The Structure of the Mandarin Syllable:
Why, When and How to Teach it

Hana Třísková

Introduction

In this article, we shall be concerned with the isolated segmental syllable of Mandarin,[1] and will be leaving aside the suprasegmental component of tone. A traditional model of the Chinese syllable (see Figure 2) splits up a syllable into two basic subsyllabic components: an initial (*shēngmǔ* 声母, i. e. an initial consonant), and a final (*yùnmǔ* 韵母, i. e. the rest of the syllable). This traditional analysis is not necessarily accepted by all phonologists (for an alternative solution, see Figure 4). However, in teaching Mandarin it has been employed more or less universally, since it was adopted as the basis for the widely accepted *pīnyīn* spelling system (*pīnyīn* is a major spelling tool used in current textbooks). Since we are concerned with methods of language teaching in this article, we shall follow this line and stick to *pīnyīn* in all respects. It is used for the notation of syllables, initials, finals, or particular sounds (it is always presented in italics). *Pīnyīn* symbols (not IPA symbols) are also used for phonological representation. Why do we need it? Since the *pīnyīn* system was not designed solely as a phonetic or phonological transcription tool but was fashioned to serve a wide-range of purposes in society, its orthography does not always faithfully reflect the underlying phonological structure of the syllables. For example, let us look at the morpheme 贵 "expensive:" its standard *pīnyīn* notation is *guì*; yet the final is phonologically /uei/, not */ui/ – the main vowel /e/ is left out in standard orthography. Should we need to provide or clarify the underlying phonological form of a particular *pīnyīn* item (especially in the case of a discrepancy between a phonological structure and standard orthography), we provide its phonological representation, enclosed in slanted brackets. For example, the sound form of the above mentioned morpheme 贵 "expensive" can be represented in three clearly distinguished ways: a standard *pīnyīn* orthography would be *guì* (italics, no brackets), an underlying phonological representation would be /*guei*/[4] (italics, slanted brackets), and a transcription in IPA would be **[kweɪ]**[4] (IPA symbols, square brackets). When there is no discrepancy between the *pīnyīn* orthography of an item and its phonological representation (e. g. *huai* vs. /*huai*/), or no need to point out the phonological representation of an item, slanted brackets are generally not used. Note that such a discrepancy may happen only in the notation of finals; there are no such cases in the notation of initials.

In Chinese, a syllable is a linguistic unit of utmost importance: among other functions, it serves as the carrier of a morpheme. In other words, one syllable corresponds to one morpheme in most cases. Mandarin syllabary is rather meager: it comprises of only about 400 segmental syllables, or 1,300 tonal syllables. In learning Mandarin Chinese, mastering the basic syllabary represents a crucial step in the process of mastering the spoken language. A poor command of the pronunciation of the isolated syllables at the very beginning of the process of studying the language can undermine the whole spoken language competence of the student in the future. Furthermore, problems with the production of syllables may go hand in hand with problems of perception – both on the side of a language student himself/herself, and on the side of a native Chinese attempting to comprehend his/her speech. Thus, it is rather obvious that reaching maximum efficiency in teaching Mandarin syllabary is a highly important pedagogical goal. The standard procedure[2] is roughly as follows:

– A student is told that every Mandarin syllable is composed of three components: an initial, a final, and a tone.

– The inventories of the initials (21 items), finals (cca 35 items) and tones (4 items) are listed.

– The whole inventory of syllables is introduced step by step, each newly taught segmental syllable being presented as a combination of a particular initial and a particular final. The finals are traditionally organized into three groups: the "simple finals," *dān yùnmǔ* 单韵母 (simple vowels), the "compound finals," *fù yùnmǔ* 复韵母 (diphthongs and triphthongs), and the "nasal finals," *bí yùnmǔ* 鼻韵母 (all finals ending with a nasal consonant). The combinations of initials with simple finals are naturally taught first.

As part of the process of learning a basic syllabary, various types of mistake appear in the students' pronunciation. If we leave aside the suprasegmental feature of tone and focus our attention on the mistakes of a segmental sort, the errors may be divided into two major groups: mispronunciations of initials, and mispronunciations of finals.

Addressing the mistakes of the first group is relatively simple, since an initial is always identical with a single consonantal segment in *pīnyīn*. Of course, there are a number of difficult cases among the initial sounds (e. g. the so called retroflex initials *zh, ch, sh, r*, or the so called palatal initials *j*, *q*, *x*). Their incorrect pronunciation may arise from the absence of a similar sound in the native language of the student. This is often the case with the affricate *zh*, or the fricative *x*, or the aspirated affricates *q*, *ch* (note that the aspirated affricates are rather rare in languages). Yet, on the other hand, the presence of a similar, but not entirely identical sound in the native language of the student may distort his/her pronunciation as well. For example, Americans tend to pronounce the Mandarin initial consonant *r* with rounded lips, since in English the initial **r** is more or less labialized[3] (the remaining features of both sounds are basically identical: both are apical postalveolar approximants). Whatever the case, the difficulties with the pronunciation of initials never surpass the difficulties with

pronouncing a single segment. Mistakes can be cured by proper instruction about the articulation of the consonant in question (let us remind ourselves that most of the initial consonants have only one allophone and that the assimilations of initials to the following vocalic element are rather subtle). For instance, for the so called retroflex initials ***zh, ch, sh, r*** the teacher must make clear the following points: first, their place of articulation is *postalveolar*; second, they should be articulated with the very *tip* of the tongue, i. e. apically (not with the underside of the tip, i. e. subapically, nor with the blade of the tongue, i. e. laminally); third, the tip of the tongue is only raised upwards, not bent backwards (as is the case of truly retroflex consonants of Tamil or other Dravidian languages).

The situation with the component of a final is much more complicated. A final is a complex structure: it can consist of a single segment (such as in the syllables *bi, ma, de* etc.), of two segments (such as in the syllables *bin, mai, die* etc.), or of three segments (such as in the syllables ***bian, miao, shuang*** etc.). First, let us find out how difficult the particular segments which may occur within a final actually are. The range of options is /*i*/, /*u*/, /*ü*/, /*a*/, /*e*/, /*o*/, /*n*/, /*ng*/. [4] Most of them are common vowels. Two segments are nasal consonants, the former one being extremely common across the languages. The only two segments which could potentially be viewed as difficult are /*ü*/ and /*ng*/. Furthermore, if we are to do full justice to the situation, we need to acknowledge that the simple sounds traditionally called "apical vowels" can also cause certain difficulties to the student (*pīnyīn* interprets them as allophones of /*i*/ occurring after the apical sibilants, namely in the syllables *zi, ci, si, zhi, chi, shi, ri*; however, the pronunciation of "apical vowels" is rather different from **[i]**). The pronunciation of a simple vowel /*e*/ (as in the syllables *she, ze* etc.) is also somewhat tricky: it is **[ɤˑ]**. Yet – taking everything into consideration – most of the problems with pronouncing the finals do not arise from the fact that this or that particular segment is difficult as such (as is the case with certain initials). The major cause of mispronunciations of the finals lies somewhere else. But *where* exactly? This article attempts to find both an answer to this question, as well as the solution. The stimulus for writing the present text came from some ponderings related to my Ph.D. dissertation, *Segmentální struktura čínské slabiky* [segmental structure of the Mandarin syllable], defended in September 2010 at Charles University, Prague.[5]

In the previous paragraph, an initial and a final were established as the two major subsyllabic components of the Mandarin syllable. In order to engage fully in the following discussion, we need first to introduce the syllable structure in more detail.

The Syllable as a Configuration of Vowels and Consonants

A syllable of any language can be viewed as a certain configuration of vowels, **V**, and consonants, **C** (we shall return to the notions of vowels and consonant later). For example, *dog* is CVC, *frog* is CCVC etc. The particular configurations are called syllable types (sometimes alternative terms are used, e. g. syllable patterns, syllable structures, syllable templates etc.; we prefer the term "syllable types").

Every language has its own set of permissible syllable types. For Mandarin, twelve syllable types can be established: V (*a*), CV (*ma*), VV (*ai*), VC (*an*), GV (*ya*), GVV (*wai*), GVC (*yan*), CVV (*mai*), CVC (*man*), CGV (*mie*), CGVV (*kuai*), CGVC (*mian*). The symbol G stands for "a glide," which is a prenuclear vowel. Accepting this symbol makes it possible to distinguish between two kinds of VV sequence: GV as a rising diphthong, VV as a falling diphthong. Note that V (a nuclear vowel) is present in all syllable types – it is an obligatory item, at least in the view of *pīnyīn* (some authors, on the other hand, work with the concept of syllabic consonants).[6] The vowelless syllables such as *hng*, *hm*, *ng* do not need to be considered – they are used only as interjections and remain outside the system of Mandarin syllables.

The Syllable as a Structure

In the previous paragraph a syllable was presented as a linear sequence of vowel and consonant segments, chained together like beads on a string. However, nowadays, linguists accept the view that a syllable (of any language) is a hierarchically organized structure, with subsyllabic components of more than one level. The commonly accepted general model of a syllable is presented in Figure 1. It reminds us of an upside down tree:
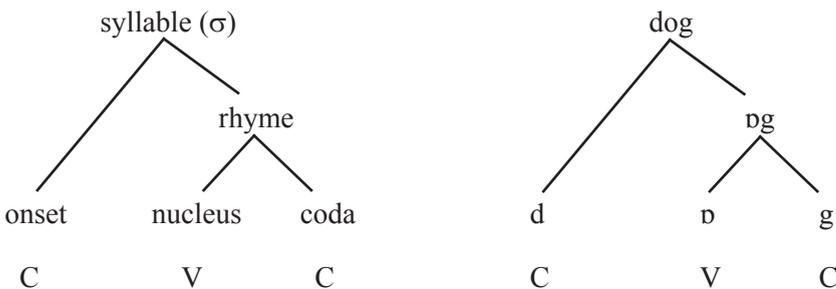
Figure 1. A general model of syllable structure, and an English example *dog*

In this model a syllable consists of a consonantal *onset* (one or more consonants), followed by a *rhyme* (sometimes written as *rime*). A rhyme consists of a *nucleus* (typically a simple vowel, sometimes a diphthong, less frequently a syllabic consonant), which may be followed by a consonantal *coda* (one or more consonants). The concrete options for completing this abstract model differ in various languages. For example, some languages prohibit the presence of more than one consonant in an onset (Mandarin), some languages prohibit empty onsets (Hebrew), some languages prohibit the use of a coda (Hawaiian), etc. Generally, every syllable of every language must have a nucleus.

The Syllable Structure in Mandarin

As far as the Mandarin syllable is concerned, there are two major models, i. e. two major ways of looking at its structure:[7]

a) The *Initial-Final model*   声母—韵母样式   *shēngmǔ – yùnmǔ yàngshì*
b) The *Onset-Rime model*   节首—韵基样式   *jiéshǒu – yùnjī yàngshì*

The *Initial-Final model* is a traditional one. The Chinese spelling system *pīnyīn* is based on it. The model can be portrayed in the following way (neglecting the tone component):

syllable 音节

initial 声母   final 韵母

medial 韵头   rime 韵

nucleus 韵腹   ending 韵尾

C = consonant
G = glide
V = vowel
X = consonant or vowel

C	G	V	X
k	w	a	n
k	w	a	i
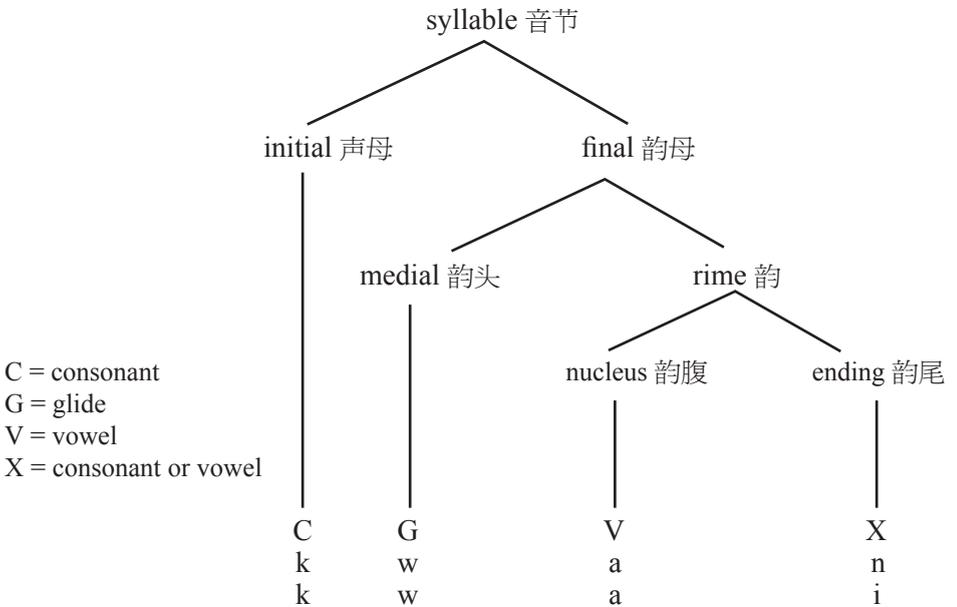
Figure 2. The *Initial-Final model* [8]

The *Initial-Final model* can be alternatively represented in the form of a table:

tone			
	final		
initial	medial	rime	
		nucleus	ending vocalic / consonantal

Figure 3. The *Initial-Final model* [9]

At the lowest level, that is at the level of segments, four components appear:

C	initial	声母 *shēngmǔ*	(initial consonant)
G	medial	韵头 *yùntóu*, 介音 *jièyīn*	(glide, prenuclear vowel)
V	main vowel	韵腹 *yùnfù*	(nucleus, nuclear vowel)
X	ending	韵尾 *yùnwěi*	(a vowel or a nasal consonant)

The element called an ending seems to remind us of the coda element found in the model illustrated in Figure 1. However, there is an important difference: while a coda must be a *consonant*, the ending of a Chinese syllable (*yùnwěi*) may be either a *consonant*, or a *vowel*. That means the ending element of a Chinese syllable cannot be simply identified with a coda element of a general syllable scheme. We shall touch upon this problem later. Furthermore, note that the component of a medial has no counterpart in the general model of a syllable (Figure 1).

There are eight permissible combinations of the elements C, G, V, X in Mandarin. We call these combinations syllable templates. They are listed here: V (*a*), CV (*ma*), VX (*ai, an*), GV (*ya*), GVX (*wai, yan*), CVX (*mai, man*), CGV (*mie*), CGVX (*kuai, mian*). Note that the number of syllable templates is smaller than the number of syllable types. The reason is that X may be either V, or C.

Let us make a brief excursion into the world of terminology at this point. In Russian, there exist three alternative terms worthy of our attention in relation to several subsyllabic components: централъ ("a central") for the main vowel, терминалъ ("a terminal") for an ending, and субфиналъ ("a subfinal") for a rime. The first two were coined by a Russian phonologist and orientalist V. B. Kasevich in the 1960s. The third term was coined by another Russian linguist – A. A. Moskalev, also in the 1960s. These three terms have a number of advantages: they share the same suffix -*al* with the terms initi*al*, fin*al*, medi*al*; they are short, unambiguous, and they are language specific: they refer to the components of a *Chinese* syllable, not of a general abstract syllable scheme (as is the case with the expressions *main vowel, nucleus, coda* or *rhyme/rime*, which are regularly used in syllable theory). The terms централъ, терминалъ and субфиналъ appear for instance in a Russian textbook on Mandarin phonetics by A. N. Speshnev (published in 1972, 1980 and 2003). Their conversion into Czech is quite smooth: *centrála, terminála, subfinála* (it is immediately clear by their morphology that they are nouns; the same holds for their Russian versions). The translation of централъ, терминалъ and субфиналъ into English is less advantageous: the words *central, terminal* and *subfinal* may be either nouns, or adjectives; however, the same is also true for the words *initial* and *final*. We presume that the terms a *terminal* and a *subfinal* may be employed without much trouble; we accept them. The remaining term a *central* is questionable, however – it can possibly be confused with an adjective referring to the tongue position during the articulation of a vowel, as in the expression "a central vowel" (cf. "a front vowel," "a back vowel"). One could possibly think

of coining a new English term for the main vowel of the Chinese syllable, also containing the suffix *-al*, e. g. "a principal." The result would be a neat set of labels with an identical ending *-al* for all components of the Mandarin syllable, except for the tone. Furthermore, we wish to make a terminological remark concerning the component of a *medial*. In the phonetic / phonological descriptions the medials /*i*/, /*u*/, /*ü*/ are commonly characterized as glides (being transcribed as [j], [w], [ɥ]), or alternatively as semivowels (being transcribed as [i̯], [u̯], [y̆]; a diacritical arch indicates a vowel is non-syllabic). We prefer to avoid the latter term (as well as the latter transcriptions) as being somewhat confusing. Another option is to use a rather modern term, *approximant*, instead of a glide (note that it may cover a slightly broader range of sounds than the term glide – depending on the author). In such a case the transcriptions [j], [w], [ɥ] should be employed (see also below).

The *Initial-Final model* draws on the traditions of medieval Chinese philology. The origins of the phonological concepts embodied in this model can be traced back to the ancient rhyming dictionaries, in particular to the oldest rhyming dictionary that has been preserved – *Qièyùn* 切韵 (A.D. 601). In this dictionary, a "spelling" system ***fǎnqiè*** 反切 was used systematically. The *fǎnqiè* system indicates the reading of an unknown character by using two other, presumably known, characters. The first one represents an initial consonant (*shēngmǔ* 声母), while the second one represents the rest of the syllable (called "a final," *yùnmǔ* 韵母 later on), including the tone. For example, the modern Mandarin syllable *huáng* can be represented by two characters pronounced as ***h(ē)*** and (*k*)***uáng***. Furthermore, the rhyming dictionaries reflected the knowledge held by their authors of the subsyllabic constituent corresponding to "rhyme" in Figure 1. It was the rhyme-carrying part of the syllable (*yùn* 韵 in Chinese; this Chinese term means either "a rhyme," or "a rhyme category"). Note that the component called *yùn* 韵 does not include a medial (cf. Figure 2); the syllables with an identical rhyme and a different medial may rhyme. For instance, the modern syllables ***dāng***, *xiāng* and *kuāng* rhyme. As far as the remaining subsyllabic components are concerned, there are signals that the medieval Chinese phonologists had some implicit knowledge of the existence of the medial element (*yùntóu* 韵头), and of the terminal element (*yùnwěi* 韵尾). However, Chinese phonological tradition never actually arrived at an explicit analysis of the syllable into segments. The analysis of *yùn* 韵 into a main vowel (*yùnfù* 韵腹) and ending (*yùnwěi* 韵尾) was achieved under the influence of Western phonology.

The second model of the Mandarin syllable mentioned above was the *Onset-Rime model*. It is much more recent than the *Initial-Final model*, being adopted by many contemporary phonologists.[10] It is not compatible with *pīnyīn*. This model is shown in the following diagram:

```
                        σ
                     /     \
                   O         R
                  / \       /  \
                 /   \     N    Co
                 |   |     |     |
                 C   G     V     X
                 k   w     a     n
                 k   w     a     i
```

σ = syllable
O = onset
R = rime
N = nucleus
Co = coda
C = consonant
G = glide
V = vowel
X = consonant or glide
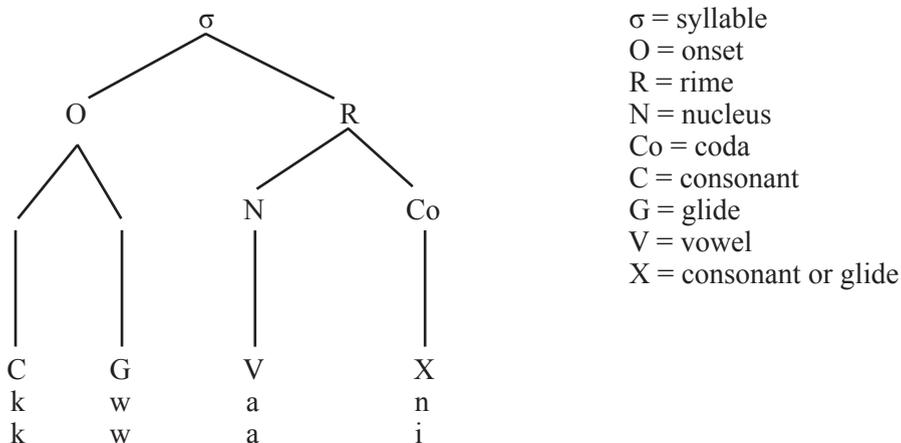
Figure 4. The *Onset-Rime model* [11]

Let us compare the models in Figures 2 and 4. A unit called *rime* is present in both. The view that a connection between a syllable nucleus and a syllable coda is generally more tight in languages than that between an onset and a nucleus is broadly accepted among phonologists (cf. Figure 1). Mandarin is no exception: the component of rime is regularly present within basically all models of the Mandarin syllable. We have established a term *subfinal* for it, marking it as VX. A totally tight connection between the segments constituting a subfinal, i. e. between a main vowel and a terminal, results in the assimilations between both segments. The phoneticians agree on the observation that a non-high main vowel tends to assimilate in backness to a terminal element: if a terminal is front, i. e. /*i*/ or /*n*/ (as in *bai, mei, ban, ben*), a main vowel is pronounced as front, or at least as non-back, e. g. *ban* → **[pan]**, *ben* → **[pən]**; if a terminal is back, i. e. /*u*/ or /*ng*/ (as in *bao = /bau/, bang, beng*), the main vowel is pronounced as back, e. g. *bang* → **[pɑŋ]**, *beng* → **[pʌŋ]** (note that assimilations of a mid vowel /ə/ are accepted less unanimously in phonological literature than those of a low vowel /**a**/; some authors have **[əŋ]**, not **[ʌŋ]**). The described assimilations run through a large part of Mandarin syllabary: they concern all syllables containing VX, namely the syllables with the finals of the VV, VC, GVV and GVC type (only types V and GV are excluded). These assimilations should be given due attention, in particular those of a low vowel /**a**/: note that the difference between a front **[a]** and a back **[ɑ]** may serve as the only perceptual cue for distinguishing between *ban* and *bang*, *zhan* and *zhang* etc. in casual speech, if a terminal nasal consonant is missing in pronunciation: *ban* → **[pã]**, *bang* → **[pɑ̃]**.

Another point shared by both models is the notion of X (an *ending element*), which may be either a nasal consonant, or a vowel (let us point out again that there is a difference between the notion of X and the notion of a coda).

On the other hand, there is a major difference between both models: in Figure 2 (the *Initial-Final model*) a glide is a part of a final. In Figure 4 (the *Onset-Rime model*) a glide belongs to an onset of a syllable; consequently, there is no final. There are two variants of the *Onset-Rime model*. The first one is illustrated in Figure 4: a glide has its own slot within an onset. Some authors are more radical, though – they hold that a glide should *not* have a separate slot within an onset.[12] In their view, although a glide still does have the status of a phoneme at the underlying level, at the phonetic level it is realized only as a secondary articulation to an initial consonant, e. g. /**sw**/ → [**sʷ**]. Thus, the underlying combination CG surfaces as a single sound, e. g. *suān* 蒜 "garlic" is pronounced as [**sʷan**], not [**swan**]; the lips are rounded already during the articulation of [**s**], while there is no separate sound [**w**].[13] This makes the situation different from the English word *swan*. There is one more argument going on among those authors who accept the *Onset-Rime model*: they may disagree about the affiliation of the terminal vowels /**i**/, /**u**/ (as in the syllables *mai, kuai, hao* = /*hau*/). Some authors place /**i**/, /**u**/ into the nucleus of a syllable. The nucleus then has two slots, while the coda remains empty. This is a case of falling diphthongs: the whole falling diphthong is placed within a *nucleus* (an analysis common in Western phonology). Such a solution disrupts the correspondence between the rhyme part of both models. Namely, it liquidates a general correspondence between the nucleus and *yùnfù* (the main vowel), and, by the same token, the correspondence between the coda and *yùnwěi* (the terminal). The cases which witness a disruption of the correspondence are the finals containing falling diphthongs or triphthongs. Other authors prefer to put /**i**/, /**u**/ into the *coda slot*, viewing them as vocalic terminals (Figure 4). This is not in accord with the general "Western" notion of a coda (which is understood as a postnuclear *consonant*). Such a solution, however, is in accord with the Chinese view of the syllable.

As we are concerned with teaching Mandarin pronunciation (rather than with various alternative phonological analyses of the Mandarin syllable), we are bound to adopt *pīnyīn*, i. e. the *Initial-Final model*. Thus we do not need to concern ourselves with the *Onset-Rime model* in any more detail. The mention of it has had an important purpose, though: we wanted to show that the views of the structure of the Mandarin syllable are not unanimous, that alternative solutions exist and that the commonly accepted initial-final approach reflected in *pīnyīn* is not "God's Truth."

Presentation of the Mandarin Syllable in (Text)books

If we look at the various treatments of the *pǔtōnghuà* sound system (be it of Chinese or Western provenance, be it a textbook or a monograph, be it at an elementary level or at an advanced level, be it a part of a general language textbook or of a publication specializing solely in Mandarin pronunciation), we discover that the majority of these treatments do not explain the structure of the Mandarin syllable at the beginning of a text. After making a general introduction plus presenting the

segmental inventory of *pǔtōnghuà* (not always), the authors usually state that a syllable is composed of an initial, a final and a tone; they then proceed to a survey of particular initials, finals, tones and the drilling of particular syllables. A chapter "Syllable" usually comes only later in the book, if at all. Let us mention a few examples of concrete publications:

- The English textbooks and treatments focused on providing a description of the sound system of *pǔtōnghuà*: Dow 1972 (syllable structure very briefly introduced on p. 98), Huang 1981 (no information), Ma 1999 (no information), Chin 2006 (gives the syllable structure somewhat more attention on p. 65), Lin 2007 (syllable structure introduced on p. 106). One of the exceptions to the rule is a textbook by Speshnev 1972 (in Russian, not in English) already mentioned above: the author sketches out a syllable scheme at the very beginning (on p. 11; note that Speshnev's treatment of finals is based upon the unique system of Dragunov & Dragunova 1955).

- The textbooks and treatments published in Chinese in the P. R. C.: below are examples of the organization of chapters in some of these books. The last example is the only one from many publications known to us which puts the chapter about syllable structure *before* the chapters dealing with the initials and finals.

  ...Initials – Finals – Syllable – Tones...[14]
  ...Initials – Finals – Tones – Syllable...[15]
  ...Initials – Finals – Tones – Modifications of sounds – Syllable...[16]
  ...Vowels – Consonants – Syllable – Initials – Finals – Tones...[17]

The next observation to make is that in the chapter "Syllable" one does not usually find a lucid hierarchical scheme of the syllable. There may be some sort of a structured table,[18] or even just a verbal explanation. To sum up, the Chinese (text)books – quite naturally – do employ the traditional concepts of an initial and a final, yet a student receives only the first two lines of the scheme presented in Figure 2. The complete analysis of the syllable structure regularly comes rather late – after treating and practicing the initials, the finals and whole syllables. Such an ordering has a major disadvantage: if a chapter "Syllable" comes late, a student has no knowledge of the inner composition of a final while learning Mandarin syllabary. The subsyllabic components within a final (i. e. a subfinal, a medial, a main vowel, and a terminal) are not introduced early enough, consequently they cannot be utilized in teaching.

It seems that Chinese authors are seldom aware of the fact that the syllable structure in the mother tongues of foreign students or readers of their books may be rather different from the syllable structure in Mandarin. Being native speakers of Chinese, they are not sensitive to the potential problems that foreign students may struggle with. Yet the differences in syllable structure between students' mother

tongues and the target language – Mandarin may trigger numerous mistakes in pronunciation. These mistakes are often endlessly repeated, as their subsurface causes have a systematical character. We shall now provide an example of a language with a considerably different syllable structure: it is the case of Czech.

A Comparison between Mandarin and Czech Syllable Structures

The structure of the Mandarin syllable reviewed above indicates that the Mandarin syllable is rather different from the syllable of many European languages. Let us make a comparison with the Czech syllable (Czech is the native language of the author of this article). The structure of the Czech syllable is highly complex. It can be portrayed in the following way:
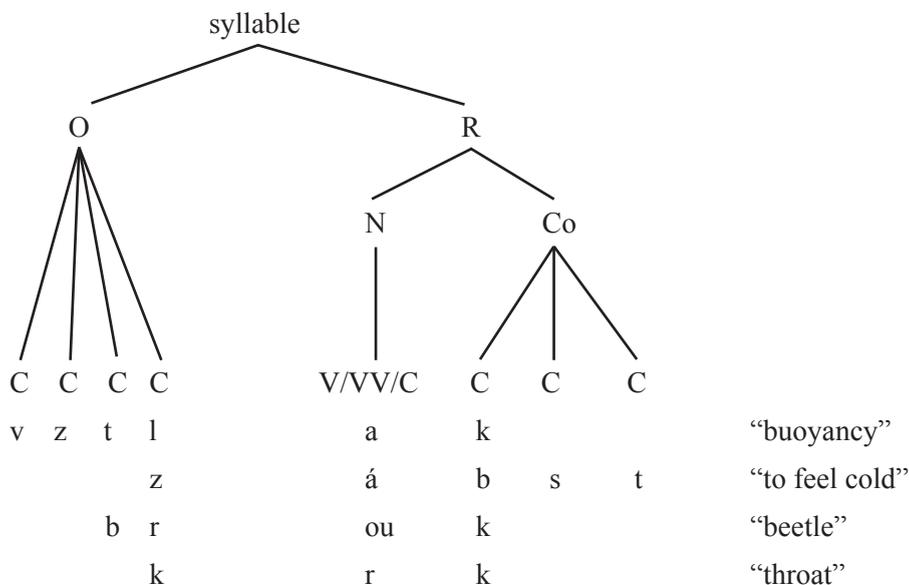
Figure 5. The syllable structure in Czech

An onset can contain up to four consonants, e. g. *vztlak* "buoyancy." A nucleus is mostly a single vowel, e. g. *med* "honey," sometimes a syllabic consonant /r/ or /l/, e. g. *krk* "throat," *vlk* "wolf," quite rarely a diphthong /ou/, e. g. *louka* "meadow" (other diphthongs than /ou/ occur only in loanwords). A coda can contain up to three consonants,[19] e. g. *zábst* "to feel cold." Let us review the major consequences of the discrepancies between the Czech and Mandarin syllable structures for the Czech learners of Mandarin and for the Chinese learners of

Czech. (We shall also mention some other differences between the phonological structures of both languages.)

- If we leave aside the words of foreign origin (e. g. *automatický* "authmatic"), Czech does not have any indigenous diphthongs except for /**ou**/. This diphthong does not contain any glide – both elements have a stable, clear articulation more or less identical to a respective monophthong. As for the glides, Czech students are familiar with the sound [**j**]. However, in Czech /**j**/ has the phonotactic characteristics of a consonant, occurring either in a syllable onset (e. g. in *jíst* "to eat"), or in a syllable coda (e. g. in *můj* "my") – not of a glide functioning as a component of a diphthong. Consequently, Czech students often mispronounce Mandarin syllables containing rising diphthongs or triphthongs. They erroneously use a medial and/or a vocalic terminal to make up the extra peak(s) of sonority, breaking a syllable into two, or even three syllables (pronouncing *xie* as [**ɕi. je**], *xiao* as *[**ɕja.ʊ**] or *[**ɕi.ja.ʊ**] etc.). The falling diphthongs are also difficult for them: they tend to overpronounce the terminal vowels /*i*/, /*u*/, wrongly attempting to reach the articulatory targets of the high vowels [**i**], [**u**]. They may even pronounce the terminal /*i*/ as an approximant [**j**], pronouncing *mai* as *[**maj**] etc. (for a correct pronunciation of the falling diphthongs, see below). Note that even the native speakers of languages which, unlike Czech, abound in diphthongs, might still not pronounce the Mandarin diphthongs well. This occurs, even though some of these diphthongs may be phonologically identical to those of Mandarin (e. g. /ai/ is found both in Mandarin and in English). This is because the articulatory and acoustic properties of diphthongs vary among languages – e. g. the articulatory targets may be slightly different, the transition between them may have different forms, a temporal structure of a diphthong may be different etc.[20]

- Although the most common syllable type in Czech (as well as in other languages)[21] is "the optimal syllable" CV, the syllables containing consonant clusters are nevertheless very common there. Consonant clusters can be found both in an onset (*vztlak*) and in a coda (*zábst*). This feature of the Czech syllable obviously does not lead to any problems for the Czechs learning Chinese, but often causes troubles for the Chinese learning Czech: they tend to make up a separate syllable for each consonant involved, as is regularly done with the words/names of foreign origin occurring in the Chinese context (e. g. the monosyllabic name **Kl**au**s** is turned into a trisyllabic form *Ke.lao.si*).

- Czech has two rather common syllabic consonants /**r**/, /**l**/ (*krk* "throat," *vlk* "wolf"). Again, this fact does not pose a problem for Czech learners of Chinese. Indeed, exactly the opposite occurs – the familiarity with the notion of syllabic consonants can be advantageous as the grounds for explaining the pronunciation of the Mandarin syllables *zi, ci, si, zhi, chi, shi, ri*: according to many authors,

the nucleus of these seven syllables is formed by a syllabic consonant (the other alternative, accepted by *pīnyīn*, is to view their nucleus as an "apical vowel"). Yet Chinese learners of Czech may have trouble with the syllables containing syllabic vowels.

- A velar nasal /ŋ/ is not present in the inventory of Czech consonantal phonemes. A sound [ŋ] does occur in Czech, but only as a product of assimilation: it is a positional variant of /**n**/ followed by a velar stop /**k**/ or /**g**/ within a word (as in the word *lanko* /lanko/ "cord": it is pronounced as [laŋko]). Consequently, Czech speakers often wrongly insert a velar stop [**k**] after [ŋ] – both in Mandarin syllables such as *bing*, mispronouncing it as *[piŋk] (the correct form is [piŋ˥]), and in English words such as *song*, mispronouncing it as *[soŋk] (the correct form is [sɒŋ˥]).

- Unlike English or Mandarin, which are considered to be stress-timed languages, Czech is a typical syllable-timed language. This means that, at least in standard pronunciation, the stressed syllables do not undergo any considerable lengthening, while the unstressed syllables do not undergo any considerable shortening, or reduction of the segments involved. Due to this fact, Czech students often fail to pronounce the unstressed Mandarin syllables properly, i. e. in a reduced way. They wrongly assume that articulating them in a "slurred" way fails to meet the standards of correct pronunciation.

- Long and short vowels are distinctive in Czech, e. g. *dal* "he gave," *dál* "further". This often hinders Czech students from the due lengthening of stressed syllables (as is a feature of stress-timed languages). On the other hand, Chinese students may not be sensitive to the differences in vowel quantity while learning Czech.

- In Czech, /**n**/ in a coda position (*pán* "gentleman") ought to be pronounced with a complete closure in the alveolar region (i. e. the airstream passing through the nasal and oral cavities is obstructed by the contact of the tongue at the alveolar ridge); after a closure an audible release may follow prepausally. In Mandarin, a terminal /**n**/ has different properties: the contact of the tongue is frequently missing in the spontaneous speech of native speakers (see below). Czech students may thus have trouble with Mandarin syllables such as *min, huan* etc.: they may consistently strive to reach the articulatory target for [**n**], which may sound unnatural and hypercorrect. Furthermore, students sometimes pronounce such Mandarin syllables with an audible release, which is definitely wrong.

- In Czech assimilations to a neighboring segment are typical for consonants, e. g. *vstup* "the entrance" is pronounced as [**fstup**]. Such assimilations may cause problems for Chinese learners of Czech (*[**vstup**]). On the other hand, vowel quality is rather stable in Czech, and not much influenced by the neighboring

segments. In Mandarin the situation is basically the opposite: it is mainly the vowels which undergo assimilations, e. g. *ban* is pronounced as [pan], while *bang* is pronounced as [paŋ] (see below). Czech students frequently ignore these assimilations, being insensitive to vowel quality changes.

Of course, the concrete problems that native speakers of various languages may struggle with vary. For example, the difficulties experienced by Czech students will certainly be different from the difficulties faced by Japanese, Thai, American or Russian students. Some of the problems may be shared by all learners, however.

Components at the Level of Segments – C, G, V, X

As shown above, four structural components can be established at the level of segments in the Mandarin syllable. We marked them as C, G, V, X. Their specific role, or function within a syllable is defined by their position within a syllable template. For simplicity's sake we assume that a particular position implies a particular function and use both terms (i. e. function and position) alternatively. Each position has its own inventory of segments – see the table below (the phonemes are in *pīnyīn*); there may be no more than one segment inserted in a particular position.

position 1	position 2	position 3	position 4
C	G	V	X
/b/, /p/, /m/, /f/, /d/, /t/, /n/, /l/…	/i/, /u/, /ü/	/a/, /o/, /e/, /i/, /u/, /ü/	/i/, /u/, /n/, /ng/

Figure 6. A Mandarin syllable template and the segmental inventories for each position

Note that these particular inventories are not strictly characterized by either vowel or consonant status: although only consonants are allowed in position 1 (C), and only vowels are allowed in positions 2 (G) and 3 (V), the segment in position 4 (X) may be either a consonant, or a vowel. This makes the notion of "X" different from the notion of a "coda," as mentioned above.

Phonotactic Rules

Phonotactic rules define the permissible sequences of speech sounds in a language. Let us examine the situation in Mandarin. There are phonotactic rules at several levels operating within the Mandarin syllable (some phonologists may view them as constraints of various ranks – cf. the so called constraint-based approaches).

We have established four positions C, G, V, X, followed by the inventories of segments for each of them. One position cannot accept more than one segment. The only obligatory item in a syllable is V. These premises outline the basic phonotactics of Mandarin syllable. Hence, we can set up the following set of *basic phonotactic rules*:

– the segments may enter only such combinations within a syllable which respect the inventories of segments for particular positions C, G, V, X

– one position may be occupied at most by one segment

– V is an obligatory item, C, G, X are optional items

This set of basic rules predetermines a theoretically possible inventory of Mandarin syllables. If we apply these rules, it is clear, for example, that no consonant clusters may occur within a Mandarin syllable, since there may be only one segment inserted in positions C and X, and there is always at least one vowel between them (syllables such as */*bla*/, */*bn*/ do not exist). Similarly, it is clear that there are no syllables starting with /*ng*/ (such as */*nga*/, */*ngai*/ etc.), although the consonant /*ng*/ is found in the inventory of Mandarin consonants; the reason for the absence of such syllables is that /*ng*/ is not allowed as C. Or, it is clear that syllables such as */*baen*/, */*baei*/ do not exist, since /*a*/ is not allowed as G (note that /*a*/ in *bai, ban* etc. is V, not G). Or, we can determine that the diphthongs */*aü*/, */*eü*/ or a triphthong */*iaü*/ do not exist, since /*ü*/ is not allowed as X. Or, we can state that except for the syllables ending with a nasal, there are no closed syllables, since the consonants /*b*/, /*p*/ etc. are allowed only as C, and so on.

Applying this set of basic rules is only the first step. *Additional rules* must be applied in order to arrive at the existing Mandarin syllabary:

Another set of phonotactic rules operates at a higher level, prohibiting combinations of segments with particular features within a final. First, a neighborhood of high vowels is prohibited (that is, a high main vowel cannot combine with a medial, nor with a vocalic terminal). This rule eliminates, as ill-formed, items such as the finals */*iui*/, */*üu*/, */*ui*/ (remember that *-ui* is phonologically /*uei*/, not */*ui*/), or */*iu*/ (remember that *-iu* is phonologically /*iou*/, not */*iu*/). Second, the marginal elements of a triphthong must differ in both features which constitute a contrast between two high vowels (i. e. in the features [±*back*], [±*round*]). For example, a triphthong */*üai*/ is eliminated as ill-formed, as its marginal elements are both [−*back*]. Or a triphthong */*üau*/ is eliminated as ill-formed, as its marginal elements are both [+*round*].

The last set of phonotactic rules operates at yet a higher level, determining the permissible combinations of initials and finals. The constraints are largely dependent on the place of articulation of an initial, and on the category a final belongs to from the four traditional categories of finals, *sì hū* 四呼. Let us remind ourselves that

the major criterion for placing a final into a particular *hū* category is the beginning element of a final: it may be either a high front unrounded vowel /*i*/ – *qíchǐ hū* 齐齿呼 category, or a back rounded vowel /*u*/ – *hékǒu hū* 合口呼 category, or a high front rounded vowel /*ü*/ – *cuōkǒu hū* 撮口呼 category, or none of them – *kāikǒu hū* 开口呼 category. Let us give some examples of the constraints: the velar initials *g*, *k*, *h* cannot combine with the finals of *qíchǐ hū* and *cuōkǒu hū* categories: */*gian*/, */*hün*/ etc. are thus eliminated as ill-formed. The palatal initials *j*, *q*, *x*, on the other hand, cannot combine with the finals of *kāikǒu hū* and *hékǒu hū* categories: */*qan*/, */*xung*/ etc. are thus eliminated as ill-formed.

All the gaps in the Mandarin syllable inventory mentioned above (e. g. */*nga*/, */*müu*/, */*gian*/) must be viewed as systematic gaps, since they are eliminated as part of a rule. Furthermore, there are accidental gaps such as the finals */*üang*/, */*ien*/, the syllables */*fai*/, */*len*/ etc. These forms do not break any rule. The reason why they are missing in the inventory has not yet been discovered, hence their absence has to be viewed as accidental. The phonotactic rules, as well as the systematic and accidental gaps, are discussed in Třísková 2010.[22] Note that Duanmu's phonological analysis of the phonotactics of the Mandarin syllable,[23] which is principally different from our approach, leads him to the following conclusion: "A puzzling fact about Standard Chinese is that the majority of expected syllables are missing." Our analysis,[24] though, which strictly respects specific inventories of segments for each of the four positions, does not by any means draw the same conclusion.

The Phonetic Features of C, G, V, X: Various Phonemes in the Same Position

Each of the components C, G, V, X displays certain phonetic features shared by all segments occurring in this particular position. In other words, all C components share certain phonetic features, and the same holds for all G components, all V components and all X components (regardless of the differences in their consonantal or vocalic status). We shall now illustrate that two different phonemes may indeed share certain phonetic properties, if they occur in the same position. Let us examine two front terminals: a vocalic terminal /*i*/, as in *kai*, and a consonantal nasal terminal /*n*/, as in *kan*. They have quite a lot in common, regardless of the fact that phonologically /*i*/ is a vowel while /*n*/ is a consonant:

– A vocalic terminal /*i*/ (as in *kāi* 开 "to open") is realized in the following way: it is rather lax, not reaching the articulatory target of a high vowel [i]: it is more or less centralized, hence it should be transcribed as [ɪ], or [ɪ̈]. In some situations it may even have the character of [e] (e. g. *kai* is sometimes pronounced not as [kʰai], but as [kʰae]). It can even disappear completely in fast casual speech if a syllable is unstressed: e. g. a diphthong /*ai*/ in *dǎkāi* 打开 can be realized as a monophthong. The syllable as a whole can be lengthened if it bears stress, /*i*/ receiving a due amount of time.

– A nasal terminal /*n*/ [**n**] (as in *kān* 刊 "publication") is also realized in a rather lax way; a closure is very often incomplete in spontaneous speech, i. e. the tip of the tongue does not touch the alveolar ridge.[25] A terminal /*n*/ may even disappear, surfacing as a mere nasalization of the preceding vowel: *an* → [**ã**]. For instance *kān* 刊 may be pronounced as [**kʰã**]. Even if a closure happens to be complete, it is not followed by an explosion. Apparently, the consonantal character of /*n*/ in the terminal position has been severely eroded.[26] Note that a nasal consonant [**n**], as well as a nasal consonant [**ŋ**] belong to the same class of sounds as vowels: it is a class of sonorants.[27]

To sum up, the terminals /*i*/, /*n*/ are both lax, their articulation often does not reach the target, and both may possibly disappear completely in connected spontaneous speech. They are both front, thus trigger analogous assimilations of the preceding main vowel (/*ai*/ → /*[ai]*/, /*an*/ → [**an**]). In both cases, the syllable as a whole can be lengthened if it bears stress, the terminal element receiving a due amount of time. The properties of a terminal /*n*/ are rather close to the properties of vowels. The shared features of the terminals /*i*/ and /*n*/ described above arise from the fact that they are both front, and share the same function in a syllable.

The Phonetic Features of C, G, V, X: The Same Phoneme in Various Positions

There are a few phonemes which can occur in more than one position (the underlined items in Figure 6), namely three high vowels and one nasal consonant: /*i*/, /*u*/, /*ü*/, /*n*/. The vowel /*i*/ can function as G or V or X, the vowel /*u*/ can function as G or V or X, the vowel /*ü*/ can function as G or V, and the nasal consonant /*n*/ can function as C or X. If we compare the realizations of one particular phoneme in various positions, we can observe that it surfaces as rather different sounds. In other words, each of the phonemes /*i*/, /*u*/, /*ü*/, /*n*/ has several distinct positional variants. Let us give an example. We shall compare /*i*/ functioning as a main vowel (V), as a medial (G), and as a terminal (X):

– /*i*/ functioning as a main vowel (as in *mì* 秘 "secret", or *mǐn* 敏 "agile"). It tends to be fully articulated, reaching the articulatory target of a high vowel [**i**]. If not followed by a nasal ending, it may be lengthened (*mi* [**miː**]); the lengthening regularly happens in stressed syllables. (For the sake of simplicity, we are leaving aside the cases of *zi*, *ci*, *si*, *zhi*, *chi*, *shi*, *ri* where /*i*/ is realized as the "apical vowel.")

– /*i*/ functioning as a medial, i. e. as the first part of a rising diphthong, such as /*ie*/ (as in *miè* 灭 "to wipe out"). The articulation is rather tense; it reaches the position of a high vowel [**i**] with a very narrow stricture, being possibly accompanied by a little friction. The tongue immediately leaves the position, making the medial sound very short. As a whole, the sound has the character of an approximant [**j**] (thus *mie* is pronounced as [**mje**], etc.). Some authors even view a medial as a

mere secondary articulation to an initial consonant: e. g. *mie* → **[mʲe]**, not **[mje]**, *suan* → **[sʷan]**, not **[swan]**.[28]

– /*i*/ functioning as a terminal, i. e. as the second part of a falling diphthong, such as /*ai*/ (as in *kāi* 开 "to open"). Its articulation was described above: it is rather lax, more or less centralized (it does not reach the articulatory target of **[i]**, nor it is pronounced as **[j]**); it can even disappear.

It is obvious that, according to its function within the final, the phoneme /*i*/ may surface as several rather different sounds.

The Phonetic Features of G, V, X – a Summary

We are now in a position to summarize the general phonetic features of G, V, X. We will leave aside the C component for the moment, since – if we neglect the possible influence of a following high vowel – the initial consonants basically have no positional variants, being allowed only in the onset of a syllable (the only exception is /*n*/ allowed also in the coda); their phonetic features are those of a consonant occurring in an onset. In other words, we can focus our attention only on the components which constitute a final, that is G, V, X. Before we proceed further, let us posit six structural types of Mandarin finals: V (-*a*), VV (-*ai*), VC (-*an*), GV (-*ia*), GVV (-*uai*), GVC (-*uan*).

The phonetic features of the medials (G), the main vowels (V), and the terminals (X) can be summed up as follows:

• The medials /*i*/, /*u*/, /*ü*/ have a tense articulation, reaching the articulatory target of the high vowels **[i]**, **[u]**, **[y]**; the tongue leaves the position immediately, making the sound very short, possibly with a little friction involved. As a whole, the medials have the character of the approximants transcribed as **[j]**, **[w]**, **[ɥ]** respectively (note that the sounds called approximants are considered to be consonants, not vowels). Another option is to transcribe the medials as the "semivowels" **[i̯]**, **[u̯]**, **[y̯]**. We prefer the former solution, i. e. **[j]**, **[w]**, **[ɥ]**. For example, the syllable *xie* we transcribe as **[ɕje]**, not **[ɕi̯e]**. We believe that such a transcription may help students to avoid a common mistake: breaking a syllable containing a rising diphthong or a triphthong into two or even three syllables (*xie* → *[ɕi.je]). Note that regardless of the fact that the medials are transcribed as approximants (e. g. *mie* **[mje]**, *duo* **[twɔ]**, *xue* **[ɕɥe]**) which belong to the phonetic consonants, phonologically we view them as vowels. That means we accept rising diphthongs /*ie*/, /*uo*/, /*üe*/ etc., and view them as consisting of two vocalic phonemes. The analogical situation exists with triphthongs.

• The main vowels /*a*/, /*o*/, /*e*/, /*i*/, /*u*/, /*ü*/ are usually pronounced as full vowels, reaching a pertinent articulatory target. If not followed by a terminal, they may

be lengthened in the stressed syllables, e. g. *ma* → **[maː]**. (We are leaving aside the influences of a loss of stress and tone. Both of them may suppress the main vowel in some situations – the main vowel may be reduced or even fully omitted: e. g. *xiū* /*xiou*/[1] can be realized as **[ɕju¹]**).

- The terminals /*i*/, /*u*/ (X = V), /*n*/, /*ng*/ (X = C) have a lax, rather unstable and variable articulation, often not reaching the articulatory target; they may disappear in pronunciation completely. The vocalic terminals /*i*/, /*u*/ are mostly realized as centralized **[ɪ]**, **[ʊ]**, or even as more open **[e]**, **[o]** (*mai* → **[maɪ]**, **[mae]**, *mao* → **[maʊ]**, **[mao]**). They may even disappear in fast casual speech, a falling diphthong being monophthongized. A nasal terminal /*n*/ **[n]** often has an incomplete closure. If the closure happens to be complete, it is not followed by an explosion (this can be expressed by a small diacritic: **[n̚]**). A nasal terminal /*ng*/ **[ŋ]** must not be followed by an explosion or by a **[k]** sound (e. g. pronouncing *bang* as ***[pɑŋk]** is wrong). Both nasal terminals /*n*/, /*ng*/ may disappear in casual speech, being realized only as the nasalization of a preceding vowel, e. g. *ban* → **[pã]**, *bang* → **[pã]**. All four terminal segments share certain features, although two of them are vowels (/*i*/, /*u*/) and two of them are consonants (/*n*/, /*ng*/): the consonantal character of /*n*/, /*ng*/ is weak.

The systematic linkage between the function of a segment within a syllable structure and some of its articulatory properties is not specific solely to Mandarin – it can be found across other languages. For instance, a consonant in a syllable coda is generally more sensitive to articulatory weakening than a consonant in a syllable onset. Another example: vowel-like (vocoid) segments occurring *before* a syllable nucleus (i. e. as the first component of a rising diphthong) often tend to have the character of an approximant, e. g. Spanish *puerta* "door" → **[pwerta]**. On the other hand, vowel-like (vocoid) segments occurring *after* a syllable nucleus (i. e. as the second component of a falling diphthong) tend to keep their vocoid character, that is to keep the properties of a vowel, e. g. Spanish *auto* "car" → **[au̯to]**.[29] Thus, we cannot claim that we are dealing with a Mandarin specialty here. Yet, in Mandarin this phenomenon is rather striking, transparent and well describable. This is due to the simplicity of the Mandarin syllable structure, the restricted inventories of segments allowed in particular positions (a mere four phonemes being allowed in more than one position) and clear phonotactic rules. Furthermore, this phenomenon runs literally through the whole inventory of Mandarin syllables and, therefore, it is definitely worth our attention.

Pīnyīn and the Traps Linked to its Orthography

The Chinese character script, by its very nature, cannot serve as a guide to pronunciation. This is the task of the *pīnyīn* spelling system (the official name is *Hànyǔ pīnyīn fāng'àn* 汉语拼音方案). It was created for the notation of the sound form of standard Mandarin – or, strictly speaking, initially just for the notation of the standard readings of characters (*pīnyīn* was approved back in 1958). *Pīnyīn*, after considering various other solutions, was eventually conceived as a romanization system: it uses the letters of the Latin alphabet. This is undoubtedly a great advantage for a Western learner of Mandarin. Note that in the process of designing a new phonetic alphabet after 1949, it was suggested that the Cyrilic or even the newly designed graphemes inspired by the strokes of Chinese characters could be used – cf. the *Zhùyīn zìmǔ* system, which was conceived in 1913 (this system was not purely alphabetic, though; see below).

In an alphabetic script such as the Latin script, ideally one letter should represent one speech sound – either a vowel, or a consonant. In fact, there need not be (and hardly ever is) a consistent one-to-one correspondence between the graphemes of a script and the phonemes of the language, or between the graphemes of a script and the particular sounds of the language. The degree to which the written form of a language reflects its phonological system or its actual sounds depends on the orthographic rules of the language in question. Since the sound systems of languages which use the Latin alphabet differ, the concrete ways in which these languages employ Latin letters are wide-ranging (including the use of modifications such as various digraphs, diacritical marks etc.). The divergences among the orthographies of various languages regularly cause trouble for the students of foreign languages, especially at the beginning of their studies: adult learners of a foreign language are to a large extent at the mercy of its written form. *Pīnyīn* orthography, having been designed by linguists on a green field, is fortunately much simpler than, for example, English orthography, which, for various historical reasons, is very complex and often inconsistent. Even so, *pīnyīn* orthography has its own pitfalls.

When students begin to study a foreign language, they tend to make intuitive assumptions about the orthography of the studied language, which sometimes turn out to be right and sometimes turn out to be wrong. These assumptions may be of the following sort: that a particular letter represents the same or similar sound as in their native language, that one letter regularly represents one sound, that what is written by using the same letter sounds more or less similar, that all letters are "equally important," that each letter has to be pronounced, that each sound they can hear is represented in the script by a separate letter, etc. Let us give some contrary examples from the English and *pīnyīn* orthography:

– sometimes one letter represents two sounds: Engl. *six* [sɪks], *pīnyīn bo* [pwɔ]
– sometimes two letters represent one sound: Engl. ***that*** [ðæt], *pīnyīn shu* [ʂuː]

– sometimes a letter is not pronounced as a separate sound: Engl. *know* [nəʊ], *pīnyīn ri* [ʈʂ]

– sometimes one letter represents several rather different sounds: Engl. *all* ([ɔː]), *bat* ([æ]), *ate* ([eɪ]), *arrive* ([ə]), *pīnyīn mi* ([iː]), *mie* ([j]), *mai* ([ɪ]), *si* ([ɻ̩ː]), *shi* ([ʐ̩ː]), or *pīnyīn mu* ([u]), *xu* ([y])

– a letter often represents rather different sounds in both languages: Engl. *class* ([k]), *century* ([s]), *pīnyīn cao* [tsʰ], or Engl. *buy* ([b]), *pīnyīn bang* ([p])

– some letters may be "less important" than others, that is they represent a sound which is not acoustically / perceptionally prominent, or may be even absent in pronunciation; cf. the weak forms of English function words such as *for, do, are, would, has* (e. g. the English *and* may be pronounced as [ən] or even [n̩] if unstressed: *fish and chips*), or the monophthongization of Mandarin falling diphthongs in fast speech (e. g. the unstressed preposition *gěi* 给 "to" may sound as [ke]), or the devoicing of Mandarin high vowels following the fricatives in unstressed syllables (e. g. *dòufu* 豆腐 "bean curd" may sound as [toʊfu̥], or even as [toʊf]).

– some sounds are not represented in the script at all: Engl. *cure* [kjʊə], *pīnyīn gui* [kweɪ], etc.

Let us focus on *pīnyīn* orthography. For some of the problematic cases there is no other solution than to learn the orthographic rules mechanically, as they are more or less arbitrary. For example, omitting /*e*/ in a final /*uei*/ (e. g. *gui* is phonologically /*guei*/, not */*gui*/), or omitting /*o*/ in a final /*iou*/ (e. g. *xiu* is /*xiou*/, not */*xiu*/), or writing a main vowel /*u*/ with a letter "**o**" in some cases (e. g. *dong* is /*dung*/, not */*dong*/), or writing a terminal /*u*/ with a letter "**o**" (e. g. *gao* is /*gau*/, not */*gao*/). These rules probably have a practical motivation, such as saving one letter, or enhancing visual clarity. Another rule established for the sake of economy is omitting an umlaut on "**ü**" after the palatals /*j, q, x*/ (*xu* is phonologically /*xü*/, thus it is pronounced as [ɕyː], not as *[ɕuː]). This rule utilizes a complementary distribution: /*u*/ is forbidden after /*j, q, x*/, thus a confusion of /*u*/ and /*ü*/ cannot happen. Some other tricky orthographic rules are also phonologically motivated. The "apical vowels" [ɻ̩], [ʐ̩] may occur only after the apical sibilants /*z, c, s, zh, ch, sh, r*/, while a high front vowel [i] and an approximant [j] are allowed only after /*b, p, m, d, t, n, l, j, q, x*/. The sound of the "apical vowels" is rather different from [i]. Nevertheless, the authors of *pīnyīn*, taking into account the common historical origin and the fact of complementary distribution, eventually decided to interpret the "apical vowels" as positional variants of /*i*/ (not as a separate phoneme). Both the "apical vowels" and [i] / [j] are thus written with a letter "*i*." For instance, *mi* is pronounced as [miː], while *si* is unexpectedly pronounced not as *[siː], but as

[ʂʐ̩ː], etc. This rule has to be learned mechanically, too, although explaining its underlying logic to a student is undoubtedly advantageous (the mispronunciation of the syllables *zi, ci, si, zhi, chi, shi, ri* is a very common mistake in beginners). Another pitfall is the graphical solution of three phonological pairs of unvoiced stops that contrast in aspiration (/*p-b*/, /*t-d*/, /*k-g*/). To avoid a diacritical mark, *pīnyīn* writes the aspirated stops without indicating aspiration: /*p, t, k*/, while the unaspirated stops are written with the symbols commonly used for the voiced stops: /*b, d, g*/. Beginners frequently misread /*p, t, k*/ as the unaspirated [p], [t], [k] (this mistake is rather serious, as aspiration is a distinctive feature in Mandarin: e.g. *dùzi* 肚子 "belly" vs. *tùzi* 兔子 "rabbit"). A student should undoubtedly know the underlying phonology of the consonant system, yet in order to eradicate the described mistakes, mechanical drills are the only effective remedy.

What *Pīnyīn* Does Not Tell Us

An inherent property of an alphabetic script such as the Latin script is that it presents the words of a language as "beads on a string." Particular "beads" are represented either by consonant letters (let us mark them by a circle ○), or by vowel letters (let us mark them by a black dot ●). For instance:

$$\text{/d\textcaron{v}g/} = \circ{-}\bullet{-}\circ, \quad \text{/fr\textscripta g/} = \circ{-}\circ{-}\bullet{-}\circ$$

As far as Mandarin word, or rather the Mandarin syllable is concerned, the situation is analogous: in standard *pīnyīn* notation, a syllable looks like a linear sequence of vowels and consonants, arranged according to certain phonotactic rules. For instance:

$$\text{/\textbf{kuan}/} = \circ{-}\bullet{-}\bullet{-}\circ, \quad \text{/\textbf{kuai}/} = \circ{-}\bullet{-}\bullet{-}\bullet$$

A "beads-on-a-string" image of the Mandarin syllable, although providing far more phonetic information than a Chinese character (*kuān* vs. 宽, *kuài* vs. 快), is of course only a very rough image of how a syllable actually sounds. *Pīnyīn* notation captures in a more or less accurate way the phonemes constituting a syllable, *not* their particular allophones (the same applies in principle to any Latin orthography). The allophones must be "deciphered" by the reader (of course, they can be written down in IPA if necessary).

The realizations of phonemes in connected speech are influenced by a large number of linguistic factors operating at various levels of the language (not to mention the paralinguistic factors). The examples include: the neighboring segments, the position of a word within an utterance, stress, speech tempo, style etc. In this article we are concerned, in particular, with the following factor: the function of a segment within a syllable, decided by its position within a syllable template. We have observed that there are four phonemes allowed in more than one position: /*i*/,

/*u*/, /*ü*/, and /*n*/, and that the realization of these phonemes varies according to their position within a syllable. Yet beginners tend to ignore these variations, sticking to the pronunciation suggested by a *pīnyīn* letter. For instance, they vaguely assume the letter "**i**" (e. g. in *mi, mie, mai*) should be always pronounced in the same way – namely as **[i]**. In fact the phoneme /*i*/, written as a letter "*i*," is realized variously as **[i]**/**[iː]**/**[ɻ̩ː]**/**[ʐ̩ː]** (if functioning as a main vowel), or as **[j]** (if functioning as a medial), or as **[ɪ]** (if functioning as a terminal). Indiscriminate adherence to the **[i]** pronunciation may lead to the serious error mentioned above: the emergence of one more syllabic peak (or even two more peaks). One can overhear *mie* pronounced as **[mi.je]*, *mai* as **[ma.ji]*, etc. Another mistake is realizing /*n*/ in the terminal position with a meticulously performed closure and an audible release. Fortunately, the information about the function of a segment (and thus about the proper choice of a particular allophone) is retrievable by the reader: it is encoded in the location of the letter within a linear sequence of items which constitute a *pīnyīn* syllable (e. g. in *kuan*, "*k*" is in the first position, "*u*" in the second, "*a*" in the third, and "*n*" in the fourth). Yet to decode the information correctly a reader must have a preliminary knowledge of syllable structure (e. g. in *kan*, "*k*" is in the first position, "*a*" is actually in the third position, as the second position is unoccupied, and "*n*" is in the fourth position), and of the *pīnyīn* orthography ("*i*" in *kui* is in the fourth position, as /*e*/ is omitted).

Furthermore, although the image of a syllable is linear in script, we have already pointed out that each particular segment is entangled within a net of hierarchical relationships amongst units of several different levels. These relationships determine the closeness of the connection between the neighboring segments (e. g. in *kuan*, the relationship between "*a*" and the following "*n*" is tighter than that between "*a*" and the preceding "*u*," as the relationship between a main vowel and a terminal is, as a rule, tighter than the relationship between a medial and a main vowel; this setting results in assimilations within a subfinal, while the assimilations between a medial and a main vowel basically do not occur).

Naturally, a standard script cannot reflect such features. It follows that students must not expect that *pīnyīn* notation as such provides sufficient guidelines for pronouncing the Mandarin syllables correctly. Learning the *pīnyīn* notation of syllables is just the first step, which has to be followed by a diligent training of the correct phonetic shapes of syllables. Certain methodological instruments can be found which may facilitate the students' task. These instruments are, in our opinion, a preliminary explanation of the syllable structure and the employment of the traditional concepts of an initial, (*shēngmǔ*), a medial (*yùntóu*), a main vowel (*yùnfù*), and a terminal (*yùnwěi*).

The Options for Phonological Interpretation

While attempting to record the sound forms of languages in ancient times, Western culture arrived at analyzing speech into segments; the Westerners developed the notions of *vowels and consonants* (old Indian phonology had been familiar with them long before this time). These notions are reflected in various types of phonetic script of an alphabetic kind (such as the Latin script, or Cyrilic). Vowels and consonants form an inherent part of Western phonological analysis; they appear as the basic phonological components, the vocalic and consonantal subsystems together establishing a segmental inventory of a language. A standard analysis of the sound systems of languages such as English, French, Czech etc. starts from their segmental inventory.

Pīnyīn accepts the Western analytical approach to the sounds of speech: it works with the vowel and consonant phonemes, using the letters of the Latin alphabet. The *pīnyīn* inventory of Mandarin segments can be posited in the following way:

– the vocalic subsystem: /*a*/, /*o*/, /*e*/, /*i*/, /*u*/, /*ü*/
– the consonantal subsystem: /*b*/, /*p*/, /*m*/, /*f*/, /*d*/, /*t*/, /*n*/, /*ng*/, /*l*/... etc.

The speakers of languages which use the Latin script absorb the notions of vowels and consonants while learning to read and write. They work with them quite intuitively thereafter. For instance, a Western learner of Mandarin, when coming across the syllable *kuan*, recognizes a consonantal or vocalic affiliation of a particular segment without much hesitation, identifying the letters *u*, *a* as vowels and the letters *k*, *n* as consonants. This does not hold for native speakers of Chinese, however. The notion of vowels and consonants is by no means something familiar to them. The Chinese character script, as well as the centuries old Chinese phonological tradition are based on rather different principles. Old Chinese phonology did not analyze a syllable into discrete segments. It did not develop a notion of vowels and consonants, instead establishing rather different concepts. In China the phonological ponderings were closely related to lexicology, not to the sounds as such. Kratochvil writes:

> "...scholarly interest in language sounds in China has never been aroused simply because the sounds were there, but because they became relevant for some or other endeavour connected with the script. Sounds have thus never been considered in their own right, as raw bits of nature, so to speak, but only as abstract correlates to the units of the writing system."[30]

In other words, the basic unit of phonological interest in China was the syllable, viewed as the reading of a particular Chinese character able to rhyme with certain other syllables under specific conditions. A syllable was bisected into an initial consonant, and the rest of the syllable – the final (the *fǎnqiè* 反切 method). A specific

object of interest was the subfinal, *yùn* (because of rhyming rules). The last stage of this tradition is reflected in the phonetic system *Zhùyīn zìmŭ* 注音字母 (created in 1913, officially approved in 1918). It moved one step further, trisecting a syllable into an initial, a medial, and a subfinal. A subfinal was *not* analyzed into segments, being always represented by a single symbol (for instance the syllable *lan* was written as ㄌㄢ, i. e. l-an, while *luan* was written as ㄌㄨㄢ, i. e. l-u-an).

Some modern Chinese phonologists claim that the traditional analysis is more suitable for Chinese than the analysis into vowel and consonant segments. The article by You Rujie et al. 1980 suggests the following phonological elements along traditional lines:

– *shēngwèi* 声位   which can be translated as "the initialemes"
– *yùnwèi* 韵位   which can be translated as "the finalemes"
– *diàowèi* 调位   which can be translated as "the tonemes"

The unit of *yùnwèi* may be represented by two kinds of sound: one kind is *jìng yīnsù* 静音素, or "static speech sounds" (such as **[a]**, **[i]** etc.), another kind is *dòng yīnsù* 动音素, or "kinetic speech sounds" (such as **[aⁱ]**, **[aⁿ]** etc.). Let us point out again that both "static" and "kinetic speech sounds" always stand as a single phonological unit, a single *yùnwèi* (e. g. **[aⁱ]** is not analyzed as two units). Each of the three "*wèi*" categories has its own inventory of permitted items.

At this point we need to mention the approach to phonological analysis expounded by the British linguist, J. R. Firth (†1960). It lies outside the mainstream of Western phonology, being remarkably akin to the lines of Chinese phonological thought. It is the *"polysystemic" approach*. Linguist David Crystal characterized it in the following way:

> "An approach to linguistic analysis proposed by J. R. Firth in which different linguistic systems are set up at different places in structure, no attempt being made to identify the systems with each other."[31]

In other words, the system of contrasts at one point in a syllabic (or other) structure is established independently of the system of contrasts at other points in this structure. For example, in a CVC structure, one system is operating at the first 'C' position, while another system is operating at the second 'C' position. Sampson points out the benefits of Firth's analysis in relation to Chinese. Let us quote several lines, as they neatly make the point we are aiming at:

> "…in Mandarin Chinese the only consonants which can occur at the end of a syllable are **[n]**, **[ŋ]**, of which the latter is not among the many consonants which can occur syllable-initially. A phonemicist would presumably want to treat **[ŋ]** as an allophone of one of the initial consonants, but which? – **[k]**? **[m]**? – while Firth simply recognizes a two-member syllable-final system which is very different from the multi-member

syllable-initial system. Firth argues, correctly in my view, that phonemicists are led into error by the nature of European writing systems. A phonemic transcription, after all, represents a fully consistent application of the particular principles of orthography on which European alphabetic scripts happen to be more or less accurately based. It is natural that scholars working with Oriental cultures, many of which had scripts based on other principles and whose traditions of philological discourse were independent of European thought, should be skeptical about elevating their own tribal speech-notation system into an axiom of science. Certainly the Chinese, who had a very old-established vocabulary for discussing the pronunciations of words, would never have described, say, the syllable that we would transcribe **[nan]**, as consisting of a sequence of three segments in which the first and third are identical."[32]

Drawing both on the traditions of old Chinese phonology and on the Western approach, while taking inspiration from the "polysystemic" approach, we could possibly think of a new, combined analysis.[33] In the first step, the traditional elements *shēngmǔ* (an initial)*, yùntóu* (a medial)*, yùnfù* (a main vowel), and *yùnwěi* (a terminal) are accepted as the basis for establishing four independent systems of phonological components: "initialemes," "medialemes," "centralemes" and "terminalemes." Each system has its own inventory of segments:

– 声母位 *shēngmǔ wèi*   "the initialemes"   **/b/**, **/p/**, **/m/**, **/f/**, **/d/**, **/t/**, **/n^I/**, **/l/**... etc.

– 韵头位 *yùntóu wèi*   "the medialemes"   **/i^M/**, **/u^M/**, **/ü^M/**

– 韵腹位 *yùnfù wèi*   "the centralemes"   **/a/**, **/o/**, **/e/**, **/i^C/**, **/u^C/**, **/ü^C/**

– 韵尾位 *yùnwěi wèi*   "the terminalemes"   **/i^T/**, **/u^T/**, **/n^T/**, **/ng/**

We have to deal with several segments standardly viewed as one phoneme occurring in several positions by Western phonologists, namely **/i/**, **/u/**, **/ü/**, **/n/**. The suggested solution is as follows: they are marked by the index letters, indicating their function: ^{M} (indicating the function of a medial), ^{C} (indicating the function of a central, or a main vowel), and ^{T}, (indicating the function of a terminal). For example, **/i^M/**, **/i^C/**, **/i^T/**; they are viewed as three different phonological units, *not* as one phonological unit **/i/** with three allophones in complementary distribution.

The following table compares a standard solution accepting a single phoneme **/i/**, with our analysis, suggesting three phonological units:

	i as a medial	**i** as a main vowel	**i** as a terminal
standard analysis	phoneme **/i/** **[j]**	phoneme **/i/** **[i]**	phoneme **/i/** **[ɪ]**
suggested analysis	medialeme **/i^M/** **[j]**	centraleme **/i^C/** **[i]**	terminaleme **/i^T/** **[ɪ]**

A standard "Western" solution is phonologically economical, as it involves a single phoneme **/i/** (with three allophones **[j]**, **[i]**, **[ɪ]**). The suggested analysis is less efficient: it accepts three phonological units, each with a single allophone: a medialeme (*yùntóu wèi*) **/i^M/** with a single allophone **[j]**, a centraleme (*yùnfù wèi*) **/i^C/** with a single allophone **[i]**, and a terminaleme (*yùnwěi wèi*) **/i^T/** with a single allophone **[ɪ]**. The proposed analysis is seemingly less economical (as it needs to posit three phonological units), and less acceptable in the eyes of a Western phonologist (all three units are in complementary distribution). On the other hand, its advantage is that it gets rid of the troublesome allophonic variations of high vowel segments conditioned by the position within a syllable. Grasping hold of these variations might be hard for a beginner in the language. A solution of the sort "one phonological unit – one allophone" might be more transparent. (The "only" allophonic variations of the segments within a final which are left to be solved are those of the non-high vowels; they are caused by an assimilation to the following terminal segment, e. g. *ban* → [pan], *bang* → [pɑŋ].) A legitimate objection to a proposed model is that it complicates a phonological notation of syllables. For example, for *mi* it would be /*mi^C*/, for *mie* it would be /*mi^M e*/, for *mai* it would be /*mai^T*/. We could think of another, non-*pinyin* notation, using the IPA symbols:

i: /*i^M*/ → /j/   /*i^C*/ → /i/   /*i^T*/ → /ɪ/

u: /*u^M*/ → /w/   /*u^C*/ → /u/   /*u^T*/ → /ʊ/

ü: /*ü^M*/ → /ɥ/   /*ü^C*/ → /ü/

n: /*n^I*/ → /n/   /*n^T*/ → /N/

The above mentioned syllables would be then written as /*mi*/, /*mje*/, /*maɪ*/. It is not clear whether the suggested analysis is plausible for classroom usage. However, it may be potentially inspiring in some respects, as it offers an alternative interpretation of the Mandarin syllable, drawing on traditional concepts.

Types of Mistake

We can now sum up the preceding discussion. Initial learners of Mandarin come from many different countries. They may misread the isolated *pīnyīn* syllables in various ways and for various reasons. The major types of segmental mistake may be caused by:

● The difficulty of a particular segment. This concerns primarily the *initials*, e. g. the initials /*zh*/, /*ch*/, /*r*/, /*q*/, /*x*/ etc. A difficult segment belonging to a *final* part is /*ng*/. All of these mistakes can be removed by proper instruction about the articulation of the segment in question.

● Traps associated with the *pīnyīn* orthography. These mistakes can be eliminated mostly by mechanical training and dictation drills.

● Discrepancies between the sound system of Mandarin and the sound system of the native language of the student. In this article we were concerned with one particular topic of a sound system: syllable structure. We may conclude now that the differences in the phonological and phonetic structure of the syllable in both languages cause two major sorts of segmental mistake (both of them concern a final):

– Neglecting the specific phonetic features of the components G, V, X. It is namely through neglecting the positional variants of the segments /*i*/, /*u*/, /*ü*/, /*n*/. These variations concern a massive range of syllables – i. e. all syllables with the finals of the VV, VC, GV, GVV and GVC type (only the final type V with a non-high main vowel remains aside). This means that mistakes of this kind may occur in almost any Mandarin syllable.

– Neglecting the assimilations of non-high vowels (most importantly of a low vowel /*a*/) to the following back terminals /*u*/, /*ng*/. This sort of mistake concerns the syllables containing VX (namely the syllables with the finals of the VV, VC, GVV and GVC type) which have a back terminal. This sort of error is probably less serious than the previous sort, consequently it is commonly ignored by language teachers.

We shall call the mistakes described in the last two paragraphs "structural mistakes." The mispronunciation of a terminal /*ng*/ will be included here as well, since it is also a structural mistake in a certain sense.

Structural Mistakes – Examples

Let us review the major cases of structural mistake and provide a few examples:

– In the finals comprising the rising diphthongs /*ia*/, /*ie*/, the medial /*i*/ may be mispronounced as **[i]** (instead of correctly as **[j]**); this often leads to the breaking up of a syllable into two syllables, as an extra syllabic peak emerges. An approximant **[j]** may be wrongly inserted before the second syllabic peak (*xia* → *[ɕi.ja]*, *xie* → *[ɕi.je]*). The situation is analogical in the finals comprising triphthongs /*iau*/, /*iou*/ (*xiu* = /*xiou*/ → *[ɕi.jou]*).

– In the finals comprising the rising diphthongs /*ua*/, /*uo*/, /*üe*/, the medials /*u*/, /*ü*/ may be mispronounced as **[u]**, **[y]** (instead of correctly as **[w]**, **[ɥ]**); this often leads to the breaking up of a syllable into two syllables, as an extra syllabic peak emerges. An approximant **[w]** or **[ɥ]** may be wrongly inserted before the second

syllabic peak (*gua* → *[**ku.wa**], *xie* → *[**ɕy.ɥe**]). The situation is analogical in the finals comprising the triphthongs /*uai*/, /*uei*/ (*guai* → *[**ku.waɪ**]).

– In the finals comprising the falling diphthongs /*ai*/, /*ei*/, /*au*/, /*ou*/, the terminals /*i*/, /*u*/ may be mispronounced as [**i**] (or even [**j**]), [**u**] (instead of correctly [**ɪ**], [**ʊ**]); a second syllabic peak may emerge (*mai* → *[**maj**], *[**ma.ji**]). The terminal /*u*/ is mispronounced as [**u**] less frequently – thanks to *pīnyīn* orthography "o" (/*mau*/ = *mao*). Yet the danger of an emergence of an extra syllabic peak is here, too (*mao* → *[**ma.o**]). The situation is analogical in the finals comprising the triphthongs /*uai*/, /*uei*/ (*guai* → *[**kwaj**]).

The mistakes described above are typical for students whose native languages do not have (or hardly have) any diphthongs. Czech students can serve as an example.

– In the finals comprising a simple vowel followed by a nasal terminal such as /*an*/, /*en*/, /*ang*/, /*ing*/ etc., the terminals may be wrongly pronounced with an audible release (*gan* → *[**kan**] instead of [**kan˺**], *bang* → *[**pɑŋk**] instead of [**pɑŋ˺**]). Again, both these mispronunciations are typical for Czech students. Another frequent shortcoming is the consistent pronouncing of the terminal /*n*/ in an overly "perfect" manner (i. e. with a complete closure at the alveolar ridge), which may sound unnatural. The situation is analogical in the finals with a medial added: /*ian*/, /*uang*/ etc. (*guang* → *[**kwɑŋk**]).

– In the finals comprising VX, the assimilations of a low vowel /*a*/ (and less importantly of a vowel /*e*/) to the back terminals /*u*/, /*ng*/ may be neglected (*bang* → *[**paŋ**] instead of [**pɑŋ**], *bao* = /*bau*/ → *[**paʊ**] instead of [**pɑʊ**], *guang* → *[**kwaŋ**] instead of [**kwɑŋ**]).

In classroom teaching one can frequently observe that several types of structural mistake occur within the same syllable in the pronunciation of the students. For instance:

guang may be sounded as *[**ku.waŋk**] instead of [**kwɑŋ**]

The mistakes are as follows: In a nasal final /*uang*/ the medial /*u*/ is mispronounced as [**u**], which results in the emergence of an extra syllabic peak; an approximant [**w**] is wrongly inserted; furthermore, the main vowel /*a*/ is mispronounced as a front [**a**], neglecting the due assimilation to the back nasal terminal; finally a terminal /*ng*/ is mispronounced as [**ŋk**].

It is obvious, that a sweeping majority of the syllables contained in the Mandarin syllabary is at risk of being mispronounced in the sense of the structural mistakes described above (the cases to be left aside, such as *ba, zha, po, mo, de, ge,* account for less than 40 out of the approximately 400 segmental syllables). These errors indeed make up a good deal of the mispronunciations heard in the classroom, and – not infrequently – also in the speech of advanced speakers of Mandarin.

Summary

- The structure of the Mandarin syllable is fairly simple. This does not automatically imply, however, that it can be easily adopted by native speakers of other languages. Certain types of difficulty must be expected for learners whose native languages have a considerably different syllable structure. These are especially the languages where diphthongs and triphthongs are rare or completely absent (such as Czech).

- There is a general consensus among phonologists that the Mandarin syllable is composed of four components C, G, V, X, defined by their position within a syllable template. They correspond to four traditional components: "an initial, a medial, a main vowel, a terminal." Their vocalic/consonantal phonological status is not uniform, at least in the case of X: phonologically, X may be either a consonant, or a vowel. (Yet, phonetically the nasal terminals are rather close to vowels: they belong to the class of sonorants, their consonantal features appear to have been eroded.)

- Phonologists further agree that the closest relationship is between V and X, unanimously accepting the component of a rime/rhyme/subfinal (cf. the syllable models presented in Figures 1 and 3).

- Each of the components C, G, V, X displays certain phonetic properties shared by all segments allowed in the position in question.

- There are two crucial types of mistake related to syllable structure: first, neglecting the due phonetic features of the components C, G, V, X (in particular the positional variants of /*i*/, /*u*/, /*ü*/, /*n*/), and second, neglecting the anticipatory assimilations within a subfinal. Both types of mistake can (and frequently do) occur within the same syllable. They account for a massive proportion of pronunciation mistakes as a whole.

- Most of the treatments of the *pǔtōnghuà* sound system depart from the inventory of vowel and consonant phonemes. A syllable has been traditionally rendered as a combination of an initial, a final and a tone, while a final is chiefly presented as a string of vocalic and consonantal segments (or "*pīnyīn* letters"). Syllable structure is seldom analyzed in any more detail. It is explained rather late (if at all). The generalizations about the properties of particular components are missing. Consequently, the correct phonetic shape of each newly acquired final has to be learned by mechanical imitation. Students frequently repeat the same mistakes, being unaware of their underlying logic.

- However, the underlying logic of the major mistakes in pronouncing the isolated syllables is systematical and rather transparent. We believe this fact may be efficiently utilized in language teaching: a simple remedy is to elucidate the structure of the Mandarin syllable and the phonetic properties of its constituents at the very beginning of a study program.

Suggested Procedure

In light of the above described facts we suggest a rearrangement of the steps in the process of teaching the basic syllabary of Mandarin to foreign students in the following way:

1. A brief introduction of the Mandarin syllable as a combination of up to four C and V segments; an account of twelve syllable types: V, CV, VV, VC, GV, GVV, GVC, CVV, CVC, CGV, CGVV, CGVC.
2. An introduction into the hierarchical structure of the Mandarin syllable, using the *Initial-Final model*; outlining the subsyllabic components of different levels and their mutual relationships and pointing out that the closest relationship is between V an X.
3. An overview of the four components at the level of segments which build up a syllable template C-G-V-X: an initial (*shēngmǔ*, C), a medial (*yùntóu*, G), a main vowel (*yùnfù*, V), and a terminal (*yùnwěi*, X). Pointing out that these components are principally defined by their position within a syllable template, not by their vocalic/consonantal status. Introducing eight syllable templates: V, CV, VX, GV, GVX, CVX, CGV, CGVX.
4. Listing the inventories of phonemes allowed for each of the positions C, G, V, X. A brief account of the phonetic features of C, G, V, X. Listing the phonemes which may occur in more than one position, namely /*i*/, /*u*/, /*ü*/, /*n*/. A brief overview of their positional variants.
5. An overview of the inventory of phonological initials (C). Explaining the phonological system of Mandarin initials (stops and affricates fall into two contrasting series – one unaspirated, the other aspirated).
6. An overview of the inventory of phonological finals (i. e. the existing sequences GVX). The inventory of finals is to be presented in the following order:
- the basic finals: V (= simple vowels), VV (= falling diphthongs), and VC (= a simple vowel plus a nasal),
- the complex finals augmented by a medial: GV (= rising diphthongs), GVV (= triphthongs), and GVC (= a medial plus a simple vowel, plus a nasal).

Note that the suggested final types are different both from the groups of "simple finals," "compound finals" and "nasal finals" commonly used in Chinese textbooks, and from the four traditional categories *sì hū*.

Optionally, a brief overview of the systematic and accidental gaps in the inventory of finals may follow.

7. An overview of the inventory of phonological tones. Elucidating the phonetic properties of tones on the isolated syllables.

8. Presenting the whole inventory of syllables. First, mapping the existing combinations of initials and finals and explaining the systematic and accidental gaps in the syllable inventory (using the *sì hū* categories of finals). A well-known foldable table of syllables may serve as a useful aid here. Second, presenting the combinations of the segmental syllables with the tones: not all syllables may occur in all four tones.

9. Teaching the whole inventory of tonal syllables (the tones must always be included in what follows):

– Combinations of all initial consonants with the V finals (simple vowels). The initials are to be ordered according to the traditional rows, reflecting the place of articulation: */b, p, m, f / – /d, t, n, l / – /z, c, s/*, etc. Explaining the phonetic features of the initials within the particular rows. Practical training in relation to the syllables.

– Combinations of the initial consonants with the VV finals (falling diphthongs) and VC finals (simple nasal finals). Explaining the phonetic features of the VV finals and VC finals, pointing out the specific properties of the four terminals */i/*, */u/*, */n/*, */ng/*. Explaining the phonetic rules operating within a subfinal, in particular the anticipatory assimilations of a main vowel to a terminal. Practical training in relation to the syllables.

– Combinations of the initial consonants with the GV finals (rising diphthongs). Explaining the phonetic features of the GV finals, pointing out the specific properties of the three medials */i/*, */u/*, */ü/*. Practical training in relation to the syllables.

– Combinations of the initial consonants with the GVV finals (triphthongs). Explaining the phonetic features of the GVV finals. Practical training in relation to the syllables.

– Combinations of the initial consonants with the GVC finals (complex nasal finals). Explaining the phonetic features of the GVC finals. Practical training in relation to the syllables.

Conclusion

At the present time, students of Mandarin mostly have to rely on blindly imitating their teachers (possibly with some help of the IPA included) when learning the basic syllabary. We believe that a timely explanation of syllable structure in the course of teaching is essential for furnishing students with a solid pronunciation foundation. It could make the teaching of Mandarin syllabary progress at a faster pace, be more efficient and scientifically-based. The methodological approach suggested in this article (which has already been roughly outlined by the author of this article)[34] might produce good results, especially when teaching the speakers of languages with a notably different syllable structure (Western students in particular). Some language teachers may object that the students actually do not need this preliminary knowledge, and that supplying too much phonological and phonetic information about the Mandarin syllable at too early a stage of their studies can be counterproductive. They may also argue that a theoretical base is not needed as the students will, sooner or later, acquire the proper phonetic shapes of the syllables simply through imitation. Our conclusions are different, however. It is clear that they have to be tested in the classroom teaching in the future. Even if these assumptions turn out to be wrong, applying the suggested procedure cannot actually harm a normal course of teaching: the instructions about syllable structure do not require any large amount of time. At a rough estimate, two or three lessons may suffice.

Notes

[1] The term "Mandarin" is – among other usages – commonly used as a label for modern standard Chinese as codified in the P. R. C.; its meaning is then identical with *pǔtōnghuà*.

[2] E. g. Helen H. Shen et al., *Hanyu pinyin rumen / Introduction to Standard Chinese Pinyin System*.

[3] Allan Cruttenden, ed., *Gimson's Pronunciation of English*, 206.

[4] Note that some underlyingly present elements may be missing in a standard *pīnyīn* orthography of particular finals: e. g. *gui* is phonologically /*guei*/, *xiu* is /*xiou*/. In other cases, the underlying *pīnyīn* phoneme may be written with a different letter in standard *pīnyīn* orthography: *xu* is /*xü*/, *miao* is /*miau*/, *dong* is /*dung*/, etc.

[5] A preliminary outline of this topic was introduced in my lecture titled "How to teach the segmental structure of the Mandarin syllable" presented in October 2010 at the Phonetic Laboratory of the Institute of Linguistics, CASS, Beijing, and at the Center for Studies of Chinese as a Second Language, Beijing Language and Culture University. These two presentations were followed by inspiring discussions with attending Chinese linguists, phoneticians, language teachers and students. They helped me to get closer to the core of the problem.

[6] E. g. Duanmu San, *The Phonology of Standard Chinese*, or A. A. Dragunov, and J. N. Dragunova, "Struktura sloga v kitaiskom nacionalnom yazyke".

[7] Li Wen-Chao, *A Diachronically Motivated Segmental Phonology of Mandarin Chinese*, 75.

[8] Li Wen-Chao, *A Diachronically Motivated Segmental Phonology of Mandarin Chinese*, 75, and Lin Yen-Hwei, *The Sounds of Chinese*, 107.

[9] Cheng Chin-Chuan, *A Synchronic Phonology of Mandarin Chinese*, 11.

[10] E. g. Li Wen-Chao, *A Diachronically Motivated Segmental Phonology of Mandarin Chinese*, or Duanmu San, *The Phonology of Standard Chinese*, or Lin Yen-Hwei, *The Sounds of Chinese*, 108.

[11] Lin Yen-Hwei, *The Sounds of Chinese*, 108.

[12] E. g. Duanmu San, *A Formal Study of Syllable*, *Tone, Stress and Domain in Chinese Languages*, or Duanmu San, *The Phonology of Standard Chinese*, or Duanmu San, *Syllable Structure – the Limits of Variation.*

[13] Duanmu San, *The Phonology of Standard Chinese*, 28.

[14] Wu Zongji et al., *Xiandai hanyu yuyin gaiyao.*

[15] Huang Borong, and Liao Xudong, *Xiandai Hanyu*, or Wang Lijia et al., *Xiandai Hanyu*, or Cao Wen, *Hanyu yuyin jiaocheng.*

[16] Ma Sheng Jing-Heng, *Hanyu fayin lianxi ji ceyan / Drills and Quizzes in Mandarin Chinese Pronunciation.*

[17] Lin Tao, and Wang Lijia, *Yuyinxue jiaocheng.*

[18] Wu Zongji et al., *Xiandai hanyu yuyin gaiyao*, 128, or Huang Borong, and Liao Xudong, *Xiandai Hanyu*, 90, or Cao Wen, *Hanyu yuyin jiaocheng*, 103.

[19] Zdena Palková, *Fonetika a fonologie češtiny*, 271.

[20] Cao Jianfen, "Putonghua yumu /ao/ yu /ou/ de shipin xiebian duibi fenxi."

[21] Tomáš Duběda, *Jazyky a jejich zvuky*, 139.

[22] Hana Třísková, *Segmentální struktura čínské slabiky*, 212, 334.

[23] Duanmu San, *The Phonology of Standard Chinese*, 51.

[24] Hana Třísková, *Segmentální struktura čínské slabiky*, 215, 229.

[25] Karel Ohnesorg, and Oldřich Švarný, *Études expérimentales des articulations Chinoises*, 61.

[26] Shi Xiangdong, "Hanyu putonghua de -*n* yunwei," 218.

[27] The categorization of the sounds of speech into a class of *sonorants* (i. e. the vowels, liquids, glides, and nasals), and a class of *obstruents* (i. e. the stops, fricatives, and affricates) is based on the phonetic properties of the sounds.

[28] See above – Duanmu San, *The Phonology of Standard Chinese.*

[29] Tomáš Duběda, *Jazyky a jejich zvuky*, 52.

[30] Paul Kratochvil, "Traditions in Chinese Linguistics: Fact or Fiction?," 18.

[31] David Crystal, *Dictionary of Linguistics and Phonetics*, 298.

[32] Geoffrey Sampson, *Schools of Linguistics*, 216.

[33] Such an analysis is outlined in Hana Třísková, *Segmentální struktura čínské slabiky*, 51.

[34] namely in Hana Třísková, "The Sounds of Chinese and how to Teach them," 523.

References

Cao, Jianfen. "Putonghua yumu /ao/ yu /ou/ de shipin xiebian duibi fenxi [Comparative acoustic analysis of Chinese diphthongs /ao/ and /ou/]." In Cao, Jianfen. *Xiandai yuyin yanjiu yu tansuo* [Modern phonetic research]. Beijing: Shangwu yinshuguan, 2007. pp. 90–100.

Cao, Wen. *Hanyu yuyin jiaocheng* [The course of Mandarin pronunciation]. Beijing: Yuyan wenhua daxue chubanshe, 2002.

—. *Hanyu yuyin xunlian* [Exercising Mandarin pronunciation]. Beijing: Beijing daxue, 2008.

Chao, Yuen-Ren. *A Grammar of Spoken Chinese*. Berkeley: University of California Press, 1968.

Cheng, Chin-Chuan [Zheng, Jinquan]. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton de Gruyter, 1973.

Chin, Tsung [Jin, Cong]. *Sound System of Mandarin Chinese and English: a Comparison.* München: Lincom Europa, 2006.

Cruttenden, Allan ed. *Gimson's Pronunciation of English.* London: Arnold, 2001.

Crystal, David. *Dictionary of Linguistics and Phonetics*. Oxford: Blackwell Publishers, 1997.

Dow, Francis D. *An Outline of Mandarin Phonetics.* Canberra: Australian National University Press, 1972.

Dragunov, Alexandr Alexandrovich, and Dragunova, Yekaterina Nikolaevna. "Struktura sloga v kitaiskom nacionalnom yazyke [The syllable structure of modern standard Chinese]." *Sovetskoye vostokovedeniye* 1 (1955): 57–74. For the Chinese version see Long, Guofu.

Duanmu, San. *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages.* Ph.D. dissertation, MIT, Cambridge (Mass.), 1990.

—. *The Phonology of Standard Chinese.* Oxford: Oxford University Press, 2002.

—. *Syllable Structure – the Limits of Variation.* Oxford: Oxford University Press, 2009.

Duběda, Tomáš. *Jazyky a jejich zvuky*. *Univerzálie a typologie ve fonetice a fonologii* [The languages and their sounds. the universals and typology in phonetics and phonology]. Praha: Karolinum, 2005.

Huang, Raymond. *Mandarin Pronunciation.* Hong Kong: Hong Kong University Press, 1981.

Huang, Borong, and Liao, Xudong. *Xiandai Hanyu* [Modern Chinese] I. - III. Beijing: Gaodeng jiaoyu chubanshe, 2002.

Kratochvil, Paul. "Traditions in Chinese Linguistics: Fact or Fiction?" *Cahiers de Linguistique Asie Orientale* 1 (1977): 17–30.

Li, Wen-Chao. *A Diachronically Motivated Segmental Phonology of Mandarin Chinese.* New York: Peter Lang, 1999.

Lin, Tao, and Wang, Lijia. *Yuyinxue jiaocheng* [The course of phonetics]. Beijing: Beijing daxue, 2003.

Lin, Yen-Hwei. *The Sounds of Chinese.* Cambridge: Cambridge University Press, 2007.

Long, Guofu, and Long, Guowa [Dragunov, A. A., and Dragunova, J. N.]. "Hanyu putonghua yinjie de jiegou" [Syllable structure in putonghua]. *Zhongguo yuwen* 11 (1958): 513–521.

Ma Sheng, Jing-Heng. *Hanyu fayin lianxi ji ceyan / Drills and Quizzes in Mandarin Chinese Pronunciation*. Beijing: Beijing yuyan wenhua daxue chubanshe, 1999.

Ma, Xianbin et al. *Putonghua jichu* [The basic standard Chinese]. Guangzhou: Jinan daxue chubanshe, 2004.

Ohnesorg, Karel, and Švarný, Oldřich. *Études expérimentales des articulations Chinoises.* Rozpravy Československé akademie věd, Vol. 5. Praha: Academia, 1955.

Palková, Zdena. *Fonetika a fonologie češtiny* [The phonetics and phonology of Czech]. Praha: Univerzita Karlova, 1994.

Sampson, Geoffrey. *Schools of Linguistics.* London: Hutchinson Publishing, 1985.

Shi, Xiangdong. "Hanyu putonghua de -*n* yunwei" [The ending -*n* in putonghua]. In *Le zai qi zhong – Wang Shiyuan jiaoshou qishi huadan qingzhu wenji / The Joy of Research. A Festschrift in Honor of Professor William S.-Y. Wang on his 70th Birthday*. Edited by Shi, Feng and Shen, Zhongwei. Tianjin: Nankai daxue, 2004. pp. 213–221.

Shen, Helen H., Tsai, Chen-Hui, and Zhou, Yunong. *Hanyu pinyin rumen / Introduction to Standard Chinese Pinyin System.* Beijing: Beijing yuyan wenhua daxue chubanshe, 2006.

Speshnev, Nikolai Alexeevich. *Fonetika kitaiskogo jazyka* [The phonetics of the Chinese language]. Leningrad: Izdatelstvo Leningradskogo universiteta, 1980. 1st publ. 1972.

—. *Vvedenie v kitaiskij jazyk: fonetika i razgovornyj jazyk* [An introduction into the Chinese language: phonetics and spoken language]. Sankt Peterburg: Karo, 2003.

Třísková, Hana. The Sounds of Chinese and how to Teach them. A review essay on Lin, Yen-Hwei. *The Sounds of Chinese.* Cambridge: Cambridge University Press, 2007. *Archiv Orientální* 4 (2008): 509–543. http://www.orient.cas.cz/kontakty/pracovnici/triskova.html

Třísková, Hana. *Segmentální struktura čínské slabiky* [Segmental structure of Mandarin syllable]. Ph.D. dissertation, the Charles University, Prague, 2010.

Wang, Lijia et al. *Xiandai Hanyu* [Modern Chinese]. Beijing: Shangwu yinshuguan, 2002.

Wu, Zongji et al. *Xiandai hanyu yuyin gaiyao* [An overview of the modern Chinese phonetics]. Beijing: Huayu jiaoxue chubanshe, 1992.

Xu, Shirong. *Putonghua yuyin changshi* [Common knowledge about the phonetics of putonghua]. Beijing: Yuwen chubanshe, 1999.

You, Rujie, Qian, Nairong, and Gao, Zhengxia. "Lun putonghua de yinwei xitong" [About the phoneme system of putonghua]. *Zhongguo yuwen* 5 (1980): 328–334.