +V (c2) B: \ \ (c3) C: /+ \ \
E3.04	4.4	shuō chī-wǎnr-miàn, nuǎnhuo-nuan ³ huo-ba, chī-yi-wǎnr-miàn nuǎnhuo-nuan ³ huo	(c1) D: / ° \ \ \ \ / (c3) A: \ (c2) D: \ / ° \
E3.05	3.5	shíjìshang wó ³ -d-biǎomiàn r yì-céng shì-miàntiáor	(c1) A: (c2) D: V // ° / (c3) C: \ \ /+ \ \ \ /
E3.06	4.3	duì, ta'men-jiā kěndìng-you-ren ² zai-wàibianr gongzuo ⁴ , suo ³ yi jiù ⁴ -gei-ta-yǒu nài-zhong shǒugōng-d xiǎo-d guà-miàn.	(c1) A: \ (c2) D: \ / ° \ (c3) A: \
E3.07	3.0	dōu-shì, hěn-bu-róngyì-d wo-zhīdao, suǒyi	(c1) A: \ (c2) B: \ \ (c3) B: \ \ (c4) D: V ° \ / (c5) B: //
E3.08	3.4	zhèi-ge, kě ³ shì-díxia-ne, wó ³ -nei wǎn-lìstou, kě ³ yi-fàng sì-wǔ-ge-jīdàn, biǎomiàn-r-yì-céng miàntiáor	(c1) B: \ \ \ \ (c2) A: ° \ (c3) B: // V (c4) B: \ ° \
E3.09	3.9	jiùshì yào-gei-wo chī-d, nǐ-zhīdao-ma	(c1) B: \ \ \ (c2) A: ° \
E3.10	2.4	suǒyi, jiùshì, kěshì tāmen biéren...	(c1) A: \ (c2) A: \ (c3) B: \ \ \
E3.11	4.8	kěndìng-méi-yǒu. rang-tāmen, gei-tā yidian ³ buguǎn-shénme-chī, wó ³ -bu wo ³ -dòu bú-jīde-le	(c1) B: \ / (c2) A: ° \ (c3) B: // V (c4) B: \ ° \
E3.12	3.7	wó ³ -jiù-zhèi-ge, wó ³ -jiù tèbié-gǎndòng, làotàitai-duì-wo ³ zhēn-xin-zhen-yì-d	(c1) B: \ \ (c2) B: \ \ \ / (c3) C: \ /+ \ / °

Appendix 2: An example description of linear segmentation and syllable prominence

List of abbreviations: c = colon, s = segment, A = arsis, T = thesis

Example 1: E1.01 you³-yi-cì, zǎochēn-qǐlǎi², wó-hái-mei qǐchuáng,

Linear segmentation: 3 cola (c1, c2, c3); 4 segments (c1_s1; c2_s1; c3_s1, c3_s2)

Syllable prominence:

(c1_s1) you³-yi-cì: A weakened-tone + T void-tone + A full-tone

(c2_s1) zǎochēn-qǐlǎi²: A full-tone + T void-tone + T weakened-tone + A weakened-tone

(c3_s1) wó-hái-mei: T weakened-tone + A full-tone + T void-tone

(c3_s2) qǐchuáng: T full-tone + A full-tone

Example 2: E2.29 yin¹wei-gēn wàibianr jiēchu-bu-duō, suo³yi-tā xīn tèbié-zhēncheng

Linear segmentation: 2 cola (c1, c2); 6 segments (c1_s1, c1_s2, c1_s3; c2_s1, c2_s2, c2_s3)

Syllable prominence:

(c1_s1) yin¹wei-gēn: A weakened-tone + T void-tone + A full-tone

(c1_s2) wàibianr: A stressed-prominence + T void-tone

(c1_s3) jiēchu-bu-duō: A full-tone + T void-tone + T void-tone + A full-tone

(c2_s1) suo³yi-tā: A weakened-tone + T void-tone + A full-tone

(c2_s2) xīn: A full-tone

(c2_s3) tèbie-zhēncheng: A full-tone + T void-tone + A full-tone + T void-tone

Example 3: E3.07 dōu-shi, hěn-bu-róngyì-d wo-zhīdao, suǒyì

Linear segmentation: 3 cola (c1, c2, c3); 4 segments (c1_s1; c2_s1, c2_s2; c3_s1)

Syllable prominence:

(c1_s1) dōu-shi: A full-tone + T void-tone

(c2_s1) hěn-bu-róngyì-d: A stressed-prominence + T void-tone + T full-tone + A full-tone + T void tone

(c2_s2) wo-zhīdao: T weakened-tone + A full-tone + T void-tone

(c3_s1) suǒyì: A full-tone + T void-tone