

Sixth International Olympiad in Theoretical, Mathematical and Applied Linguistics

Bulgaria, Sunny Beach, 4–9 August 2008

Problem for the Team Contest

At the time when the dictionary *Guangyun* was compiled (1007–1011), the Chinese language was comparatively homogeneous. Since the Chinese script is not phonetic, the dictionary employed a simple system for giving the pronunciation of each character using two other characters, the pronunciation of which the reader was supposed to know (they were in common use). This system is known as *fanqie*.

Later, when Chinese dialects split apart, it was still possible to use many of the ancient *fanqie* transcriptions, but in different (and more complex) ways in different dialects.

Here are some such transcriptions. For each character its reading in Cantonese is given.

character	=	transcription
1. 倦 kyn ²	= 渠 k^hœy ²¹	★ 卷 kyn ³
2. 求 k^hau ²¹	= 巨 kœy ²	★ 鳩 kau ⁵³
3. 住 cy ²	= 持 chⁱ ²¹	★ 遇 y ²
4. 病 piŋ ²	= 皮 p^hei ²¹	★ 命 miŋ ²
5. 掉 tiu ²	= 徒 t^hou ²¹	★ 弔 tiu ³
6. 鳩 kau ⁵³	= 居 kœy ⁵³	★ 求 k^hau ²¹
7. 僖 hei ⁵³	= 許 hœy ³⁵	★ 其 k^hei ²¹
8. 朗 loŋ ¹³	= 盧 lou ²¹	★ 黨 toŋ ³⁵
9. 韶 siu ²¹	= 市 si ¹³	★ 昭 ciu ⁵³
10. 帳 cœŋ ³	= 知 ci ³	★ 亮 lœŋ ²
11. 愀 chⁱiu ³⁵	= 親 ch^{an} ³	★ 小 siu ³⁵
12. 舞 mou ¹³	= 文 man ²	★ 甫 p^hou ³⁵
13. 謏 siu ³⁵	= 先 sin ⁵³	★ 鳥 niu ¹³
14. 臼 k^hau ¹³	= 其 k^hei ²¹	★ 九 kau ³⁵
15. 斜 ch^e ²¹	= 似 chⁱ ¹³	★ 嗟 ce ⁵³
16. 葦 kau ³	= 古 ku ³⁵	★ 候 hau ²

- (a) Explain how ancient *fanqie* transcriptions could be used in modern Cantonese.
- (b) How were the *fanqie* transcriptions designed to work at the time of the compilation of *Guangyun*? The old simple rule can be applied with correct results in Cantonese to only one of the transcriptions above. Which one?

In most Chinese dialects today (including Cantonese and Mandarin) there are no voiced consonants other than sonorants (**l**, **m**, **n**, **ŋ**). At the time when *Guangyun* was compiled the language had other voiced consonants, which later merged with the voiceless ones: voiced fricatives became voiceless fricatives (e. g., **z** > **s**), voiced stops became aspirated or unaspirated voiceless stops (e. g., **d** > **t** or **t^h**). The voiced sounds have been retained in the Wu dialect of Chinese. For example, the character 徒 is pronounced [du²¹] in Wu, [t^hou²¹] in Cantonese and [t^hu³⁵] in Mandarin.

- (c) Which of the characters in the section above were pronounced with voiced initial consonants at the time of the compilation of *Guangyun*? Under what conditions did the voiced consonants become aspirated or unaspirated in Cantonese?
- (d) In Classical Chinese there were four tones, but only three of them are present in this problem. Explain how these three tones have evolved to yield the six tones of Cantonese.

Here are some more transcriptions, but with Mandarin readings only:

17.	遭 can^5	=	張 caŋ^5	★	連 lian^{35}
18.	良 liaŋ^{35}	=	呂 ly^{214}	★	章 caŋ^5
19.	遵 cun^5	=	將 kiaŋ^{51}	★	倫 lun^{35}
20.	蕭 xiao^5	=	蘇 su^5	★	彫 tiao^{51}
21.	嵌 k^hian^5	=	口 $\text{k}^h\text{ou}^{214}$	★	銜 xian^{35}
22.	先 xian^5	=	蘇 su^5	★	前 $\text{k}^h\text{ian}^{35}$
23.	巉 c^han^{35}	=	鋤 c^hu^{35}	★	銜 $\text{k}^h\text{ian}^{35}$
24.	娉 xij^{51}	=	胡 xu^{35}	★	頂 tiŋ^{214}
25.	弗 $\text{c}^h\text{an}^{214}$	=	初 c^hu^5	★	限 xian^{51}
26.	趨 $\text{c}^h\text{uei}^{214}$	=	千 k^hian^5	★	水 ŋuei^{214}
27.	初 c^hu^5	=	楚 c^hu^{214}	★	居 ky^5
28.	釧 $\text{c}^h\text{uan}^{51}$	=	尺 c^h^{214}	★	絹 kyan^{51}
29.	卷 kyan^{214}	=	居 ky^5	★	轉 cuan^{214}
30.	處 c^hu^{51}	=	昌 $\text{c}^h\text{aŋ}^5$	★	據 ky^{51}
31.	俚 p^hij^5	=	普 p^hu^{214}	★	丁 tiŋ^5
32.	蚪 tou^{214}	=	當 taŋ^5	★	口 $\text{k}^h\text{ou}^{214}$

- (e) Ignoring the tones for the moment, formulate rules for using the ancient fanqie transcriptions in Mandarin.

Given are Chinese characters with both Cantonese and Mandarin readings:

	Cantonese	Mandarin		Cantonese	Mandarin
33.	唐 t^hou^{21}	$\text{t}^h\text{aŋ}^{35}$	40.	采 pin^2	pian^{51}
34.	謨 mou^{21}	mo^{35}	41.	帝 tai^3	ti^{51}
35.	踐 c^hin^{13}	kian^{51}	42.	透 t^hau^3	t^hou^{51}
36.	少 siu^{35}	ŋao^{214}	43.	被 p^hei^{13}	pei^{51}
37.	夔 $\text{k}^h\text{wai}^{21}$	$\text{k}^h\text{uei}^{35}$	44.	囂 hiu^{53}	xiao^5
38.	你 nei^{13}	ni^{214}	45.	粉 fan^{21}	fen^{35}
39.	暫 caam^2	can^{51}			

- (f) Describe how the tones and initial voiced consonants have evolved in Mandarin. What rules for reading tones in fanqie transcriptions for Mandarin can be formulated?
- (g) Some combinations of initial consonant and tone are extremely rare in modern Mandarin. Which ones?

More characters, with their readings in both Cantonese and Mandarin, are given below. Some tones have been left out:

	Cantonese	Mandarin		Cantonese	Mandarin
46.	置 $\text{t}^h\text{uŋ}^{\dots\dots}$	$\text{t}^h\text{uŋ}^{35}$	49.	眠 min^{21}	$\text{mian}^{\dots\dots}$
47.	載 coi^3	$\text{cai}^{\dots\dots}$	50.	蛸 $\text{siu}^{\dots\dots}$	xiao^5
48.	米 $\text{mai}^{\dots\dots}$	mi^{214}	51.	亂 $\text{lyn}^{\dots\dots}$	luan^{51}

- (h) Determine what the missing tones are.

- (i) Read the following transcriptions in Cantonese:

52.	梯 ? = 土 t^hou^{35}	★	雞 kai^{53}
53.	嘯 ? = 蘇 sou^{53}	★	弔 tiu^3
54.	浪 ? = 魯 lou^{13}	★	當 toŋ^{53}
55.	憊 ? = 蒲 p^hou^{21}	★	拜 paai^3

- (j) Read the following transcriptions in Mandarin. Some transcriptions cannot be read by themselves, but this problem contains enough information to read all of them:

- | | |
|--------------------------------------|-----------------------------|
| 56. 賽 ? = 先 $\acute{x}ian^5=13A=22X$ | ★ 代 tai^{51} |
| 57. 簡 ? = 古 $ku^{214}=16A$ | ★ 限 $\acute{x}ian^{51}=25B$ |
| 58. 賞 ? = 書 $\acute{s}u^5$ | ★ 兩 $lianj^{214}$ |
| 59. 佰 ? = 普 $p^hu^{214}=31A$ | ★ 乃 nai^{214} |
| 60. 沘 ? = 胡 $xu^{35}=24A$ | ★ 畎 k^hyan^{214} |
| 61. 犬 ? = 苦 k^hu^{214} | ★ 沘 = 60X |
| 62. 下 ? = 胡 $xu^{35}=24A$ | ★ 駕 kia^{51} |
| 63. 捍 ? = 下 = 62X | ★ 赧 nan^{214} |
| 64. 紂 ? = 除 $\acute{c}^hu^{35}$ | ★ 柳 $liou^{214}$ |
| 65. 囊 ? = 奴 nu^{35} | ★ 當 $taŋ^5=32A=54B$ |
| 66. 鰓 ? = 蘇 $su^5=20A=22A=53A$ | ★ 來 lai^{35} |

NB: Mandarin is China's official language, based on the dialect of Beijing. It is spoken by approx. 850 mln people. Wu (Shanghainese) is spoken by 90 mln people, Cantonese (Yue) by 70 mln.

Each Chinese dialect has a fixed number of tones (melodies in one of which every syllable is pronounced). The system proposed by the linguist Yuen Ren Chao, which is used in this problem, denotes five levels of the voice by numbers from 1 (lowest) to 5 (highest) and transcribes the

melody as a succession of levels: a^3 , a^{214} . All the tones you need are present in this problem.

The mark ^h indicates that the preceding stop consonant is aspirated (pronounced with a puff of air). $x = ch$ in Scottish *loch*, $\eta = ng$ in *hang*. $\acute{c} \approx ts$ in *hats* (pronounced as a single consonant), $\acute{s}$ and $\acute{c}$ are hard consonants similar to English *sh* in *shut* and *ch* in *chuck*, $\acute{x}$ and $\acute{k}$ are soft consonants similar to *sh* in *sheet* and *ch* in *cheat*. $\acute{o}$ and $\acute{y} =$ French *eu* and *u* (German *ö* and *ü*).

If you do not want to write Chinese characters, you can refer to them using the number of the transcription where they occur and specifying which character you mean: X (transcribed), A (first in the transcription) or B (second in the transcription).

Note that in the Mandarin reading of character 28A there is no vowel.

—Todor Tchervakov

English text: Todor Tchervakov.

Good luck!