Alphabets in disarray: Russian transliterations from Cyrillic vs Modern Standard Chinese pseudo-transcriptions in Pinyin

Joan Julià-Muné
Universitat de Lleida
jjulia@filcat.udl.cat

Abstract

This paper falls within the field of contrastive phonetics and language teaching. It aims at analysing and explaining the romanisation processes undertaken by Russian and Modern Standard Chinese whenever used by non-native speakers learning the language. The analysis focusses on Russian transliterations from its Cyrillic alphabet and so-called Pinyin transcriptions from Putonghua – or Mandarin Chinese – into the Roman alphabet, by pointing out misinterpretations made by Western linguists, especially in English, Spanish and Catalan environments, with regard to both Russian and Chinese L2 teaching. Different strategies are put forward to improve the phonetic renderings and pronunciation accuracy of Cyrillic-Roman transliterations and Pinyin – meaning precisely and ironically 'phonetic script' – transcriptions so that L2 or FL students of these languages may overcome their mispronunciations and prevent subsequent misunderstandings. Mismatchings between the phonetic alphabet, which aims to reflect the actual pronunciation, with the adaptation of transliterations from Cyrillic for Russian and with Pinyin for Chinese will be discussed.

Keywords: Russian, Cyrillic alphabet, transliterations, Standard Chinese, Pinyin, romanisation systems

> 'Do you spell it with a *v* or a *w*?' inquired the judge. 'That depends upon the taste and fancy of the speller, my Lord', replied Sam"
> (Charles Dickens 1837).

1. Introduction: from sound to script[1]

This contribution focusses on a Russian transliteration case from its Cyrillic alphabet and a sample of so-called Pinyin transcriptions[2] from Standard Chinese transferred into the Latin alphabet. It will point out misinterpretations which have been made especially in English and Romance languages environments, with regard to both Russian and Chinese L2 teaching. Mismatchings between the phonetic alphabet, which aims to reflect the actual pronunciation, with the adaptation of transliterations from Cyrillic for Russian and from Pinyin for Chinese will be presented and discussed.

In Roman times the Greek alphabet, originated from the Phoenician syllabary, evolved into the Latin alphabet in the central peninsula of Southern Europe and, in a more conservative

[1] The research this paper is based on is related to the project on descriptive phonetics sponsored by the Research Agency of the Autonomous Government of Catalonia (AGAUR de la Generalitat de Catalunya 2009 SGR 408), which was undertaken by the Grup de Recerca en Fonètica (Grecfon) during the two-year period 2012-2014. A first unpublished version was presented at the international congress on *Conditioned Identities. Wished-for and Unwished-for Identities*, held at the University of Lleida in November 2013). Later on the project was taken over by the recently created Chair of Asian Studies, at the University of Lleida, where research on the subject has been undertaken for the period 2014-2018.

[2] This paper also aims to pay homage to the Pinyin designer, the economist by training and linguist by devotion Zhou Youguang (Changzhou, 13 January 1906 - Beijing, 14 January 2017), in spite of the flaws of this romanisation system, 60 years after this new alphabet was approved by the People's Republic of China's government in 1958. Zhou Youguang died recently at the age of 111, being the oldest person in China at the time. And so does to the codifier of Modern Catalan, Pompeu Fabra i Poch (Gràcia/Barcelona, 20 February 1868 – Prada de Conflent, 25 December 1948), an engineer by training and a linguist by devotion, in the 150th aniversary of his birth and in the centenary of the publication of his emblematic prescriptive *Gramàtica catalana* ([7]1933, [1]1918). Further details on the apparent parallelism between the processes of standardisation of Chinese and Catalan can be found in Julià-Muné (2018).

way, the Cyrillic alphabet in South-eastern and Central Europe in the 9th century. The latter was adopted by the Russian-speaking regions of Christian Orthodox faith. At present both scripts the Latin alphabet and the Cyrillic alphabet, among others, are used to transfer the Chinese logographic writing, which is made up of characters or sinograms, into phono-alphabetical scripts, as we will see below. Figure 1 shows a sample of logographic and alphabetic scripts related to the languages we are about to discuss: Chinese and Russian.

Figure 1. A sample of Chinese and Russian scripts

The Chinese characters are read as *zhong* (centre) *guó* (country, empire), i.e., *Zhongguó* (the Middle Country, in the centre of the world) as Chinese citizens name their own country. The Chinese-designed romanisation system, known as Pinyin, is used (see table 7). Phonetic transcription, by using IPA (International Phonetic Alphabet) symbols, is required in order to read it properly: [^{1}tßuN ^{2}kwo]. As Chinese is a tonal language, tonal marks should be noted:

1-high level, 2-rising. As far as alphabetic script is concerned, Fig. 1 shows a Russian noun phrase in Cyrillic alphabet, which transliterates into Latin alphabet as *Rossískaia Federàtsia* and pronounced [rå»siJskaja fidi»Ratsja] (Russian Federation).

Surprisingly, original pronunciations are sometimes misinterpreted and, therefore, wrongly transliterated or transferred into the Latin alphabet, misleading the learner of the language about its actual pronunciation. For instance, a Russian proper name such as Khru**shch**ev in English, Chru**schtsch**ow in German, Khrou**chtch**ev in French, Jru**shch**ov in Spanish, Khru**ixtx**of and Khru**sx**of in Catalan, etc. show a medial palatal sound traditionally misspelled in most languages, since it requires a single IPA symbol (a particular misunderstood letter in Cyrillic). As far as Chinese is concerned, words spelled with "b,d,g" which stand for /p,t.k/ phonemes, among other misleading orthographic items used in present-day Pinyin, will be dealt with. This mismatching also explains why Chinese native speakers learn how to pronounce English more accurately than any Romance language, including Galician, Spanish and Catalan.

In the following pages the consonant systems of English, Catalan and Russian will be contrasted, especially with regard to palatal sounds, so that a misinterpreted phoneme/Cyrillic letter of Russsian can be pronounced and transliterated more precisely and, consequently, adapted in a more rational and proper way. Later, the consonant system of Chinese will be displayed and bilabial plosive phonemes in English, Catalan and Chinese be contrasted so that the mismatching sound-script and its consequences affecting both English-speaking and Romance language-speaking learners can be better understood.

2. Russian pronunciation: transliteration from alphabet to alphabet

2.1 The pronunciation of English, Catalan and Russian in contrast

We will be focussing just on the consonant voiceless palatal sounds in the three following tables.

p	b			t	d						k	g		
		f	v	T	D	s	z		S	Z				h
						(tr	dr)		tS	dZ				
m						n						N		
						l								
									®					
											j		w	

Table 1. Inventory of consonant phonemes in English

p	b			t	d							k	g	
		f	(v)*			s	z		S	Z				
						ts	dz		tS	dZ				
m							n			‒		N		
							l			¥				
							R							
							r							
											j		w	

Table 2. Inventory of consonant phonemes in Catalan
(*Present in Valencian and Balearic varieties)

p	b			t	d								k	g
		f	v			s	z		ß*	Ω*	¸		x	
						ts					t¸			
m				n										
				l										
						R\r								
											j			

Table 3. Inventory of consonant phonemes in Russian
(*Apicopalatal or retroflex fricatives, unlike the English and Catalan palatal fricatives/affricates, such as *ca**sh**ing/ca**tch**ing* and *fa**i**x**a*** 'sash' / ***fa**t**x**a*** 'presence', shown in tables 1 and 2, respectively)

2.2. Russian consonant phonemes of palatal obstruent articulation

Both languages Catalan and Russian have four units each in this area of articulation. Catalan has a parallelism between fricatives and affricates [S, tS; Z, dZ] as in *fa**i**xa* 'sash', *fa**t**xa* 'presence'; *a**j**untar* 'to assemble', *a**dj**untar* 'to attach', respectively. Russian, however, has three fricatives and one affricate [ß, Ω, ˛, t˛], which are displayed in Table 4. Figure 2 shows the articulation of the voiceless alveolopalatal fricative involved, which is transcribed as [˛] or [˛:].

Printed Cyrillic	*Name of the letter*	*Transliteration into the Latin alphabet*	*Phonetic transcription and articulatory classification*	*Nearest equivalent (English and Catalan)*	*Russian word and translation into English*
Ш ш	ша	sh (*English*) (i)x (*Catalan*)	Voiceless apicopalatal/retroflex fricative [ß]	***sh**oe* *ei**x**ut* 'dry' ***x**ut* 'shot' (in football)	**ш**ум 'noise'
Ж ж	жэ	zh (*English*) j (*Catalan*)	Voiced apicopalatal/retroflex fricative [Ω]	*mea**s**ure* *a**j**ut* 'assistance'	**ж**ук 'beetle'
Щ щ	ща	*shch (English) *sx, *(i)xtx (Catalan)	Voiceless alveolopalatal fricative [˛(:)] *[˛t˛]	**fre**sh ch**eese* ***sh**eaf* **e**ix tx**ec* 'Czech axis' *ull**s** 'eyes'*	**щ**ётка [»˛ otkå] 'brush'
Ч ч	че	ch *(English and Spanish)* tx (*Catalan*)	Voiceless alveolo-palatal affricate [t˛]	***ch**eck* ***tx**ec* 'Czech' *any**s*** 'years'	**ч**еловек 'man'

Table 4. Chart of the four Russian palatal obstruents
(The asterisk shows inaccurate scripts or examples)

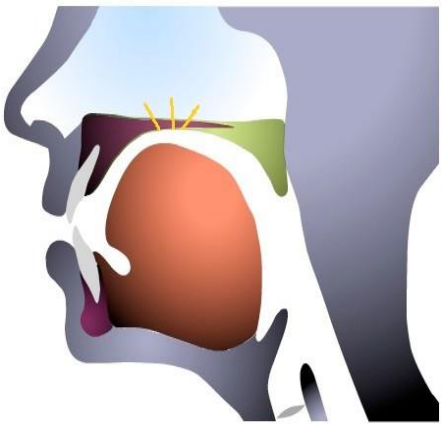

Figure 2. Approximate articulation of the Russian phoneme /˛/ [s^j]
Voiceless alveolopalatal fricative (cp. North-eastern Catalan *ull**s*** 'eyes')

Examples: **щ**ётка [»˛ otkå] 'brush'; про**щ**ание [pRå»˛ anje] 'farewell';
пла**щ** [p...5a˛] 'overcoat'

In mid-20th century, this Russian phoneme was believed to be a sequence of voiceless

alveolopalatal/palatoalveolar fricative+affricate [ˌtˌ\StS]. Such a description was included, for instance, in the popular textbook *The Penguin Russian Course.* It stated that the Cyrillic letter ***щ*** was pronounced as "The nearest English equivalent is the *shch* in *fresh cheese*" (Fennell 1974: xx), which had a "nearest Catalan equivalent" in ***eix txec*** 'Czech axis' [StS] or ***és txec*** 'he's a Czech' [stS]. The *Gran Enciclopèdia Catalana* (GEC) followed suit in its first two editions and so did the Catalan Academia, the Institut d'Estudis Catalans (IEC 1996), which adopted the surprising orthographic sequence ***sx*** for the Russian phoneme in question. However, it is not a matter of two-sound sequence – in fact, it is the result of a fusion – which required at least a two-letter sequence to be transliterated, but a single sound, articulated as an alveolopalatal fricative, with some lengthening [ˌ(:)], which is present in Catalan as an allophone of /s/ after a palatal lateral, whose articulation is close to a palatoalveolar [S] and can be adapted as such. Table 5 and 6 show various interpretations and their subsequent adaptations in Latin script in different languages. Finally, a proposal is made by the author by using the two-letter *-ix-*, being /S/ the neariest Catalan equivalent to the Russian phoneme under discussion. Therefore, *Khruixov* seems to be the most accurate adaptation of this anthroponym.

Correlation ISO-IPA in various languages (Adapted and extended from IEC, 1996)						
ISO	English	French	German	Spanish	Catalan	IPA
Transliteration	*Latin pseudotranscriptions*					*Phonetic transcription*
s&& (GEC, 1st ed.)	sh/ch	ch/tch	sch/tsch	sh/ch	sx (IEC) (i)xtx (GEC, 2nd ed.) ix (author's proposal)	S*tS (ˌ*tˌ) (*[ˌtˌ])

Table 5. Between the transliteration from Russian and the phonetic adaptation
(The column on the right shows the Catalan adaptation ([S*tS]) and the Russian sounds involved ([ˌ*tˌ]).
The asterisk shows misunderstood pronunciations)

Traditional adaptations	*Paragraphic (1) and paraphonic (2) pronunciations*	*Original script*
*Khrusxov (IEC) *Khru&&ov (GEC) *Khruixtxov (GEC) [Khruixov (as suggested by the author)]	Khrusxov 1. [kRu»Sçf] 2. [xRu»Sçf]	*Cyrillic:* Хрущёв[3] *Transliterated:* Khru&&ov (ISO/GEC) *Transcribed:* [xRu»ˌçf]

Table 6. A sample of adapted Russian proper names after having been transliterated (*Paragraphic adaptation*: according to the original script, as usually spelled or transliterated; *paraphonic adaptation*: according to the original pronunciation of the adapted name.
In both cases the adaptation will be carried out by means of the usual sounds of the adapting language)

Transliterations from Cyrillic Russian into Latin script are sometimes faulty and as a result the adapting language must distort its orthography. As we have just seen, Catalan had to supply letter sequences, such as *-(i)xtx-* or *-sx-* to preserve the misinterpreted original pronunciation. Unfortunately, this paraphonic adaptation was too successful, but needless to say, time has come so that we should aim to formalise a flawless orthology in any language.

[3] The family name of the Soviet Premier (1958-1964), was adapted paraphonically in Catalan as [k\xRu»Sçf] or as [kRu»SEf] in a paragraphic way, by following the English *Khru**shche**v* (from the Russian diphthong ё [jo]).

3. Standard Chinese/Mandarin/Putonghua pronunciation: transference from sinograms to Pinyin alphabet

3.1. Putonghua and Pinyin

The official language of the People's Republic of China[4] – the language with the highest number of speakers as L1 on Earth, around 900 million – is known as Modern Standard Chinese (formerly Mandarin), which is also named Putonghua (*putonghuà*[5] or common language). At present is presented to foreign learners in a romanised script known as Pinyin (Yin & Felley 1990), meaning precisely 'phonetic script' in Chinese. Table 7 displays a sample of Chinese proper names, toponyms and anthroponyms, in the two best known romanisation systems. The English-designed Wade-Giles, which was more adjusted to the real pronunciation of Chinese, has become obsolete by now.

Wade-Giles (1859)	*Pinyin (1958)*	*Catalan adaptation*
Mao Tse(-)tung/Tse Tung	Mao Zedong	Mao Zedong
Teng Hsiao-p'ing	Deng Xiaoping	Deng Xiaoping
Chiang Kai-shek	Jiang Jieshi	Chiang Kaishek
Kuomintang	Guomindang	Guomindang *(Nationalist Party)*
Tsinghua	Qinghua	(Universitat de) Qinghua *(Beijing)*
Hong Kong	Xiang Gang	Hong Kong
Peking	Beijing	Pequín
Kwangtung	Guangdong	Província de Canton
Kwangchow	Guangzhou	Ciutat de Canton
Tientsin	Tianjin	Tianjín

Table 7. Chart contrasting the two best known romanisation systems for Standard Chinese (Tonal marks are left out)

3.2. Chinese consonant sounds which are articulated as bilabial obstruents

Table 8 displays the Standard Chinese sounds. Pinyin script is shown in italics and single quotation marks are used, whilst Chinese sounds which are absent from the Catalan inventory are shown in clear grey and so are in dark grey those unknown by Chinese-speaking learners. Chinese learners find especial difficulties in the acquisition of some Catalan consonant phonemes, namely stops, laterals and rhotics. While the distinctive feature for the three pairs of Romance language stops is voicing ([p/b, t/d, k/g]), the distinctive feature for the Chinese stops is aspiration (p^h/p, t^h/t, k^h/k). This fact implies that Chinese speakers tend to mispronounce Romance languages rather than English. Table 9 (Julià-Muné 2011, 2014) shows one of the segmental problems that Chinese speakers must overcome when learning Catalan as L4, after having learned English and Spanish.[6] Chinese learners of Catalan usually get confused in their production and perception when using pairs such as *vi* 'wine' and *pi* 'pine tree', *puta* 'whore' and *Buda*; *dental* 'dental' and *tendal* 'awning'; *boda* 'wedding', *poda* 'pruning' and *vota/bota* '(he, she, it) votes/boot,barrel'. Perception tests show confusion in 80% of learners when listening out of context utterances. This confusion is dramatically reduced to 50% in production tests (reading), since they can count on spelling support.

[4] Further detailed information on other (oral) Chinese languages, can be found in Ramsey (1987).

[5] Chinese terms such as *putonghua* and *pinyin*, with no tonal marks, are transferred into Pinyin and pronounced as if they were stressed on the last syllable. Therefore, they are spelled as *putonghuà* and *pinyín* in their Catalan adaptation.

[6] In fact, they have to deal with the same problem in Spanish, since they have hardly overcome this major hindrance when learning their first Romance language.

	POINT OF ARTICULATION																					
WAY OF ARTICULATION	*Bilabial*		*Labio-dental*		*Inter-dental*		*Dental (apico-dental)*		*Alveolar /lamino-alveolar)*		*Post-alveolar*		*Retroflex (apico-palatal)*		*Palatal (alveolo/ dorso-palatal)*		*Velar (dorso-velar)*		*Uvular (dorso-uvular)*		*Glottal*	
	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Plosive (aspirated)	p *'b'* pH *'p'*						t *'d'* tH *'t'*										k *'g'* kH *'k'*					
Fricative			f *'f'*						s6 *'s'*				ß Ω 'sh' 'r'		˛ *'x'*		x *'h'*					
Affricate (aspirated)									tÉs6 *'z'* tÉs6H *'c'*				tÉß *'zh'* tÉßH *'ch'*		tÉ˛ *'j'* tÉ˛H *'q'*							
Nasal	*'m'* m								*'n'* n								*'ng'* N					
Lateral									*'l'* l				¥									
Rhotic									r R													
Semivocalic approximant/ Glide															*'y'* j *'u'* Á		*'w'* w					

Table 8. Chart of Standard Chinese consonants (Norman 1988; Lin 2007)

As Table 9 displays, the two Chinese phonemes are almost correlated with the allophones of the English phoneme /p/ together with the devoiced variant of /b/ in initial position. However, Catalan and Spanish /p/ correlates with the unaspirated Chinese /p/, but the plosive and approximant allophones of Catalan and Spanish /b/ fall outside the perceptive scope of the Chinese learner. As a result, by attributing the Pinyin letters 'p' and 'b' to Chinese phonemes /p^h/ and /p/, respectively, Chinese learners of Catalan and Spanish as well as radio and TV speakers trying to adapt Chinese terms, especially proper names, to our Romance languages, are not favoured in their tasks (see Table 10).

	/p/		**/b/**		
	pH	p	b8	b	B
English	**p**in	s**p**in	**b**in	du**bb**in	
Catalan/ Spanish		**p**a / **p**an 'loaf, bread' ca**p**ar 'to castrate'		**v**a '(he,she,it) goes; vain'	ca**v**ar 'to dig'
	/pH/	/**p**/			
Standard Chinese	**p**í 'skin' **p**a# 'lie down'	**b**í 'nose' **b**a# 'number 8'			

Table 9. Contrasted obstruent bilabial phonemes in English, Catalan, Spanish and Standard Chinese

If we add lateral and rhotic confusion [l\¥\R\r > l/R][7] to the stop confusion [p/b > p, ʦ/ʣ > ʦ, k/g > k] we conclude, after perception and production tests, that they are unable to distinguish the following eight Catalan words: *pare* 'father', *parra* 'vine', *pala* 'spade', *palla* 'straw'; *vara* 'stick', *barra* 'bar, jaw', *bala* 'bullet', *balla* 'he/she dances'. Once again, spelling may help them when reading them out. As a result, teaching strategies and their consequent didactic drills should be designed carefully so that Chinese learners could be able to tell apart every lexical unit of the above mentioned four pairs, among other similar cases.

Traditional adaptation C-Central Catalan B-Balearic, V-Valencian N-W North-western	*Paragraphic (1) and paraphonic (2) pronunciations*	*Original version (in romanised and phonetic script)*
Pequín C, B [p´»kin] NW, V [pe»kin]	Beijing 1. [bej»ZiN éb´j»] 2. [pej»tSiN ép´j é»dZiN]	*Pinyin:* Beijīng *Phonetic transcription (IPA):* [^{3}p´j^{1}dʑ˧•iN]

Table 10. A sample of Chinese proper names in Latin script and phonetic transcription (IPA)

In the situations referred to so far it is likely that the language needs being catered for are so wide, although predictable, that the overall aim will be to establish a fairly generalised linguistic and communicative competence in the learners which allow them overcome their shortcomings. After all, having rationally worked out this opinion, any new Pinyin designer aiming to enhance the system, as well as accurate interpreters of Russian transliterations from Cyrillic and subsequent adapters would wisely not share Dickens's report as cited at the beginning of this article: 'That depends upon the taste and fancy of the speller, my Lord'.

References

Dickens, C. (1837). *The Posthumous Papers of the Pickwick Club.* London: Chapman and Hall.

Fabra, P. ([7]1933, [1]1918). *Gramàtica catalana.* Barcelona: Institut d'Estudis Catalans.

Fennell, J. L. I. (ed.) (1974). *The Penguin Russian Course*. Harmondsworth: Penguin.

Institut d'Estudis Catalans (1996). "Proposta sobre el sistema de transcripció i transliteració dels noms russos al català". In *Documents de la Secció Filològica* III. Barcelona: IEC, 55-89.

Julià-Muné, J. (2009). "Les llengües de la Xina. El contrast fonètic i l'adaptació al català de l'onomàstica xinesa". In Joan Julià-Muné (ed.) *Visions de la Xina: cultura multimil·lenària*. Lleida: Institut d'Estudis Ilerdencs, 59-80.

Julià-Muné, J. (2011). "La transferència fonotàctica en l'aprenentatge d'L2: el cas de parlants de mandarí que aprenen català", in M.-R. Lloret & C. Pons (eds.) *Noves aproximacions a la fonologia i la morfologia del català. Volum d'homenatge a Max Wheeler*. Alacant: Institut Interuniversitari de Filologia Valenciana, 295-319.

Julià-Muné, J. (2014). "Catalan", in C. Fäcke (ed.) *Manual of Language Acquisition.* Berlin: De Gruyter, 345-370.

Julià-Muné, J. (2017). "Propuesta de adaptación ortológica al español de la onomástica china". *Onomàstica*, 3, 29-68.

Julià-Muné, J. (2018). "The Standardisation of Chinese: A process run alongside Catalan for a century?". X Col·loqui Internacional d'Història de la Llengua (University of Girona, 27-30 June 2016). Amsterdam: John Benjamins (Forthcoming).

Lin, Y.-H. (2007). *The Sounds of Chinese*. Cambridge: Cambridge University Press.

Norman, J. (1988). *Chinese*. Cambridge: Cambridge University Press.

Ramsey, S. R. (1987). *The Languages of China.* Princeton, N.J.: Princeton University Press.

Yin, B. and M. Felley (1990). *Chinese romanization: Pronunciation & orthography*. Beijing: Sinolingua.

[7] The alveolar tap [R] does not occur in their L1, but Chinese tend to use it, as a hypercorrection, instead of /l/.