Conal Boyce

贾伯康

CHINESE AS IT IS

A 3D Sound Atlas with First 1000 Characters

Chinese As It Is

A 3D Sound Atlas
with First 1000 Characters[‡]

[‡]Sound Atlas: 1120 characters

First-Year Student Subset: 903 characters ("First 1000")

plus 1274 auxiliary characters, for a grand total of 2394 汉字 in this volume

Conal 贾伯康 Boyce

Table of Contents

LIST OF RADICALS

部首目录

部首索引表

1 Stroke Radicals

一部

一 **92**

+1 stroke
丁 **48**
七 **75**

+2 strokes
才 **42**
干 **52** *(1, 4)*
三 **79**
上 **81**
下 **89**
万 **88**
丈 **96**

+3 strokes
不 **41** *(2, 4)*
丑 **44**
夫 **51**
互 **56**
井 **59**
开 **60**
卅 **79**
天 **86**
屯 **87**
无 **89**
牙 **92**
专 **98**

+4 strokes
东 **48**
平 **74**
且 **75**
正 **97**

+5 以上
亚 **92**
更 **52** *(1, 4)*
来 **63**
两 **64**
求 **76**
事 **82**
歪 **88**
面 **68**
甚 **81**
奏 **99**
孬 **69**

丨部

中 **97**
北 **40**
出 **44**
电 **47**
旧 **59**
串 **45**

丿部

川 **45**

久 **59**
千 **75**
长 **43, 96**
午 **89**
乐 **63**
生 **82**
丢 **48**
尴 **51**
年 **70**
乓 **72**
乒 **74**
卵 **66**
我 **89**
垂 **45**
乖 **53**
靠 **60**

丶部

头 **87**
为 **88**
良 **64**
农 **71**
举 **59**

乙部
（乚乛）

乜 **68**
了 **64**

飞 **50**
也 **92**
习 **89**
书 **82**
以 **92**
民 **68**
买 **66**

2 Stroke Radicals

二部

二 **50**
亏 **62**
云 **94**
些 **90**

十部

古 **53**
华 **56**
卖 **66**
丧 **80**
直 **97**
真 **96**

厂部

厂 **43**
仄 **95**
厦 **80**

匚部

区 **76**
叵 **74**

刂部

刊 **60**
创 **45**
刚 **52**
列 **65**
别 **41** *(2, 4)*
刷 **82** *(1, 4)*
制 **97**
剖 **74**

卜部

卡 **60, 75**
外 **88**

冂部

内 **70**
肉 **78**
同 **87**
网 **88**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

亻部

2-4 strokes

仍	**78**
什	**81**
他	**85**
传	**45**
份	**50**
价	**58**
伟	**88**
休	**91**
优	**93**

5 strokes

伯	**41**
作	**100** *(1, 4)*
但	**47**
低	**47**
佛	**51**
你	**70**
体	**86**
住	**97**

6 strokes

侧	**42**
佩	**73**
使	**82**

7 strokes

便	**41, 73**
俄	**49**
俩	**64**
俗	**84**

信	**90**

8 以上

倒	**47**
候	**56**
借	**58**
停	**86**
债	**96**
假	**58**
偷	**87**
偶	**72**
偏	**73**
傀	**62**
傻	**80**
催	**46**
僧	**80**
像	**90**

八部

八	**39**
分	**50**
半	**40**
共	**53**
兴	**90** *(1, 4)*
兵	**41**
单	**47**
卷	**59** *(3, 4)*
养	**92**
黄	**57**
尊	**100**

人部

人	**78**
入	**78**
个	**52**
从	**45**
今	**59**
令	**65**
会	**57**
全	**76**
伞	**79**
含	**55**
余	**93**
命	**68**
禽	**76**

勹部

勺	**81**
句	**59**
包	**40**
够	**53**

几部

几	**57**
凳	**47**

儿部

儿	**50**
兄	**91**

先	**90**
免	**68**

亠部

六	**65**
玄	**91**
产	**43**
交	**58**
高	**52**
离	**64**
旁	**72**
商	**81**

冫部

冲	**44** *(1, 4)*
次	**45**
决	**59**
冷	**64**
净	**59**
准	**98**
凑	**45**
凛	**65**

冖部

冗	**78**
写	**90**
军	**60**
罕	**55**

冤	**94**

讠部

2-3 strokes

计	**57**
认	**78**
让	**77**
讨	**85**
训	**91**

4 strokes

访	**50**
讽	**51**
讲	**58**
论	**66**
设	**81**
许	**91**

5-6 strokes

诉	**84**
该	**51**
话	**56**
诗	**82**
详	**90**

7 strokes

说	**83**
误	**89**
语	**93**

8 以上

谁	**81**
谆	**98**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

拼	**74**	搂	**65** *(1, 3)*	茶	**43**	**弋部**	**尢部**
挑	**86** *(1, 3)*	搔	**80**	草	**42**		
挖	**88**	提	**86**	荒	**57**	**廾部**	**弋部**
指	**97**			莽	**67**		
拽	**97** *(1, 4)*	10 以上		莫	**68**		

7 strokes		摁	**49**	菜	**42**	**大部**
挨	**39**	搞	**52**	萌	**67**	
换	**56**	摸	**68**	菩	**74**	

捐	**59**	搜	**84**	葵	**62**	大	**46**	**小部**
捆	**62**	摊	**85**	落	**66**	太	**85**	
捞	**63**	摇	**92**	蒿	**55**	夯	**55**	小	**90**
捋	**66**	撇	**73** *(1, 3)*	蓝	**63**	央	**92**	少	**81** *(3, 4)*
捏	**70**	摔	**83**	蒙	**67**	夺	**49**	光	**53**
损	**84**	摘	**96**	蕊	**79**	尖	**58**	省	**82**
挺	**86**	播	**41**	薄	**40, 41**	夸	**61**	党	**47**

8 strokes		撑	**44**	藏	**42**	奔	**40**	
掐	**75**	撅	**59**	蘑	**68**	奉	**51**	
接	**58**	撩	**64**	孽	**70**	美	**67**	
控	**61**	撒	**79** *(1, 3)*			奇	**75**	
描	**68**	操	**42**	**寸部**		套	**85**	
排	**72**	擦	**42**			爽	**83**	
捧	**73**	攘	**69**	寸	**46**			
探	**85**	攒	**46, 94**	对	**49**			
掏	**85**			耐	**69**			
推	**87**	**艹部**						

9 strokes		苍	**42**
揍	**43**	花	**56**
揣	**44** *(1, 3)*	劳	**63**
搓	**46**	苦	**61**
搁	**52**	茄	**75**
		英	**93**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

忄 部

忙	**67**
怀	**56**
快	**62**
怄	**72**
怪	**53**
怕	**72**
恨	**55**
恰	**75**
悔	**57**
惨	**42**
情	**76**
惰	**49**
惯	**53**
愣	**64**
愧	**62**
慢	**67**
懂	**48**
懒	**63**

门部

门	**67**
闩	**83**
闪	**81**
闯	**45**
问	**88**
闷	**67** *(1, 4)*
闲	**90**
闹	**69**

氵 部

2-4 strokes

汉	**55**
污	**89**
沟	**53**
沆	**55**
没	**67**
沁	**76**
汤	**85**
汪	**88**

5 strokes

法	**50**
泡	**73**
浅	**75**
油	**93**
沾	**96**

6 strokes

活	**57**
派	**72**

7 strokes

浮	**51**
海	**54**
浪	**63**
流	**65**
润	**79**
消	**90**

8 以上

混	**57**
涮	**83**

淆	**90**
港	**52**
清	**76**
湍	**87**
湾	**88**
温	**88**
滚	**54**
溜	**65**
满	**67**
漏	**65**
漂	**73** *(1, 3)*
潮	**43**
澳	**39**

宀部

宁	**71**
它	**85**
安	**39**
宅	**96**
宠	**44**
定	**48**
审	**81**
宗	**99**
宣	**91**
宾	**41**
害	**54**
家	**58**
宽	**62**
宰	**94**
寝	**76**

寡	**53**

辶部

边	**41**
过	**54**
迁	**93**
迟	**44**
还	**54, 56**
进	**59**
连	**64**
迎	**93**
远	**94**
运	**94**
这	**96**
进	**40**
迷	**67**
送	**84**
退	**87**
选	**91**
追	**98**
逦	**41**
逞	**44**
逢	**51**
逛	**53**
通	**87**
透	**87**
造	**95**
逼	**40**
道	**47**

彐部

归	**54**
当	**47** *(1, 4)*
寻	**91**

尸部

尺	**44**
层	**43**
局	**59**
尿	**70**
展	**96**

己部
（巳）

弓部

张	**96**
弱	**79**
强	**75**

子部

子	**99**
存	**46**
孙	**84**
孝	**90**
孤	**53**
学	**91**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

孪 66	给 52	弄 65, 71	松 84	横 55 *(2, 4)*
	结 58	玩 88	枕 96	
女部	绕 77	现 90	5 strokes	**犬部**
	统 87	班 40	标 41	
女 71	绷 40	理 64	某 69	犬 76
奶 69	绸 44	望 88	柳 65	状 98
奴 71	绿 66	璀 46	柠 71	臭 44
好 55	绳 82	瑕 89	染 77	
妈 66	续 91		柔 78	**歹部**
如 78	缓 56	**韦部**	6 strokes	
她 85	缩 84		柴 43	歹 46
妹 67		**木部**	梗 52	死 83
耍 82	**马部**		桥 75	残 42
娃 88		木 69	桑 80	
要 92	马 66	1-2 strokes	桃 85	**车部**
娘 70	驮 87	本 40	样 92	
婚 57	驴 66	机 57	桌 98	车 43
婆 74	验 92	杀 80	7 以上	轮 66
嫩 70		朽 91	渠 76	软 78
嫖 73	**幺部**	杂 94	梯 86	转 98 *(3, 4)*
		3 strokes	棒 40	轴 97
纟 部	幺 90	杯 40	棍 54	轰 56
		村 46	棱 64	辍 45
红 56	**巛部**	杠 52	棉 68	
约 94		4 strokes	森 80	**戈部**
纯 45	4 Stroke Radicals	极 57	椰 92	
纵 99		林 65	楷 60	我 89
练 64	**王部**	果 54	楼 65	戳 45
细 89		枪 75	概 51	
组 99	王 88		模 69	
绑 40	主 97			

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

钅部

钮	**71**
铆	**67**
钱	**75**
铁	**86**
钻	**99** *(1, 4)*
锐	**79**
银	**93**
锅	**54**
铺	**74**
错	**46**

矢部

知	**97**
短	**49**
矮	**39**

禾部

和	**55**
科	**61**
秒	**68**
秋	**76**
种	**97** *(3, 4)*
秤	**44**
秘	**67**
租	**99**
移	**92**
稳	**88**

白部

白	**39**
百	**39**

瓜部

瓜	**53**
瓢	**77**

用部

用	**93**
甩	**83**
甬	**40**

鸟部

鸟	**70**
鸭	**92**

疒部

病	**41**
疼	**86**
疵	**45**
痕	**55**
痤	**46**
痛	**87**
瘪	**41**
瘸	**77**
癣	**73**

立部

亲	**76**
站	**96**
竣	**60**
端	**49**

穴部

究	**59**
穷	**76**
空	**61**
窃	**75**
突	**87**
窍	**75**
容	**78**
窄	**96**
窨	**59**
窝	**89**
窗	**45**

衤部

补	**41**
袄	**39**
被	**40**
袍	**73**
袜	**88**
袖	**91**
裤	**61**
裸	**66**

褪	**87**

疋部

皮部

皮	**73**
皱	**97**

矛部

6 Stroke Radicals

耒部

老部
（耂）

老	**63**
考	**60**

耳部

耳	**50**
取	**76**
耸	**84**
聊	**64**
聘	**74**
聪	**45**

臣部

卧	**89**

西部
（覀）

西	**89**
票	**73**

页部

顶	**48**
顺	**83**
预	**93**
领	**65**
颊	**58**
颓	**87**
颠	**47**
颤	**43**

虍部

虎	**56**
虏	**65**
虐	**71**

虫部

虫	**44**
虽	**84**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

虾 **89**
蛤 **54**
蠢 **45**

缶部

缺 **77**

舌部

舌 **81**
乱 **66**
舍 **81**
舔 **86**

竹部

竹 **97**
笨 **40**
答 **46** *(1, 2)*
等 **47**
筐 **62**
筛 **80**
简 **58**
管 **53**
算 **84**
篡 **46**
簪 **94**
纂 **99**

臼部

自部

自 **99**

血部

血 **91**

舟部

衣部

衣 **92**
表 **41**
装 **98**
囊 **69**

羊部

着 **96, 98**
群 **77**
善 **81**

米部

米 **67**
粉 **50**
料 **64**

粗 **46**
精 **59**
糊 **56**
粥 **97**
糖 **85**
糟 **95**

艮部

羽部

翁 **89**
翻 **50**

糸部

紧 **59**
累 **63** *(3, 4)*

7 Stroke Radicals

麦部

走部

走 **99**
起 **75**
超 **43**
趁 **44**
越 **94**

赤部

赤 **44**

豆部

豆 **48**

酉部

酥 **84**
酪 **68**
酿 **70**
酸 **84**
醋 **46**
醒 **90**

辰部

卤部

里部

量 **64**

足部

足 **99**
趴 **72**
跫 **49**

跋 **41**
跑 **73**
跟 **52**
跨 **61**
路 **65**
跳 **86**
踏 **85**
踹 **44**
蹭 **43**
蹲 **49**

身部

身 **81**
躲 **49**
躺 **85**

采部

谷部

豸部

貌 **67**

角部

解 **58**

Legend: Annotations such as '*(2, 4)*' and '(dōu, dū)' are explained in the Notes on page **xviii**.

言部

（言 ）

言	**92**
譬	**73**

辛部

辣	**63**

8 Stroke Radicals

青部

雨部

雪	**91**
雷	**63**
零	**65**
需	**91**
霓	**70**

齿部

黾部

隹部

雅	**92**

雕	**48**

金部

鱼部

9 Stroke Radicals

革部

革	**52**
勒	**63**
靴	**91**
鞋	**90**

骨部

骰	**80**
髓	**84**

鬼部

鬼	**54**

食部

音部

10 以上

髟部

鬓	**41**

麻部

麻	**66**

鹿部

黑部

黑	**55**

鼠部

鼠	**82**

鼻部

鼻	**40**
鼾	**55**

Notes:

Split readings may occur within the same row of the matrix, on different rows, or on different pages. To indicate a split reading, one of the following three notations is used:

1. A pair of tone numbers if both occur in the same row. Ex: 干 *(1, 4)*
2. A pair of romanizations if occurring in different rows. Ex: 色 (sè, shǎi)
3. A pair of page references if occurring on different pages. Ex: 长 **43, 96**

A few of the indexed items happen to be "secondary" entries in the dictionary:

一 (one), 她 (she), 它 (it).

我 under 戈 is a "courtesy" entry in the index. Its "new radical" is this: 丿 .

The slang character 肏 I've sorted into the 月 bin by analogy with 氽 in the 水 bin.

PROLOGUE

Chinese As It Is endeavors to carve the subject "at its joints."[1] To accomplish this, we will use three unusual devices in formulating our dictionary:

1. The dictionary is presented as a matrix with four columns and 396 rows. The four columns correspond to the four tones of Mandarin. They help combat the westerner's proclivity for relegating a character's tonal identity to the status of "something extra that I should learn." The four columns provide a degree of *situational awareness* about the tonal landscape.

2. The romanized heading for each cell of the matrix is color coded to indicate *tonal weighting*. For example, there is a SINGLE place in the whole language where jiao acquires second tone, but one is faced with a PLETHORA of fourth-tone jiao words to differentiate; thus, to second-tone jiáo I apply one color and to fourth-tone jiào a different color. Given this second layer of information, as provided by a four-part color coding scheme, a student has the opportunity for yet more *situational awareness* of the tonal landscape, and a fighting chance of glimpsing the terrain as experienced by a native speaker.

3. A typographical convention is used to discriminate between (a) characters that work as *stand-alone* words, and (b) characters that are merely *halves* of words, in the way that *wal* alone is not an intelligible stand-in for *walnut*. For more about this, see **The Hazelnut Fallacy** on page **8**.

Regarding the title and cover graphic: Tonal weighting (item 2 above) is what makes the dictionary 3-dimensional. The cover art is based on rows 283–286 of the actual matrix (which occur on pages **81–82** below). The cells depicted on the cover differ from the actual cells in the following two ways:

(a) All romanizations and definitions are excluded, for aesthetic considerations.
(b) The graphic contains curved lines and a suggestion of 3-D objects that can rise above

---

1. This rather Chinese-sounding phrase is a favorite of occidental philosophy professors. (The author learned it from his daughter, whose B.A. was in that field.) Apparently it originates with Plato, but for us it will surely recall Butcher Ding and his method of carving oxen, as related in the Zhuāngzǐ, Chapter III.

(RED) or sink below (GREEN) a notional sea level (represented by Hànzì with BLACK type face).

Whimsy aside, the graphic does provide a glimpse of how the whole 62-page table will be perceived, once the user of the dictionary has invested the small effort required to learn the meaning of the color coding and typographic conventions. Note in passing that there is an implicit *fourth* color, as promised, represented by the three BLANK cells in the second row of the graphic.

There is one final aspect of the color coding to mention: Per the established convention, **blue** denotes a hyperlink (in the PDF of this volume). Note that the hyperlinked page numbers in the **Radical Index** on pages **vii–xviii** are more than a navigational convenience. They allow each of the 1120 entries in the **Radical Index** to double as a character flash card.

Scope: There are several ways to judge the size of this dictionary. Depending on one's perspective and interests, the count is anywhere from 2394 characters down to 903 characters. The various numbers are explained in the annotations to **Figure 2** on page **18** and **Figure 3** on page **24**. A suitable nickname for the dictionary might be "The 903" (nine-oh-three) since it contains a clearly marked subset of **903** characters for the beginner, and is thus also a species of "First 1000 Chinese Characters."

For the general reader (i.e., a beginning or intermediate student of the language)

THE PROLOGUE ENDS HERE

In early drafts, I referred to this work as *A Shared Topography for Written and Spoken Chinese: 1171 lexemes loaded to the morphemic space*. Reacting to that title, one might have asked: "To make a *conventional* dictionary, when we round up vocabulary items and sort them A-to-Z by pronunciation, isn't that *already* a kind of merger of the lexicon and sound system, in one 'shared space'?" My response: "Yes, but where Chinese is concerned, a dictionary in A-to-Z format lets us down: Its linear contents need to be decanted into a curved bottle, so to speak." And this is only one of various issues I address in this work, most of which are of the following variety: so obvious they are easily missed. Please bear with me while I use a page or two to try articulating them.

A discipline such as linguistics or particle physics excels at breaking down the object of study into its constituent parts. In principle, the discipline should also offer guidance in the opposite direction: some notion of building back *up* to the entity that was analyzed to death, whether that be the fleeting spoken word or a living atom, robbed of its eternity by being smashed in an accelerator. In practice, the 'building up' gets short shrift. (For more about this, see **Appendix A: Aufbau (Building Up)**, page **101**.)

For Chinese specifically, there are two special problems or distractions beyond the standard breaking-down activities of grammar, morphology, and phonology, i.e., beyond what we would expect for, say, Quechua or Romanian. Those problems are

- Tones
- Characters

In the table that follows, I've placed those two topics in a broader linguistic context. In the first column, on a scale of 1 to 10, I assign impressionistic "grades" to the various topics within linguistics. In the second column, I juxtapose the student's perception of which topics are prominent, and which remain in the background, in his/her subjective experience of Chinese.

The takeaway is the gross mismatch (four circled numbers) between the two sets of rankings, the details of which one may dispute without making a dent in the overall picture that emerges:

		The Linguist's Relative Success at Theorizing the Item	Item's Subjective Prominence in the Eyes of the Student
Phonology	Tones & Stress	③	⑨
	Initials & Finals	8	7
Morphology		6	5
Grammar and Syntax		4	3
Hànzì (Chinese characters)		①	⑩

In short, the two features of Chinese most conspicuous to the student, filling up, say, two-thirds of the student's field of vision, are precisely those two topics where the linguistics establishment is the weakest.

Yes, after a fashion, linguistics has handled the challenge of the tones "successfully," bolting them on as a cerebral afterthought to mainstream phonology. While some students are able to handle the tones by jumping into the deep end and swimming, most students need (or think they need) an analytical approach, which, if done thoroughly, would entail these six steps:

[i] learning (or reviewing and really understanding for the first time) *what* a phoneme is;

[ii] learning how even a tone can play the *role of* phoneme;

[iii] learning the tonal *contours* themselves;

[iv] immediately *un*learning some tones, in deference to the sandhi rules and neutralization principles;

[v] learning how to make tones an *integral* part of each vocabulary item in one's long-term memory;

[vi] learning to apply *stress* as an overlay to all the above (as 2-1 or 2-3-1 or 2-3-3-1, after Chao [1968a] p. 35, where the three stress levels are analyzed as allophones of a single predictable stress phoneme).

Not only is this standard approach daunting but it is myopic, as a fixation on "the four tones of Mandarin" provides not the slightest hint of the larger context where we find the tones of a "real" Chinese language such as Yue 粤 [Cantonese] following the old *eight*fold path of píngshǎngqùrù 平上去入 , traversed as a yīn series, then as a yáng series. (My heretical opinion: If students had a clear understanding of this context early on, a significant number of them would probably elect to use Hànyǔ only as a gateway subject, leading directly into

Wu, Min or Yue, these three supposed "dialects" being in fact the only real Chinese languages extant; for more about this, see **Appendix B: Dialects** on page **105**.)

The other topic that looms large on the horizon of the student is Hànzì (alias *kanji*, in Sino-Japanese). But from the hard-line linguist-purist viewpoint, Hànzì have scarcely any existence, save as 50,000 scraps of hieroglyphic noise, threatening to distract students from the real thing, the Sacred Sound System. Thus, Hànzì become, from the outset, the elephant on the table, since the *student* will not let them go, never mind what the linguists' party-line may be: To the student (as to many natives), the characters *are* the language. Thus, when students think about dictionaries they tend to cast phonology, morphology and grammar — even lexicology as a topic in linguistics — to the four winds, focusing instead on various low-level logistical concerns which might collectively (and facetiously) be termed 'dictionary-ology' (to distinguish this dictionary-*user* viewpoint from that of the dictionary-*maker* or lexicographer):

How *many* individual Hànzì are covered by this dictionary I'm about to buy? How *many* compound entries? What *kind* of characters are used, traditional or simplified? How are the characters and their compounds *sequenced* — by romanization; by *BoPoMoFo*; by four-corner code? How well are the *split readings* (pòdú 破读) handled, as between cháng 长 and zhǎng 长 , for instance? What kind of *indexing*, if any, is provided — via traditional radicals; new radicals; apparent radicals;[2] stroke count; four-corner method? And so on.

Then this *too* becomes a world unto itself. (And, up to a point, rightly so: It's a big subject. In the ensuing pages we won't neglect it, only try to rein in some of the digressions in that direction.)

---

2.   Split readings are discussed in Chao (1968a), p. 172; for apparent radicals, see *Ibid.* p. xxv–xxvii.

Magnifying these internal forces that may draw one into addictive byways of the language (the tones, the characters) are external "anthropological" forces which must also be managed; e.g., the pressure in academia to carve out a specialty;[3] or, a feeling among the Chinese, often picked up and shared by the foreign student, that Chinese literacy — knowing how to read and write *characters* — is tantamount to Chinese *education*;[4] and so on.

Yet all the while, far beneath the surface eddies of worry over this topic-comment structure or that elusive radical or that tricky sandhi pattern, there is simply one thing: *the Chinese language*. Or, more properly, the 'one thing' of interest should be understood as the great Chinese Language *Family*, inclusive of Wu 吳 , Min 閩 , Yue 粵 and others, surely a force of Nature if ever there was one, running deep and silent through a billion-plus dreams each night.

Back to the topic at hand: 'a shared topography' for Hànyǔ. My original aim was simply to build a modified kind of dictionary in which the tones would enjoy better representation than usual. (For more about this aspect, see section [b] on page **117**.) But the materials soon forced me to give equal attention to the realm of morphology, lexicology and Hànzì; thus, a new kind of shared space with some functional resemblance to a conventional dictionary, but using a different format. (One might conclude that this approach was motivated by a desire to address the disparities shown in the Linguist/Student table on page **4**. However, the Linguist/Student table came very late in the project, and was therefore not the impetus it might appear to be, rather a kind of "validation," if anything, of the approach already adopted.)

---

3. A typical result of such forces would be my own article on 'Min sandhi in verse recitation', written several decades ago when I was still "part of the establishment" (Boyce, 1980).

4. In one of the 'Cultural Notes' in *Speaking Chinese in China*, the authors state flatly that "How a person makes his characters has importance far beyond his need to communicate," and they go so far as to say that Mao TseTung's accent, equivalent to Cockney or Brooklynese, would have registered with his constituency as a nonevent; quoted and paraphrased from Hsu and Brown (1983), p. 288. I think they've overstated their case, but still it is interesting food for thought. In a different but related lesson about the primacy of the character, note the discussion of gā and gá on page **19** below, where it would be literally impossible to talk only about our several *words* of interest, exclusive of the *character* 嘎 that works as a de facto linchpin binding them all together from above, as it were. From such examples, one might easily conclude that when push comes to shove, the venerable character really *is* king. But see also footnote **24** on page **32** below.

A particular kind of sharing and unity is one of the novelties[5] I claim for the work. But this aspect is not fully defined until we look also at 'morphemic space', next.

As noted above, the original subtitle for this work included the term 'morphemic space'. That was intended as shorthand for the following:

> Take the 400-odd headings in your typical medium-sized Chinese-English dictionary, multiply them out by the four tones of pǔtōnghuà (alias Mandarin alias Hànyǔ), and you have a grid of 1600 cells. Toss out the *numerous* nonsense combinations such as 'min' in fourth tone, and the remaining 1100-odd cells constitute a large 'canvas' for forming words, the place where all the valid morphemes must reside.

The actual numbers that we will be working with are 396 multiplied by 4 to give 1584. Above I've rounded off the numbers to 400 and 1600, as the intent is just to give the general flavor. (For more about this, please refer to **Appendix D: Defining the Morphemic Space**, page **117**.)

Regarding the word 'space': Out of the total number of cells available (as described above), I populate 1171 cells, in ways that may suggest varying "depth" of the data they contain: Sometimes a cell contains a single-character word (alias monosyllabic word); sometimes a cell is populated by a single character whose use is illustrated by a *two*-character word; possibly a single character will be accompanied by as many as 19 small-font characters signalling an 'overloaded' cell; and so on. Saying it another way, each cell is assigned a kind of "tonal weighting." And, as mentioned earlier, this is what gives the dictionary a notional third dimension.

Note that the term 'morphemic space' denotes our "approved" but as yet *un*populated table. To refer to the *populated* table, I use the term 'matrix'. The matrix begins on page **39** and runs through page **100**. This word 'matrix' is admittedly overworked, but it has the virtue of being short and evocative on its own, without much need for qualifiers or footnotes.

---

5. By the law of large numbers, it seems probable that someone else has written a similar essay on Hànyǔ, but I personally have not yet encountered such. In the meantime, it is interesting to note that "obvious" and "original" (or "novel") are not mutually exclusive: Even today, one of the most startling moments in all of western music is when the string section starts trudging up and down the chromatic scale, halfway into the first movement of Berlioz' *Symphonie fantastique*. At the same time one thinks, "But this is also such an *obvious* card to play. It's just that no one else did it before him." Like it or not, the idea is "uniquely" his, even though it has surely been conceived in the head of many a four-year old while pounding on a keyboard. Similarly, my treatment of Chinese is possibly unique, but it is also quite obvious; indeed, an important part of my thesis is that this way of looking at Chinese is not arbitrary or personal; rather, it is absolutely natural and *inevitable*, and in that sense, "obvious."

<u>The Hazelnut Fallacy</u>

Suppose an American children's dictionary contained the following two entries, accompanied by a picture of fancy mixed nuts:

hazel: something crunchy and delicious; can be added to chocolate bars
wal: something crunchy and delicious; can be added to oatmeal cookies

Putting the picture and text together, an adult perusing the dictionary would infer that these were the entries for 'hazelnut' and 'walnut', respectively; but she would refrain from buying such a book as it could only fill her child's mind with bewildering half-truths, not to say nonsense. Who would foist such a volume on the public? you ask. And yet, in the haphazard tradition of Chinese pedagogy in the west, nearly *all* the books we see with titles such as "First 1000 Chinese Characters" contain examples every bit as ridiculous as 'wal' trotted out as an English word. Along with its 3-dimensional approach to the tonal landscape, what sets this volume apart (when taken in its role as "First 1000 Chinese Characters") is its scrupulous attention to morphology — its refusal to pass off the numerous Mandarin analogues to 'wal' as independent words in the spoken language. This layer of information is handled by a simple typographical convention (introduced on page **22**), not by the color coding mentioned earlier.

This brings us to the phrase 'lexemes loaded to ...' which was also part of the original subtitle. That phrase was intended as shorthand for the following:

We are not simply going to paste Chinese characters into the cells of the empty table (our morphemic space) like so many stamps onto the blank pages of a collector's album; rather, we are going to carefully consider and *process* each entry from a lexicological perspective. Briefly, this means asking: "Does this *character* work as a monosyllabic *word*? If not, which character should follow or precede it to form a two-syllable word that *is* a valid utterance?"

I realize this business about Chinese being not *always* monosyllabic is only a rehash of the warnings in Chao (1968a) p. 143f., and in Packard (2000) *passim*.[6] But the individual zì 字 *is* seductive, and the myth dies hard. (This discussion is continued in **Appendix K: Free and Bound Morphemes** and **Appendix L: Muchengxue and the Monosyllabic Mystique**.) Thus, it is important to keep fighting the misconception on all fronts.

The pledge above to honor *lexemic* loading[7] (in preference to "stamp-album pasting") is my way of carrying the battle forward, while attempting also a kind of *Aufbau* (Building Up), as alluded to earlier.

A twist in the bookkeeping scheme: 1171 is both the number of lexemes *and* the basic character count for this project. On page **16**, I explain how I arrived at the number '1171' itself, and on page **32** I explain the curve I've thrown for counting lexemes vs. characters, such that '1171', paradoxically, is the tally for both. Meantime, for pedagogical purposes, there are still other ways to count the characters in this project, the high number being 2394, and the low number 903. (For an overview of these various numbers, see **Figure 2** on page **18**.)

It is now high time that we move on to the project itself. Since 一画胜千言 (one picture beats a thousand words), let's begin with a graphical, sneak preview of the matrix:

---

6. "... when we say a morpheme is free we mean it is sometimes free, whereas if a morpheme is bound it is always bound"; Chao (1968a) p. 144. *Language and Symbolic Systems* includes a discussion of free and bound *forms* [of words, in languages generally]; Chao (1968b) p. 53. In *Mandarin Primer*, there is a section that discusses 'morphemes' as distinct from 'syntactic words'; also, 'free *words*' versus 'bound *words*'; Chao (1948) p. 33, emphasis added. <u>Comment</u>: In substituting the terms 'free words' and 'bound words' for 'free morphemes' and 'bound morphemes' in the 1948 work, Chao probably sought language that would be agreeable to the student, as distinct from his fellow linguists. But for me, the substituted terms don't work. One appreciates the good intention behind them, but they entail some level-shifting sleight of hand that only hinders comprehension rather than helping it (because 'free' modifies *all* of 'word' while 'bound' implicitly modifies something *inside* of 'word', namely its two constituent halves).

7. Long after developing this structure (in relative isolation), it occurred to me that the more common way to talk about such terrain would be to say, "The [loaded] *lexicon* [is] the inventory of the morphemes of a language," as in Chao (1968b) p. 57. Not to be contrary, but my view of the terrain happens to work the other way around: "Let's use the loaded morphemic space as the inventory of lexemes of the language." It's another way of saying the same thing.
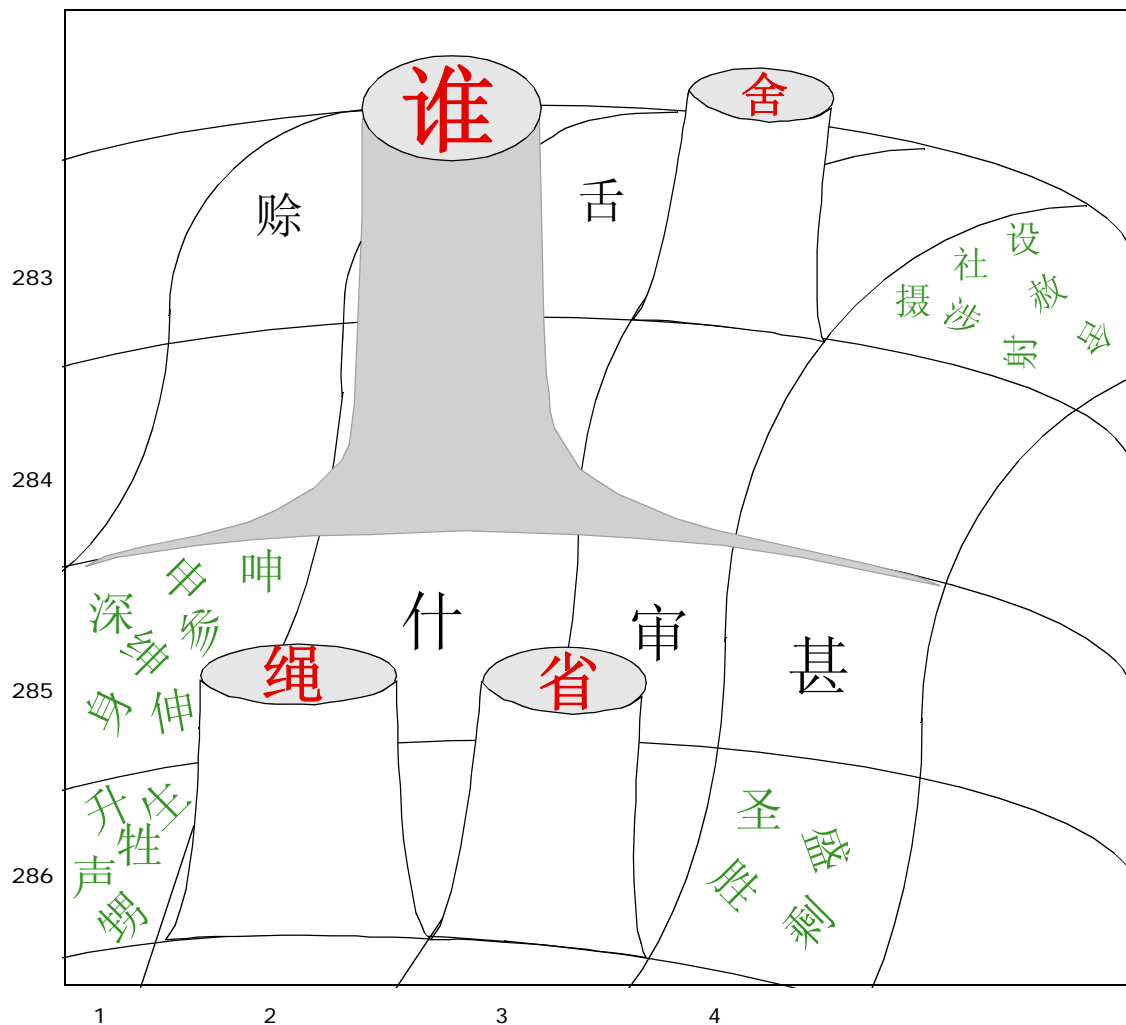
FIGURE 1: First glimpse of planet Hànyǔ

Figure 1 above is a fanciful 'zoom' on the following 16 cells, excerpted from the matrix (at page **81**), *sans* characters and definitions. This random snippet, boiled down to its color coded

skeleton, is meant to attune one to the implied "topographical" features that are present *throughout* the whole matrix:

Row#	Tone 1	2	3	4
283	shē	shé	**SHĚ**	**shè**
284	---	**SHÉI**	---	---
285	**shēn**	shén	shěn	shèn
286	**shēng**	**SHÉNG**	**SHĚNG**	**shèng**

In **Figure 1**, these same four rows are represented as notional lines of latitude, with tones 1 through 4 distributed horizontally, something like meridians for an imaginary topographical surface. With pseudo-geographic mountains, we endeavor to suggest an implied third dimension of the ostensibly flat table. (The domination by **SHÉI** 谁 of one whole 'latitude' in **Figure 1** is deliberate, and will be explained in due course.) As for the color coding, for now one may think of the black characters as being "at sea level"; the red ones "above sea level"; and the burgeoning green ones as "filling little depressions below sea level." A less whimsical role for the color coding will be explained shortly.

USING THE DICTIONARY: STUDENT'S VIEW AND SCHOLAR'S VIEW

The matrix that runs from page **39** through page **100** is potentially useful to both students and scholars (specialists). In the **Prologue**, I only allude to these two different ways of using the dictionary. Here they will be spelled out, each in turn:

The student's view

Let's say this notion of pouring flat vocabulary items into a tone-centric 'vessel with curves' may be novel or even original (as discussed in footnote **5**), but would it be *useful* to a student?

Yes. One advantage to the student would be a heightened awareness of where one is in the tonal dimension, since every entry falls in a column labeled 1st, 2nd, 3rd or 4th Tone.[8] Moreover, the tonal identity of a given character is always reinforced by that of the neighboring characters above and below, along the y-axis, as it were. So, in a very simple but effective way, the matrix can help bring the four tones to life. Used this way, it may be regarded as a kind of Chinese Sound Atlas (this being one of my earliest names for the whole project).

For those with a good ear, I believe the tone-centric (or tone-enforcing) aspect of the matrix will have direct appeal. But what about those students who are made uncomfortable by the tones in Chinese, which are then just "those extra inflectional things I'm supposed to attempt"? To those students I would say, "but tones can be your friend, too" in the following sense: Suppose you wanted to learn the phrase zhōngyīng yǒuhǎo.de cháoliú ("the trend of good relations between China and Britain").[9] The phrase would be learned more

---

8. For more about this, see **Appendix D: Defining the Morphemic Space**, page **117**.

9. Source: Circa 02/02/09, after pausing four seconds to process the fact that a British protester had hurled a shoe in his general direction, Premier Wēn Jiābǎo said calmly to his audience at Cambridge University: 老师们，同学们，这种卑鄙的伎俩阻挡不了中英友好的潮流 Lǎoshī.men, tóngxué.men, zhèzhǒng bēibǐ.de jìliǎng zǔdǎng.buliǎo zhōng-yīng yǒuhǎo.de cháoliú "[My fellow] teachers and students, this sort of base trick [伎俩] cannot halt the progress [lit. 'trend' 潮流] of Chinese-British friendship."(I've included "My fellow" based on his tone of voice, not the actual words.)

quickly and would be more likely to stick in one's memory if one happened to notice its tonal pattern, and distilled it to this: '11-33-22' (setting aside .de with the neutral tone, and the tone sandhi on yǒuhǎo, which results ultimately in '33' being modified to '23'). Even as an abstraction (i.e., for someone with a tin ear who is resigned to bungling many of the tonal contours), I should think this might serve as a handy framework on which to hang the words, and so much the better if one has a good ear: then '11-33-22' becomes a kind of jingle to strengthen one's memory.

With numbers such as the following bandied about, the beginner can find it difficult simply to get a handle on "What *is* Chinese?" never mind "Is this the right major for me?":

"Here's a book called *501 Essential Chinese Characters*"

"But they say you need two to three thousand characters for literacy ..."

Item: There are 4,500 characters in the *Zhongda* dictionary of 1999

Item: There are 5,000 characters in the C.H. Fenn dictionary of 1926

Item: There are 7,773 characters in the Mathews dictionary of 1931

Item: There are 20,000 characters [est.] in the *Gwoyeu Tsyrdean* of 1947

Item: There are 40,545 characters in the *Kāngxī Dictionary* of 1716

Item: There are 50,000 characters in the Morohashi[10]

Shall one conclude that the domain is finite or infinite, so to say? Might "Chinese major" be a quixotic path, rather along the lines of "Fools rush in where Angels fear to tread"? Speaking of angels, if Morohashi Tetsuji finished his thirteen-volume labor of love in 1960, when did he *begin*, I wonder? (Answer: He began work on the *Dai Kan-Wa jiten* in 1919.)

And so on. These are some of the questions that might arise.

---

10. Sources: The *Zhongda* total is on the back cover of that dictionary. Mathews: The sequence terminates at 7773 on p. 1164 (but the actual total is 7785 per p. VI). *Gwoyeu Tsyrdean* total: Sampling 40 pages across its 4 volumes, I obtained an average of 4.3 character-headings per page; then 4485 pages times 4.3 characters/page = 19,285, which I've rounded to 20,000. *Kangxi Zidian* 康熙字典 total: cited from Chao (1948), p. 15. The final figure, 50,000, is from tribal knowledge; the actual number of head entries appears to be 49,964 per the Wikipedia entry for 'Dai Kan-Wa jiten'. For the record, Chiang Yee (p. 24) weighs in with "half-a-million" as his estimate. His number (equivalent to *ten* Morohashis!) thus becomes an interesting bit of sinological trivia, but one cannot take it seriously, so I do not cite it above. (This faux pas aside, Chiang Yee's monograph on Chinese calligraphy is a valuable resource; see page **115** below, e.g.)

And somewhere in this wilderness, the voice of reason? Yes. Contemplating, in a *practical* way, *all* of the Chinese language — both ancient and modern together — the famous linguist Y.R. Chao once had this to say:

> As a subjective estimate, I should say that some 3,500 Basic Characters ... would probably be a fair representation of the content of the Chinese language.   — Chao (1948), p. 16

That sounds reasonable. As of 1948, it might even have stood as the last word on the subject. But see **Appendix C: Simplified Chinese and the Second Law** on page **109**, where I explore the idea that for us regular folks, as distinct from bureaucrats who launch "simplification" campaigns, the Chinese character count can only ever increase, just like the amount of entropy in the universe. (And ditto in the matter of 214 traditional radicals nowadays simplified to 182 "new" radicals, plus or minus a few per the whim of a moment. An instance of this travesty, to which I've acquiesced in grin-and-bear-it mode, can be observed at the front of this volume. There's nothing "new" about it; the whole scheme is, in fact, *medieval*.)

Given the vagaries of the character count as sampled above, for *some* students the notion of our matrix could be comforting, the reaction being: "Here is something that resonates with Y.R. Chao's assessment: A domain that has a *boundary* of some kind! The suggestion of something finite, if vast!" Granted, for some other students, a glance at our matrix, 396 rows deep, might only feel like a glimpse into the abyss, confirming all their worst fears about the language. So be it. There are plenty of alternative avenues for approaching the language.

Viewed from a distance, our matrix might also remind one of the many books entitled *My First 1000 Chinese Characters*, and such. Certainly it is organized in such a way (alphabetically by PīnYīn) that one can quickly look up the definition for any of its 1171 primary entries (which includes a core of 903 characters suitable for the beginner). But there the resemblance ends. In books with titles such as *My First 1000 Characters* or *501 Essential Chinese Characters*, one often finds a character listed as if it were a monosyllabic word, accompanied perhaps by some etymological nugget that is often bungled in the telling, when what the student *needs* to know is that (very often) s/he is looking at only *one half* of a syntactic word, a valid utterance. By contrast, our matrix is at pains to show the missing half of a word whenever needed. On the downside, our matrix (by its nature) is *somewhat* constrained in the matter of presenting reasonably high-frequency characters to the student that won't "waste his or her time." For instance, in one of the more obscure corners of the realm we find the word zuǎn 纂 'to compile, to edit'. Would the first-year student think it worthwhile learning

that word (and *that* 20-stroke character, which — as of 2010 at least — has no simplified form)? Probably not.

To address this issue, here is my recommendation for segregating the advanced items, to get at the core of rudimentary material: In using the matrix, simply skip over ("read past") the 217 cells that I've flagged with *double borders*, and which I list out here for quick reference as well:

```
aī aí ǎng àng āo áo ǎo ào bāi bǎng bàng bēng béng běng bèng biē biě bìn
bò bū càn cāng cáo cào cēn cén cèng chān chě chēn chěng chèng chǒng chòng
chuāi chuǎi chuài chuǎn chuǎng chuò cī cuān cuán cuàn cuǐ cuó diān dǔn
duò ē fěi fén gà gàng gěng gǒng gǔn hā há hǎ hāi hān hāng hàng hāo háo
hén hēng hèng hōng hǒng hòng hǒu huǎn huǎng huàng huī hún hùn jiá jiáo jū
juān juē jùn kēng kōu kuā kuǎ kuāng kuí kuǐ kuì kūn kǔn lǎ làn lǎng lāo
lào lēi léng lèng liāo liē liě līn lǐn lìn lǒng lòng lōu lǔ luán luǎn lūn
luō mǎng mǎo mēng méng mī miāo miē mǐng miù náng nǎng nāo náo něi ní niǎn
niàng niē niè niù nú nù nüè nuò ó ǒ ò ōu pā pēi pèn pǐ piáo piǎo pìn pǒ
pū qiā qiǎ qià qiàng qiào qǐn qìn qú qué rāng ráng rě ruǐ ruì sà sài shá
shāi shǎi shē shuà shuāi shuàn shǔn shuò sǒng sū tiè tuān tuí tún tùn tuó
tuò wǎi wāng wēng wèng xiá xiáo xiǔ xuàn xūn yē yōng yū yǔn zā zǎi zān
zǎn záo zè zèn zhǎ zhǒu zhòu zhuì zhūn zòng zuān zuǎn zuàn
```

Character count details: From 1171, eliminate the double readings to get a unique character count: 1120.[11] From that group, subtract the 217 characters represented immediately above: 1120 – 217 = 903 net characters suitable for the beginning/intermediate level. Thus, as a book within the book, we have defined a version of "My First 1000 Characters" (literally 903 in our case, as depicted in **Figure 2** on page **18**). So much for the *what*; as for *why* I allow those 217 items into the matrix in the first place, please refer to "**The Boundary Problem: inclusion/ exclusion rules for the matrix**" on page **19**. That's a much longer story.

---

11. Miscellaneous bookkeeping notes: [1] Within the corpus of syllables that are *not* excluded, there are cases such as bié/biè and dā/dá where I allowed the same character to occupy two cells, because the split reading (pòdú 破读) in question seemed important to me (more important than forcing variety on the two cells). After taking all such double-counted characters into account, we arrive at the number 1120 as the number of *unique* characters represented by the 1171 cell headings. [2] In its meaning of 'old, dated', the character chén is not high-frequency, yet I do not place it on the list of exclusions. Why? Because Chén is a common surname. [3] In its reading lòng ('lane, alley'), I exclude 弄, but in its reading nòng ('to make, play with, fiddle with') I keep it. [4] The word ruǐ 'pistil' I place on the list of exclusions; however, regarding its (new) borderline status, see discussion on page **142**. [5] It may seem odd at first that I would exclude haó, huī and xiá, since these three are headings for *overloaded* cells, but in my rough estimation, *all* the characters associated with these cells may be regarded as somewhat "advanced," hence the blanket exclusions.

Does our '903' strike you as a fairly respectable number for the foreign student to tackle? If so, savor the moment! The bubble will burst as soon as you stumble upon a booklet with toddlers depicted on the cover and a title such as 学前 600 字 or *[Your] Preschool [Child's First] Six Hundred Characters*. It appears that our list of 903, aimed presumably at adolescents or adults, will put its user only a few hundred characters ahead of the Chinese toddler ...

In summary, for the beginning student this tool may serve as the list of My First 1000 Characters (figuratively speaking), but it is far *more* than a list since it pays heed to the distinction between free and bound morphemes; and, as a kind of Chinese Sound Atlas, it fosters a sense of how the lexicon is distributed across the tonal landscape. (Just don't imagine that 903 counts as a very "large" number of characters.)

The scholar's perspective on the dictionary

An "object in nature" will be our shorthand for the second potential use of the matrix. The reference frame is now the thinner air of sinology, a place for doing heuristics, in contrast to the rough-and-tumble of Chinese language pedagogy. Let's begin by (conceptually) rebuilding the whole matrix from the ground up:

In the tried and true tradition of sinology, before we can have syllables, we must define their building blocks, called 'initials' and 'finals'. I subdivide the latter into 'plain finals' and 'compound finals'. To produce the compound finals, we start by taking the cartesian product of two series which yields 39. But of those 39 theoretical compound finals, only 22 are real.[12] Next, multiplying the 21 shēngmǔ (initials) by all 39 of the yùnmǔ (finals) we get 819 theoretical combinations, but of those 819, only 396 emerge as actual syllables in the language, after we eliminate the numerous nonsense cases[13] such as f + –iang => @fiang or s + –ei => @sei. (So, this is another derivation of '396', perhaps a better way to understand "what the number means" than simply paging through a dictionary, A to Z, and counting that many romanized headings — plus or minus a half-dozen — as mentioned in the **Prologue**.)

We then multiply those 396 syllables by 4 tones to form a *potential* morphemic space comprising 1584 cells. But of these, we exclude 290 cases of 'absolute nonsense'; then we also

---

12. For details, please refer to **Appendix J: Finals in BoPoMoFo, à la Gwoyeu Tsyrdean, with PinYin Updates**, page **161**.

13. I use the '@' symbol to flag a nonexistent form, saving '*' for other uses. (The convention is to use '*' in this role.)

exclude 123 cases of 'judgement-call nonsense' (i.e., rare or archaic or recondite cases) to arrive finally at 1171.[14] At this stage, we sort the 396 rows of the grid by PīnYīn alphabetical order, then populate the 1171 cells that are the "real" morphemes of Chinese, leaving 290 + 123 = 413 cells empty. (With respect to our "object in nature" theme, the ideal sorting order for the matrix would surely be *BoPoMoFo*, but that would be off-putting to too many readers; bowing to practicality, I use the PīnYīn sorting order instead of *BoPoMoFo* sorting order.)

---

14. Time for a reality check. Here is a point of comparison: "... Mandarin has about 1,300 different syllables (counting tones)"; Chao (1948), p. 14. Reconciliation: If we add back my 123 cases of 'judgement-call nonsense', the total comes to 1171 + 123 = 1294, which closes the gap with Chao's "about 1,300." In a later work, he speaks of "1277 syllables"; Chao [1968a], p. 185, which bolsters my independently calculated number from the other side. All is well.

Figure 2 presents a high-level view of the project in terms of character counts and pedagogical levels. This graphic also provides context for the Boundary Problem section to follow, on page **19**f. (In that connection, see also **Figure 3** on page **24**.)

Total cells in matrix: 4 columns x 396 rows = 1584

Outer crescent: 1584 – 1171 = 413 empty cells
(These 413 cells encompass two flavors of 'nonsense'.)

1584

The dashed circle represents the Boundary Problem discussed next, on page **19**f.

1171

1171 is the number of *populated cells*

1120

1120 is the *unique* character count: Namely, a subset of the 1171 populated cells, adjusted for 51 cases of "broken reading" [pòdú], such as bú, bù, sè, shǎi, etc.

903

Core list of 903 characters for beginning students. This is a subset of the 1120 unique characters, after elimination of the 217 items judged to be "intermediate or advanced or obscure," i.e., those listed on page **15** and flagged by double borders around their cells.

1274

Auxiliary characters in small font, i.e., all of those appearing at the bottom of the various green cells. These characters carry no definition; neither are they represented in the Radical Index. In that sense, they are 'auxiliary'. As for their pedagogical level, this group of 1274 is a mixture of beginning and intermediate level characters. See discussion on pages **28**–**29**.

High number for total character count in this volume:
1120 + 1274 = **2394**
(This is represented graphically as the sum of the two shaded circles, one of which is mostly hidden.)

Technical note: Is the above a Venn diagram? It is, but it does not map well to the matrix itself. Why not? Strictly speaking, the 1274 'auxiliary' characters reside *inside* the 1584-cell matrix, but from a character count perspective I needed to show them as a separate circle, seemingly "outside" the other four circles, the largest of which represents the matrix.
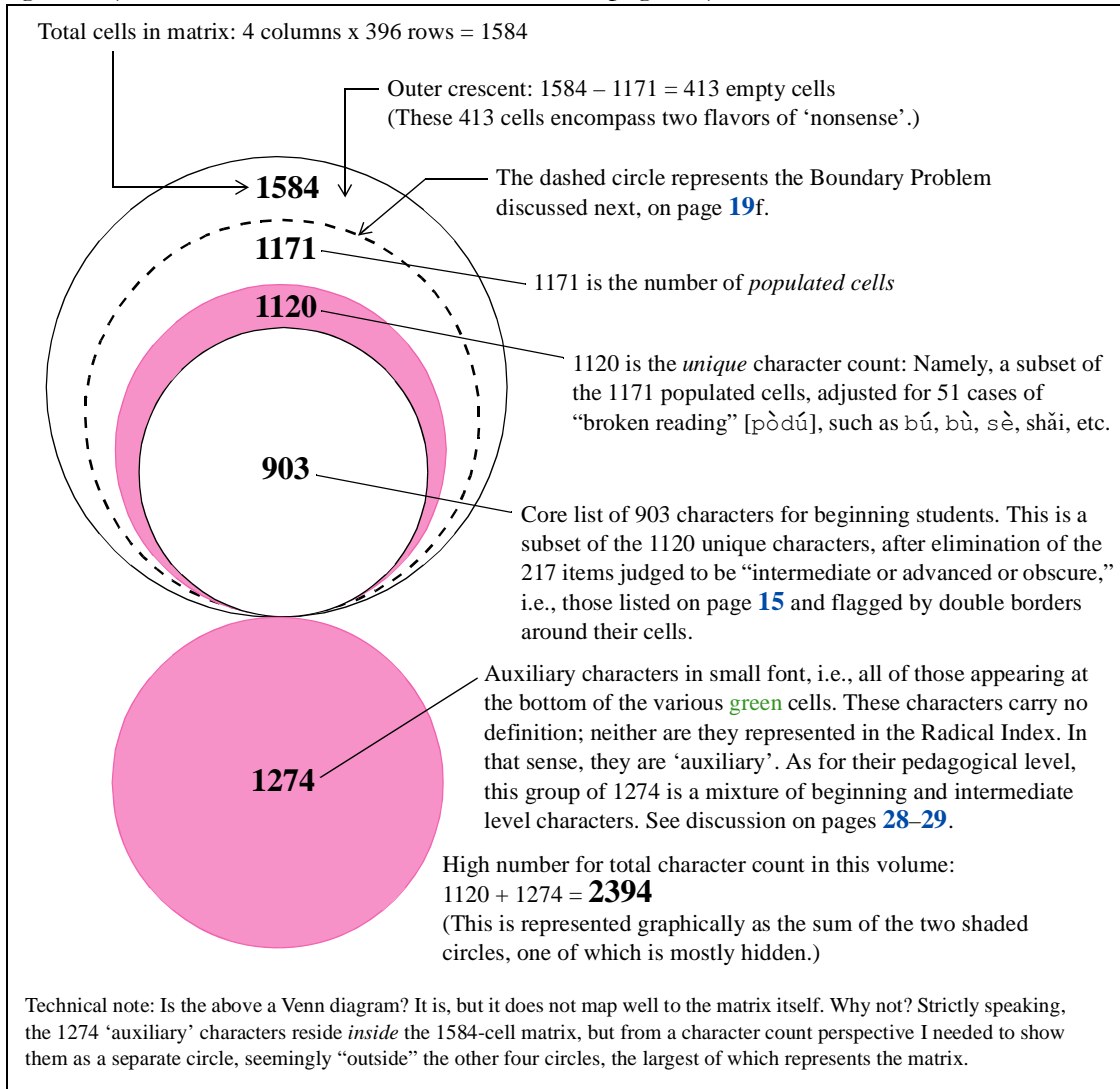
FIGURE 2: Character Counts and Pedagogical Levels

The Boundary Problem: inclusion/exclusion rules for the matrix

As presented earlier, our process includes the following step: "… exclude 123 cases of judgement-call nonsense." Another way to describe that step would be to say: "We had to create a border for the matrix, out of thin air, with no objective rules to guide us." In the abstract, this may sound mysterious, so I will work through one concrete example, to demystify the process:

Take the case of '*ga*': this initial-final combination has a rather queer sound in Mandarin, I think, just as '*gah*' does in English. So, as I step through its four theoretical tonal forms, gā gá gǎ gà, to create row 88 of the matrix, my first question is: "How many of these four slots are occupied by *real* morphemes?" As it happens, the answer is "all four," not that you will see the whole set in a medium-sized dictionary such as the *Zhongda*, but for sure you will in a large dictionary such as the *Gwoyeu Tsyrdean*. At this point, I know that none of the four cells in row 88 will remain *entirely* blank. At the very least, all four will be plugged with asterisks. By contrast, the slots for mìn and shěi are *absolutely* blank because these sounds are pure unmitigated nonsense, thus creating a hard boundary for that part of the matrix (as shown schematically in **Figure 3** on page **24**) since they entail nothing problematic to discuss.

Next question: "How many of these potential entries, gā gá gǎ gà, shall we judge *suitable* for the matrix?" Certainly we want the first one because it is used in transliterating 'curry' as gālí 咖喱 (also 'coffee' as kāfēi, throwing us a bit of a curve with the switch from 'g' to 'k'). What about the second one? *As* a second-tone word, the character 嘎 occurs in older dictionaries (e.g., the *Gwoyeu Tsyrdean*) in terms such as gádiào 嘎调 'an abruptly sharp or animated phrase in a Chinese aria' and gá.gar-tiān 'a day that is cool in the morning and evening, but hot at high noon'. Meanwhile, the primary pronunciation for 嘎 is *first*-tone gā, an allegedly onomatopoeic syllable for representing a screeching or creaking sound. The character 嘎 is reasonably high frequency when read in the *first* tone; also, in the second-tone slot, there exists a "real" word, gá 钆 'gadolinium' (*Gd*, the 64th element). All told, this makes for a rather uncomfortable judgment call. But in the end, I exclude gá from the matrix, populating its would-be cell by a lone asterisk. As always per my convention, the asterisk means: "If you look in a large enough (and/or old enough) dictionary, you'll find something low-frequency or obscure or recondite for this morphemic slot; nothing you need to worry about just now. Maybe when you're in graduate school."

The *Zhongda* dictionary and the *HarperCollins* are both silent on third-tone gǎ: no entry at all. But the *Gwoyeu Tsyrdean* has gǎ 玍 , a character that is shown occurring only in a few

pejoratives such as gǎ.gǔ 乾古 'quirky, eccentric' and gǎzá.zi 乾杂子 'a term for chastising someone who is quirky (in an unpleasant way)'. The character 乾 is wonderfully weird looking (in accord with its meaning). Plus, it is used in the colorful phrase gǎqīmǎbā 乾七马八 'improper, unsavory (referring to local vagrants and villains)'. I find myself quite drawn to this morpheme, gǎ 乾 , because of its colorful nature, but I know that on balance, it is just plain *rare*. Accordingly, it's *not* much of a judgment call this time: for the gǎ cell (too), I enter only an asterisk.

At first glance, gà 尬 might seem almost as peculiar as gá and gǎ, but then one remembers that gà is the second half of gān-gà 尴尬 meaning 'awkward, an embarrassing situation', which is a fairly common word (occurring also as gāngan-gagà.de). So, even though gà is a *bit* of an oddity,[15] it definitely belongs in the matrix. Thus, gā 咖 and gà 尬 are the two survivors of my vetting process, while the slots for gá and gǎ carry only asterisks. Done. The row looks like this:

gā 咖 curry = gālí	*	*	**GÀ** 尬 awkward = gāngà

In summary: As I consider each of the 1584 cells, I allow myself to be pulled simultaneously in two directions: [1] I would like to populate this cell with *something* if possible; [2] if that 'something' feels too recondite or outlandish, then I must exclude it, and populate the cell

---

15. For a century or more, gān-gà has been a regular part of Mandarin, although it appears to have entered Mandarin originally as a borrowing from the Wu language, i.e., from Shanghainese: The Mathews dictionary gives *ka⁴* (= gà in PīnYīn) first as "A staggering gait," followed by "*kan¹-ka⁴* in an embarrassing situation. (*Wu-dial.*)"; Mathews (1943 [1931]), p. 478. Meanwhile, in the *Zhongda* dictionary (2003 [1999]), we find no *separate* entry for gà, but its importance is acknowledged indirectly via the entry for gān-gà on p. 231.

with an asterisk instead. Little by little, as each such tension is resolved in favor of [1] or [2], a notional "border" for the matrix is created.[16]

It is not an easy process, but at least it can be done, and in a reasonable amount of time. Then, in effect, one has a *comprehensive* model of the language. (At least subjectively, that's the effect.) Just for fun, one might want to contrast that situation with the out-and-out mess that characterizes a language such as English, as sketched out in **Appendix E: In Search of a 'Taming Mechanism' for English**, page **119**.

To go further in this direction where Chinese is taken as an "object in nature," we'll need to see definitions of the per-cell notation conventions, based on font, on color, and on other such graphical indicators. First there will be a two-page synopsis of all six conventions, on pages **22–23**. Later, the notation conventions are presented in detail, on pages **28–38**. Then the matrix itself follows, on pages **39–100**.

---

16. Only *after* that border is firmly established can we apply the logic of "1171 − 217 − 14 = 903" mentioned earlier, which attempts to segregate advanced vocabulary items from beginning/intermediate vocabulary, and thus defines another kind of boundary, inside the matrix, as it were. I discussed that process out of sequence because it was needed in the context of My First 1000 Characters and "could not wait." Ideally, I would not have introduced that aspect of the matrix before this page.

THREE KINDS OF PREVIEW / QUICK REFERENCE

In this section, I provide **[a]** a two-page synopsis of the notation conventions; **[b]** a pie chart synopsis of the matrix; **[c]** a bird's-eye view of the whole language, compressed to three pages, as color coded syllables. (All three previews can serve later as Quick Reference guides.)

Notation Conventions — a two-page synopsis

1. Color coding of the PĪnYĪn cell headings as green, **black** or RED.

<u>dié</u>	bié	GĚI	bǐ	ān
谍	别	给	比	安
espionage = diébào	don't; other	give	compare	peaceful = ānjìng(de)
——————— 迭谍牒叠碟蝶				

When a cell heading is underscored and green, this denotes an 'overloaded syllable' alias 'overloaded cell'. A heading in an ordinary **black** font, illustrated above by the entry for bié (also bǐ and ān), indicates a 'regular' syllable — one that has only two or three characters associated with it. If the heading is ALL-CAPS and RED, this denotes a 'prime syllable', meaning that in the entire language there is exactly *one character* associated with this syllable. *Note*: The terms 'RED' and 'prime' are used interchangeably in this book.

2. Stand-alone words vs. "half-words"

Illustrated by bǐ versus ān above: The English definition for bǐ is centered immediately beneath the character and it has *no* trailing equals sign. I use this format to indicate that the PĪnYĪn *syllable* in question works as a stand-alone *word* (alias free morpheme). The difference between this unadorned notation and the next one, with an equals sign, is crucial: The trailing *equals* sign after 'peaceful' leads to a multisyllabic Chinese word immediately below. This notation means, "YES, there is a *character* 安 which means 'peace' or 'peaceful', immediately understood in a written context, BUT ān by itself will very likely not work as a *word* in conversation." (In one of the appendices, we explore at length

the reasons why this simple notion of a 'bound morpheme' is ultimately inadequate, requiring further analysis.)

3. *Italicized* definition: This typeface denotes a case of *character-overload*. This is roughly the opposite of the overloaded syllable type: Instead of having a cluster of characters all tied to a single syllable, we have only "one word" ostensibly but it can mean many different things; illustrated by TÀO below:

TÀO	běn	NÁNG	*
套	本	囊	
cover, trick into, copy, MEASURE *for suits/ stamps/ volumes*	book	used in bēináng backpack, book bag	
ānquán-tào condom	bǐjìběn notebook	dǎn-náng gall bladder	

4. Some cells contain an illustrative example, in *smaller font*, *left justified.* Three of the cells reproduced immediately above happen to illustrate this feature, e.g., bǐjìběn in the běn cell. However, many cells in the matrix contain *no* such illustrative example. When present, an illustrative example may have been thrown in only as an "extra" or it may be an implied warning about how to use (or not use) the word.

5. Double border means 'advanced vocabulary', as discussed on page **15**.
   Illustrated by NÁNG above.

6. A cell may be empty except for ' * ' as a generic pointer to a larger dictionary.
   See discussion of gǎ, etc., on page **19**.

Pie chart

THE REGIONS

The matrix: 4 x 396 = 1584 cells (= everything encompassed by the outer curve)

EMPTY region: 290 blank cells (18% of 1584)

ASTERISKED: 123 cells that are empty but asterisked (8% of 1584)

RED: 263 prime syllables (17%)

BLACK: 646 regular syllables (40%)

GREEN: 262 overloaded syllables (17%)

THE BOUNDARIES

The matrix: hard boundary, by definition.

EMPTY: hard boundary (absolute nonsense)

ASTERISKED: soft boundary (dashed line): Judgment calls: Which characters (syllables) are deemed "too recondite to be worth mentioning"?

BLACK/RED: hard boundary, by definition of 'prime' (but see the caveat in footnote 23 on page 30).

GREEN/BLACK boundary: soft boundary (dashed line). Judgment call: By fiat, I set the threshold at 4 characters. I.e., if 4 or more characters are associated with a given syllable, I classify the syllable as overloaded (GREEN).

Alternatively, one may base the percentages on the populated cells, which number 1171 (= 263 + 646 + 262). This creates proportions that are more psychologically meaningful: RED 22%, BLACK 55%, GREEN 22%. See discussion of RED on page 29.
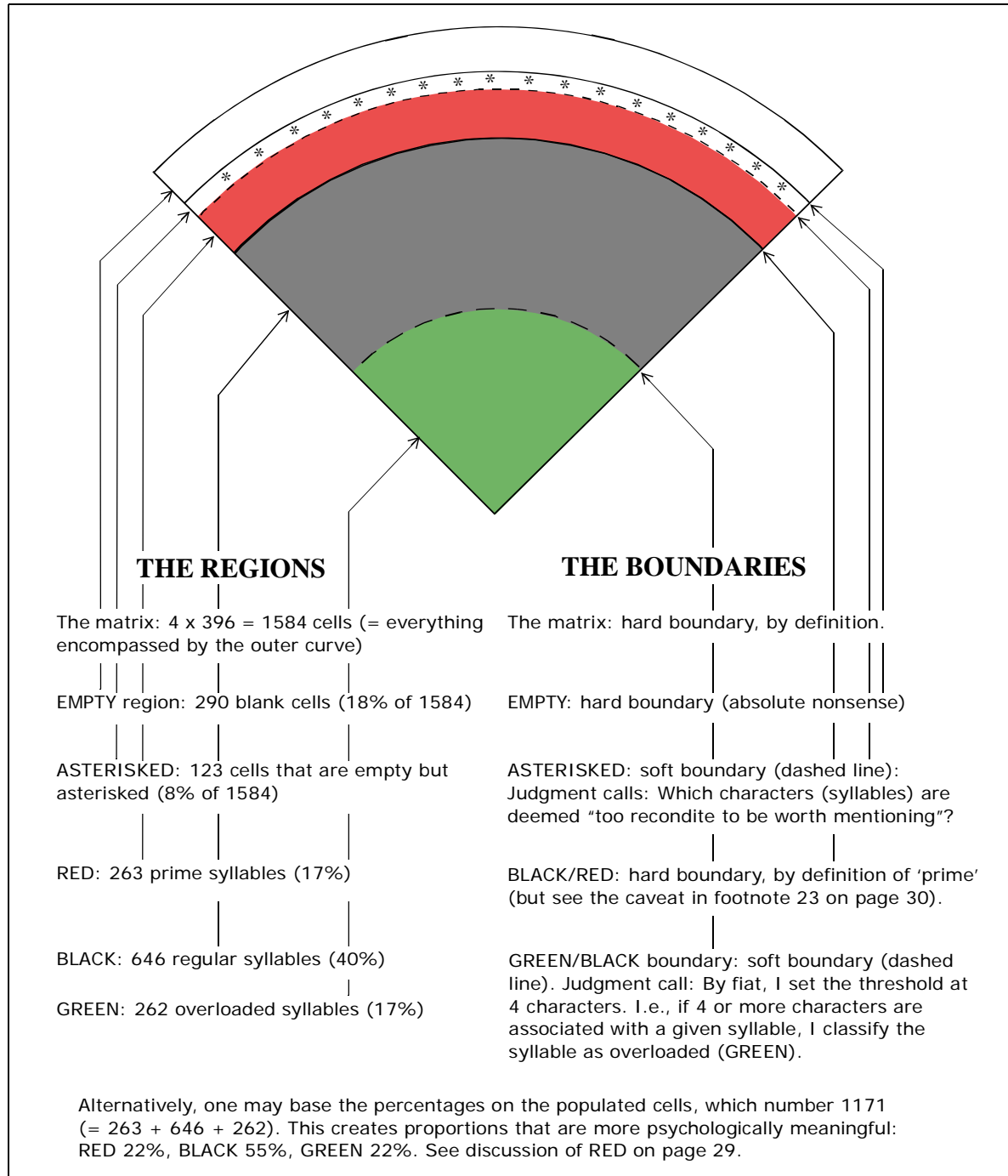
FIGURE 3: Pie Chart of Empty, Prime, Regular, and Overloaded Cells

Notes: The color coding summarized in **Figure 3** is what makes the dictionary 3-dimensional.
Compare **Figure 2** on page **18**, which is a different way of slicing the same pie.
For more about the empty cells, specifically, see the preface to **Appendix F: Not Even Wrong**, page **125**.

The 1171 headings distilled to a three-page summary

On these three pages I present a bird's-eye view of the whole language, by boiling the matrix down to a straight enumeration of its 1171 morphemic headings (or, less formally, its inventory of "syllables"). From the color coding, one gets a general sense of how the 'overloaded' syllables (green), 'prime' syllables (red), and 'normal' syllables (black) are distributed across the 1171 populated cells.

A aī aí aǐ aì ān (*second tone, third tone?*)[17] àn āng áng àng AO áo AO aò

B bā bá bǎ bà bāi BÁI bǎi bài bān bǎn bàn BĀNG bǎng bàng bāo báo bǎo bào bēi BĚI bèi BĒN běn bèn bēng BÉNG BĚNG bèng BĪ bí bǐ bì biān biǎn biàn biāo biǎo biē bié BIĚ biè bīn bìn bīng bǐng bìng bō bó bǒ bò bū bú bǔ bù

C CA CAI cái cǎi CÀI cān cán CǍN càn cāng CÁNG cāo cáo CǍO CÀO cè CEN CÉN céng CÈNG chā chá CHǍ chà chāi chái chān chán chǎn chàn chāng cháng chǎng chàng chāo cháo chǎo chē CHĚ chè chēn chén chèn chēng chéng chěng CHÈNG chī chí chǐ chì chōng chóng CHǑNG chòng CHŌU chóu chǒu CHÒU chū chú chǔ chù chuāi CHUǍI CHUÀI chuān chuán chuǎn chuàn chuāng CHUÁNG CHUǍNG chuàng chuī chuí CHǓN chún CHǓN CHUŌ chuò cī cí CǏ cì cōng cóng CÒU CǓ cù cuān cuán cuàn cuī cuǐ cuì cūn CÚN CÙN cuō cuó cuò

D dā dá DǍ DÀ dāi dǎi dài dān dǎn dàn dāng dǎng dàng dāo dǎo dào dé DĚI dēng DĚNG dèng dī dí dǐ dì diān diǎn diàn diāo diǎo diē dié dīng dǐng dìng DIŪ dōng dǒng dòng dōu dǒu dòu dū dú dǔ dù DUĀN DUǍN duàn DUĪ duì dūn dǔn dùn duō duó duǒ duò

E ē é ě è EN ÈN ér ěr èr

F fā fá fǎ fà fān fán fǎn fàn fāng fáng fǎng FÀNG fēi féi fěi fèi fēn fén FÉN fèn fēng féng fěng fèng FÓ FǑU fū fú fǔ fù

---

17. What happened to 2nd tone án and 3rd tone ǎn? Recall that we are showing only the "real" syllables here, while omitting those that are nonexistent in pǔtōnghuà (*or,* are so arcane that I choose by fiat to *treat* them as nonexistent). Accordingly, for a given syllable, you will not necessarily see all four tones represented in this list. Only the pertinent ones. (By the way, everything that is missing here *does* appear somewhere else: in the mirror-image table of **Appendix F: Not Even Wrong**, page **125**.)

G gā **GÀ** gāi **GǍI** gài <u>gān</u> <u>gǎn</u> gàn <u>gāng</u> gǎng gàng <u>gāo</u> gǎo gào <u>gē</u> <u>gé</u> gè **GěI** gēn gēng gěng **GèNG** <u>gōng</u> gǒng gòng <u>gōu</u> gǒu <u>gòu</u> <u>gū</u> <u>gǔ</u> <u>gù</u> guā guǎ guà **GUAI** **GUǍI** **GUÀI** <u>guān</u> guǎn <u>guàn</u> guāng guǎng guàng <u>guī</u> guǐ <u>guǐ</u> gǔn **GÙN** guō guó guǒ guò

H hā há hǎ hāi hái **HǍI** hài hān <u>hàn</u> hěn <u>hàn</u> hāng háng hàng hāo <u>háo</u> **HǍO** <u>hào</u> hē <u>hé</u> <u>hè</u> **HEI** hén hěn **HèN** hēng héng hèng hōng <u>hóng</u> hǒng hòng hóu **HǑU** hòu hū <u>hú</u> hǔ <u>hù</u> huā <u>huá</u> <u>huà</u> huái **HUÀI** huān huàn huǎn <u>huàn</u> huāng <u>huáng</u> <u>huǎng</u> huàng <u>huī</u> huí huǐ <u>huì</u> hūn hún hùn huó huǒ <u>huò</u>

J <u>jī</u> <u>jí</u> <u>jǐ</u> <u>jì</u> <u>jiā</u> jiá jiǎ <u>jià</u> <u>jiān</u> <u>jiǎn</u> <u>jiàn</u> <u>jiāng</u> jiǎng <u>jiàng</u> <u>jiāo</u> **JIÁO** <u>jiǎo</u> <u>jiào</u> <u>jiē</u> <u>jié</u> jiě <u>jiè</u> <u>jīn</u> <u>jǐn</u> <u>jìn</u> <u>jīng</u> <u>jǐng</u> <u>jìng</u> jiǒng jiū jiǔ <u>jiù</u> jū jú jǔ <u>jù</u> juān **JUǍN** juàn juē <u>jué</u> <u>jūn</u> jùn

K kā kǎ kāi kǎi kān kǎn **KÀN** kāng kàng kǎo kào kē ké kě <u>kè</u> <u>kěn</u> kēng kōng kǒng kòng kōu **KǑU** kòu kū **KǓ** kù **KUA** **KUǍ** kuà <u>kuài</u> kuān kuǎn kuāng **KUÁNG** <u>kuàng</u> kuī kuí kuǐ kuì kūn **KǓN** **KÙN** kuò

L lā lá lǎ <u>là</u> **LÁI** lài <u>lán</u> lǎn làn láng lǎng **LÀNG** **LǍO** láo lǎo lào lè **LEI** léi lěi <u>lèi</u> léng lěng lèng <u>lí</u> <u>lǐ</u> <u>lì</u> **LIǍ** <u>lián</u> liǎn liàn <u>liáng</u> **LIǍNG** <u>liàng</u> **LIāO** liáo **LIǍO** liào **LIE** liě <u>liè</u> līn <u>lín</u> lǐn lìn <u>líng</u> lǐng lìng liū <u>liú</u> liǔ liù <u>lóng</u> lǒng lòng **LǑU** lóu lǒu lòu lú lǔ <u>lù</u> **LǓ** **lǚ** <u>lǜ</u> luán **LUǍN** **LUÀN** lüè **LȔN** lún **LÙN** luō <u>luó</u> **LUǑ** luò

M mā **MÁ** mǎ **MÀ** **MÁI** **MǍI** <u>mài</u> mán **MǍN** màn máng mǎng **MāO** máo **MǍO** <u>mào</u> <u>méi</u> měi mèi **MEN** mén mèn **MĒNG** méng měng mèng mī mí mǐ mì mián miǎn **MIÀN** **MIāO** miáo miǎo miào **MIE** miè **MÍN** mǐn míng **MǏNG** **MÌNG** miù mō <u>mó</u> **MǑ** <u>mò</u> móu **MǑU** **MÚ** mǔ <u>mù</u>

N **NÁ** **NǍ** nà nǎi nài nán **NÀN** **NÁNG** **NǍNG** **NǍO** náo nǎo **NÀO** **NěI** **NèI** **NÈN** **NÉNG** ní nǐ <u>nì</u> niān nián niǎn **NIÀN** **NIÁNG** **NIÀNG** niǎo **NIǍO** **NIE** niè **NÍN** <u>níng</u> nìng **NIÚ** niǔ **NIÙ** nóng **NÒNG** **NÚ** nǔ **NÙ** **NǓ** **NUǍN** nüè **NUÓ** nuò

O **Ó** **Ǒ** **Ò** ōu ǒu òu

P pā pá pà **PAI** pái **PÀI** pān pán pàn pāng páng **PÀNG** pāo páo **PǍO** pào pēi péi pèi **PEN** **PÉN** **PÈN** pēng péng **PěNG** **PèNG** pī <u>pí</u> pǐ <u>pì</u> <u>piān</u> pián piàn piāo piáo piǎo piào piē **PIĚ** pīn pín **PǏN** **PÌN** **PĪNG** <u>píng</u> pō **PÓ** **PǑ** pò **PŌU** pū <u>pú</u> pǔ pù

Q qī qí qǐ qì QIA QIǍ qià qiān qián qiǎn qiàn qiāng qiáng qiǎng qiàng qiāo qiáo qiǎo qiào QIE QIÉ QIĚ qiè qīn qín QǏN QÌN qīng qíng qǐng qìng qióng qiū qiú qū qú qǔ qù quān quán QUǍN quàn quē QUÉ què qún

R rán rǎn RANG ráng rǎng RÀNG ráo RǍO RÀO RĚ RÈ rén rěn rèn RENG RÉNG RÌ róng RǑNG róu RÒU rú rǔ rù ruǎn RUǏ ruì rùn ruò

S sā sǎ sà sāi sài sān sǎn SÀN sāng SǍNG SÀNG sāo sǎo sào sè SEN SENG shā SHÁ SHǍ shà SHAI SHǍI SHÀI shān shǎn shàn shāng shǎng shàng shāo sháo SHǍO shào shē shé SHĚ shè SHÉI shēn shén shěn shèn shēng SHÉNG SHĚNG shèng shī shí shǐ shì SHOU shǒu shòu shū shú shǔ shù SHUA SHUǍ SHUÀ shuāi SHUǍI shuài shuān SHUÀN shuāng SHUǍNG SHUǏ shuǐ SHǓN shùn SHUO shuò sī SǏ sì SONG sǒng sòng sōu SÒU sū SÚ sù suān suàn suī suí SUǏ suì SUN sǔn suō suǒ

T tā tǎ tà TAI tái tài tān tán tǎn tàn tāng táng tǎng tàng tāo táo TǍO TÀO TÈ téng tī tí tǐ tì tiān tián tiǎn TIAO tiáo TIǍO tiào tiē tiě TIÈ tīng tíng tǐng TONG tóng tǒng tòng TOU tóu TÒU tū tú tǔ tù TUAN TUÁN tuī TUÍ TUǏ tuì TUN tún TÙN tuō tuó tuǒ tuò

W wā WÁ WǍ wà WAI WǍI WÀI wān wán wǎn wàn WANG wáng wǎng wàng wēi wéi wěi wèi wēn wén wěn WÈN wēng wèng wō WǑ wò wū wú wǔ wù

X xī xí xǐ xì xiā xiá xià xiān xián xiǎn xiàn xiāng xiáng xiǎng xiàng xiāo XIÁO xiǎo xiào xiē xié xiě xiè xīn xìn xīng xíng xǐng xìng xiōng xióng xiū xiǔ xiù xū XÚ xǔ xù xuān xuán XUǍN xuàn xuē xué xuě xuè xūn xún xùn

Y yā yá yǎ yà yān yán yǎn yàn yāng yáng yǎng yàng yāo yáo yǎo yào yē YÉ yě yè yī yí yǐ yì yīn yín yǐn yìn yīng yíng yǐng yìng yōng yǒng yòng yōu yóu yǒu yòu yū yú yǔ yù yuān yuán YUǍN yuàn yuē yuè YUN yún yǔn yùn

Z zā zá ZǍ zāi zǎi zài zān ZÁN ZǍN zàn zāng zàng zāo ZÁO zǎo zào zé ZÈ ZÉI ZĚN ZÈN zēng zèng zhā zhá zhǎ zhà zhāi zhái ZHǍI zhài zhān zhǎn zhàn zhāng zhǎng zhàng zhāo ZHÁO zhǎo zhào zhē zhé zhě zhè zhēn zhěn zhèn zhēng zhěng zhèng zhī zhí zhǐ zhì zhōng zhǒng zhòng zhōu zhóu zhǒu zhòu zhū zhú zhǔ zhù ZHUA ZHUǍ ZHUAI ZHUÀI zhuān ZHUǍN zhuàn zhuāng zhuàng zhuī zhuì zhūn ZHǓN zhuō zhuó zī zǐ zì zōng ZǑNG zòng ZǑU zòu ZU zú zǔ ZUAN ZUǍN ZUÀN ZUǏ zuì zūn zuō zuó zuǒ zuò

NOTATION CONVENTIONS — DETAILED PRESENTATION

Notes: [a] For clarity, and to minimize the need for page-turning, I allow a certain degree of overlap (redundancy) between this section (numbering about ten pages) and its synopsis on pages **22–23**. [b] Subtopics pertaining to pedagogy and subtopics pertaining to linguistics are *commingled* in this section.

1. Color coding of the PīnYīn cell headings as green, **black** or RED.[18]

    When a cell heading is underscored and green, this denotes an 'overloaded syllable' — one that has 4 or more reasonably high-frequency characters associated with it (and possibly as many as 19; see the entry for yì on page **92**). A supplemental list of characters is given in smaller font in a rectangular box at the bottom of the cell (without definitions). I refer to these collectively as the 'auxiliary characters'. Sample cell in context (replica of row **66**):

Tone 1	2	3	4
diē 爹 dad	dié 谍 espionage = diébào ————————— 迭谍牒叠碟蝶		

    Note that whenever there is such a list, it *reiterates* the large-font character — the one that "occupies" the cell in question. (For instance, in the example above, the main entry in the cell is reiterated as the second character 谍 in the supplemental list.) My rationale for this overlap is that there is often "nothing special" about the one I chose to represent the cell. In terms of importance or frequency, it is often on a par with one or more of the other characters in the list. It just happens to be *the one I chose*, arbitrarily, to be the notional "occupant" of the cell.

---

18. Rationale for the unsightly underscores and ALL-CAPS: In the event of a black-and-white photocopy, the underscores and CAPS may serve as proxy for the color coding. It was an uncomfortable judgment call. By a thin margin, the guaranteed survival of the coding trumps aesthetic considerations, I decided.

Bookkeeping aspect: Taken all together (and with each list adjusted by minus-one[19] to prevent double-counting), the small-font items aggregate to 1274 characters. In my judgement, these are all reasonably high-frequency characters, hence "worth knowing" in the context of an undergraduate Chinese curriculum; but since these are *not* given the full lexicological treatment[20] as described in the **Prologue**, I track them separately, as 'auxiliary characters', distinct from the primary entries that number 1120 unique characters (as a subset of the 1171 populated cells). When these two figures — 1274 and 1120 — are summed we obtain the "high number" shown in **Figure 2** on page **18**: 2394.

(For purposes of this work, the supplemental characters in small font are meant only to show where and how and to what *extent* the 'overloaded syllable' phenomenon occurs. For that reason, I stop short of appending their definitions. The supplemental list is not meant to be exhaustive for the syllable in question, only suggestive of how overloaded it is.)

I said I set the threshold for inclusion in this 'overloaded' category at 4 or more. Why that number as the cutoff point? Because it feels right: When a syllable has only two or three characters mapped to it, that seems a "reasonable" mental/psychological burden for the student to handle, whereas the burden seems "excessive" to me when a syllable has four or more high- or medium-frequency characters all mapped to it — especially when this means a staggering 19 characters, as is the case for the syllable yì! (But the more difficult question is how to build the list of those "high- or medium-frequency characters." This part of the discussion I defer to **Appendix G: Frequency of Occurrence**, page **141**.)

---

19. On the other hand, for a character that has two readings, such as 咽 with readings yān (pharynx or throat) and yàn (to swallow), I treat each (re)appearance of the character as substantive, and therefore make no such "minus one" adjustment in the tally. There are several dozen instances of this 咽 type.

20. Thus, for example, 萝 and 逻 are listed in small font at the bottom of the 'overloaded' luó cell, with no warning that they are, respectively, only the first *half* of luó.bo 'radish' and first *half* of luójí 'logic'. See row 182 of the table, on page **66**.

A heading in an **ordinary black** font indicates a 'regular' syllable. To distinguish this category from the 'overloaded' category (one syllable mapped to 4 or more characters), I define it arbitrarily as including all syllables with which only two or three characters are associated. In other words, it's our catchall or default category. Sample cell:

```
        bié
        别
    don't; other
```

An anomaly: For all such cases of a 'regular' syllable, I do *not* show the missing one or two characters (the way I show the other 3 to 18 characters for an overloaded syllable in the 4 to 19 range). In this instance, for example, there exists another bié character not shown. That missing character looks like this 鳖 and it is used in the word biéjiǎo 鳖脚 'shoddy'. But I cite the character 鳖 nowhere in the matrix. Rather, the student is on her own to discover such missing pieces. The missing character(s) might be of rather low frequency (as in the case of 鳖) or of relatively high frequency, comparable in importance to the character that occupies the cell. As currently formulated, this is an inconsistency and shortcoming of the matrix, perhaps[21] to be addressed in a later version.

Opposite extreme from the overload case: If the PīnYīn heading in a cell is ALL-CAPS and RED, this denotes a 'prime syllable', meaning that in the entire language there is exactly *one* character associated with this syllable (i.e., this cell). It has no homophones. Sample cell in context (replica of row **94**):

Tone 1	2	3	4
		GĚI 给 give	

Here we find a given cell "occupied" by a single character; *and*, that single character is often a high-frequency vocabulary item to boot. I was pleasantly surprised by this category, to say the least, as I saw it take shape as a whopping 22% of the total. Before compiling the matrix, I had been aware of wǒ as an example of the phenomenon (and in fuzzy retrospect, gěi and děi were in my peripheral vision, so to say, although by rights

---

21. I say "perhaps" because one must also consider the danger of making the matrix too busy. In that regard, it already seems near the breaking point; see footnote **24**. Or, saying it another way, the matrix is possibly *already* optimal.

they should have been *more* noticeable than wǒ since they are members of the exclusive super-RED category, introduced below). What caught me by surprise during this project was the sheer number of such "prime real estate" sites strewn across the landscape, accounting for 263 of the 1171 populated cells. Especially when it is viewed against the framework of the overloaded syllables discussed earlier, we begin to see how remarkable it is that a 'prime' category even exists.[22] (See also **Figure 3** on page **24**.)

Also, note that there are some cases of RED thrown into especially high relief by the absence of any sister syllables in the other three tones across a given row, namely:

cā, còu, děi, fó, fǒu, gěi, hēi, liǎ, miù, néng, nín, nǔ, nuǎn, pōu, rì, sēn, sēng, shéi, tè, zéi

This subset I refer to as the super-RED cells, to mark their status as being not only 'cell-wise prime' but 'row-wise prime', so to speak.

And, similar to the situation for the larger category of (plain) RED, half of *these* are likewise high-frequency words: děi, gěi, hēi, liǎ, néng, nín, nǔ, nuǎn, rì, shéi, tè.[23] Something to boost student morale.

---

22. This category has 'ontological status', by the way; it is not an artifact subject to one of my arbitrary cutoff points, although it does involve some very small judgment calls. (See **Footnote to Appendix H: The Prime Syllables Only**, page **155**.) An aside: I can't help noticing a certain resemblance here to the zeta function "landscape," at least as filtered through the words Marcus du Sautoy in *The Music of the Primes* (2003), pp. 97–99. Indirectly, the "landscape" of **Figure 1** on page **10** above and my use of the term 'prime syllable' were influenced by du Sautoy's treatment of the subject, although I've inverted the meaning of 'sea level' as proposed in his analogy: His actual primes are *at* sea level, while our figurative primes follow the commonsense notion of residing *above* sea level. (I hasten to add that any serious comparison between the venerable zeta function and our rough-hewn matrix would only inspire derisiveness from the mathematician — something about moving from the sublime to the absurd.)

23. One of the super-RED syllables that is not especially high frequency but interesting in its own right is fó. If we focus on the historical vista prior to the 20th century, the morpheme fó must strike us as a rarity: evidence of a *lasting* influence from the outside on *mainstream* Chinese writing and speech. See remarks on fó 佛 and sēng 僧 in **Appendix E: In Search of a 'Taming Mechanism' for English**, page **119**.

2. Stand-alone words vs. "half words":

bǐ
比
compare

free morpheme
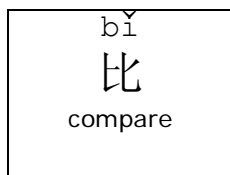
ān
安
peaceful =
ānjìng(de)
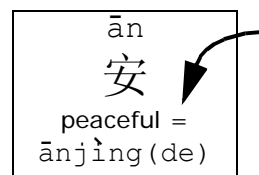
bound morpheme(s)

- The English definition is centered immediately beneath the character and it has *no* trailing equals sign (illustrated by bǐ above). I use this format to indicate that the PīnYīn *syllable* in question works as a stand-alone *word*, alias free morpheme.

- A trailing *equals* sign leads to a multisyllabic word on the next line down in the cell (illustrated by ān above). This latter notation means

"YES, there is a *character* 安 which means 'peace' or 'peaceful', immediately understood in a written context, BUT ān by itself will very likely not work as a *word* in conversation. At best it will sound 'funny'; at worst it will convey nothing. This is a bound morpheme, and the word you want appears on the next line down: ānjìng (or ānjìng.de to call out its adjectival flavor)."

Some further examples: See the entries for āng (āngzāng), xiá (xiácī), and so on. On a few occasions, I use the word 'in' or the phrase 'used in' instead of the equals sign to flag this type. E.g., in the cell for NÁNG (which also illustrates Convention #5), and in the cell for quǎn.

Note that I do not supply the Hànzì that goes with the second syllable in cases such as ānjìng or āngzāng. This is an "executive decision" I made, [a] to reduce clutter in the matrix and [b] to keep the focus where it needs to be, on *words*, not the pretty characters.[24]

Important note about the <u>character count</u>(s) for this project (nonintuitive aspect): Because of the practice regarding second syllables described immediately above, the figure 1171 (derived already on pages **16–17**) is the tally both for lexemes *and* for basic characters in this project, *even though* so many of the lexemes are bisyllabic.

---

24. In a similar vein, Chao writes, "One of the disadvantages of using a character text for beginning students is ... the obscuring of the distinction between syntax and morphology." Chao (1948), p. 77. I appreciate primacy-of-the-word as a philosophy, but I also feel, as most Chinese do, that the characters *are* the language, so far as "Chinese education" is concerned. Consequently, I'm willing to go only partway in that direction. For more about this, see also footnote **4** on page **6**.

Representation of verbs in the English definitions (verb/noun disambiguation)

I use the infinitive form sparingly, only where there is an immediate danger of ambiguity. For instance, 'force' would be ambiguous: verb or noun? Therefore, I forestall the ambiguity by defining bī as 'to force'. Conversely, 'compare' presents itself already as a verb. Therefore, I definine bǐ as 'compare', without the word 'to'. This minimizes clutter.

(Granted, 'compare' *may* be a noun in English. But I judge it improbable that the student will see 'compare' and think "This is a noun, as in 'The moon is beyond compare'." Still, I must at least acknowledge the existence of this fringe area of grammar, and this is why I include the qualifier "immediate danger" above. A second example: Presented with the gloss 'give' for gěi, the student will likely take gěi as a verb, as intended, not as a noun. I see no immediate danger that she will complain of ambiguity, citing "There is not enough give in the trampoline" as evidence that " 'give' can be either a verb *or* a noun." Commonsense should dictate that 'give' is being used to identify gěi as a verb. Therefore, I dispense with 'to' in this case as well. Third example: On page **99**, I use the glosses 'to drill' and 'a drill' for zuān and zuàn, respectively: two readings of the same character, 钻 , the former a verb and the latter a noun.)

At this point, we have barely scratched the surface of the following questions: "How valid is the concept of 'monosyllabic or multisyllabic' when applied to Chinese?" and "Can a generalization be made?" Short answer: No. For a longer answer, please refer to **Appendix K: Free and Bound Morphemes** (which includes a discussion of chéngyǔ) and to **Appendix L: Muchengxue and the Monosyllabic Mystique**.

3. *Italicized* definition: This typeface denotes a case of *character-overload* (not to be confused with the phenomenon of an 'overloaded syllable', introduced in connection with Notation Convention #1 already). Reproduced below are two examples of what I mean by character-overload:

<table>
<tr>
<td align="center">
kě
可
<i>can; a prefix meaning '-able'; but; also used as an intensifier (before any verb)</i>
</td>
<td align="center">
TÀO
套
<i>cover, trick into, copy,</i>
MEASURE<i>for suits/ stamps/ volumes</i>
ānquán-tào
condom
</td>
</tr>
</table>

Here we have roughly the "opposite" of the overloaded syllable phenomenon: Instead of seeing a bevy of different characters all associated with the same syllable, we have only "one word to learn" but that one word has as many as a dozen distinct definitions. Some of these overloaded characters become so dense with barnacles of meaning that they seem literally able to mean *anything*.

Next level of detail: I use the "italicized definition" convention to flag two cases: [i] a character that has several meanings that seem only *vaguely related* (or frankly unrelated, or even diametrically opposed); [ii] a character that has accumulated an "unreasonable" number of *extended meanings*, piled on down through the Ages as it were.

This notation convention is meant to warn you that the list of definitions does not entail a clean pair or trio of meanings (as when cù means 'vinegar' but also 'jealousy') but something more challenging. For example, kě can mean 'can or may' (in kě.yi). It can mean 'but' (in kě.shi). Or, as a prefix to a verb, it can function like our suffixes '–able' and '–ful', as in kěài 'lovable' and kěchǐ 'shameful'. But its most *common* use (it seems to me) is as an all-purpose intensifier before a verb, although this major role for kě is usually ignored in dictionaries (presumably because it is "only" a *function* of the word, not strictly speaking one of its *definitions*?). Anyway, here is kě in action in its most popular role, as an *intensifier*:

[1] … , nǐ dā.ying.guo.wo "… , you promised me" (sounding not too emotional), contrasted with
    … , nǐ kě dā.ying.guo.wo "… , you *prom*-ised me!"

[2] `Wǒ xìkàn.le` "I scrutinized it" (ho-hum, matter-of-fact tone),
contrasted with the following, where we see `xìkàn` inverted, with adverbial particle
`.de` added as an infix (`kàn.de xì`), and `xì` qualified by prefix `kě`:
`Wǒ kàn.de kě xì.le` "But I'm the one who *scrutinized* it" or "But *I* looked more
closely than *you* ever did!" Two possible translations, depending on context.

In passing, note how example [1] concludes with three neutral tones (`.ying` and `.guo`
and `.wo`) all strung together. Once you've encountered such a pattern, a light goes on over
your head and you think: "Yes, of course one would want to use a neutral tone not only
on `yíng` (the second half of `dā.ying`, which simply happens to be pronounced that
way) and on `guò` (which is always neutral when used in this role as an aspect marker), but
also on the `wǒ` (the object 'me'), in such a context."

This is how real people talk, but I doubt you'll ever see it in a textbook for Beginning
Chinese or even Intermediate Chinese.[25] One might be forced to conclude it was
"advanced stuff," except that the underlying principle is so simple and natural, once you
see what it is: "When a highly predictable word falls toward the end of a sentence, there's
often no need to make a fetish of its tone." That's how I would generalize it. So, keep an
eye out for the many analogous situations in Chinese, and tone *down* those tones if you
want to sound natural! Also, note this special case where, conversely, one *would* want to
emphasize 'nǐ' at the end of a sentence: `Dé.la.ba nǐ!` This has the flavor of "Enough
of that, *you*!" or "Stop that, *you*!" In both English and Chinese, the flavor of intimacy and
playfulness comes mainly from the syntactic inversion (verb-subject instead of
subject-verb) but can certainly be intensified if one places some stress on 'you'.
(My source for these various examples: the TV series called 一年又一年 .)

Returning to the topic of character-overload: A natural implication of my italicized
notation is that *context* will be important, or a second character will be needed to form a
compound, and only then will the intended meaning be clear, e.g., `pǐncháng` 'savor' vs.
`wùpǐn` 'article of merchandise' vs. `pǐnxiāo` 'play the vertical flute' (more likely to occur
in *double-entendre* than in its archaic musical sense).

---

25. In the Introduction to *A Grammar of Spoken Chinese*, we encounter the flip side of the observation I just made,
the exception that proves my rule: "[The style to be studied in this book] is definitely not the style of speech a
foreigner is expected to use in talking to a native speaker of Chinese. Rightly or wrongly, reasonably or
unreasonably, a native speaker is often surprised or even disconcerted when a foreigner talks exactly as he does
instead of talking more stiffly, preferably with a little accent, *accenting syllables that are usually slurred, giving a full
tone when it should be neutral*, and so on." Chao (1968a) p. 14, emphasis added.

Admittedly, my way of defining character-overload entails some uncomfortable judgement calls: Do the definitions for qì constitute a "reasonable" set of extended meanings, or do they range so widely (covering 'gas', 'air', 'manner', 'anger', 'to provoke', etc.) as to slip across the line into my "character-overload" category? Conceivably qì is just an extreme case of the "extended meaning" concept, but I opted to treat it rather as a "character-overload" type. Ditto for shēng, whose definitions include 'give birth', 'grow', 'student', 'living', 'raw', 'unfamiliar', 'stiff', and 'very'. These judgment calls have no particular consequence linguistically; the issue here is how best to introduce certain words to the poor beleaguered student!

4. Illustrative example in smaller font, left justified:
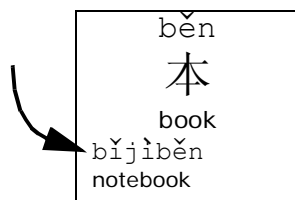
Note that my illustrative examples (left-justified and in smaller font) vary across a broad spectrum from

[a] cases where I *merely* illustrate usage in passing (under āo and under běn, for example) to

[z] cases where I'm *warning* you (once again) that the constituent parts are *almost* as tightly "bound" (as a linguist might say) as in a compound of the āngzāng variety.

Consider the word áozhōu under áo for instance: Yes, áo by itself means 'to boil or stew something' BUT it isn't used freely as a *general* cooking term. In its cooking definition (ignoring its other definitions here), it is followed almost always by zhōu (congee, gruel) or yào (medicinal herbs). So my example (áozhōu) is not a juxtaposition of the verb plus any random object; rather, it is carefully chosen to convey a hint of how áo is used — along with a hint, albeit faint, of how *not* to use it.

So, the million-dollar question:
How do we know where we are on this great broad spectrum from **[a]** to **[z]**?

Answer: By way of one's individual reading and listening experience.

I considered devising a set of subclassification labels for showing my opinion of where a given item is positioned on the spectrum, but decided it would be too fiddly, also too vulnerable to contention and hairsplitting. Already, in deciding to treat something as

an ānjìng type or as a case of **[z]**, I had judgement calls to make. For example, in poetry, there is no question that cǎi 彩 works just fine as a stand-alone word meaning 'colored'. And if you happen to be quoting a famous poem to someone about multicolored clouds at dawn, then temporarily it may seem to you that cǎi should work in ordinary conversation as well. But that would be a imprudent assumption (I think). Accordingly, I treat cǎisè as an ānjìng or āngzāng type, meaning you should always say the entire two-syllable word (to convey 'colored' or 'in color', as in connection with a TV set or a movie), not the first syllable by itself, even though in literature (and in advertisements) the meaning of the character cǎi, all by itself, is perfectly obvious and unambiguous. This is a borderline case that required a judgment call for purposes of getting it shoehorned into the matrix.

The 'cake' words bǐng and gāo illustrate an entirely different kind of problem. Typically the student encounters bǐng first in yuèbǐng 'mooncake'. Around the same time, one becomes aware of another term for cake: dàngāo (literally 'egg cake', suggesting the western type). Questions arise: How tightly bound is gāo to dàn and how tightly bound might be bǐng to yuè? I.e., is either of these an ānjìng type word, or can the part meaning 'cake' be broken off and used freely? In short, which one, if either, is a *general* term similar to English 'cake'? All good questions, but as it happens, neither bǐng nor gāo has the same *semantic* scope as English 'cake' and *that's* where one needs to get focused. It turns out that bǐng refers to things that are small and hard and dry (thus tiěbǐng = 'discus' and bǐnggān = 'biscuit'), while gāo refers to things that are large and soft and wet (or even squishy, as in zāugāo "messy + cake" = "What a mess!" or "Oh no!"). Thus, rather than worrying about how tightly "bound" bǐng or gāo might be (which in terms of linguistics would have been a question of morphology), we should have been worrying about what the two words actually *denote* (which switches us off into the realm of semantics, instead of morphology). In this particular case — which is not all that unusual — there's no easy answer. The best we can hope for is some kind of Venn diagram (written or mental) that might crudely show the three-way semantic overlap between 'cake' and bǐng and gāo. Not to say similar problems don't arise in studying European languages, but given the deeper time-line for China and the rarity of East/West contacts during those long millennia, the patterns of semantic mismatch and the need for "Venning" can be more vexing when studying Chinese than, say, French or German. In fact, anything might happen. (And it does!)

5. Advanced vocabulary:

<div style="text-align:center;">

NÁNG
囊
used in bēináng
backpack, book bag
dǎn-náng
gall bladder

</div>
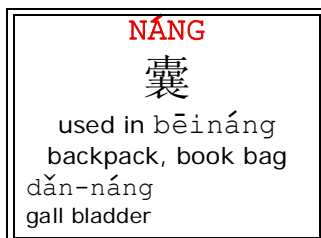
Skip over the 217 double-bordered cells, illustrated by NÁNG above, to traverse the basic vocabulary, more in accord with the needs of a first-year student. This category was covered in detail on page **15**. (Reminder: the arithmetic for reconciling this category to its notional inverse — "My First 903 Characters" — has two steps, not one.)

General observation: Many of the cells with double borders contain RED items (ALL-CAPS). When we think back on how the matrix was created — about the nature of the beast — this makes sense. See **Figure 3** on page **24**.

6. A cell may be nominally 'empty' but with a generic pointer ' * ' to a medium- or large-sized dictionary, possibly something on the order of the *Gwoyeu Tsyrdean*:

<div style="text-align:center;">

*

</div>
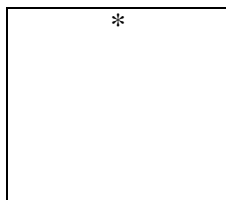
As for the philosophy *behind* the asterisk notation, that we've covered in detail via the discussion of gá and gǎ on page **19**f.

THE MATRIX

	1st	2nd	3rd	4th Tone
1				**A**
	aī 哎 oh no! / surprise = aīyā / aīyō	aí 挨 suffer, receive (scolding)	aǐ 矮 short (not tall)	aì 爱 to love
2	ān 安 peaceful = ānjìng(de)	*	*	àn 按 according to —————————— 岸按案暗
3	āng 肮 dirty = āngzāng	áng 昂 lift		àng 盎 abundant = àngrán
4	ĀO 凹 concave (opp.tū) āobù dent	áo 熬 boil or stew something áozhōu make congee	Ǎo 袄 coat mián-ǎo/jiá-ǎo padded/lined jacket	ào 澳 Macao (abbrev.) GǎngÀoTái HK, Macao, & Taiwan —————————— 拗傲奥澳懊
5				**B**
	bā 八 eight —————————— 八巴扒芭疤	bá 拔 pull up	bǎ 把 take; BA-construction	bà 罢 stop bàgōng to strike
6	baī 掰 break, pick baī yùmǐ pick corn	BAÍ 白 white	baǐ 百 hundred	baì 拜 visit = baìfǎng

	1st	2nd	3rd	4th Tone
7	bān 班 class, team shàngbān go to work ———————— 扳班般斑搬		bǎn 版 edition	bàn 半 half ———————— 办半扮伴瓣
8	BĀNG 帮 to help		bǎng 绑 tie up bǎngjià kidnap	bàng 棒 terrific, "great!"
9	bāo 包 include = bāokuò	báo 薄 thin (of things)	bǎo 饱 full (after eating)	bào 报 report (n. or v.) = bàodào ———————— 报抱豹暴爆
10	beī 杯 cup (as in "three cups of…") ———————— 杯卑背悲碑		BĚI 北 north	bèi 被 by (one of the several PASSIVE constr. markers) ———————— 背被倍辈
11	BĒN 奔 to speed bēnchí to speed along		běn 本 book bǐjìběn notebook	bèn 笨 stupid
12	bēng 崩 collapse xuěbēng avalanche	BÉNG 甭 no need to… Contraction of búyòng, represented by a sort of joke character I'd say.	BĚNG 绷 scowl běngliǎn pull a long face	bèng 迸 spout bèngfā burst out (laughing, e.g.)
13	BĪ 逼 to force, to pressure someone	bí 鼻 nose = bí.zi	bǐ 比 compare	bì 必 must ———————— 必闭毕避

	1st	2nd	3rd	4th Tone
14	biān 边 side, edge		biǎn 贬 reduce biǎnjià reduced price	<u>biàn</u> 便 convenient = fāngbiàn —————————— 变便遍辨辩
15	biāo 标 mark (n. or v.) biāodiǎn punctuation		biǎo 表 watch shǒubiǎo wristwatch	
16	biē 憋 suppress biēqì feel suffocated	bié 别 don't; other	BIĚ 瘪 dented; shriveled zhuàngbiě to dent (a car, e.g.)	biè 别 difficult = biè.niu (it is) difficult (working for this kind of person)
17	bīn 宾 guest		*	bìn 鬓 sideburns = bìnfà
18	bīng 兵 army, soldier		bǐng 饼 cake See discussion on page 37.	bìng 病 illness
19	<u>bō</u> 播 broadcast (n. or v.) gǒangbō broadcast/broadcasting —————————— 波拨玻剥播菠	<u>bó</u> 伯 uncle = bófù —————————— 伯博箔薄	bǒ 跛 lame = bǒ.de	bò 薄 peppermint = bò.he
20	bū 逋 abscond	bú 不 not changes to 2nd tone before 4th tone, as in: búhùi	bǔ 补 mend, supplement	<u>bù</u> 不 not —————————— 不布步部簿

	1st	2nd	3rd	4th Tone
21	**C**			
	CĀ 擦 wipe, rub			
22	CĀI 猜 guess	cái 才 not…until	cǎi 彩 multicolored / in color = cǎisè	CÀI 菜 vegetable; dish (on a menu)
23	cān 参 join cānjiā participate	cán 残 defective cánquē deficient	CǍN 惨 tragic, inhuman cǎnwúréndào simply inhuman	càn 灿 glorious = cànlàn
24	cāng 苍 blue, gray liǎnsè cāngbái pale (from illness)	CÁNG 藏 hide cánglóngwòhǔ hidden talent (lit. hidden dragon, sleeping tiger)	*	
25	cāo 操 control, engage in cāozuò diànnǎo operate a computer	cáo 嘈 noisy = cáozá(de)	CǍO 草 grass, grassy, (very) cursive, illegible cǎogǎo or cǎonǐ draft of a document. (To avoid a profanity, be careful of the tones on the latter compound.)	CÀO Like the character for béng (甭), the one for cào (肏) strikes me as a joke character, not quite real: enter + flesh = carnal knowledge. The word is profane, but its 'prime' status helps minimize collisions with regular words. But see cǎo.
26				cè 侧 side _____ 册厕侧测策

	1st	2nd	3rd	4th Tone
27	CĒN 参 uneven = cēncī (of appearance or quality)	CÉN 岑 a small spired mountain		
28	*	céng 层 floor, layer (lit. or fig.)	*	CÈNG 蹭 rub, smear, creep, scrounge
29	chā 差 difference also chābié, chāyì	chá 茶 tea	CHĂ 叉 to cross chǎ.zhe tuǐ with legs crossed	chà 差 to differ chà.buduō just about right
30	chāi 拆 tear open chāixìn open a letter	chái 柴 firewood = mùchái	*	*
31	chān 搀 *support by the arm; mix*	chán 禅 Zen = chánzōng	chǎn 产 to produce	chàn 颤 tremble / shiver = chàndǒu
32	chāng 昌 flourishing / thriving = chāngmíng	<u>cháng</u> 长 long ———————— 长肠尝常偿	chǎng 厂 factory = gōnchǎng	chàng 唱 sing = chànggē
33	chāo 超 exceed, super chāoshì supermarket	cháo 潮 tide, Zeitgeist sīcháo trend	chǎo 炒 stir-fry chǎomiàn chowmein, fried noodles	*
34	chē 车 car / automobile = qìchē		CHĚ 扯 pull, drag into	chè 彻 penetrate chèdǐ thorough

	1st	2nd	3rd	4th Tone
35	chēn 嗔 rebuke / scold = chēnguài	chén 陈 old, dated; a surname chéncí làndiào cliches	*	chèn 趁 take, ride on (e.g., an opportunity) chèn.zhe jīhùi take the opportunity to
36	chēng 撑 prop up	chéng 城 city —————————— 成诚承城乘程	chěng 逞 flaunt (one's ability) = chěngnéng	CHÈNG 秤 scales, balance
37	chī 吃 eat	chí 迟 late, tardy chízǎo sooner or later —————————— 池驰迟持匙	chǐ 尺 ruler (for measuring) = zhíchǐ	chì 赤 red; bare chìjiǎo.de barefoot
38	chōng 冲 rush forward (like a wave) chōnglàng to surf	chóng 虫 insect = kūnchóng	CHǑNG 宠 spoil, dote upon chǒngwù pet (n.)	chòng 冲 pungent
39	CHŌU 抽 take out, inhale chōuyān to smoke (i.e., take smoke out of cigarette)	chóu 绸 silk = sīchóu	chǒu 丑 ugly = chǒulòu	CHÒU 臭 smelly
40	chū 出 go out	chú 除 except = chú.le	chǔ 处 deal with = chǔlǐ	chù 处 place, department
41	chuāi 揣 conceal..in one's clothes	*	CHUǍI 揣 surmise / speculate = chuǎimó	CHUÀI 踹 kick, tread

	1st	2nd	3rd	4th Tone
42	chuān 川 *river; abbrev. for Szechuan province*	chuán 传 hand…down	chuǎn 喘 gasp = chuǎnxī	chuàn 串 string…together; conspire
43	chuāng 窗 window	CHUÁNG 床 bed	CHUǍNG 闯 cause trouble = chuǎnghuò	chuàng 创 create = chuàngzuò
44	chuī 吹 blow	chuí 垂 hang down		*
45	CHŪN 春 spring (season)	chún 纯 pure = chúnjìng	CHǓN 蠢 stupid, clumsy	
46	CHUŌ 戳 a seal or "chop" = chuō.zi			chuò 辍 stop (of painting, writing, studying…)
47	cī 疵 flaw or blemish = xiácī	<u>cí</u> 瓷 china(ware) = cíqì —————————— 词瓷辞慈磁	CǏ 此 this, here, now (formal)	cì 次 next
48	cōng 聪 clever = cōngmíng	cóng 从 from		
49				CÒU 凑 gather, encounter còuciǎo lucky, by (good) coincidence

	1st	2nd	3rd	4th Tone
50	**CŪ** 粗 coarse, crude, sketchy cūzhīyīèr have only a sketchy understanding of smthg	*		cù 醋 vinegar, jealousy chīcù be jealous (lit. ingest vinegar)
51	cuān 汆 to blanch, to stew	cuán 攒 assemble, huddle cuándòng to huddle together		cuàn 篡 falsify = cuàngǎi
52	cuī 催 to hurry someone or something		cuǐ 璀 dazzling = cuǐcàn	cuì 脆 crisp, brittle
53	cūn 村 village = cūnzhuāng	**CÚN** 存 exist = cúnzài	*	**CÙN** 寸 unit of length approx. 3 centimeters
54	cuō 搓 rub cuōshǒu rub hands together	cuó 痤 acne = cuóchuāng	*	cuò 错 wrong, wrongly
55	# D			
	dā 答 to answer / promise = dā.ying	dá 答 answer (n.) = huídá	**DǍ** 打 hit, strike	**DÀ** 大 big
56	daī 待 stay		daǐ 歹 evil dǎitú gangster	<u>daì</u> 带 take along _____ 贷怠袋逮戴

	1st	2nd	3rd	4th Tone
57	dān 单 *single; a list* ———————— 丹单担耽		dǎn 胆 courage	dàn 但 but ———————— 诞淡蛋弹氮
58	dāng 当 *should; act as*		dǎng 党 political party, faction	dàng 当 regard as = dàngzuò
59	dāo 刀 knife dāochā knife & fork	*	dǎo 倒 fall, lose, spoil shàngxiàdiāndǎo upside-down ———————— 导岛倒捣祷蹈	dào 道 road, the Way Dào Dé Jǐng The Tao Te Ching ———————— 到倒悼盗道稻
60	*	dé 得 get		
61			DĚI 得 must	
62	dēng 灯 lamp		DĚNG 等 wait	dèng 凳 stool = dèng.zi
63	dī 低 low = dī.de	dí 敌 enemy = dírén	dǐ 底 bottom = dǐbù	dì 地 Earth = dìqiú ———————— 地弟的帝递第
64	diān 颠 fall, upside-down shàngxiàdiāndǎo upside-down		diǎn 点 a little, some = yīdiǎn (yìdiǎr)	diàn 电 electric; lightning ———————— 电店垫淀惦殿

	1st	2nd	3rd	4th Tone
65	diāo 雕 carve diāokè carving (n.)		*	diào 调 melody / tone (lit. or fig.) = diào.zi —————————— 吊钓调掉
66	diē 爹 dad	dié 谍 espionage = diébào —————————— 迭谍牒叠碟蝶		
67	dīng 丁 cube of diced meat = ròudīng		dǐng 顶 extremely	dìng 定 settled, definite yídìng definitely
68	DIŪ 丢 lose			
69	dōng 东 east		dǒng 懂 understand	dòng 动 move
70	dōu 都 all (same character is used for 'capital' immediately below)		dǒu 斗 cup, pipe, funnel	dòu 豆 bean dòu.fu tofu
71	dū 都 capital = shǒudū	dú 读 read = dúshū —————————— 毒独读渎犊	dǔ 赌 gamble = dǔbó	dù 肚 belly = dù.zi —————————— 杜肚妒度渡

	1st	2nd	3rd	4th Tone
72	DUĀN 端 *ending; beginning; carry (in both hands, i.e., properly)*		DUĂN 短 short (not long)	duàn 断 break (arm, leg) = nòngduàn _____ 段断缎锻
73	DUĪ 堆 to pile duījī to pile up		*	duì 对 *correct, answer, pair*
74	dūn 蹲 to squat = dūnxià		dǔn 趸 wholesale = dǔnpī	dùn 盾 shield = dùnpái _____ 炖盾钝顿遁
75	duō 多 much	duó 夺 seize duóqǔ capture	duǒ 躲 hide / avoid = duǒbì	duò 惰 lazy = lǎnduò.de
76				E
	ē 阿 pander to ēyúfèngchéng creep & crawl to others	é 俄 Russia éyǔ Russian language _____ 讹俄鹅蛾额	ě 恶 nauseating / to feel nauseous = ěxīn	è 饿 hungry _____ 厄扼轭恶饿遏愕鳄呃
77	ĒN 恩 kindness ēnhuì favor			ÈN 摁 to press (e.g., a button or switch)

	1st	2nd	3rd	4th Tone
78	*	ér 儿 child, son nǚ-ér daughter (comes across as "female son")	ěr 耳 ear	èr 二 two
79	**F**			
	fā 发 send, develop, spread	<u>fá</u> 罚 fine / penalty = fákuǎn _____ 乏伐罚阀筏	fǎ 法 way / method = fāngfǎ	fà 发 hair = tóu.fa
80	fān 翻 translate = fānyì	fán 烦 trouble (you to…) = má.fan(nǐ...)	fǎn 反 oppose fǎn gémìng counter revolutionary	<u>fàn</u> 犯 commit (a crime or offense) _____ 犯饭泛范贩
81	fāng 方 direction nánfāng south	fáng 房 house = fáng.zi	fǎng 访 to call on fǎngwèn to visit	**FÀNG** 放 to put, to place fàngpì [1] to fart; [2] "You're talking nonsense!"
82	fēi 飞 to fly	féi 肥 fertilizer = féiliào	fěi 诽 slander (n. or v.) = fěibàng	<u>fèi</u> 费 spend _____ 吠肺废沸费
83	fēn 分 divide, assign, apportion	fén 坟 grave = fénmù	**FĚN** 粉 *powder, vermicelli, pulverize, pink*	<u>fèn</u> 份 a part, a share _____ 分份奋粪愤

	1st	2nd	3rd	4th Tone
84	fēng 风 wind —————— 丰风枫封疯锋蜂	féng 逢 come across féngchǎngzuòxì enjoy the fun while it lasts	fěng 讽 satirize = fěngcì	fèng 奉 *to present, believe in* fèngxiàn to dedicate —————— 凤奉俸缝
85		FÓ 佛 Buddha, Buddhism		*
86	*	*	FǑU 否 deny fǒujué veto	*
87	fū 夫 man fūzǐ scholar, old fogey	fú 浮 float = piāofú —————— 伏芙扶服俘浮符幅福	fǔ 腐 rotten fǔbài corrupt —————— 抚斧府俯辅腐	fù 复 duplicated, complex fùzá complicated —————— 付负妇附赴复副富腹覆
88				G
	gā 咖 curry = gālí	*	*	GÀ 尬 awkward = gāngà See discussion, page 20.
89	gāi 该 ought = yīnggāi		GǍI 改 change, alter	gài 概 *outline, probably* dàgài probably

	1st	2nd	3rd	4th Tone
90	**gān** 干 *have to do with; dry, dried...* mángguǒgān dried mangoes (lit. mango dried-things) ———————— 干甘肝柑尷		**gǎn** 感 feel ———————— 赶敢感橄擀	gàn 干 do
91	**gāng** 刚 just (now) ———————— 冈刚肛纲钢缸		gǎng 港 *harbor, Hong Kong* Gǎng Ào Tái HK, Macao, & Taiwan	gàng 杠 thick line, cross out
92	**gāo** 高 tall ———————— 高羔膏糕		gǎo 搞 do	gào 告 tell
93	**gē** 搁 put ———————— 哥胳搁割歌	**gé** 革 change gémìng (The) Revolution ———————— 革阁格搁蛤隔嗝	*	gè 个 unit for counting wǔ.ge five (of...)
94			**GĚI** 给 give	
95	gēn 跟 with, and	*	*	*
96	gēng 更 alter = gēnggǎi		gěng 梗 stalk (of flower or leaf)	**GÈNG** 更 even more...

	1st	2nd	3rd	4th Tone
97	<u>gōng</u> 工 work = gōngzuò gōng.fu free time; kungfu ―――――――― 工弓公功攻供宫恭躬		gǒng 拱 to join one's hands together = gǒngshǒu (cf. pěng)	gòng 共 together / altogether = yígòng
98	<u>gōu</u> 沟 ditch, gully gōutōng to communicate, speak openly ―――――――― 勾沟钩篝		gǒu 狗 dog	<u>gòu</u> 够 enough ―――――――― 勾构购垢够
99	<u>gū</u> 孤 alone, orphaned ―――――――― 估沽姑孤菇	*	<u>gǔ</u> 古 ancient (traditional) ―――――――― 古谷股骨鼓	<u>gù</u> 故 on purpose, therefore ―――――――― 固故顾雇痼
100	guā 瓜 melon		guǎ 寡 widow = guǎfù	guà 挂 hang, affix guàduàn hang up (telephone)
101	GUĀI 乖 well-behaved		GUǍI 拐 turn (into an alley, e.g.)	GUÀI 怪 *strange, to blame*
102	<u>guān</u> 观 look, view guānniàn concept ―――――――― 关观官冠棺鳏		guǎn 管 manage, control, govern, be in charge of bùguǎn… no matter…	<u>guàn</u> 惯 be accustomed to… = xíguàn ―――――――― 冠惯灌罐
103	guāng 光 light		guǎng 广 broad goǎngkuò(de) vast	guàng 逛 stroll guàngjiē go window shopping

	1st	2nd	3rd	4th Tone
104	guī 归 *return, belong to* —————————— 归龟规闺瑰		guǐ 鬼 ghost yángguǐ foreigner (It means "foreign devil" but nowadays it tends to be used only in jest.)	guì 贵 expensive —————————— 柜刽贵桂跪
105			gǔn 滚 *to roll; "scram!"* yáogǔnyuè rock and roll (music)	GÙN 棍 stick gùnbàng stick as weapon (club)
106	guō 锅 wok	guó 国 country, nation(al) wǒguó China (lit. our nation)	guǒ 果 fruit = shuǐguǒ	guò 过 pass through; to live; the past
107	H			
	hā 哈 laughter: ha-ha	há 蛤 toad = há.ma If pronounced as gé, then it means 'clam': gé, gélí, géchú	hǎ 哈 occurs in: hǎ.ba-gǒu pekingese dog; sycophant Note that the ShihTzu [shīzǐ] is a separate breed	
108	hāi 咳 or 嗨 'oh' or 'hey' to draw someone's attention	hái 还 still	HǍI 海 ocean	hài 害 to harm

	1st	2nd	3rd	4th Tone
109	hān 鼾 snore = dǎhān	hán 含 contain hányì implication —————————— 含函涵寒	hǎn 罕 rare = hǎnjiàn	hàn 汉 the Han people; Chinese (adj.); person hànzì Chinese characters (kanji) hǎohàn a real man, good guy —————————— 汉汗旱焊憾
110	hāng 夯 a rammer; to pound	háng 行 profession, line of work		hàng 沆 used in: hàngxièyīqì to be in cahoots with
111	hāo 蒿 artemesia = àihāo	háo 毫 fine hair háomǐ millimeter —————————— 蚝毫豪壕	HǍO 好 good	hào 号 number = hàomǎ —————————— 号好耗浩
112	hē 喝 to drink	hé 和 and, with —————————— 合何和河荷核		hè 贺 congratulate = zhùhè —————————— 贺荷喝赫褐鹤
113	HĒI 黑 black		*	
114	*	hén 痕 trace (n.) = hénjì	hěn 很 very	HÈN 恨 to hate
115	hēng 哼 to hum (also pronounced 'hng' for 'hmmph!')	héng 横 horizontal		hèng 横 harsh, unexpected hèngcái ill-gotten gains

	1st	2nd	3rd	4th Tone
116	hōng 轰 boom hōngzhà to bomb	hóng 红 red _____ 弘红宏虹洪鸿	hǒng 哄 to cheat = hǒngpiàn	hòng 哄 cause an uproar
117	*	hóu 猴 monkey = hóu.zi	HǑU 吼 to roar, to howl hǒujiào to howl	hòu 候 wait shí.hou time
118	hū 呼 exhale hūxī breathe	hú 糊 paste hú.tu confused (in mind) or chaotic (of things) _____ 狐胡壶葫湖煳蝴糊	hǔ 虎 tiger	hù 互 mutually = hùxiāng _____ 互户护糊
119	huā 花 *flower (n.), spend (v.)*	huá 华 flourishing; China, sino- huárén the Chinese people _____ 划华哗滑猾		huà 话 words, talk (n.) shuōhuà to talk _____ 化划画话桦
120		huái 怀 *keep in mind; become pregnant* huáiyùn become pregnant		HUÀI 坏 bad
121	huān 欢 happy xǐhuān to like	huán 还 return	huǎn 缓 slow, relaxed huǎnmàn.di slowly	huàn 换 exchange huànsuàn conversion (of currency) _____ 幻宦换唤患焕

	1st	2nd	3rd	4th Tone
122	huāng 荒 desolate, famine jīhuāng famine	huáng 黄 yellow huángsè yellow; smutty _____ 皇黄惶煌蝗磺簧	huǎng 谎 lie (n.) = huǎngyán _____ 恍晃谎幌	huàng 晃 shake yáohuàng to rock
123	huī 灰 ash, dust huīchén indoors dust (distinguish chéntǔ = outdoors dust) _____ 灰恢挥辉徽	huí 回 return	huǐ 悔 regret = hòuhuǐ	huì 会 *be able to (to know how); meet* huìyì meeting _____ 汇会讳绘贿彗惠喙慧
124	hūn 婚 marriage jiéhūn get married	hún 馄 wonton = hún.tun	*	hùn 混 mix; pass off a fake hùnluàn chaotic
125	*	huó 活 *to live; work (n.)* shēnghuó life	huǒ 火 fire	huò 货 currency; goods huòwù goods, merchandise _____ 或货获祸惑霍豁
126				J
	jī 机 machine = jīqì jīchǎng airport _____ 几击饥机肌鸡奇积基激	jí 极 extreme (adj.) = jíduān(de) _____ 及吉级极即急疾集嫉籍	jǐ 几 how many? or several _____ 几己挤给脊	jì 计 calculate = jìsuàn _____ 计记伎纪技系忌际妓季 剂迹既继祭寄寂髻

	1st	2nd	3rd	4th Tone
127	jiā 家 family, home ———————— 加夹茄佳家嘉	jiá 颊 cheek = miànjiá	jiǎ 假 false, pretend (adj.)	jià 价 price jiàgé biāoqiān price tag ———————— 价驾架假嫁
128	jiān 尖 pointed, sharp ———————— 尖奸歼坚间肩艰监兼煎		jiǎn 简 simple = jiǎndān jiǎnjié terse (like Classical Ch.) ———————— 拣茧柬捡检减剪简	jiàn 见 see ———————— 见件间建荐贱剑健渐践 腱溅鉴键
129	jiāng 将 is going to jiānglái future ———————— 江将姜浆僵缰疆		jiǎng 讲 speak	jiàng 降 drop, reduce jiàngwēn to cool, to wane (e.g. of a craze) ———————— 匠降将酱
130	jiāo 交 hand…in, pay ———————— 交郊浇娇骄胶教椒焦蕉 礁	JIÁO 嚼 chew	jiǎo 脚 foot ———————— 角狡饺矫脚搅	jiào 教 teach ———————— 叫觉校轿较教窖酵
131	jiē 接 *connect, receive* ———————— 阶疖皆结接揭街	jié 结 to tie, weave jiéhūn to get married ———————— 节劫杰洁结捷睫截	jiě 解 untie, solve jiěfàng liberate, the Liberation	jiè 借 borrow OR lend ———————— 介戒届界借

	1st	2nd	3rd	4th Tone
132	**jīn** 今 present, now ———————— 巾斤今金津筋		**jǐn** 紧 tight, pressing jǐnjí urgent ———————— 仅尽紧锦谨	**jìn** 进 enter ———————— 尽进近劲浸禁
133	**jīng** 精 essence wèijīng MSG (lit. flavor essence) ———————— 京经惊晶睛兢精鲸		**jǐng** 井 well jǐngdǐzhīwā limited perspective (lit. frog at bottom of well) ———————— 井阱颈景警	**jìng** 净 clean = gānjìng(de) ———————— 径净胫痉竞竟敬静境镜
134	*		jiǒng 窘 hard-up, embarrassed	
135	jiū 究 *investigate; actually*		jiǔ 久 a long time	**jiù** 旧 old, old-fashioned ———————— 旧咎救就舅
136	jū 拘 arrest jūliú detain; custody	jú 局 *chessboard, situation,* *authorities* dāngjú authorities (e.g. as quoted in a news item)	jǔ 举 raise jǔlì give an example	**jù** 句 sentence, a few words ———————— 巨句拒具俱剧距飓锯聚
137	juān 捐 donate		**JUǍN** 卷 roll up; a roll	juàn 卷 exam paper, document, volume
138	juē 撅 pout = juēzuǐ	**jué** 决 decide = juédìng ———————— 决诀角觉绝爵嚼	*	*

60

	1st	2nd	3rd	4th Tone
139	jūn 军 army jūnshì.de military _____ 军均龟君菌鞍		*	jùn 竣 to complete = jùngōng
140	**K**			
	kā 咖 used in transliteration kāfēi coffee		kǎ 卡 used in transliteration for 'calorie', 'card', 'cassette', e.g. shuākǎ swipe a [credit] card	
141	kāi 开 to open, switch on		kǎi 楷 model kǎishū characters in "printed" style, opp. cursive style	*
142	kān 刊 publish = kāndēng		kǎn 砍 to chop kǎndǎo to fell (a tree)	KÀN 看 look at, read, see; it depends on…
143	kāng 康 healthy = jiànkāng(de)	*	*	kàng 抗 resist / defy = kàngjù
144	*	*	kǎo 考 have an exam kǎolǜ consider, mull over	kào 靠 lean on, trust

	1st	2nd	3rd	4th Tone
145	kē 科 science `kējì` science & technology	ké 咳 cough = `ké.sou`	kě 可 *can; a prefix meaning '-able'; but; also used as an intensifier (before any verb)*	kè 课 subject, class, lesson ———————— 克刻客课
146			kěn 肯 willing ———————— 肯垦恳啃	*

147	kēng
	吭
	huffing & puffing =
	`kēng.chi`
	`kēng.chi.de`
			*	

148	kōng 空 empty		kǒng 恐 afraid that… = `kǒngpà`	kòng 控 to control = `kòngzhì`
149	kōu 抠 stingy = `kōu.ménr`		KǑU 口 mouth, opening	kòu 叩 knock `kòubài`, `kòutóu` kowtow
150	kū 哭 cry		KǓ 苦 bitter, difficult `chīkǔ` endure hardship	kù 裤 pants
151	KUĀ 夸 exaggerate `kuākǒu` boast		KUǍ 垮 collapse	kuà 跨 step, surpass

	1st	2nd	3rd	4th Tone
152	*		*	<u>kuài</u> 快 fast, soon _____ 会块快筷
153	kuān 宽 wide, broad, lenient		kuǎn 款 *funds, style* xiànkuǎn cash (in the sense of funds that are liquid) kuǎnshì style	
154	kuāng 筐 basket	KUÁNG 狂 crazy, wild, -phile huákuáng sinophile		<u>kuàng</u> 矿 mine (re ore, etc.) _____ 旷况矿框
155	kuī 亏 *be in the wrong; how could you...?*	kuí 葵 sunflower = kuíhuā or xiàngrìkuí	kuǐ 傀 puppet (political) = kuǐlěi Dist. mù-ǒu = puppet	kuì 愧 ashamed = xiūkuì
156	kūn 昆 insect = kūnchóng		KǓN 捆 tie...up; a bundle	KÙN 困 difficult, trap, predicament kùn.nan difficulties, troubles
157				kuò 扩 expand

	1st	2nd	3rd	4th Tone
158				**L**
	lā 拉 pull	lá 拉 tear (v.)	lǎ 喇 trumpet / flared = lǎ.ba	là 辣 hot, spicy _____ 落腊辣蜡
159		LÁI 来 come, bring		lài 赖 *depend; blame wrongly*
160		lán 蓝 blue = lánsè.de _____ 兰拦蓝篮	lǎn 懒 lazy	làn 烂 soft, rotten
161	*	láng 狼 wolf	lǎng 朗 bright, clear lǎngdú read...aloud	LÀNG 浪 wave; wasteful
162	LĀO 捞 take, wangle	láo 劳 to work; to trouble someone	lǎo 老 old	lào 烙 branding iron làoyìn to brand (cattle); make a deep impression
163	*			lè 乐 joy, happy lèjíshēngbēi out of extreme joy comes sorrow
164	LĒI 勒 tie tightly	léi 雷 thunder	lěi 累 accumulate lěilěi countless	lèi 累 tired _____ 肋泪类累

	1st	2nd	3rd	4th Tone
165	*	léng 棱 edge, corner seen also with 禾 radical sānléngjìng prism	lěng 冷 cold	lèng 愣 staggered (e.g., by news)
166	*	lí 离 be separated from (i.e., in different place) ———————— 厘离梨犁篱	lǐ 理 reason, logic wùlǐxué physics ———————— 李里俚理鲤	lì 力 force, strength ———————— 力历厉立沥丽励利例隶 荔砾粒痢
167			LIǍ 俩 the two of us; some	
168	*	lián 连 *connect, including, even* ———————— 连帘怜莲联	liǎn 脸 face diūliǎn lose face	liàn 练 to practice
169		liáng 良 good liángxīn conscience ———————— 良凉梁量粮	LIǍNG 两 two	liàng 量 capacity, measure cíhuìliàng one's vocabulary ———————— 亮凉谅辆量
170	LIĀO 撩 raise (curtain or skirt); put up (one's hair); sprinkle	liáo 聊 chat	LIǍO 了 finish	liào 料 *material, talent, expect* cáiliào material or talent yùliào expect

	1st	2nd	3rd	4th Tone
171	LIĒ 咧 babble, gossip = liē.lie	*	liě 咧 grin = liězuǐ	liè 列 to set…out, list, rank; measure for trains ————————— 列劣烈猎裂
172	līn 拎 carry	lín 林 woods, forest ————————— 邻林临鳞	lǐn 凛 cold, stern lǐnrán stern	lìn 赁 to rent fángwūchūlìn room for rent
173	*	líng 零 or 0 zero ————————— 伶灵凌铃陵菱零	lǐng 领 *collar, neck, lead, occupy, belong to* lǐngtǔ territory	lìng 令 to order, command
174	liū 溜 slide	liú 流 flow liúxíng popular, fashionable liúgǎn bìngdú flu virus ————————— 刘留流硫瘤	liǔ 柳 willow	liù 六 six
175		lóng 龙 dragon ————————— 龙聋笼隆	lǒng 垄 *ridge, monopolize* lǒngduàn monopoly, monopolize	lòng 弄 lane / alley = lòngtáng
176	LŌU 搂 to rake, extort	lóu 楼 tall building; floor (of a building)	lǒu 搂 to embrace	lòu 漏 to leak, divulge, leave out
177	*	lú 炉 stove	lǔ 虏 captive fúlǔ POW, to capture	lù 路 road, route ————————— 陆录鹿禄路露

	1st	2nd	3rd	4th Tone
178		Lǘ 驴 donkey	lǚ 旅 travel = lǚxíng ———————— 侣旅铝履	lǜ 绿 green = lǜsè.de ———————— 律虑绿率滤
179		luán 孪 twin (bros./sisters) = luánshēng(de)	LUǍN 卵 ovum = luǎn.zi	LUÀN 乱 disorderly, confused hùnluàn.de chaotic
180				lüè 略 summary, simplify
181	LŪN 抡 to brandish (a knife or one's fist, etc.)	lún 轮 wheel		LÙN 论 essay, theory; to discuss
182	luō 捋 to stroke	luó 罗 net for catching birds luózhī to frame someone ———————— 罗萝逻锣骡螺	LUǑ 裸 expose luǒtǐ naked	luò 落 fall luòhòu fall behind
183	M			
	mā 妈 mom	MÁ 麻 *hemp, numb, marijuana* dàmá marijuana	mǎ 马 horse	MÀ 骂 to tell…off áimà to be told off (lit. to suffer a scolding)
184		MÁI 埋 bury	MǍI 买 buy	mài 卖 sell ———————— 迈麦卖脉

	1st	2nd	3rd	4th Tone
185	*	mán 瞒 hide the truth bù mán nǐ shuō to tell you the truth…	MǍN 满 full, satisfied	màn 慢 slow
186		máng 忙 busy	mǎng 莽 *coarse grass; rash*	
187	MĀO 猫 cat	máo 毛 body hair; surname of Máo Zédōng	MǍO 铆 to rivet mǎodīng rivet (n.)	mào 貌 appearance = xiàngmào _____ 冒贸帽貌
188		méi 没 not have, did not Distinguish from shè! _____ 没玫眉梅媒煤霉	měi 美 *beautiful,* *American (adj.)*	mèi 妹 younger sister = mèi.mei
189	MĒN 闷 stuffy; shut oneself in	mén 门 door		mèn 闷 low, bored mènmènbúlè to feel really low
190	MĒNG 蒙 deceive = mēngpiàn	méng 萌 to sprout	měng 猛 fierce = měngliè(de)	mèng 梦 dream (n.) mèngxiǎng to dream about… or aspire to…
191	mī 眯 to squint	mí 迷 be lost	mǐ 米 rice (this is uncooked rice while fàn is cooked rice)	mì 秘 secret (adj.) = mì.mi.de

	1st	2nd	3rd	4th Tone
192		mián 棉 cotton = miánbù	miǎn 免 avoid miǎnfèi.de free of charge	MIÀN 面 face, aspect
193	MIĀO 喵 meow	miáo 描 describe = miáoshù	miǎo 秒 second (of time) miǎobiǎo stopwatch	miào 庙 temple = miàoyǔ
194	MIĒ 乜 to squint = miē.xié			miè 灭 go out (fire)
195		MÍN 民 the people = rénmín or mínzhòng	mǐn 敏 sensitive = mǐn-gǎn	
196		míng 名 name = míngzì	MǏNG 酩 get very drunk = mǐngdǐng	MÌNG 命 *life, fate; order (n.)* jiùmìng.a! Help!
197				MIÙ 谬 false miùlùn fallacy
198	MŌ 摸 to stroke, fish out (of pocket)	mó 蘑 mushroom = mógū ————————— 摹模膜摩磨蘑魔	MŎ 抹 to apply, wipe, erase	mò 莫 nobody, nothing, not mòmíngqímiào baffling (lit. cannot comprehend its wonder) ————————— 末没茉陌莫墨默磨

	1st	2nd	3rd	4th Tone
199		móu 谋 to plan, consult, plot	MǑU 某 a certain (person); some (way)	*
200		MÚ 模 appearance = múyàng	mǔ 母 mother = mǔqīn	mù 木 wood / timber = mùcái _____ 木目牧墓幕暮
201				N
	*	NÁ 拿 hold, with	NǍ 哪 which?	nà 那 that (one), those, in that case…
202			nǎi 奶 breast, milk dúnǎi the tainted milk scandal of 2009 (lit. poisoned milk)	nài 耐 endure = rěnnài
203	*	nán 难 difficult	*	NÀN 难 trouble zāinàn disaster
204	*	NÁNG 囊 used in bēináng backpack, book bag dǎn-náng gall bladder	NǍNG 攮 stab nǎng.zi dagger	*
205	NÃO 孬 cowardly, bad	náo 挠 scratch náo zìjǐ scratch oneself	nǎo 脑 brain	NÀO 闹 noisy rè.nao lively

	1st	2nd	3rd	4th Tone
206			**NĚI** 馁 disheartened, famished	**NÈI** 内 inside
207				**NÈN** 嫩 delicate, tender
208		**NÉNG** 能 can		*
209	*	**ní** 霓 neon = níhóng (borrowed from an old word for 'rainbow')	**nǐ** 你 you	<u>**nì**</u> 腻 oily, fed-up —————————— 逆匿腻溺
210	niān 拈 take, pick up	nián 年 year	niǎn 碾 grind; roller	**NIÀN** 念 read, study, idea
211		**NIÁNG** 娘 mom, auntie		**NIÀNG** 酿 brew = niàngzào
212			niǎo 鸟 bird	**NIÀO** 尿 urine
213	**NIĒ** 捏 to hold, make, mold	*		niè 孽 sin = zuìniè
214		**NÍN** 您 you (respectful form)		

	1st	2nd	3rd	4th Tone
215		<u>níng</u> 柠 lemon = níngméng ——————————— 宁柠柠凝	*	nìng 宁 rather = nìngkě or nìngkěn or nìngyuàn
216	*	NIÚ 牛 cow, ox, beef	niǔ 钮 button or knob	NIÙ 拗 obstinate
217		nóng 农 agricul-ture/-tural = nóngyè		NÒNG 弄 make, play with
218		NÚ 奴 slave (n.) = núlì 奴隶	nǔ 努 to try hard = nǔlì 努力	NÙ 怒 furious = nǎonù
219			Nǚ 女 woman = fùnǚ	
220			NUǍN 暖 warm	
221				nüè 虐 tyrannical = bàonüè(de)
222		NUÓ 挪 move, shift to one side	*	nuò 诺 promise; "here you are" (to a customer)

	1st	2nd	3rd	4th Tone
223	O			
	*	ó 哦 oh…also…	ǒ 嚄 oh…actually…	ò 哦 oh…now…
224	ōu 欧 European Ōu-Měi shèhuì Western society Used in transliteration: Obama = Ōubāmǎ (or Àobāmǎ in PRC w/ 奥)	*	ǒu 偶 *puppet; by chance* mù-ǒu puppet ǒurán by chance	òu 怄 to be annoyed = òuqì
225	P			
	pā 趴 lie prone = pāfú	pá 爬 crawl, climb, rise (lit. and fig.)		pà 怕 be afraid (of); may be
226	PĀI 拍 *to beat, shoot (a film)* qīngpāi to pat	pái 排 arrange, row	*	PÀI 派 faction, school of thought
227	pān 攀 climb = pāndēng, pānpá	pán 盘 tray = tuōpán chápán tea tray		pàn 叛 betray pàntú traitor
228	pāng 乓 ping-pong = pīngpāngqiú	páng 旁 side	*	PÀNG 胖 fat

	1st	2nd	3rd	4th Tone
229	pāo 抛 throw	páo 袍 gown / cheongsam = qípáo	**PǍO** 跑 run	pào 泡 light bulb = dēngpào
230	pēi 呸 bah!	péi 陪 go with, accompany		pèi 佩 admire = pèifú
231	**PĒN** 喷 gush = pēnchū	**PÉN** 盆 basin, pot pénzāi bonsai		**PÈN** 喷 delicious = pènxiāng
232	pēng 烹 to cook = pēngtiáo	péng 朋 friend = péngyǒu	**PĚNG** 捧 hold in both hands; handful (cf. gǒng)	**PÈNG** 碰 hit
233	pī 批 criticize/comment on = pīpíng	<u>pí</u> 皮 skin = pífū _____ 皮疲啤琵脾	pǐ 癖 favorite hobby or "addiction" = pǐhào	<u>pì</u> 譬 for example… = pìrú shuō... _____ 屁辟僻譬
234	<u>piān</u> 偏 slanting, biased _____ 片扁偏篇	pián 便 cheap = pián.yi	*	piàn 片 piece, slice
235	piāo 漂 float = piāofú	piáo 嫖 visit prostitutes	piǎo 漂 to bleach = piǎobái	piào 票 ticket, note, bill
236	piē 撇 cast aside		**PIĚ** 撇 leave aside, fling piěkāichéngjiàn leave aside one's prejudices	*

	1st	2nd	3rd	4th Tone
237	pīn 拼 put together (e.g., sounds) pīnyīn zìmǔ phonetic alphabet pīnmìng to go all out	pín 贫 poor = pínqióng	PǏN 品 *article, product, quality,* *to taste, to play the* *vertical flute (lit. or fig.)* wùpǐn article of merchandise pǐnxíng character and conduct	PÌN 聘 *betroth, engage, employ* But a more common term for 'to employ' is: gùyòng
238	PĪNG 乓 ping-pong = pīngpāngqiú	píng 平 flat, calm, even, equal píngjìng.de calm (not stormy) —————————— 平评坪苹凭屏瓶		*
239	pō 坡 gradient = pōdù	PÓ 婆 old lady pópomāmā fussy	PǑ 叵 impossible, un- xīnhuái pǒcè harbor the unfathomable (= dark designs)	pò 破 broken, worn out
240	PŌU 剖 dissect = jiěpōu	*	*	
241	pū 扑 rush at	pú 菩 bodhisattva = pú.sà —————————— 仆匍菩葡	pǔ 普 common = pǔtōng (pǔtōnghuà)	pù 铺 shop = pù.zi

	1st	2nd	3rd	4th Tone
				Q
242	qī 七 seven ——————— 七妻戚漆	qí 奇 strange = qíguài ——————— 齐其奇祈歧脐骑棋旗鳍	qǐ 起 rise qǐshēn rise (as from a chair) ——————— 乞岂企启起	qì 气 *gas, air, manner, anger, provoke* ——————— 气弃汽器
243	QIĀ 掐 pinch, throttle, nip	*	QIĀ 卡 *get stuck, wedge, block, fastener, hairpin, checkpoint*	qià 恰 appropriate = qiàdàng
244	qiān 千 thousand yíhuà shèng qiānyán A picture 'beats' 1000 words. ——————— 千迁牵铅谦签	qián 钱 money ——————— 前虔钱钳乾	qiǎn 浅 shallow	qiàn 歉 *poor, apology* bàoqiàn apologize
245	qiāng 枪 spear, gun, pistol	qiáng 强 strong	qiǎng 抢 rob qiǎngqián "highway robbery"	qiàng 呛 irritate
246	qiāo 敲 knock, blackmail qiāozhà to extort ——————— 跷锹敲橇	qiáo 桥 bridge ——————— 乔侨桥瞧	qiǎo 巧 *nimble, skillful; coincidental; false*	qiào 窍 hole, crux juéqiào knack
247	QIĒ 切 cut	QIÉ 茄 eggplant = qié.zi (See App. F)	QIĚ 且 *for the time being; even; and*	qiè 窃 steal = tōuqiè qiètīng to eavesdrop, to bug

	1st	2nd	3rd	4th Tone
248	qīn 亲 parent, relative, marriage qīnqī relative	qín 禽 avian, birds qínliúgǎnbìngdú avian flu virus (bird flu) —————————— 芹琴禽勤噙秦	QǏN 寝 sleep, bedroom, tomb (common in poetry, not in speech)	QÌN 沁 ooze, exude
249	qīng 清 clear = qīng.chu/qīngxī —————————— 青轻倾清蜻	qíng 情 feelings/emotions = gǎnqíng	qǐng 请 ask, invite, "please…"	qìng 庆 celebrate = qìngzhù
250	*	qióng 穷 the poor = qióngrén	*	
251	qiū 秋 autumn = qiūtiān	qiú 求 request (v.), demand (n.) —————————— 囚求酋球	*	
252	qū 区 *distinguish, area* qūbié to distinguish one thing from another —————————— 区曲驱屈躯趋	qú 渠 ditch, channel	qǔ 取 take, get, choose qǔxiāo cancel	qù 去 go
253	quān 圈 circle = quān.zi	quán 全 complete = wánquán(de) —————————— 权全泉拳颧	QUǍN 犬 dog, used in: quǎnkē canine	quàn 劝 advise

	1st	2nd	3rd	4th Tone
254	quē 缺 lack (n. or v.) = quēfá	QUÉ 瘸 be lame		què 却 *step back, decline; however*
255	*	qún 群 crowd = rénqún		
256				R
		rán 然 *correct, so, however* zìrán nature, naturally	rǎn 染 to dye, get the habit of	
257	RĀNG 嚷 to shout = rāngrāng	ráng 瓤 pulp of fruit, kernel of nut	rǎng 壤 soil = tǔrǎng pìrǎng remote backwater	RÀNG 让 make allowances, let, give way, transfer, by (PASSIVE construction)
258		ráo 饶 forgive ráomìng spare a life	RĂO 扰 disturb / interrupt = dǎrǎo	RÀO 绕 to wind, go around, make detour, confuse ràoxíng dàolù detour (n.)
259			RĔ 惹 provoke, cause rěyǎn showy (lit. provokes the eyes)	RÈ 热 hot

	1st	2nd	3rd	4th Tone
260		rén 人 human being	rěn 忍 endure rěnshòu/róngrěn tolerate / accept	<u>rèn</u> 认 know of, admit, accept rènmìng recognize one's limits, one's fate, etc. _____ 刃认任韧
261	RĒNG 扔 throw	RÉNG 仍 remain, still réngrán still	*	*
262				RÌ 日 day, sun; Japanese (adj.) rìqī date
263		<u>róng</u> 容 *facial expression,* *appearance, fit, tolerate* _____ 荣绒容溶熔融	RǑNG 冗 superfluous rǒngcháng long-winded	
264		róu 柔 soft, gentle qīngróu(de) gentle	*	RÒU 肉 meat, flesh
265		rú 如 be like, as, if	rǔ 乳 breast, milk, suckling rǔzhìpǐn dairy products	rù 入 enter
266		*	ruǎn 软 soft, gentle róuruǎn soft	

	1st	2nd	3rd	4th Tone
267		*	RUǏ 蕊 pistil Used in transliteration; see discussion on page 142.	ruì 锐 sharp(ly), vigour ruìlì astute
268		*		rùn 润 *sleek, lubricate,* *polish, profit*
269		*		ruò 弱 weak, young xūruò(de) frail, weak
270				S
	sā 撒 let go, lose control sāyě have a tantrum		sǎ 撒 scatter, sprinkle sǎbō to sow	sà 卅 thirty, thirtieth wǔsà May 30, 1925
271	sāi 塞 stuff into, cork huósāi piston (lit. "live cork")			sài 赛 strategic pass/location= yàosài
272	sān 三 three		sǎn 伞 umbrella, parasol, parachute sǎnbīng paratrooper	SÀN 散 disperse, scatter sànbù go for a stroll

	1st	2nd	3rd	4th Tone
273	sāng 桑 mulberry (tree) sāngyè mulberry leaves		SǍNG 嗓 throat, voice = sǎng.zi	SÀNG 丧 lose sàngqì be pessimistic sounding; unlucky
274	sāo 搔 to scratch (an itch)		sǎo 扫 to sweep	sào 扫 broom = sào.zhou
275				sè 色 color = yánsè
276	SĒN 森 dense, forest-like sēnlín forest			
277	SĒNG 僧 Buddhist monk = sēnglǚ			
278	shā 杀 kill —————————— 杀沙纱刹煞鲨	SHÁ 啥 contraction of shén.ma	SHǍ 傻 stupid, foolish	shà 厦 skyscraper = mótiāndàshà
279	SHĀI 筛 to sieve shāi.zi sieve (n.)		SHǍI 色 or 骰 dice = shǎi.zi	SHÀI 晒 to sun-dry or tan shài tàiyáng to sunbathe

	1st	2nd	3rd	4th Tone
280	<u>shān</u> 山 mountain _____ 山删衫珊舢煽		shǎn 闪 dodge, sprain, flash	<u>shàn</u> 善 kind(-hearted) = shànliáng _____ 扇善擅膳缮
281	shāng 商 discuss, commerce, businessperson gǎngshāng a Hong Kong businessman		shǎng 赏 award (n.), admire, appreciate xīnshǎng appreciate	shàng 上 upper part, go up, top shàngxué go to school
282	shāo 烧 burn, have a temperature fāshāo have a fever	sháo 勺 ladle	SHĂO 少 few, lack	shào 少 young, teenager kuòshào rich kid
283	shē 赊 buy on credit shēzhàng on credit	shé 舌 tongue = shé.tou	SHĚ 舍 abandon, give (altruistically)	<u>shè</u> 设 to set up, plan, suppose shèbèi equipment Distinguish from méi! _____ 设社舍射涉赦摄
284		SHÉI 谁 who?		
285	<u>shēn</u> 身 body = shēntǐ _____ 申伸身呻参绅深	shén 什 what = shén.me (shém.me) ...shén.me.de ...and so on (or what-not)	shěn 审 careful, review, try(case) shěnchá censor or examine	shèn 甚 extreme shènzhì (so...that) even...

	1st	2nd	3rd	4th Tone
286	shēng 生 *give birth, grow, student, living, raw, unfamiliar, stiff, very* xìng shēnghuó sex life ――――――― 升生声牲甥	SHÉNG 绳 rope = shéng.zi	SHĚNG 省 save, omit; province shěngshì save (avoid) trouble	shèng 胜 victory/victorious = shènglì yíhuà shèng qiānyán A picture 'beats' 1000 words ――――――― 圣胜盛剩
287	shī 诗 poetry, a poem ――――――― 尸失师诗虱狮施湿嘘	shí 时 hour = xiǎoshí time = shíjiān ――――――― 十石时识实拾食蚀	shǐ 使 *use, make (cause), ambassador* dàshǐguǎn embassy ――――――― 史矢使始驶屎	shì 事 thing, matter, accident, problem ――――――― 士示世仕市式事势侍饰试视是适释
288	SHŌU 收 *put away, harvest, arrest, gain*	*	shǒu 手 hand	shòu 受 receive, suffer, bear shòuhuìzhě bribee ――――――― 寿受狩兽授售瘦
289	shū 书 write, book, letter, document qíngshū love letter ――――――― 书叔殊梳舒疏输蔬	shú 熟 ripe, cooked, familiar shúshí cooked food	shǔ 鼠 mouse OR rat shǔbiāo mouse (on computer) ――――――― 属暑署鼠数薯	shù 数 number shùjù data (also zīxùn) ――――――― 术束述树竖恕数漱
290	SHUĀ 刷 brush (n.), scrub shuākǎ swipe a (credit) card; pay by credit		SHUĂ 耍 play, juggle, mess…around shuǎnùng.le taken for a "ride"	SHUÀ 刷 pale = shuàbái

	1st	2nd	3rd	4th Tone
291	shuāi 摔 fall, break, throw shuāijiāo fall over; come to grief; wrestling		SHUǍI 甩 to swing shuǎishǒubùguǎn wash one's hands of…	shuài 帅 commander-in-chief shuài.qi handsome
292	shuān 闩 bolt (n. or v.)			SHUÀN 涮 rinse, instant boil, be tricked (bèi shuàn.le)
293	shuāng 双 two, even, pair, double shuāngguān(yǔ) double-entendre		SHUǍNG 爽 clear, frank, refreshing	
294		*	SHUǏ 水 water shuǐdiàn utilities (water & electr.)	shuì 睡 be asleep
295			SHǓN 吮 suck	shùn 顺 with, along the way, follow, obey …shùnbiàn on the way to…
296	SHUŌ 说 say			shuò 硕 large shuòshì master's degree, M.A.
297	sī 思 think, thought sīxiǎng thought _____ 司丝私思撕嘶		sǐ 死 die; extremely lèisǐ.le totally exhausted	sì 四 four (an inauspicious number because of sǐ)

	1st	2nd	3rd	4th Tone
298	**SŌNG** 松 pine tree; relax, loose, "soft on" sōngkāi loosen (clothing)	*	sǒng 耸 to tower, alarm sǒngréntīngwén sensationalize	sòng 送 deliver, see off sòngbié wǎnhuì farewell banquet
299	sōu 搜 search sōusuǒ yǐnqíng search engine		*	**SÒU** 嗽 cough = ké.sou
300	sū 酥 biscuit, limp sūruǎn to feel weak	**SÚ** 俗 custom, vulgar rùxiāng suísú When in Rome…		sù 诉 tell = gàosù or gào.su —————————— 诉肃素速宿粟塑
301	**SUĀN** 酸 sour suānlàtāng hot & sour soup		*	suàn 算 calculate suànpán abacus
302	suī 虽 although = suīrán	suí 随 go along with, follow one's inclination suíbiàn do as one likes	**SUǏ** 髓 marrow = gǔsuǐ	suì 岁 year (of age or reign) …wànsuì long live… —————————— 岁碎隧穗
303	**SŪN** 孙 grandchild sūn.zi grandson		sǔn 损 damage = sǔnhuài	*
304	suō 缩 contract, withdraw suōxiě abbreviation	*	suǒ 所 place, reason for suǒyǐ so, therefore —————————— 所索琐锁	*

	1st	2nd	3rd	4th Tone
305				T
	tā 他 he, she (她), it (它)		tǎ 塔 pagoda = bǎotǎ	tà 踏 step on, tread
306	TĀI 胎 foetus, tire lúntāi tire (n.)	tái 台 tower, stage, Taiwan táifēng typhoon		tài 太 highest, too, so
307	tān 摊 *spread out, fry, share, stall (n.)*	tán 谈 to talk, chat ―――――――― 坛谈弹潭檀	tǎn 毯 carpet = dìtǎn	tàn 探 explore, scout (v. or n.) tànjiū to probe (investigate)
308	tāng 汤 hot water, soup	táng 糖 sugar ―――――――― 唐堂膛糖螳 2	tǎng 躺 lie down	tàng 烫 hot, scalding, iron
309	tāo 掏 take out, dig, steal	táo 桃 peach táosè peach colored; sexual ―――――――― 逃桃陶淘	TǍO 讨 demand, discuss tǎolùn discuss	TÀO 套 *cover, trick into, copy,* MEASURE *for suits/ stamps/ volumes* ānquán-tào condom
310	*			TÈ 特 special tèbié specially, peculiar

	1st	2nd	3rd	4th Tone
311	*	téng 疼 sore, dearly yáténg toothache Compare tòng		
312	tī 梯 ladder diàntī elevator	tí 提 raise, mention huàtí topic ———————————— 提啼题蹄	TǏ 体 body = shēntǐ tǐwēn temperature	tì 替 replace, for tìzuìyáng scapegoat
313	tiān 天 *sky, heaven, day, season, weather, nature, God, overhead*	tián 田 field	tiǎn 舔 lick	*
314	TIĀO 挑 carry on pole, choose tiāotì nitpick	tiáo 调 harmonize, blend, provoke tiáopí naughty	TIǍO 挑 raise, prick, incite	tiào 跳 jump
315	tiē 贴 stick, subsidize tiēshēn yīfú underclothes		tiě 铁 iron tiělù railway	TIÈ 帖 a book of model calligraphy; a rubbing from an inscription
316	tīng 听 listen	tíng 停 stop ———————————— 廷亭庭停	tǐng 挺 straighten, very	*

	1st	2nd	3rd	4th Tone
317	TŌNG 通 connect with, workable	tóng 同 be the same tóngbāo compatriot; brother & sister _____ 同铜童瞳	tǒng 统 system = xìtǒng _____ 统捅桶筒	tòng 痛 ache tóutòng headache Compare téng
318	TŌU 偷 steal	tóu 头 head	*	TÒU 透 penetrate, thorough
319	tū 突 break out tūrán sudden, suddenly	tú 图 picture, plan dìtú map _____ 图徒途涂屠	tǔ 土 *soil, land, territory, opium, local, folk* tǔchǎn local product	tù 兔 rabbit = tù.zi
320	TUĀN 湍 torrential, rapids	TUÁN 团 ball, group tuánjié unite	*	*
321	tuī 推 push	TUÍ 颓 decline, dejected	TUǏ 腿 leg	tuì 退 retreat, fade tuìxiū retire
322	TŪN 吞 to swallow = tūnxià	tún 屯 hoard = túnjī	*	TÙN 褪 take off clothes
323	tuō 脱 take off clothes, shed	tuó 驮 carry on one's back	tuǒ 妥 appropriate tuǒxié compromise	tuò 唾 saliva, spit = tuòyè

	1st	2nd	3rd	4th Tone
324	**W**			
	wā 挖 dig	WÁ 娃 baby, doll = wá.wa	WǍ 瓦 tile	wà 袜 sock(s) = wà.zi
325	WĀI 歪 slanting		WǍI 崴 to sprain	WÀI 外 outside, foreign
326	wān 湾 bay Hǎiwān Zhànzhēng The Gulf War	wán 玩 play, have fun, visit _____ 丸完玩顽	wǎn 晚 late wǎn.shang evening _____ 挽晚婉碗	wàn 万 ten thousand wànyī if by any chance…
327	WĀNG 汪 expanse of water, watery, tearful	wáng 王 king	wǎng 网 net wǎngmí internet fan	wàng 望 to watch, to hope juéwàng.de zhǔfù Desperate Housewives _____ 妄忘旺望
328	wēi 危 dangerous = wēixiǎn wēijī crisis	wéi 为 act as, be wéinán embarrassed (lit. "be in difficulty") _____ 为违围帷惟维	wěi 伟 great = wěidà _____ 伟伪苇尾纬委萎猥	wèi 味 flavor wèijīng MSG (lit. "flavor essence") _____ 卫为未位味胃谓喂慰
329	wēn 温 warm wēndù temperature	wén 文 writing, culture wényán Classical Chinese _____ 文纹闻蚊	wěn 稳 steady, firm	WÈN 问 ask

	1st	2nd	3rd	4th Tone
330	wēng 翁 old man bǎiwàn fùwēng millionaire		*	wèng 瓮 jar
331	wō 窝 nest		wǒ 我 I, our wǒguó China (lit. "our nation")	wò 卧 lie down, sleeping wòshì bedroom
332	wū 污 dirt, dirty wūrǎn pollute, pollution —————————— 乌污巫诬屋	wú 无 not have wúxíng invisible, imperceptible	wǔ 午 (at) noon = zhōngwǔ —————————— 五午武侮舞	wù 误 mistake wùhuì misunderstand(ing) —————————— 勿务物误恶悟雾
333				X
	xī 西 west (direction) = xīfāng xīyáng The West —————————— 夕西吸希昔析牺息惜晰 稀熄	xí 习 to practice = liànxí —————————— 习席袭媳	xǐ 喜 to like = xǐ.huan	xì 细 thin, fine, detailed xìjūn germ —————————— 戏系细隙
334	xiā 虾 prawn xiǎoxiā shrimp lóngxiā lobster	xiá 瑕 flaw or blemish = xiácī xiábùyǎnyú See page 172 for tr. —————————— 匣侠峡狭瑕辖		xià 下 lower part or level, go down

	1st	2nd	3rd	4th Tone
335	<u>xiān</u> 先 first, earlier xiānshēng Mr., teacher —————————— 仙先纤锨鲜	<u>xián</u> 闲 leisure = xiánxiá —————————— 闲贤弦舷嫌	xiǎn 显 display, be apparent xiǎnrán obviously	<u>xiàn</u> 现 *present, ready (money),* *on the spot, reveal* —————————— 现限线宪陷羡献
336	<u>xiāng</u> 乡 countryside, home —————————— 乡相香箱	xiáng 详 detailed = xiángxì(de)	xiǎng 想 think	<u>xiàng</u> 像 portrait, look like, as if —————————— 向项象像橡
337	<u>xiāo</u> 消 disappear, spend, take xiāoxī news —————————— 削消宵销	XIÁO 淆 confused	xiǎo 小 small	<u>xiào</u> 孝 filial piety xiàoshùn show filial obedience —————————— 孝校哮笑效
338	xiē 些 some / a few = yīxiē	<u>xié</u> 鞋 shoe —————————— 协邪胁挟偕斜携鞋	xiě 写 write	<u>xiè</u> 谢 thank —————————— 泄卸械亵谢蟹
339	<u>xīn</u> 心 heart, mind —————————— 心辛欣新薪	*	*	xìn 信 believe, be religious, letter, information xìnxī shídài Information Age
340	xīng 兴 prosper, rise, popular	<u>xíng</u> 行 walk, travel, OK bùxíng that won't do —————————— 刑行形型	xǐng 醒 come to, sober up wèixǐng not yet sobered up	<u>xìng</u> 兴 happy = gāoxìng —————————— 兴杏幸性姓

	1st	2nd	3rd	4th Tone
341	xiōng 兄 brothers (elder & younger) = xiōngdì	xióng 熊 bear xióngmāo panda		*
342	xiū 休 rest = xiū.xi		xiǔ 朽 rot, go senile	xiù 袖 sleeve = xiù.zi _____ 秀袖绣锈嗅
343	xū 需 need = xūyào _____ 须虚墟需嘘	XÚ 徐 slowly, gently = xúxú	xǔ 许 maybe / perhaps = yěxǔ	xù 续 continue = jìxù _____ 序叙畜绪续絮婿蓄
344	xuān 宣 announce = xuānbù	xuán 玄 mystery = xuánxū	XUǍN 选 choose, vote, anthologize xuǎnjǔ elect	xuàn 眩 dazzle = xuànmù
345	xuē 靴 boot = xuē.zi	xué 学 to study	xuě 雪 snow	xuè 血 blood = xuèyè
346	xūn 熏 smoked (of meat) = xūnzhì(de)	xún 寻 search = xúnzhǎo _____ 旬寻巡询循		xùn 训 teach, train, rule (n.) xùnliàn training _____ 训讯驯徇殉

	1st	2nd	3rd	4th Tone
347	**Y**			
	yā 鸭 duck ――――――――― 丫压押鸦鸭	yá 牙 tooth = yáchǐ	yǎ 雅 elegant	yà 亚 Asia = yàzhōu
348	yān 烟 smoke ――――――――― 咽烟胭阉淹腌	yán 言 words yǔyán language ――――――――― 延严言岩炎沿研盐颜檐	yǎn 眼 eye = yǎn.jing	yàn 验 a trial / to test = shìyàn ――――――――― 厌砚咽宴唁验谚焰燕赝
349	yāng 央 plead, center zhōngyāng central	yáng 阳 The Yang of Yin & Yang ――――――――― 扬羊阳杨佯洋	yǎng 养 foster	yàng 样 style, type yīyàng the same (lit. one type)
350	yāo 要 ask for something, demand, threaten ――――――――― 幺夭妖要腰邀	yáo 摇 shake, sway yáolán cradle (lit. and fig.) ――――――――― 窑谣摇遥	yǎo 咬 bite	yào 要 important, to want, to order (in restaurant) ――――――――― 药要钥耀
351	yē 椰 coconut = yē.zi	YÉ 爷 grandfather = yé.ye	yě 也 also	yè 业 industry = gōngyè or hángyè ――――――――― 业叶页曳夜液腋
352	yī 衣 clothing = yī.fu ――――――――― 一衣伊医依壹	yí 移 move, change yízhí transplant ――――――――― 仪宜姨移遗疑	yǐ 以 *use, by, for, within* yǐwéi to think that…, be under the impression that… ――――――――― 乙已以蚁倚椅	yì 意 meaning, wish, expect yì.si meaning, idea, interest ――――――――― 亿义艺忆议亦异译抑役 易疫益谊翌裔意溢翼

	1st	2nd	3rd	4th Tone
353	yīn 阴 The Yin of Yin & Yang ———————— 因阴音姻殷	yín 银 silver yínháng bank	yǐn 隐 conceal, secret ———————— 引饮隐瘾	yìn 印 to stamp or print yìnxiàng impression Distinguish yǐngxǐang
354	yīng 英 flower, hero, Britain yīngxióng hero ———————— 应英婴樱鹦鹰	yíng 迎 welcome = huānyíng ———————— 迎荧盈萤营蝇赢	yǐng 影 shadow = yǐng.zi yǐngxiǎng affect, influence Distinguish yìnxiàng	yìng 应 respond huíyìng respond, answer
355	yōng 拥 embrace = yōngbào	*	yǒng 永 eternal = yǒngjiǔ ———————— 永泳俑勇涌	yòng 用 use
356	yōu 优 excellent; abundant; give preferential treatment yōuyǎ elegant ———————— 优忧幽悠	yóu 油 oil yóumén accelerator (pedal) ———————— 尤由邮犹油鱿游	yǒu 有 to have	yòu 右 right (right-hand) ———————— 又右幼诱
357	yū 迂 circuitous yūfǔ pedantic	yú 余 remain / be left = shèngyú ———————— 于余鱼娱渔愉愚	yǔ 语 language yǔwén Chinese (as an item in THEIR curriculum: one of the 3 R's) ———————— 与予宇羽雨语	yù 预 in advance ———————— 玉芋郁育狱浴预域欲遇 寓愈

	1st	2nd	3rd	4th Tone
358	yuān 冤 injustice `yuān.jia` enemy (also lover)	<u>yuán</u> 圆 round —————————— 元园员原圆援缘猿源	YUǍN 远 far	<u>yuàn</u> 愿 wish, desire, promise `yuànyì` be willing to —————————— 苑怨院愿

359	yuē
	约
	about (approx.)
				<u>yuè</u>
				越
				Viet Nam =
				`yuènán`
				`yuè...yuè...`
				the more…the more…
				——————————
				月乐岳阅悦越

360	YŪN 晕 dizzy = `tóuyūn`	yún 云 cloud	yǔn 陨 fall from the sky `yǔnxīng` meteorite	<u>yùn</u> 运 move, transport, luck `yùndòng` exercise, sport —————————— 孕运晕酝韵熨
361	# Z			
	zā 咂 make noises or clucks of appreciation = `zāzuǐ`	zá 杂 complicated = `fùzá`	ZǍ 咋 why?	
362	zāi 灾 disaster = `zāinàn`		zǎi 宰 rule, dominate, slaughter	zài 在 live, rest with, at `xiànzài` now
363	zān 簪 hairpin	ZÁN 咱 we (inclusive) = `zán.men`	ZǍN 攒 save	zàn 赞 assist, commend `zànchéng` approve

	1st	2nd	3rd	4th Tone
364	zāng 脏 dirty zāngzì swear word		*	zàng 脏 organ, as in: xīnzàng, shènzàng heart, kidney, etc.
365	zāo 糟 dregs, messy zāogāo Oh no!	ZÁO 凿 to chisel	zǎo 早 early zǎochén morning	zào 造 make, concoct, harvest chuàngzào create _____ 皂灶造燥躁
366		zé 责 responsibility = zérèn _____ 则责择泽	*	ZÈ 仄 slanting; the oblique tones in prosody (shǎng, qù, rù)
367		ZÉI 贼 thief		
368	*		ZĚN 怎 how?	ZÈN 谮 slander = zènyán
369	zēng 增 increase (n. or v.) = zēngzhǎng			zèng 赠 to present, donate juānzèng donate
370	zhā 扎 *prick, set up camp, plunge into*	zhá 炸 fry in oil	zhǎ 眨 wink	zhà 炸 blow up zhàyào explosive (n.)

	1st	2nd	3rd	4th Tone
371	zhāi 摘 pick, take extracts from zhāiyào summarize; an abstract	zhái 宅 house = zhùzhái	**ZHǍI** 窄 narrow	zhài 债 debt huánzhài repay a debt
372	**zhān** 沾 moisten, stick, touch zhānguāng benefit from association w/ _____ 占沾毡粘瞻		zhǎn 展 develop, exhibition zhǎnlǎnhuì exhibition	**zhàn** 站 stand up _____ 占战站颤蘸
373	zhāng 张 *open, extend, open for business, M for desk, surname Zhang* Zhāngguān Lǐdài confuse things (lit., Zhang's hat, Li wears it)		zhǎng 长 oldest, head; form, grow, acquire jǐngzhǎng police chief	**zhàng** 丈 husband = zhàng.fu _____ 丈仗帐账胀障
374	zhāo 招 beckon = zhāoshǒu	**ZHÁO** 着 *touch, be affected by, be lit, fall asleep* zháojí worried	zhǎo 找 look for	**zhào** 照 light up, reflect zhàoxiàng take a picture _____ 召兆照罩
375	zhē 折 toss and turn = zhē.teng	zhé 折 break, lose, convert, fold, discount zhékòu discount (n.)	zhě 者 the one who… jiàodìngzhě editor	zhè 这 this (also zhèi)
376	**zhēn** 真 true, real, genuine zhēnzhèng true _____ 贞针侦珍真斟		zhěn 枕 pillow = zhěn.tou	**zhèn** 阵 battle formation, burst zhènwáng killed in action _____ 阵振震镇

	1st	2nd	3rd	4th Tone
377	zhēng 争 contend that… = zhēngbiàn dào.. ———————— 争征症蒸		zhěng 整 whole	zhèng 正 straight, right, pure, authentic, just (right) ———————— 正证政挣症
378	zhī 知 know = zhīdào ———————— 只汁芝枝知肢织脂蜘	zhí 直 straight, vertical zhíjiē direct ———————— 执直侄值职植殖	zhǐ 指 point to, point out, refer to ———————— 只旨纸指趾	zhì 制 manufacture = zhìzào ———————— 至志制质治挚致智痣置 雉稚
379	zhōng 中 center zhōngguó China ———————— 中忠终钟衷		zhǒng 种 species, seed, race gèzhǒng… all kinds of …	zhòng 种 sow, plant zhòngdòu vaccinate ———————— 中仲众种重
380	zhōu 粥 porridge, congee ———————— 州舟周洲粥	zhóu 轴 axle, axis	zhǒu 肘 elbow	zhòu 皱 wrinkles (on skin) = zhòuwén
381	zhū 猪 pig ———————— 朱珠株诸猪蛛	zhú 竹 bamboo = zhú.zi	zhǔ 主 *host, God, main, take charge, idea, be in favor of*	zhù 住 reside (at)… ———————— 助住注驻祝著筑
382	ZHUĀ 抓 grab		ZHUǍ 爪 claw = zhuǎ.zi	
383	ZHUĀI 拽 throw, hurl		*	ZHUÀI 拽 pull

	1st	2nd	3rd	4th Tone
384	zhuān 专 concentrate, dominate zhuānyè major (at school)		**ZHUǍN** 转 turn zhuǎnyǎn before you know it	zhuàn 转 turn zhuàndòng turn part of body
385	zhuāng 装 to put on (lit. or fig.), to install jiǎzhuāng pretend		*	zhuàng 状 shape, state, certificate zhuàngtài condition
386	zhuī 追 chase, follow, seek zhuībǔ pursue and capture			zhuì 坠 fall, droop, pendant
387	zhūn 谆 earnest = zhūnzhūn		**ZHǓN** 准 allow, standard, accurate zhǔnbèi prepare	*
388	zhuō 桌 table = zhuō.zi	zhuó 着 *wear, come into contact with, send* ———————— 灼桌酌啄着琢镯		

	1st	2nd	3rd	4th Tone
389	zī 资 *money, capital, postage, ability, qualifications, to aid, data* zīxùn data		zǐ 子 son, seed, "penny", affiliated zǐgōngsī subsidiary company _____ 子仔姊籽紫	zì 自 oneself / myself = zìjǐ
390	zōng 宗 ancestor, clan, school zōngjiào religion _____ 宗综棕踪鬃		ZǑNG 总 gather, total, always	zòng 纵 north-south
391	*		ZǑU 走 walk, leave zǒusī smuggle, smuggled	zòu 奏 play (music), produce zòuxiào have effect (e.g. of medication)
392	ZŪ 租 to rent	zú 足 ample, enough, foot zhīzú be content, satisfied	zǔ 组 to form, set up; group (n.) zǔzhī organize _____ 诅阻组祖	*
393	ZUĀN 钻 to drill		ZUǍN 纂 compile, edit	ZUÀN 钻 a drill
394	*		ZUǏ 嘴 mouth	zuì 最 most

	1st	2nd	3rd	4th Tone
395	zūn 尊 senior, respect zìzūn self-respect (dignity)		*	*
396	zuō 作 workshop = zuō.fang	zuó 昨 yesterday = zuótiān	zuǒ 左 left (left-hand)	zuò 作 make, do _____ 作坐座做

APPENDIX A: AUFBAU (BUILDING UP)

> The idea of breaking the world into pieces and then explaining the pieces in terms of smaller pieces is called reductionism. It would be perfectly justified to consider Gell-Mann, the father of the quark, to be the century's arch-reductionist. But very early on, long before mushy notions of holism became trendy, Gell-Mann appreciated an important truth: While you can reduce downward, that doesn't automatically mean you can explain upward. People can be divided into cells, cells into molecules, molecules into atoms, atoms into electrons and nuclei, nuclei into subatomic particles, and those into still tinier things called quarks. But, true as that may be, there is nothing written in the laws of subatomic physics that can be used to explain higher-level phenomena like human behavior. There is no way that one can start with quarks and predict that cellular life would emerge and evolve over the eons to produce physicists. Reducing downward is vastly easier than explaining upward — a truth that bears repeating.
>
> — Johnson[26] (1999) p. 8–9

Linguistics and physics are among the several disciplines where one should care about both the 'reducing downward' and 'explaining upward', but as suggested in the passage quoted above, the latter is likely to get short shrift. In linguistics, given the interest in semantics and grammar and morphology and phonology, one senses a proficiency in reducing downward (not to say a constant danger of reductionism) but no particular interest in explaining upward, unless we include transformational grammar (alias generative grammar) in the mix.[27] In physics, though, we at least have a term that expresses the idea very directly, and without the awkwardness of Johnson's 'explaining upward' or 'mushy ... holism'. The term is *Aufbau* (German for 'building up'). Its history in German speaking areas seems straightforward, but

---

26. References for this appendix will be found in a *separate list* on page **104**. (All references outside of this appendix are keyed, as expected, to the Bibliography proper on page **186**.)

27. At first blush, transformational grammar (hereafter TG) is refreshing as a subdiscipline that professes an interest in 'explaining upward'. But upward from what? Not from primitives, not from first principles, but from arbitrary equations in a subterranean fantasyland whose sole constraint is that it must coincide with reality all along its top surface only. In effect, TG says, "WHEREAS some humans speak using Japanese syntax and other humans speak using English syntax and other humans speak using Turkish syntax, THEREFORE we hereby proclaim the existence of a magical Ur-Japanese-English-Turkish Seed in the infant brain that can sprout any one of those syntactical crops on demand." Sorry, I'm not buying. One might just as well say, "Oxygen molecules are found in the mouths of both Farsi and Danish speakers, and I have written the rules to generate a transformational phonology for either language, based on thermodynamic properties of those selfsame oxygen molecules. With my equations, I can prove that the second derivative of oxygen's vibration is equivalent to the 'deep structure' of the Farsi phonemes, likewise the Danish phonemes, and so on for the world." Who wouldn't laugh?

in English-language materials for teaching physics and chemistry, the term 'Aufbau Principle' (which explains the filling up of the various atoms' electron shells) has an oddly chequered career that I will try to summarize here. The Aufbau Principle is one of those topics that sounds like "physics" but is actually more of a chemistry topic nowadays (partly because it is so important to chemistry, partly because physics textbooks exhibit a pro-technology bias, which means apathy toward science in general and toward particle physics in particular, such that there simply isn't room for that "old and unimportant topic," so to say). Accordingly, our focus will be on chemistry textbooks, not physics as expected.

If you look in the index of a chemistry textbook published before 2000, you have a very good chance of finding 'Aufbau Principle' (e.g., in Zumdahl p. A78 or Ebbing p. I-4); if you look in one published after 2000, you're much less likely to find it.[28] Fashions change, I guess. But even when the term is (was) used in U.S. chemistry pedagogy, it is framed it improperly: It comes floating in vaguely like a mist around the heads of Pauli (as author of the Exclusion Principle) and Hund (famous for his Rules), wholly disembodied from the names 'Bohr' and 'Sommerfeld' where it originated — as though the notion of 'Aufbau' were just a commodity like air, for all the public to enjoy breathing.[29] From the following brief quote, we can see where the term originated, not in the air we breathe, so to say, but in a very particular and exciting moment in the history of particle physics:

> ... in order to justify the Bohr-Sommerfeld Aufbau (building-up) principle of the periodic system of the elements, [Pauli] came up with his famous *exclusion principle*.
>
> Marian (2001) p. 108

In other words, not only was the Aufbau Principle attached specifically to the names Niels Bohr and Arnold Sommerfeld, before being imported to the U.S. as a faceless, vapor-y thing, but better than that, it was the context for and *motivation behind* the famous Pauli principle.

The purpose of this appendix on Aufbau was to round out the discussion of 'breaking down' and 'building up' (on page **3** of the **Prologue**), which we've accomplished at this point. But since I raised the subject of electron shell-filling behavior, I will now take it a step further for the reader who might be curious about some of its details. Understand that this part is strictly optional, though, as it will have no direct bearing on the passage referred to above.

---

28. It occurs in neither the index to Kotz *et al.* nor the index to Moore *et al.*; the term is mentioned in passing, though, in Kotz *et al.*, p. 293.
29. Examples of this approach can be seen in Zumdahl p. 291–292 and Ebbing p. 296–299.

To present the idea of Aufbau properly requires several dense pages,[30] but if we are willing to anthropomorphize a bit (as one of my chemistry teachers delighted in doing), the whole idea can be summed up swiftly in three bullets, the cartoon imagery[31] of which will require only minimal exegesis:

- *lazy* — per the Bohr-Sommerfeld Aufbau Principle
- *antisocial* (BUT with laziness trumping the antisocial tendency) — per Hund's Rules
- *exclusive* — per the Pauli Principle

Elaboration on the above three bullets:

Per the original formulation of the Bohr-Sommerfeld Aufbau Principle, by assuming the electron is "lazy" (desirous of occupying the orbital with lowest possible energy state, i.e., closest to the nucleus of the atom), one can use the quantum numbers to build up reasonable models for all the elements, step by step, for the whole periodic table.

A refinement: The electron is "antisocial"[32] in the sense that it will occupy a separate orbital of its own rather than buddy up and share an orbital (thus bringing the orbital to its maximum legal occupancy, which is two electrons). This is the first of Hund's Rules, the one that chemists care about. It has a twist that is especially fun to consider when engaging in the anthropomorphizing game: If finding a separate orbital to occupy means moving away from the nucleus to one that is higher-energy, then the electron's reaction is, "No thanks, I'll stay here relaxing in this lower-energy orbital, and I'll even share it grudgingly with what's-his-face over there opposite me as I whirl around" (not in a circular or elliptical orbit, more likely along the lines of a three-dimensional four-leaf clover-cloud, we hasten to add). But now the Pauli Principle kicks in:

A further refinement (motivated by the need to explain certain anomalies of the Bohr-Sommerfeld Aufbau Principle thus far): The electron is "exclusive" in the sense that it cannot possess the same set of four quantum numbers as the electron with which it shares an orbital. One of those four numbers must differ. (I find it difficult not to think of the scenario where someone frets, "It would be a disaster if I wore the same dress to the party as *her*!")

---

30. As in 'Der Aufbau der Elemente im periodischen System', a section in Sommerfeld's classic, *Atombau und Spektrallinien*; see Sommerfeld 1944 [1919], p. 168–177. (Bohr and Pauli are both mentioned, though not Hund.)

31. For a straightforward account of Hund's Rule, without my anthropomorphizing fluff, see Kotz *et al.*, p. 300 or Moore *et al.*, p. 295 or, for that matter, almost any college chemistry textbook.

32. Potential point of confusion: Pauli himself is the one who started using this metaphor, in connection with *his* own principal! (See Arabatzis [2006] p. 219.) But really it is far more suitably applied to Hund's Rules, I think.

Enter 'up spin' and 'down spin', which some authors warn us are not *really* spins, just an abstract tweedle-dee or tweedle-dum-ness of the electron.[33] In any event, 'up spin' and 'down spin' provide the necessary infrastructure through which the final rule, known as the Pauli Principle, may be implemented: If one electron is wearing a 'down spin' dress, the other shows off her 'up spin' dress. All is well.

References for Appendix A: Aufbau (Building Up)

Arabatzis, Theodore. *Representing Electrons: A Biographical Approach to Theoretical E ntities.* University of Chicago Press (2006) 295 pp.

Ebbing, D.D. *General Chemistry*, Houghton Mifflin Company, 4th edn., Boston (1993) 1085 pp.

Johnson, George. *Strange Beauty: Murray Gell-Mann and the Revolution in Twentieth-Century Physics.* Vintage (1999) 434 pp.

Kotz, J.C.; Treichel, P.M.; Harman, P.A. *Chemistry & Chemical Reactivity*, Brooks/Cole — Thomson Learning, 5th edn., Belmont, CA. (2003) 997 pp.

Marian, Christel M. 'Spin-Orbit Coupling in Molecules'. In: Kenny B. Lipkowitz and Donald B. Boyd, eds. *Reviews in Computational Chemistry*, Vol. 17. New York: Wiley-VCH (2001) 99–204.

Moore, J.W.; Stanitski, C.L.; Jurs, P.C. *Chemistry: The Molecular Science*, Brooks/Cole — Thomson Learning, 2nd edn., Belmont, CA (2005) 1087 pp.

Serway, R.A.; Jewett, J.W. *Physics for Scientists and Engineers*, Brooks/Cole — Thomson Learning, 6th edn., Belmont, CA (2004) 1283 pp.

Sommerfeld, Arnold. *Atombau und Spektrallinien* [Atomic Structure and Spectral Lines]. Friedr. Vieweg & Sohn, Braunschweig (1944 [1931] [1919]) Vol. 1, 734 pp.

Zumdahl, S.S. *Chemistry*, D.C. Heath and Company, 2nd edn., Lexington, MA (1989) 1091 pp.

---

33. "Do not be misled; the electron is *not* physically spinning. It has an intrinsic angular momentum *as if it were spinning,* but the notion of rotation for a point particle is meaningless." Serway and Jewett, p. 945 (their emphases).

APPENDIX B: DIALECTS

For correcting the lay notion of "the dialects of China," an English/Dutch analogy is often used: "The difference between Mandarin and Cantonese is comparable to the difference between E nglish and Dutch" — words to that effect, e.g., in Chao (1968b) p. 96. That helps somewhat, but it doesn't go far enough. To paint a better picture of how the so-called dialects work in China, we need an analogy at least as rich as the following:

Suppose that Idaho, Nevada, Arizona, and all points east constitute the English-speaking portion of the United States. This will be our analogue for the Mandarin-speaking area of China. You live in Nevada, say, and when you visit the West Coast, you find that residents of California speak mainly Dutch (and some English); the residents of Oregon speak Danish (and some English); and the residents of Washington state speak Swedish (and some English) — three separate languages arrayed along the coast. And yet, for all that, these three coastal states are all unquestionably "part of America." They are welded into a seamless whole by Latin, the official language for all written communication. Thus, Latin plays the same unifying role that the Hànzì play in China.

Here we have our first inkling of how immense and complicated the "dialect" question is in China. But there is more. The actual languages I have in mind as the analogues of Dutch, Danish, and Swedish (moving south-to-north along the West Coast of America), are
Yue 粵 (Cantonese), Min 閩 (Fukienese), and Wu 吳 (Shanghainese), moving south-to-north along the east coast of China. Now for anyone musical, the respective tonal systems of these languages have a special relationship. If we take the Mandarin lexicon as a pool of two-syllable arbitrary "themes," then the Yuè lexicon may be imagined as a repository of possible "variations" on those themes; and the Mín lexicon likewise; and so on. For example, if we take wèntí 问题 'question' as a Mandarin "theme" (high-falling + high-rising), then būn-tôe 问题 (mid-level + low-rising => low-level + low-rising, per sandhi rules) is the musical Min "variation on it"; or if fāyán 发炎 'become inflamed' is our "theme" (high-level + high-rising), then hoat-iām 发炎 (mid-clipped + mid-level) is the Min variation on it.[34]

Caveat: The potential for such bisyllabic "theme and variation pairs" is certainly there, between Mandarin and Southern Min, for instance, but the reality is substantially curbed by the tonal richness of the latter. Why would that be? Because the very reason that Mandarin has so many bisyllabic words in the first place is *its* tonal *poverty*; conversely, a language such as Southern Min is willing to *tolerate* a fair amount of bisyllabic words — under the influence of Mandarin as the

---

34. *Speak Taiwanese Hokkien*, Taipei Language Institute (1969) Vol 2, pp. 113, 123.

official pǔtōnghuà (on the mainland) or guóyǔ (in Taiwan) — but it doesn't really *need* them in the same way Mandarin does for resolving homophone collisions.

And even where compounds exist, we will not necessarily find the same compounds occurring in two sentences that are notionally "parallel" (hence the discussion about "different languages" versus "different dialects"). Consider the following snippet of Cantonese (which I've transcribed from Pimsleur *Conversational Cantonese Chinese*, Unit 5, using ad hoc spelling *sans* tonal diacritics):

> 我　识　听　　一点（不是）好　识　听
> Ngoh shik-theng yat-di.　M-hai　ho　shik-theng.
> I understand a little [Cantonese; but] I don't understand [it very] well.

The only clear candidate for the theme-and-variation game would be the Mandarin/ Cantonese pairs wǒ/ngoh 我 and yìdiǎn/yatdi 一点 . Meanwhile, most of the words in the example would seem to run *counter* to my notion of theme-and-variations: For 'understand', Mandarin would use the vastly different phrase tīng.dedǒng 听得懂 , not @shí-tīng 识听 . Nor would Mandarin place hǎo 好 before the verb (unless to express an entirely different idea, "good for [verbing]"). And for the negation of 'to be', Mandarin would employ standard characters búshì 不是 , whereas Cantonese must fall back on dialect characters for m-hai (a very exotic and beautiful-sounding variant, I think, harking back to Classical Chinese perhaps). Nevertheless, a kind of indirect theme-and-variations effect occurs *even here*, two steps removed, as if by a "ghost theme." I.e., while saying ho shik-theng one is vaguely aware of how @hǎo shí-tīng 好识听 *would* sound, as the "theme," *if* that particular sequence were not Mandarin gibberish in real life

Conversely, in Southern Min we have the word bat (meaning 'did x; ever did x'), which plays a syntactic role close enough to that of guò 过 or céng 曾 in Mandarin[35] to feel like a "variation" on that "theme" even though, clearly, it is cognate to neither of those two Mandarin words.

Such is the unifying power of the Hànzì, weaving these two *languages* together *as if* they were merely two *dialects*. That's the take-away, since I realize many readers won't care about the theme-and-variations idea, per se.

---

35. E.g., Lí bat lâi chia bô? (in *Speak Taiwanese Hokkien,* Vol. 1, p. 47) meaning 'Have you ever come here [before]?' This can be partially rendered in characters as 你（曾）來 遮 勿？(using 遮 for 這).

In summary, this is yet another of the gravitational forces that may draw one away from Hànyǔ into the orbit of another "dialect": the allure of a beautiful tonal system, in which someone with a musical ear could get lost for days, months, decades.

The dark side of dialects:

As an outsider, I tend to see the so-called dialects as vehicles for preserving the "true" tonal nature of Chinese,[36] which in turn supports a higher degree of monosyllabicity, which in turn makes these languages seem "more Chinese" than Mandarin, more likely to provide a window on the classical literature, the poetry in particular. (Also, to my ear, Wu, Yue and Southern Min are all simply prettier than Mandarin, but as the English chéngyǔ, "there is no accounting for taste," so we needn't go there.) Meanwhile, from the viewpoint of an insider, the dialects have very different connotations: To a young person in Taipei, the sound of Southern Min may well connote ruralism, superstition, racism and bigotry — something along the lines of a Deep South accent in the U.S.

The Southern Min speakers of Taiwan think of themselves as běnshěng, literally "of this province," by contrast with those who might intrude from the various "outside" provinces of the mainland, posing a threat of cultural dilution by intermarriage. (Yes, they are *relatively* early inhabitants of Taiwan, though by no means the earli-*est*, i.e., they are not indigenous to the island; they are the descendents of transplanted Fukienese.) And the irony is that *their* cherished "race" (if we may call it that for moment, just to make a point) *their* own cherished race is in some danger of dying out eventually, precisely because of their own blatant racism and cultural bigotry, which each successive generation of younger people from the city will find more and more absurd and repulsive. (I'm thinking of the foul-mouthed Min-spewing matriarch in the 2009 film *Artemesia*, for instance, a type whom many of us have encountered in real life.) On the other hand, there is that "force of nature" aspect of the great Chinese Language Family mentioned earlier, and this will likely keep Southern Min (and the other "real" Chinese languages) afloat until eventually those attitudes will have softened, and any self-induced dangers faded to the vanishing point.

---

36. By contrast with the degraded state of Hànyǔ. If translated back into terms of the historical scheme, with yīn and yáng flavors of píngshǎngqùrù 平上去入 , we would see that Hànyǔ has kept Tones 1, 2, 3, and 5 while dropping Tones 4, 6, 7, and 8. Another way of saying "Tones 7 and 8 are gone" is this: All of the -p -t -k endings, crucial both to ancient prosody and to the overall "music" and general Chinese "soul" of the vernacular, are gone.

APPENDIX C: SIMPLIFIED CHINESE AND THE SECOND LAW

Before I indulge in my half-serious half-kidding venting over simplification committees and entropy (the Second Law of thermodynamics), let me work through two objective examples that show "simplification" in action, and how it can run immediately into some surprisingly insidious and complex issues.

1. The happy story of cí 辭 and cí 詞

Under the heading of 'skewed differentiation', Chao mentions, in 1968, the case of cí 辭 and cí 詞 . Their definitions are as follows:

cí 辭 1. 'words, expression' 2. 'to resign, to refuse'

cí 詞 'word, phrase; a type of poetry; used for 辭 '

Thus, there is substantial overlap, and even a long-standing 'used for' relation (as documented in Fenn [1926] p. 588); *but*, the 'used for' substitution is exercised selectively, only when the meaning is 'words', etc., not when the meaning is 'to resign'. (This nuance is spelled out clearly in Mathews p. 1033, where 辭 in its first definition *only* is marked "Inter[changeable with] 詞 .") Therefore, the one form cannot be regarded as a candidate for replacing ("simplifying") the other. Thus far, I am paraphrasing and elaborating on Chao (1968a) p. 170–171.

So, some forty years on (in 2009), are these two characters still distinct, as Chao says they must be, or have simplification zealots allowed them to slime together, using the partial overlap in their definitions as a vague, bureaucratic rationale? No, they are still distinct:

cí 辭 is simplified to 辞 (one of its long-standing alternate forms, seen in Mathews, e.g.).

cí 詞 is simplified to 词 (an approximation of one of its long-standing xíngshū [cursive] forms).

For the moment anyway, the distinction holds, nor have any brand-new forms been invented to handle either or both of these two characters. Let's say this counts as a "happy ending." The next story takes us in quite a different direction.

2. Forensics on fǎng.fú 彷彿

Ironically, this story does a good job of showing both [a] the insidious dangers of simplification *and* [b] the impetus for simplification. (Against my nature, it almost brings *me* over to the side of the simplification zealots! Almost.)

Various bisyllabic morphemes are cited by Y.R. Chao. (See Chao [1948] p. 39 and Chao [1968a] pp. 139 and 190–191. See also page **174** below.) Almost all of them are nouns; one of the few exceptions is fǎng.fú 'apparently' (or 'seemingly'). Chao writes it this way:
彷佛 (1968a, p. 139). And this immediately raises a question: In choosing to write it that way, with the two-stroke radical on both characters, does he mean *that* as the "real thing" or only as an 'abbreviated character' (a term he introduces on p. xxv of the same book), a form for convenience, meant to evoke "this form, 彷彿 , as the real thing, of course" — where the radical in both characters has three-strokes?

To see why/how such a question arises, please refer to the diagram below where I summarize the history of this term along the lines of CONTEXT A and CONTEXT B:

Viewed in terms of dictionaries, CONTEXT A means "THEN — back in the first quarter of the twentieth century" and CONTEXT B means "NOW — circa 2009."
Viewed geographically, CONTEXT A means "Taiwan, Hong Kong, Singapore, etc., where the traditional forms are still used" and CONTEXT B means "Mainland China, where the simplified characters are official (but where countless traditional forms must be recognized, simply by 'osmosis', and are no doubt even used [!] from time to time for various and sundry reasons)."

In the diagram below, 'GYTD' stands for *Gwoyeu Tsyrdean*.

CONTEXT A (Taiwan, HK, Singapore, etc.)

彷 **[1]** fǎng: see 彷彿 (also used for 倣)
  彷彿 fǎng.fú ***

彷 **[2]** páng: see 彷徉 and 彷徨
  彷徉 pángyáng: like 徘徊 páihuái
  彷徨 pánghuáng: see 徘徊 páihuái
    (俳佪 : same as 徘徊)

倣 fǎng to imitate; resembling
  倣效 fǎngxiào in imitation of
  → Alt. 仿

CONTEXT B (Mainland)

彷徨 pánghuáng
to walk back and forth;
to hesitate

徘徊 páihuái
pace up and down;
hesitate, waver

仿 [倣] to imitate; resembling
仿佛 seem; as if; apparently
仿效 in imitation of

*** In Fenn (1926) and in Mathews (1931), the form using the three-stroke radical (彷彿 fǎng.fú) is treated as a ***primary entry with its own identity***, AND ... there is *no* alternate form (such as 仿佛) even mentioned. (*Although*, in Mathews there is an *indirect* linkage this way: " 彷 [is used for] 倣 ," and the latter in turn has an alternate form, 仿 . This path I've represented by dashed arrows.)

In stark contrast to the above, by the time we get to the GYTD (1947), fǎng is treated as a secondary reading *under* páng 彷 , AND the entry for 彷彿 says *only* "same as 仿佛 ," thus strongly implying that the latter is the "real thing," since that is where a definition is (finally) given: 'generally similar to'. See pp. 200, 221, 471, 472 in the GYTD.

Sources: The upper half of the diagram is based on Mathews entries #1805 and #1808 (except the 俳佪 line in parentheses I've appended per GYTD p. 200). The lower half of the diagram is based on the *Zhongda* dictionary, pp. 570, 567, 206, and the *HarperCollins* dictionary pp. 245 and 94.

Comment: The dashed arrows suggest how the identity of fǎng.fú has changed: It is now subordinated to 仿 fǎng. But, 仿 really means 倣 , the same as ever. Thus, the juxtaposition of these three is "fake," and it causes the identity of fǎng.fú as a bisyllabic morpheme to be muddied, even lost. (Meanwhile, pánghuáng and páihuái have not experienced this problem. Their status as bisyllabic morphemes remains unambiguous. Note in passing a semantically similar but genetically unrelated bisyllabic morpheme in English: 'pingpong'. That was surely one of the most comfortable of all the foreign loans entering Chinese during the 20th century.)

In the lower right corner of the diagram, I have given a near replica of the *Zhongda* entry, where a second version of the character is shown in brackets, this way: 仿 [俲]. This is something I like very much about the *Zhongda* dictionary, its characters in brackets. But one must not assume this notation always means "*new* simplified form on the left; *old* version on the right." In this instance, for example, it means, "old *alternate* version on the left; old *primary* version on the right." Neither version is new. And this is where the whole business about fǎng.fú starts to unravel, alas. Note that 仿 traditionally was *not* an alternate for 彷 , only for 俲 , which means 'imitate, copy, like'. (This clear-cut primary/alternate relation can be easily confirmed by looking at older dictionaries such as Fenn or Mathews.) Meanwhile, we know that fǎng.fú has never been written as 俲彿 , only as 彷彿 . But *now* we have the trio 仿 and 俲 and 彷 (佛) thrown into very close proximity in the dictionary, with literally only 5 mm of separation on the page, strongly implying (to the untutored eye) that we are looking at superficial variants of a single *mono*syllabic morpheme, fǎng. At this point, the most fragile and interesting of Chao's bisyllabic morpheme examples has flown right out the window (or more aptly, has been crushed by a boot). That's why I say "simplification" can be dangerous, rather like other kinds of do-gooder legislation, with unintended consequences.

  * * * * * * * * * * * * * * * * * * *

So, who actually benefits from a "simplification" campaign anyway? A committee of mainland bureaucrats may decree that certain xíngshū forms shall replace their fántǐzì counterparts, or that brand-new simplified forms replace certain fántǐzì forms, but for real people on the ground every such simplification initiative actually multiplies the overall complexity of the language. A real person wants to have access to both mainland literature *and* the writings of someone such as Pai Hsien-yung (Bái Xiānyǒng) 白先勇 , whose nuances would risk being lost in jiǎntǐzì. Waking up to the reality of it, one might be reminded of Entropy and the Second Law: Nature tells us that entropy only ever increases. Similarly, the Big Chinese Character Count in the Sky can only ever increase over time. Any supposed decrease in the character count is just a politician's ego trip, a bureaucrat's pipe dream. Meanwhile, real people know that they must learn both jiǎntǐzì and fántǐzì, thus doubling the language effort, as it were, since the relation between the two is often not as straightforward as between a fántǐzì form and its traditional "cursive equivalent" in xíngshū (as illustrated by the incredibly convoluted fǎng.fú story above).

Another kind of problem arises: In its ostensibly "simplified" form, the character for fēng has its mountain radical on the left side: 峰 . By contrast, in the traditional form of this

character, its mountain radical is piled atop the fēng phonetic — *except* that on computer screens, even at a site that uses traditional forms (based in Taiwan for instance), this character appears with its mountain radical on the left side, not on top "where it should be." Why? Two reasons: The left-radical form has been around forever as an alternate (it is *not* a recent invention), *and* on a computer screen it happens to be the more legible of the two. So the whole issue in this case is really not about simplification at all. That's a canard.[37]

So far we have been taking the ant's eye view of the terrain. For an example of how the problem plays out on the macro scale instead, consider my own situation: In this volume, my policy is: "I use jiǎntǐzì." Not because I like the jiǎntǐzì forms, only because they are *au courant*. But I find myself breaking my own rule, whenever I cite the title of a book that I know was published using fántǐzì (in which case it would feel rather uncomfortable, not to say immoral and/or unscholarly, swapping in jiǎntǐzì after the fact), or when I quote a line from Li Po, or when I cite the name of the Wu 吳 or Min 閩 dialect, so-called (because those are both such pretty characters, and I refuse to "defile" them). Meanwhile, most of the materials in Chinese that I might think of picking up from my coffee table to read for pleasure are likewise in fántǐzì. *But*, if I happen to be thinking about chéngyǔ, for instance, and I see a good reference published on the mainland in jiǎntǐzì, I'll probably order it on the spot, rather than wait until I can find a similar book published in one of the fántǐzì regions. Thus, like countless others of us "on the ground," I live with a foot on each side of the Taiwan Strait, not very comfortably. Again, every "simplification" campaign of the politician only ever augments the *effective* size of the language, as experienced by real people.

In a similar vein, it is amusing to note that Chao 1968a uses GR romanization (cutting edge, he hoped, although that pet spelling system of his was too cerebral to ever really catch on; my opinion), while Chao 1968b (in the very same publication year) uses the hoary Wade-Giles romanization, presumably not to ruffle the feathers of "the general reader" or his publisher perhaps. To me, this is all just grist for the big Chinese Entropy Mill in the sky.

Along the same lines, one observes that Wikipedia romanizes every name twice these days, in a fashion that might be called (in jest) WahduhGillus-Peenyeen-Lomazi. Here we see yet

---

37. As mentioned earlier, a pleasant feature of the *Zhongda* dictionary is that each of its jiǎntǐzì entries is accompanied by the corresponding fántǐzì form in square brackets. This is especially appreciated in cases that involve peculiarities along the lines of fēng as just described; also, in cases such as 雕 which is not an alternate or simplified form *per se* but which now, by fiat, *subsumes* several traditional characters, identified clearly on p. 168 of the dictionary as 彫 , 琱 and 鵰 .

another doubling, so to say, of the brain cells required to process the language. It's nobody's fault. It just has to be done (as I did just a moment ago, in referencing Pai Hsien-yung above, for instance, because he was known *first* in literary circles as 'Pai Hsien-yung' and might very well *still* spell his name that way even today for all we know).

It gets worse: We tend to think of traditional/simplified fántǐzì/jiǎntǐzì as a binary split, but really there is a *third* flavor to consider: 'abbreviated characters'. (See Chao, 1968a, p. xxv, and footnote **40** on page **115** below.) Also, there are those graphical variations in printed characters that are almost invisible to the native speaker but troubling to the first-year Chinese student; also, the phenomenon of 'skewed differentiation', covered already via cí 辭 and cí 詞 in the prologue to this appendix. Also, there is the *BoPoMoFo* spelling system — not just an "extra" but something quite basic to one's Chinese education, in my opinion; see **Appendix J: Finals in BoPoMoFo, à la Gwoyeu Tsyrdean, with PinYin Updates** below. All of these are tributaries, major and minor, of the great River Entropy that we are tracking here, if whimsically.

Another kind of example: In *NTC's New Japanese-English Character Dictionary* (NTC Publishing Group, 1993 [1990], 1992 pages), every character is given twice, using its Japanese form on the left and modern Chinese (jiǎntǐzì) form on the right. Since not all Japanese *kanji* are simplified, and since not all Chinese Hànzì are simplified, this juxtaposition often results in redundancy, with the *kanji* on the left facing an *identical* Hànzì on the right, both in the traditional form. (The one on the right side of the page is accompanied by its reading in Mandarin, incidentally, which makes this monumental work much more useful than its title ["Japanese-English ..."] would suggest; i.e., it is equally a Japanese-*Chinese* dictionary and a magnificent *Chinese*-English dictionary.[38]) Sometimes the juxtaposition has the effect of showing a fántǐzì/jiǎntǐzì pair side by side, although the intent is, as always, to show a *kanji*/Hànzì contrast, e.g., for Japanese *ben* 辦 beside Chinese bàn 办 (p. 731).

Sometimes the juxtaposition really *does* show a Japanese/Chinese contrast, e.g., for Japanese *jitsu*, simplified the Japanese way with 8 strokes, 実 beside Chinese shí, simplified the Chinese way using a very different looking set of 8 strokes: 实 (p. 1011). In short, while

---

38. It is also notable for its lookup method called System of Kanji Indexing by Patterns (SKIP). The SKIP method may remind one slightly of the Four-Corner method, but it is easier to learn — a scheme to which the overworked term 'intuitive' actually applies. Using a scale of 1 to 10, with 1 meaning medieval and painful to use, and with 10 meaning modern and efficient and fun to use, I would rank traditional radicals as 2, new radicals as 3, SKIP as 9.

the nominal character count per the back cover of this NTC dictionary is 4421, from another perspective the character count is actually 8842. And if we think about it, how else *could* one hope to show the three-way crazy-quilt of mappings between [a] traditional forms and [b] Japanese simplified forms (of the 1940s) and [c] Chinese simplified forms (of the 1950s)?

Beyond the kinds of aggravation suggested above, which can be enumerated objectively (though never exhaustively!), there is a subjective side, too, to all this — those constantly nagging questions of aesthetics. True, many times it is only a matter of replacing a kǎishū 楷書 form[39] by a (boxy approximation of a) long-standing xíngshū 行書 form, as when shéi 誰 is replaced by 谁 , in which case one has only limited grounds for complaint; the "new" character has been around for centuries, and is merely being formalized or canonized by the government. (Also, closely related to the xíngshū forms, with which they overlap substantially, there is the concept of súzì 俗字 forms; see Chao 1968a p. xxv, and footnote **40** below.)

But sometimes the substitution is so ugly and gratuitous as to make one wince, as when yè 葉 is supplanted by 叶 , even though the former character, one of the most beautiful in the language, can be easily written by any Chinese person, left-handed while sound asleep and standing on his head, as it were, so when was it ever in need of "simplification"?

Not to say I *blame* anyone, exactly. As I try to look at it from the viewpoint of a civil servant, I can see good reasons to *think* about simplification, at least. Consider the energy expended by a quarter of a billion school children writing the sixteen-stroke version of lóng 龍 'dragon', one hundred times, let's say, during the course of one school year. That's 400 billion strokes for the year. Whereas, by writing the five-stroke simplified version instead (龙),[40] they could save 2.75 billion strokes. Those 2.75 billion saved strokes could represent an opportunity to study something useful, such as chemistry, instead of the belabored character for dragon, or in physical terms those 2.75 billion saved strokes surely represent enough energy to light a small town for the year.

---

39. Historical note: According to Chiang Yee, the point of reference for xíngshū was actually lìshū 隸書, not kǎishū as one would tend to assume today (Chiang, p. 80).

40. Incidentally, this form, 龙 , appears on the list of abbreviated characters in Chao (1968a) pp. xxvi–xxvii, which is prefaced by the following interesting remark: "Most of the abbreviated characters used in the book are found in 刘復 , 李家瑞 , 宋元以来俗字谱 ([Liu Fu and Li Jiarui, *Song Yuan Yilai Suzi Pu*], Shanghai: Academia Sinica, 1930). Of the newest abbreviations, such as those listed in Ronald Hsia and Peter Penn, *Dictionary of Simplified Chinese*, Hong Kong, 1959, relatively few are used here ..."; *Ibid.*, p. xxv.

One sympathizes somewhat with the language simplification committees, and yet … Is there really any denying the eternal existence of *this* elephant on the table (dragon in the room): 龍 ?

Stepping back a bit from the specific cases, there is a general sensation of dumbing down that makes one uneasy about simplification campaigns. The whole business might remind one of the case of 'begs the question' in English. Circa 2000, 'begs the question' was still used this way:

"We would like to know about X, but that person keeps changing the subject to Y."
"Yes, I know. He begs the question, doesn't he?"

In other words, the phrase meant: to beg *off*; to *evade* the question at hand; to *divert* a questioner; to *refuse* to answer the question properly; to *fail* to take it seriously and to heart. By the turn of the century, however, the phrase had morphed somehow into a different expression entirely. Now everyone believed that 'begs the question' meant '*raises the issue of …*' One pictures a dog sitting up to beg for a bone, and the human, by analogy, sitting up to beg for an answer: woof-woof. This new understanding of 'begs the question' has a kind of pop-etymology flavor about it, as though someone on a TV game show were mulling the phrase over, never having heard it before but trying valiantly on the spot to guess its meaning, through the haze of a hangover. In a halfway sane, halfway intelligent world, that subversion of the expression's meaning should never have occurred, as it did practically "overnight" — i.e., somewhere during the period 1998–2002, let's say.

I know. "Language changes, so deal with it. Don't be a schoolmarm." Etc. But *some* language changes don't even register as such. This is a case in point. My reaction here is *not* to the language change *per se*. What disturbs me is the glimpse it provides into the zombie-like stupidity and apathy of the culture *behind* the language. Two different subjects.

APPENDIX D: DEFINING THE MORPHEMIC SPACE

This appendix contains a continuation of the 'morphemic space' discussion on page . My rebuttals to two hypothetical objections are straightforward but space-consuming, so I've moved them down here to avoid interrupting the flow of the exposition earlier.

[a] First possible objection: "Isn't this a circular process that takes us right back to a laundry list of conventional dictionary 'head entries', the very place where we began?"

Not exactly. In the step called "Toss out … nonsense" (on page), I reserve the right to do so in a specially controlled manner that tracks the conventional dictionary but does not slavishly follow it. After all, a dictionary such as the *Zhongda* contains 4500 characters, whereas we will wind up with only 1171 primary entries in our table. So, yes, there is a suggestion of circularity, but not in a way that is detrimental to the project.

[b] Or, approaching it from the other direction, one might object that "this tone-acknowledging structure that you claim to have built exists in every dictionary already."

But that would be only a half-truth. A conventional dictionary is really a 1-dimensional realm, with tonal information merely mapped *onto* its single, syllable-centric dimension, as demonstrated below by the syllables `huai` and `huan`:

The conventional dictionary, when modeled properly, is a long, thin, *1-dimensional* place (where two dimensions have been "folded" into one):

```
              .
              .
          huai2
          huai4
          huan1
          huan2
          huan3
          huan4
              .
    y ↓       .
```
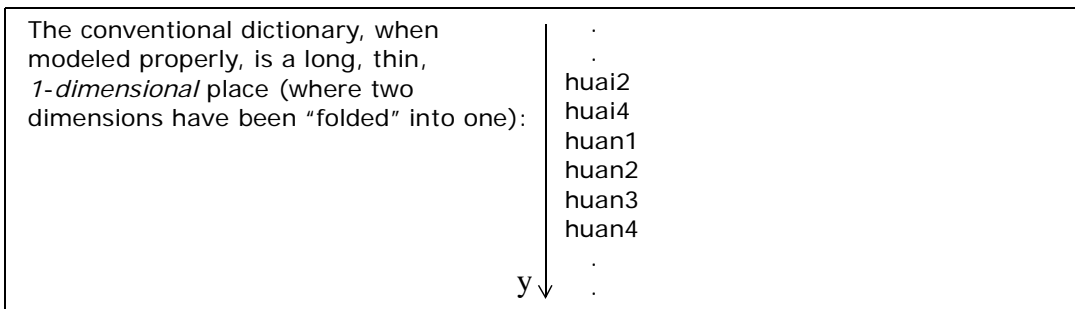
By contrast, we allow the four tones to occupy their own separate axis, as warranted by their importance to the overall sound system, and this results in a symmetric, 2D space that actually "means something." (Here we show 'huai' and 'huan' schematically, not as actually implemented in rows **120**–**121** on page **56**.)

	1	2	3	4
huai				
huan				

In our scheme the tonal identity of a given character is always reinforced by that of the neighboring characters above and below, along the y-axis, as it were. (Finally, we add a third dimension, as described on page **1** and elsewhere.)

Another observation about conventional dictionaries: Their designers endeavor to fill the space on each page as efficiently as possible. This means that the tonal signposts, such as they are, are *further* downplayed by the layout: The tonal skeleton of the language is effectively lost in the "noise" of the dictionary's competing design principles, according to which any blank space on the page is deemed undesirable, an instance of "wasted space." To the contrary, I would argue that in this context every blank space carries valuable information.

So far in this appendix, I've merely stated the obvious — some tedious but perhaps necessary arguments to forestall misunderstanding of the relation between my matrix and a garden-variety dictionary. Now for something more interesting, an actual subtlety (!) to note: Early on, our 'unit of currency' is not the morpheme *per se*; rather, we are identifying and tallying abstract slots in the sound system that we guarantee are *occupied by* one or more morphemes. Also, as detailed in the main body of this text, we are interested only in slots that contain at least one 'practical' morpheme, such as the slot shé which holds, among others, the morpheme that means 'snake'; we are not interested in a slot such as chěn that holds only one obscure morpheme, represented by the character 磣 meaning 'food with sand in it'. By fiat, I banish chěn and many others like it from the morphemic space.

APPENDIX E: IN SEARCH OF A 'TAMING MECHANISM' FOR ENGLISH

Having distilled the entire Chinese language down to a fairly manageable matrix (on pages **39**-**100**), there is a sense of having 'tamed' the language somewhat, as discussed, using slightly different language, on page **14** already.

Given that success with Chinese, it is natural to wonder: Could we do something similar for our own language, English? For instance, suppose we revisit our starting point, the 'morphemic space', which is certainly not unique to Chinese; that term of mine is derived from generic ideas in linguistics, equally applicable to English. Well, not *quite* equally. In fact, my inclination would be to run screaming from anything called 'the morphemic space of English' because morphology straddles semantics and phonology, and somewhere in that realm must lurk the infamous *ghoti* pronounced 'fish', and other such horrors that make English such a difficult language. That doesn't sound like a comfortable place to try building a matrix or 'configuration space' (to borrow a term from classical mechanics). Think also of the people who insist on writing 'tho' for 'though', and 'thru' for 'through'. The professional phonologist might demur, saying, "*That's* not phonology, that's only spelling. Not my problem." Well, whatever it is, it's a sign of a deep malaise in the language, when you compare it to, say, the spelling of Italian or Spanish or German, and it simultaneously says something vaguely unpleasant about the very small minority of people who stubbornly write 'tho' and 'thru' while all the rest of us do not — a particular flavor of denial or narcissism perhaps? At any rate, there is sufficient reason to flee phonology and morphology, and take our search elsewhere.

What if we could find in the lexicon a manageable core of *ancestral* English? That might be a safer facet of the language to focus on instead. One gathers that English is descended from a language called Anglo-Saxon. And from tribal knowledge, one already has a notion of what Anglo-Saxon is: a language characterized by short gruff words, such as 'moon', 'good', 'beard', 'wolf', and 'stone' in the mouths of our cave-dwelling ancestors (oh no, could this be another monosyllable trap?!), which, in time, were supplemented by infusions of multisyllabic

French words such as 'avenue', 'pantaloon', and 'elegance', also by Islamic borrowings such as 'algebra', 'algorithm' and 'alchemy'. Isn't that how it happened? (Actually, no! But it doesn't matter for our purposes here.)

As a native speaker of English, suppose I try to "exercise" my Anglo-Saxon linguistic roots, harking back to 5th century West German, by studying some modern German. As expected, I find *Mond*, *gut, Bart, Wolf*, and *Stein*, obviously cognate to my list of "cave-dweller" monosyllables cited above. So far the relation seems to be working as advertised. But soon that part of the vocabulary is swamped by another that feels every bit as exotic as Mongolian or Wolof. How so? The trouble with German (or the appeal of German, depending how one looks at it) is that a huge percentage of its everyday vocabulary has *no* cognate relation with *anything* in English, old or new. I'm thinking of words such as *Obst* (fruit), *Pilz* (mushroom), *Farbe* (color), and *Grenze* (border),[41] not to mention *Brunnen, Dachs, Gesellschaft, Geige, Gemüse, genau, Heft, Schlange, Schloß, ziemlich* and *zart*; nor should we forget the mystic tribe of *k*-words that includes *kaum, kein, kernig, klein, Klage, Kampf, Kraft, Kreis, Krieg, Kugel*, and *künftig*; nor the kind of word favored by cartoonists to signal "German" such as *Strafzettelschreiber, Vertrauensfrage, Meinungsverschiedenheit*, or *Schädlingsbekämpfungsmittel*. And should you have a legal claim to settle in Germany, you might want to know that the *Angeklagte* must go to *Gericht*, where a *Rechtsanwalt* (or *Anwalt*) will claim there was no *Meineid* but the *Richter* might yet hand down a stiff *Urteil*, based upon *Gegenfragen* and upon data from the *Schöffe*. That none of those legal terms has even the ghost of an inkling of a relationship with anything in English should tell us something. Thus, the putative 'Anglo-Saxon core of English', as reflected in German, is carved out in retrospect only with great difficulty, *even if* one focuses only on the lexicon and tries to ignore other German exoticisms such as its labyrinthine case system, and its quirky zigzag between SVO and SOV syntax. So much for one's dream of finding the core of English across the channel in Germany.

Meanwhile, waves of foreign borrowed words make the boundaries of English as vague and ever-changing as those of a tide pool. Yes, English is rich and vibrant (perhaps rivaling Japanese in its slutty eagerness to borrow?), but it is also an unsightly mess, with neither its borders *nor* its core well defined.

---

41. See for example the entry for *Grenze* 'border' in the *Duden* (Volume 7, 2006) p. 301, and note the absence of any Old English or Middle English cognates cited there. Similarly for the other words that I cite, if one has a working knowledge of German, s/he can confirm my assertions by perusing etymologies in the *Duden*, which may be described as a kind of German OED. Volume 7 is the one devoted to etymology.

A digression: Has China no such linguistic "mess" of her own to consider, at least in passing? No. China may have had her day in the sun as a (somewhat!) cosmopolitan nation, in touch with and at the center of all others, but in some regards the Silk Road was marked One-Way: So far as I'm aware, that moment of cosmopolitan glory, the Tang dynasty, resulted in close to zero impact upon the language. In fact, the only slight influence of another language on Chinese is seen rather in the turbulent period just *before* the Tang. By the start of the Tang, Buddhism was already flourishing, not to say rampant. And far from being emblematic of the "cosmopolitan" side of Tang culture, Buddhism was severely repressed[42] during the mid and late Tang. First, because it was something detestably foreign, hence de facto "inferior." Secondly, because it was a vehicle for tax evasion. The repression of Buddhism in that period was so severe that it never fully recovered.

In the novel *Honglou-meng* 紅樓夢 there is conspicuous Buddhist content, including multiple references to the jié 劫 (*kalpa*) as unit of time, also to jiéshù 劫數 (predestined calamity) and to yīn-guǒ 因果 (an indigenous term for 'cause and effect', whose context gives it a Buddhist connotation of *karma*). Never is it *pure* Buddhism, as a self-sufficient religion. Always it is framed in terms of a Buddhist/Taoist pair of monks: yīsēng yīdaò. Still, one must allow that *Honglou-meng* contains a serious *thread* of Buddhism, at the very least, even if it is not a "Buddhist novel" but ultimately a quintessentially Chinese novel. The question is: Outside of that important thread of Buddhist religion and philosophy in *Honglou-meng*, especially noticeable at the very beginning and end of the story, how does Buddhism fare in all the *rest* of Chinese culture? Rather poorly in my opinion. One thinks of stories where the appearance of a Buddhist monk or nun means only the opportunity for a subplot involving amorous buffoonery, nothing to do with religion per se, save as an object of ridicule.

---

42. Is it only coincidence that the fǎlún-gōng 法輪工 movement is likewise repressed today? While this movement has some aspects of a cult, it is also true that Dharma-wheel (fǎlún) is a mainstream Buddhist term from long ago.

Or, one thinks of the expression bào-fójiǎo 抱佛腳 , lit. 'to embrace the Buddha's feet', i.e., to seek desperate measures in the eleventh hour.[43] Hardly a dignified image. Meanwhile, Buddhism's overall impact on the Chinese *language* is quite limited, even though the culture may be said to contain "strong elements of Buddhism" if viewed a certain way.

At any rate, here is my attempt to list out the lasting linguistic influences of Buddhism on the Chinese language today:

Āmítófó阿彌陀佛	'the Amida Buddha' Not unlike 'Jesus Christ', this term seems more likely to occur as an exclamation than for devotional reasons.
zuòchán 坐禪	'sitting doing Zen' This topic seems far more intriguing to the foreigner (especially Japanese and Americans) than to the average Chinese.
sēng 僧	'Buddhist monk'
pú.sà 菩薩	'a bodhisattva', notably the female GuānYīn, seen often in paintings and statues.
luó.hàn 羅漢	'an arhat, a disciple of Buddha'
chà.nà 刹那	'an instant', from Sanskrit *kshana*, per Mathews p. 11; but nowadays it is a generic term for specifying the very instant when something occurs (shorter than a second).

Note that the final three terms can be added to our collection of bisyllabic morphemes (a phenomenon mentioned on page **174**; also in **Appendix C: Simplified Chinese and the Second Law**). This in turn reminds one that various *non*-Buddhist bisyllabic morphemes are also likely to be explained as foreign loan words (as noted in Chao [1948] p. 39, for example). Thus, one might say the case for the Tang being a cosmopolitan dynasty is enhanced slightly by the

---

43. Yang Bu-wei uses this expression (adjectivally: 抱佛腳的辦法) to convey the air of frantic catchup activity in U.S. universities after we declared war against Japan, and decided it was important to set up some intensive classes in Japanese and Chinese. Paraphrased from Yang (1972) p. 136. (Thus, indirectly, was born the U.S. sinology establishment overnight, having been represented heretofore by only a trickle of Ur-sinologists, too rare a breed to register on the radar as a proper academic discipline! In this connection, it is interesting to note that the *Dictionary of Spoken Chinese* published by Yale University Press in 1966 is identified on page v as "a revision of *War Department Technical Manual* TM 30-933 ... published ... on November 5, 1945.")

presence of these foreign loan words in the language, e.g., `luó.bo`. 'turnip or radish', `bō.lí` 'glass' and `méi.guì` 'rose'. But for me this does not change the overall picture substantially. The "cosmopolitan" tag feels somewhat forced or shaky to me. Or at any rate it's a one-way street, as though a rock star were to profess "an interest in literature" meaning one of the groupies in his hotel room last night happened to be a librarian: Yes, *she* (symbolizing the foreign visitor to Tang China) may enter the hotel room where he "holds court" if the time is right, but *he* won't be going to the library ever.

Note that the `fó` of `Āmítófó` is the sole occupant of row 85 in the matrix, *i.e.*, it is an instance of what I call 'super-RED'. About half the time, our super-RED classification coincides with a word that enjoys a very high frequency in the language (`děi`, `gěi`, `néng`, `nín`, `nǔ`, `rì`, `shéi`) or a moderately high frequency (`hēi`, `liǎ`, `nuǎn`, `tè`). When it comes to `fó` 佛 and `sēng` 僧 , yes, these words have a reasonably high frequency, but in their case the salient message of 'super-RED' seems to me to go in the opposite direction from bread-and-butter words such as `děi` and `gěi`, rather something to this effect: "This, believe it if you will, is a foreign *insertion* into our phonological grid!" As such, one might wish to covet the two morphemes, `fó` and `sēng`, as something more rare than hen's teeth.

What about the various sutras that were translated, in which there must be hundreds, if not thousands, of Indian names and technical religious terms, transliterated from Pali or Sanskrit into Chinese? Yes, everyone knows these *exist*. (I even remember studying excerpts from several such sutras myself, in my senior year as an East Asian Languages major at UCLA, in 1969. Looking back, the sutra seems a strange animal to present to the undergraduate. Could it be that I brought this on myself by special request, thinking, "Chinese Buddhism is important" because the art history books say it is? That part I can't recall.) They exist. But does the blizzard of specialized vocabulary in those translated sutras produce a living *presence* in the Chinese language itself? Certainly not. No more than the Latin mumbo-jumbo of the Catholic church can claim to have had any long-term *substantive* impact (outside of a few stand-up comedy routines) on the English language. In short, while one might set up plausible parallels between, say, the French component of English and the Persian component of Urdu and the Chinese component of Japanese, there is no such "back story" to tell for Chinese itself. All the way until 1921 when the letter 'Q' appeared in 阿 Q 正傳 , the title of a short story by Lu Hsun 魯迅 , the Chinese language maintained a very high standard for purity. So much so, that to my eye, at least, the 'Q' in that title still looks jarring on the page,

almost a hundred years on: For all those eons leading down to the letter 'Q', China had been the very paragon of 'ethnic cleanliness'.[44]

Now reconsider the Chinese case and contrast it with our fumblings above regarding English as a descendant of 5th century West German:

- Form a grid comprised of 396 rows and 4 columns.
- Devise a set of rules for handling three flavors: prime (RED), regular (BLACK), and overloaded (GREEN).
- Done. The realm has been defined.

I hope the stark contrast between the English case and Chinese case helps explain why I find the latter so compelling and appealing as an object of study.

---

44. Then the floodgates opened. Opening a novel at random at the public library, today one might find a reference to apartment 'C' and apartment 'D' (in a high-rise in Hell where two characters picture themselves living one day as neighbors), or pages peppered with emoticons (in Liang [2005], pp. 90, 198). Elsewhere, there are influences that make themselves known in subtler ways: In addition to the term Hànyǔ, which has a rather inward-looking flavor, similar to that of Guóyǔ, one starts to see the term Zhōnghuá yǔyán 中华语言 as well, which to me has a more self-conscious tone, one that resonates, in fact, with the phrase 'the Chinese language' *in* English, thus completing the circle. Another subtle case: The term fāngyán 方言 is routinely translated as 'dialect'. But the term itself is a very old one, and before the influence of western-trained linguists, it was probably understood to mean simply 'regional speech', i.e., "how they talk in that area," which is broad and ambiguous compared to saying "their local dialect." And yet, nowadays, 方言 really *does* mean 'dialect', having gradually changed its colors over the decades (as did cí 词 , in a different kind of morphing, likewise driven by the western-trained linguists).

APPENDIX F: NOT EVEN WRONG

Here we build a distant relative of the primary table, by skipping every real item and showing only the *non*existent sound/tone combinations instead, which number 413.

The number 413 is composed of two constituent groups:
• 290 hard-core cases of absolute nonsense;
• 123 cases that carry an asterisk. (These are "judgement-call nonsense," as explained on page **19**f.)

For anyone aspiring to write Chinese in the style of Edward Lear (who gave us "The little Fish swam / Over the syllabub sea" and "they ate with a runcible spoon")[45] or Lewis Carroll ("And the mome raths outgrabe"), this mirror-image table provides a one-stop Chinese Nonsense Emporium. On a more sobering note, this appendix also makes the following point, about four hundred times: When someone tortures the tones in Chinese, it's not *just* that they might land on the wrong word, they might instead be uttering literal nonsense — shades of Wolfgang Pauli's bon mot: "... a friend showed him the paper of a young physicist which he suspected was not of great value but on which he wanted Pauli's views. Pauli remarked sadly, 'That's not right. It's not even wrong.' "[46]

---

45. As it happens, 'runcible' is genuine Learian nonsense, but with 'syllabub' he has cheated us a bit: 'syllabub' only *looks* like nonsense because it is a rare word, of 'unknown origin, circa 1537'. The dictionary defines it as "a drink made by curdling milk or cream with an acid beverage (as wine or cider)," and with no reference to Lear. I.e., it is a "real" word, not his coinage.
46. R. Peierls. "Wolfgang Ernst Pauli, 1900-1958." (Royal Society, Great Britain) *Biographical memoirs of fellows of the Royal Society* 5: (1960) 174–192.

But don't imagine this is the whole story on Chinese nonsense. We've already seen an "even worse" kind of Chinese nonsense on page **16**, the fabrications @fiang and @sei, which come across as nonsensical in whatever tone: fiāng, fiáng, fiǎng, fiàng, sēi, séi, sěi, sèi. Any way you slice it, it's still nonsense — with one famous exception: Y.R. Chao harvested two such holes from the sound system to devise a unique name for his third daughter, Lensey.[47] Better yet, in the context of the tonal spelling system called *Gwoyeu Romatzyh*, of which he was the chief designer, her name even carries a strong hint of what its "correct"(!) tones would be, namely, tone 2 and tone 4. (Clarification: But in the table that follows, will we see lén and sèi? No! In this appendix the main topic is tone-*dependent* nonsense, to which the [fundamentally] *syllabic* nonsense of the Lensey story is only tangential.)

Like the primary table above (alias 'the matrix'), this one contains 396 rows. Some of the rows (e.g., rows number 1, 4, 5, 6, 9, etc.) are now blank all the way across. This eyesore I permit for the sake of preserving the relation between this fanciful, mirror-image grid and the "real" grid of the primary table above.

---

47. Note that the name Lensey must be rendered as a "romanized" item even inside a Chinese text. See for example Yang Bu-wei 楊步偉 (1972) p. 64, where her birth is recorded by the author, her mother. (Eventually, having married a Japanese classmate, this third daughter would be known as Lensey Namioka 波岡 ; *Ibid.*, p. 168.)

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
1	A			
2		án*	ǎn*	
3			ǎng	
4				
5	B			
6				
7		bán		
8		báng		
9				
10		beí		
11		bén		
12				
13				
14		bián		
15		biáo		biào
16				
17		bín	bǐn*	
18		bíng		
19				
20				
21	C			
		cá	cǎ	cà
22				
23				
24			cǎng*	càng
25				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
26	cē	cé	cě	
27			cěn	cèn
28	cēng*		cěng*	
29				
30			chǎi*	chài*
31				
32				
33				chào*
34		ché		
35			chěn*	
36				
37				
38				
39				
40				
41		chuái*		
42				
43				
44			chuǐ	chuì*
45				chùn
46		chuó	chuǒ	
47				
48			cǒng	còng
49	cōu	cóu	cǒu	
50		cú*	cǔ	
51			cuǎn	
52		cuí		
53			cǔn*	
54			cuǒ*	
55	D			

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
56		daí		
57		dán		
58		dáng		
59		dáo*		
60	dē*		dě	dè
61	dēi	déi		dèi
62		déng		
63				
64		dián		
65		diáo	diǎo*	
66			diě	diè
67		díng		
68		diú	diǔ	diù
69		dóng		
70		dóu		
71				
72		duán		
73		duí	duǐ*	
74		dún		
75				
76	E			
77		én	ěn	
78	ēr*			
79	F			
80				
81				
82				
83				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
84				
85	fō		fǒ	fò*
86	fōu*	fóu*		fòu*
87				
88	**G**			
		gá*	gǎ* See discussion on p.**19–20**	
89		gái		
90		gán		
91		gáng		
92		gáo		
93			gě*	
94	gēi	géi		gèi
95		gén*	gěn*	gèn*
96		géng		
97		góng		
98		góu		
99		gú*		
100		guá		
101		guái		
102		guán		
103		guáng		
104		guí		
105	gūn	gún		
106				
107	**H**			
				hà
108				
109				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
110			hǎng	
111				
112			hě	
113		héi	hěi*	hèi
114	hēn*			
115			hěng	
116				
117	hōu*			
118				
119			huǎ	
120	huāi		huǎi	
121				
122				
123				
124			hǔn*	
125	huō*			
126	# J			
127				
128		jián		
129		jiáng		
130				
131				
132		jín		
133		jíng		
134	jiōng*	jióng		jiòng
135		jiú		
136				
137		juán		
138			juě*	juè*
139		jún	jǔn*	

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
140	K			
		ká		kà
141		kái		kài*
142		kán		
143		káng*	kǎng*	
144	kāo*	káo*		
145				
146	kēn	kén		kèn*
147		kéng	kěng*	kèng
148		kóng		
149		kóu		
150		kú		
151		kuá		
152	kuāi*	kuái	kuǎi*	
153		kuán		kuàn
154			kuǎng	
155				
156		kún		
157	kuō	kuó	kuǒ	
158	L			
159	lāi		lǎi	
160	lān			
161	lāng*			
162				
163	lē*	lé	lě	
164		lén? No! See page **126**		
165	lēng*			

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
166	lī*			
167	liā	liá		lià
168	liān*			
169	liāng			
170				
171		lié*		
172				
173	līng*			
174				
175	lōng			
176				
177	lū*			
178	lǖ			
179	luān			
180	lüē	lüé	lüě	
181			lǚn	
182				
183	M			
184	māi			
185	mān*			
186	māng			màng
187				
188	mēi			
189			měn	
190				
191				
192	miān			
193				
194		mié	miě	
195	mīn			mìn

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
196	mīng			
197	miū	miú	miǔ	
198				
199	mōu			mòu*
200	mū			
201	N nā*			
202	nāi	nái		
203	nān*		nǎn*	
204	nāng*			nàng*
205				
206	nēi	néi		
207	nēn	nén	něn	
208	nēng		něng	nèng*
209	nī*			
210				
211	niāng		niǎng	
212	niāo	niáo		
213		nié*	niě	
214	nīn		nǐn	nìn
215	nīng		nǐng*	
216	niū*			
217	nōng		nǒng	
218	nū			
219	nū	nú		nù
220	nuān	nuán		nuàn
221	nuē	nué	nuě	
222	nuō		nuǒ*	
223	O ō*			

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
224		óu*		
225	P			
			pǎ	
226			pǎi*	
227			pǎn	
228			pǎng*	
229				
230			pěi	
231			pěn	
232				
233				
234			piǎn*	
235				
236		pié		piè*
237				
238			pǐng	pìng*
239				
240		póu*	pǒu*	pòu
241				
242	Q			
243		qiá*		
244				
245				
246				
247				
248				
249				
250	qiōng*		qiǒng*	qiòng
251			qiǔ*	qiù

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
252				
253				
254			quě	
255	qūn*		qǔn	qùn
256	**R**			
	rān			ràn
257				
258	rāo			
259	rē	ré		
260	rēn			
261			rěng*	rèng*
262	rī	rí	rǐ	
263	rōng			ròng
264	rōu		rǒu*	
265	rū			
266	ruān	ruán*		ruàn
267	ruī	ruí*		
268	rūn	rún*	rǔn	
269	ruō	ruó*	ruǒ	
270	**S**			
		sá		
271		sái	sǎi	
272		sán		
273		sáng		
274		sáo		
275	sē	sé	sě	
276		sén	sěn	sèn
277		séng	sěng	sèng
278				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
279		shái		
280		shán		
281		sháng		
282				
283				
284	shēi		shěi	shèi
285				
286				
287				
288		shóu*		
289				
290		shuá		
291		shuái		
292		shuán	shuǎn	
293		shuáng		shuàng
294	shuǐ	shuí*		
295	shūn	shún		
296		shuó	shuǒ	
297		sí		
298		sóng*		
299		sóu	sǒu*	
300			sǔ	
301		suán	suǎn*	
302				
303		sún		sùn*
304		suó*		suò*
305	T			
		tá		
306			tǎi	
307				
308				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
309				
310	tē*	té	tě	
311	tēng*		těng	tèng
312				
313				tiàn*
314				
315		tié		
316				tìng*
317				
318			tǒu*	
319				
320			tuǎn*	tuàn*
321				
322			tǔn*	
323				
324	W			
325		wái		
326				
327				
328				
329				
330		wéng	wěng*	
331		wó		
332				
333	X			
334			xiǎ	
335				
336				

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
337				
338				
339		xín*	xǐn*	
340				
341			xiǒng	xiòng*
342		xiú		
343				
344				
345				
346			xǔn	
347	Y			
348				
349				
350				
351				
352				
353				
354				
355		yóng*		
356				
357				
358				
359		yué	yuě	
360				
361	Z			
				zà
362		zái		
363				
364		záng	zǎng*	

	1st Tone Nonsense	2nd Tone Nonsense	3rd Tone Nonsense	4th Tone Nonsense
365				
366	zē		zě*	
367	zēi		zěi	zèi
368	zēn*	zén		
369		zéng	zěng	
370				
371				
372		zhán		
373		zháng		
374				
375				
376		zhén		
377		zhéng		
378				
379		zhóng		
380				
381				
382		zhuá		zhuà
383		zhuái	zhuǎi*	
384		zhuán		
385		zhuáng	zhuǎng*	
386		zhuí	zhuǐ	
387		zhún		zhùn*
388			zhuǒ	zhuò
389		zí		
390		zóng		
391	zōu*	zóu		
392				zù*
393		zuán		
394	zuī*	zuí		
395		zún	zǔn*	zùn*
396				

APPENDIX G: FREQUENCY OF OCCURRENCE

Disclaimer. In constructing the matrix, I had to deal with frequency-of-occurrence issues on several different fronts, and in all cases I used only educated guesses based on experience (modulated by other considerations, as detailed below), not a computer-generated ranking. What follows is yet another tour of the landscape, now with the focus on frequency of occurrence:

First, there was the question of which character to use for populating a given cell. Where I had several to choose from, I usually selected a character based on frequency. (Example: In populating the cell devoted to the syllable ān, I chose the character 安 'peaceful' in preference to the characters for 'nunnery' or 'saddle', and so forth, since one's overriding goal is to help a hypothetical beginner build toward his/her First 1000 Characters, let's say.) But other considerations sometimes defeat that general rule. For example, we might find ourselves at an intersection in the matrix where the sound *itself* happens to be represented by only one character, and that character in turn might be a fairly unusual specimen, for one of two distinct reasons: [1] it is unusual in the semantic sense (e.g., one might make the case that 'pistil' is simply an unusual word, regardless of which character might be used to represent it in Chinese; but as it happens, ruǐ is a special case, so the argument falls flat, as we'll see shortly); [2] unusual because it is only a secondary or tertiary way of expressing something, not the main way. Here I'm thinking of cases such as qué 瘸 'be lame' and ráng 瓤 'pulp of fruit or kernel of nut' where the *character* is of low frequency, and where the associated *word* is likewise obscure — i.e., not the main way of expressing the idea in question. This anticipates the next item:

Second, there were decisions to be made about marking certain cells with a *double border*, my way of flagging the 217 advanced vocabulary items that are enumerated on page **15**. Take the case of quǎn 犬 'dog'. On the one hand, we might wish to flag it as somewhat 'advanced', as a way of warning students "This is not the normal way to say 'dog' "; i.e., one would want to steer them, albeit indirectly, away from quǎn 犬 and toward gǒu 狗, the usual word for 'dog'. On the other hand, quǎn occurs in some garden-variety compounds such as 'police dog' and

'canine', so we can hardly call it rare or archaic. In the end, I decided not to apply the double border to quǎn, but I felt the decision could have gone either way.

Third, there were those cases where I left a cell empty *not* because the sound/tone combination in question is absolutely nonexistent, rather because its *representation* in the written language is so arcane or outlandish that I didn't want it cluttering up the matrix. These are the cases I flagged by entering an asterisk in a notionally "empty" cell, a decision process that I described in detail on page **19**f.

Fourth, for a syllable I had deemed 'overloaded' (e.g., for the syllable dié as illustrated on page **28**), there was the question of how exactly to populate the lower portion of the cell with "4 or more reasonably high-frequency characters." Also, I would sometimes judge a character *not* on its likelihood of popping up in Chinese specifically but rather from the outside, judging its presumed importance simply from a *semantic* standpoint in the mind of someone trying to express himself/herself in *whatever* language, English or Chinese or Icelandic.

One more angle to consider: Suppose a character is historically of low frequency, in terms of the language itself, yet high frequency in current-event reporting, because it happens to be used in transliterating the name of a person whose fame is on the rise. How should one classify such a character?

Take the case of ruǐ 蕊 'pistil'. This may seem like the poster child for "obviously rare characters" or "characters I could never possibly need." But as it happens, ruǐ 蕊 is used not only for 'pistil' but also in transliteration. Consequently, it comes up in thousands of news items wherever Hillary is mentioned, i.e., that particular Xīlāruǐ whose surname is Clinton. So, even if you don't literally "need it" (because you're not interested in botany), the character 蕊 is still *essential* for "newspaper literacy," or its Internet equivalent. I think of this as the *de facto high frequency* phenomenon. Here is a similar case: the word qín 禽 'avian', as distinct from niǎo 鸟 'bird'. The character 禽 shows up regularly in news stories of late (circa 2004–2009) because of concerns over the avian flu virus (qínliúgǎnbìngdú 禽流感病毒), but would otherwise be a fairly unusual character, outside of the classics or a zoology textbook. By happenstance, qín has *become* a high-frequency character overnight. The moral of the story: Be circumspect in assuming a given character is rare or unusual.

Final note about frequency:

I've extracted all the RED items into **Appendix H: The Prime Syllables Only** so that you can see them assembled in one place. Note that I do *not* differentiate between the low- and

high-frequency items within the RED category. My sole criterion for marking an item RED is its unique "ownership of" or "claim to" a given *cell* in the morphemic space. Still, my estimate is that about half of the 263 RED items are first-year Chinese vocabulary items, which is good news for the student. Moreover, as mentioned in "**Notation Conventions — Detailed Presentation**," half of the *super*-RED group are *also* high-frequency words, namely: děi, gěi, hēi, liǎ, néng, nín, nǔ, nuǎn, rì, shéi and tè. More good news for the student.

APPENDIX H: THE PRIME SYLLABLES ONLY

Excerpted from the primary table above, the following table is a recap of all the 'prime' syllables. This table is followed by notes pertaining to borderline cases such as BĒN, DIŪ, etc.

ĀO	ǍO	BAÍ	BĀNG
凹	袄	白	帮
concave (opp.tū)	coat	white	to help
āobù	mián-ǎo/jiá-ǎo		
dent	padded/lined jacket		
BĚI	**BĒN**	**BÉNG**	**BĚNG**
北	奔	甭	绷
north	to speed	no need to...	scowl
	bēnchí	Contraction of búyùng,	běngliǎn
	to speed along	represented by a sort of	pull a long face
		joke character I'd say.	
BĪ	**BIĚ**	**CĀ**	**CĀI**
逼	瘪	擦	猜
to force,	dented, shriveled	wipe, rub	guess
to pressure someone	zhuàngbiě		
	to dent (a car, e.g.,)		
CÀI	**CǍN**	**CÁNG**	**CǍO**
菜	惨	藏	草
vegetable; dish (on a menu)	tragic, inhuman	hide	grass, grassy, (very) cursive, illegible
	cǎnwúréndào	cánglóngwòhǔ	cǎogǎo or cǎonǐ
	simply inhuman	hidden talent	draft of a document.
		(lit. hidden dragon, sleeping tiger)	(To avoid a profanity, be careful of the tones on the latter compound.)

CÀO Like bếng (甭), cào (肏) strikes me as a joke character, not quite real: enter + flesh = vulgar term for carnal knowledge. Note that its 'prime' status helps minimize collisions with regular words. But see cǎo.	**CĒN** 参 uneven = cēncī (of appearance or quality)	**CÉN** 岑 a small spired mountain	**CÈNG** 蹭 rub, smear, creep, scrounge
CHǍ 叉 to cross chǎ.zhe tuǐ with legs crossed	**CHĚ** 扯 pull, drag into	**CHÈNG** 秤 scales, balance	**CHǑNG** 宠 spoil, dote upon chǒngwù pet (n.)
CHŌU 抽 take out, inhale chōuyān to smoke (i.e., take smoke out of cigarette)	**CHÒU** 臭 smelly	**CHUǍI** 揣 surmise / speculate = chuǎimó	**CHUÀI** 踹 kick, tread
CHUÁNG 床 bed	**CHUǍNG** 闯 cause trouble = chuǎnghuò	**CHŪN** 春 spring (season)	**CHǓN** 蠢 stupid, clumsy
CHUŌ 戳 a seal or "chop" = chuō.zi	**CǏ** 此 this, here, now (formal)	**CÒU** 凑 gather, encounter còuciǎo lucky, by (good) coincidence	**CŪ** 粗 coarse, crude, sketchy cūzhīyīèr have only a sketchy understanding of smthg
CÚN 存 exist = cúnzài	**CÙN** 寸 unit of length approx. 3 centimeters	**DǍ** 打 hit, strike	**DÀ** 大 big

DĚI 得 must	**DĚNG** 等 wait	**DIŪ** 丢 lose	**DUĀN** 端 *ending; beginning; carry (in both hands, i.e., properly)*

DUǍN
短
short (not long)
	DUĪ
	堆
	to pile
	`duījī`
	`to pile up`
		EN
		恩
		kindness
		`ēnhuì`
		`favor`
			ÈN
			摁
			to press (e.g., a button or switch)

FÀNG
放
to put, to place
`fàngpì`
`[1] to fart; [2] "You're`
`talking nonsense!"`
	FĚN
	粉
	powder, vermicelli, pulverize, pink
		FÓ
		佛
		Buddha, Buddhism
			FǑU
			否
			deny
			`fǒujué`
			`veto`

GÀ 尬 awkward = gān-gà	**GǍI** 改 change, alter	**GĚI** 给 give	**GÈNG** 更 even more…

GUĀI
乖
well-behaved
	GUǍI
	拐
	turn (into an alley, e.g.)
		GUÀI
		怪
		strange, to blame
			GÙN
			棍
			stick
			`gùnbàng`
			`stick as weapon (club)`

HǍI 海 ocean	**HǍO** 好 good	**HĒI** 黑 black	**HÈN** 恨 to hate

HǑU
吼
to roar, to howl
`hǒujiào`
`to howl`
	HUÀI
	坏
	bad
		JIÁO
		嚼
		chew
			JUǍN
			卷
			roll up; a roll

KÀN	KŎU	KŬ	KUĀ
看	口	苦	夸
look at, read, see; it depends on…	mouth, opening	bitter, difficult chīkǔ endure hardship	exaggerate kuākǒu boast
KUĂ	**KUÁNG**	**KŬN**	**KÙN**
垮	狂	捆	困
collapse	crazy, wild, -phile huákuáng sinophile	tie…up; a bundle	difficult, trap, predicament kùn.nan difficulties, troubles
LÁI	**LÀNG**	**LĀO**	**LĒI**
来	浪	捞	勒
come, bring	wave; wasteful	take, wangle	tie..tightly
LIǍ	**LIǍNG**	**LIĀO**	**LIǍO**
俩	两	撩	了
the two of us; some	two	raise (curtain or skirt); put up (one's hair); sprinkle	finish
LIĒ	**LŌU**	**Lǘ**	**LUǍN**
咧	搂	驴	卵
gossip = liē.lie	to rake, extort	donkey	ovum = luǎn.zi
LUÀN	**LŪN**	**LÙN**	**LUǑ**
乱	抡	论	裸
disorderly, confused hùnluàn.de chaotic	to brandish (a knife or one's fist, etc.)	essay, theory; to discuss	expose luǒtǐ naked
MÁ	**MÀ**	**MÁI**	**MǍI**
麻	骂	埋	买
hemp, numb, marijuana dàmá marijuana	to tell…off aímà to be told off (lit. to suffer a scolding)	bury	buy

MǍN 满 full, satisfied	**MĀO** 猫 cat	**MǍO** 铆 to rivet `mǎodīng` rivet (n.)	**MĒN** 闷 stuffy; shut oneself in
MĒNG 蒙 deceive = `mēngpiàn`	**MIÀN** 面 face, aspect	**MIĀO** 喵 meow	**MIĒ** 乜 to squint = `miē.xié`

MÍN
民
the people =
`rénmín or`
`mínzhòng`
	MǏNG
	酩
	get very drunk =
	`mǐngdǐng`
		MÌNG
		命
		life, fate; order (n.)
		`jiùmìng.a!`
		Help!
			MIÙ
			谬
			false
			`miùlùn`
			fallacy

MǑ 摸 to stroke, fish out (of pocket)	**MǑ** 抹 to apply, wipe, erase	**MǑU** 某 a certain (person); some (way)	**MÚ** 模 appearance= `múyàng`
NÁ 拿 hold, with	**NǍ** 哪 which?	**NÀN** 难 trouble `zāinàn` disaster	**NÁNG** 囊 used in `bēináng` backpack, book bag `dǎn-náng` gall bladder
NǍNG 攮 stab `nǎng.zi` dagger	**NÁO** 孬 cowardly, bad	**NÀO** 闹 noisy `rè.nao` lively	**NĚI** 馁 disheartened, famished
NÈI 内 inside	**NÈN** 嫩 delicate, tender	**NÉNG** 能 can	**NIÀN** 念 read, study, idea

NIÁNG 娘 mom, auntie	NIÀNG 酿 brew = niàngzào	NIĂO 尿 urine	NIĒ 捏 to hold, make, mold
NÍN 您 you (respectful form)	NIÚ 牛 cow, ox, beef	NIÙ 拗 obstinate	NÒNG 弄 make, play with
NÚ 奴 slave (n.) = núlì 奴隶	NÙ 怒 furious = nǎonù	NǓ 女 woman = fùnǚ	NUĂN 暖 warm
NUÓ 挪 move, shift to one side	Ó 哦 oh…also…	Ǒ 嚄 oh…actually…	Ò 哦 oh…now…
PĀI 拍 *to beat, shoot (a film)* qīngpāi to pat	PÀI 派 faction, school of thought	PÀNG 胖 fat	PǍO 跑 run
PĒN 喷 gush = pēnchū	PÉN 盆 basin, pot pénzāi bonsai	PÈN 喷 delicious = pènxiāng	PĚNG 捧 hold in both hands; handful (cf. gǒng)
PÈNG 碰 hit	PIĚ 撇 leave aside, fling piěkāichéngjiàn leave aside one's prejudices	PǏN 品 *article, product, quality,* *to taste, to play the* *vertical flute (lit. or fig.)* wùpǐn article of merchandise pǐnxíng character and conduct	PÌN 聘 *betroth, engage, employ* But a more common term for 'to employ' is: gùyòng

PĪNG 乒	**PÓ** 婆	**PǑ** 叵	**PŌU** 剖
ping-pong = `pīngpāngqiú`	old lady `pópomāmā` fussy	impossible, un- `xīnhuái pǒcè` harbor the unfathomable (= dark designs)	dissect = `jiěpōu`
QIĀ 掐	**QIǍ** 卡	**QIĒ** 切	**QIÉ** 茄
pinch, throttle, nip	*get stuck, wedge, block, fastener, hairpin, checkpoint*	cut	eggplant = `qié.zi`(See App. F)
QIĚ 且	**QǏN** 寝	**QÌN** 沁	**QUǍN** 犬
for the time being; even; and	sleep, bedroom, tomb (common in poetry, not in speech)	ooze, exude	dog used in `quǎnkē` canine
QUÉ 瘸	**RĀNG** 嚷	**RÀNG** 让	**RǍO** 扰
be lame	to shout = `rāngrāng`	make allowances, let, give way, transfer, by (PASSIVE construction)	disturb / interrupt = `dǎrǎo`
RÀO 绕	**RĚ** 惹	**RÈ** 热	**RĒNG** 扔
to wind, go around, make detour, confuse `ràoxíng dàolù` detour (n.)	provoke, cause `rěyǎn` showy (lit. provokes the eyes)	hot	throw
RÉNG 仍	**RÌ** 日	**RǑNG** 冗	**RÒU** 肉
remain, still `réngrán` still	day, sun; Japanese (adj.) `rìqī` date	superfluous `rǒngcháng` long-winded	meat, flesh

RUǏ 蕊 pistil Used in transliteration; see discussion on page **142**.	**SÀN** 散 disperse, scatter `sànbù` go for a stroll	**SǍNG** 嗓 throat, voice = `sǎng.zi`	**SÀNG** 丧 lose `sàngqì` be pessimistic sounding; unlucky
SĒN 森 dense, forest-like `sēnlín` forest	**SĒNG** 僧 Buddhist monk = `sēnglǚ`	**SHÁ** 啥 contraction of `shén.ma`	**SHǍ** 傻 stupid, foolish
SHāI 筛 to sieve `shāi.zi` sieve (n.)	**SHǍI** 色 or 骰 dice = `shǎi.zi`	**SHÀI** 晒 to sun-dry or tan `shài tàiyáng` to sunbathe	**SHǍO** 少 few, lack
SHĚ 舍 abandon, give (altruistically)	**SHÉI** 谁 who?	**SHÉNG** 绳 rope = `shéng.zi`	**SHĚNG** 省 save, omit; province `shěngshì` save (avoid) trouble
SHŌU 收 *put away, harvest, arrest, gain*	**SHUā** 刷 brush (n.), scrub `shuākǎ` swipe a (credit) card; pay by credit	**SHUǍ** 耍 play, juggle, mess…around `shuǎnùng.le` taken for a "ride"	**SHUà** 刷 pale = `shuàbái`
SHUǍI 甩 to swing `shuǎishǒubùguǎn` wash one's hands of…	**SHUàN** 涮 rinse, instant boil, be tricked (`bèi shuàn.le`)	**SHUǍNG** 爽 clear, frank, refreshing	**SHUǏ** 水 water `shuǐdiàn` utilities (water & electr.)

SHǓN 吮 suck	SHUŌ 说 say	SǏ 死 die; extremely `lèisǐ.le` totally exhausted	SŌNG 松 pine tree; relax, loose, "soft on" `sōngkāi` loosen (clothing)
SÒU 嗽 cough = `ké.sou`	SÚ 俗 custom, vulgar `rùxiāng suísú` When in Rome…	SUĀN 酸 sour `suānlàtāng` hot & sour soup	SUǏ 髓 marrow = `gǔsuǐ`
SŪN 孙 grandchild `sūn.zi` grandson	TĀI 胎 foetus, tire `lúntāi` tire (n.)	TǍO 讨 demand, discuss `tǎolùn` discuss	TÀO 套 *cover, trick into, copy,* MEASURE *for suits/* *stamps/ volumes* `ānquán-tào` condom

TÈ
特
special
`tèbié`
specially, peculiar
	TǏ
	体
	body =
	`shēntǐ`
	`tǐwēn`
	temperature
		TIĀO
		挑
		carry on pole, choose
		`tiāotì`
		nitpick
			TIǍO
			挑
			raise, prick, incite

TIÈ 帖 a book of model calligraphy; a rubbing from an inscription	TŌNG 通 connect with, workable	TŌU 偷 steal	TÒU 透 penetrate, thorough
TUĀN 湍 torrential, rapids	TUÁN 团 ball, group `tuánjié` unite	TUÍ 颓 decline, dejected	TUǏ 腿 leg

TŪN 吞 to swallow = tūnxià	**TÙN** 褪 take off clothes	**WÁ** 娃 baby, doll = wá.wa	**WǍ** 瓦 tile
WāI 歪 slanting	**WǍI** 崴 to sprain	**WÀI** 外 outside, foreign	**WāNG** 汪 expanse of water, watery, tearful
WÈN 问 ask	**WǑ** 我 I, our wǒguó China (lit. "our nation")	**XIÁO** 淆 confused	**XÚ** 徐 slowly, gently = xúxú
XUǍN 选 choose, vote, anthologize xuǎnjǔ elect	**YÉ** 爷 grandfather = yé.ye	**YUǍN** 远 far	**YŪN** 晕 dizzy = tóuyūn
ZǍ 咋 why?	**ZÁN** 咱 we (inclusive) = zán.men	**ZǍN** 攒 save	**ZÁO** 凿 to chisel
ZÈ 仄 slanting; the oblique tones in prosody (shǎng, qù, rù)	**ZÉI** 贼 thief	**ZĚN** 怎 how?	**ZÈN** 譖 slander = zènyán
ZHǍI 窄 narrow	**ZHÁO** 着 *touch, be affected by, be lit, fall asleep* zháojí worried	**ZHUā** 抓 grab	**ZHUǍ** 爪 claw = zhuǎ.zi

ZHUĀI	ZHUĀI	ZHUǍN	ZHǓN
拽	拽	转	准
throw, hurl	pull	turn zhuǎnyǎn before you know it	allow, standard, accurate zhǔnbèi prepare
ZǑNG	ZǑU	ZŪ	ZUĀN
总	走	租	钻
gather, total, always	walk, leave zǒusī smuggle, smuggled	to rent	to drill
ZUǍN	ZUÀN	ZUǏ	
纂	钻	嘴	Total Prime Syllables: 263
compile, edit	a drill	mouth	

Footnote to Appendix H: The Prime Syllables Only

Several of the entries in the above table are borderline cases (judgment calls) in the following sense: Another character exists that carries the same pronunciation, but it is so obscure or so new and technical (e.g., a character recently invented as the label for a chemical element or compound) that I discount it for purposes of building the 'prime' category. In the list that follows, each all-caps syllable refers back to a table entry, and the lowercase item that accompanies it is a word (or words) whose existence I have noted but discounted for purposes of constructing the table:

BĒN: bēn.zi 锛子 adz

CHŪN: chūnxiàng 椿象 stinkbug

DIŪ: diū 铥 thulium **(Tm)**

ĒN: ēn 蒽 anthracene **($C_{14}H_{10}$)**

HĒI: hēi 嘿 'hey! or wow! or oh!'

LÁI: lái 铼 rhenium **(Re)**, and láisè 莱塞 (alternate word for
    'laser', for which jīguāng 激光 is seen more often).

LIǍO: liǎo 钌 ruthenium **(Ru)**, and liǎo as second reading of liáo 燎 'to burn',
        for the nuance 'to singe'.

LUǑ: luǒlì 瘰疬 scrofula

MÀ: mà.zha 蚂蚱 a kind of locust

MǍN: mǎn 螨 acarid (mites and ticks as a class)

NÁ: ná 镎 neptunium **(Np)**
NÀO: nào 淖 'slush, mud, mire'
NIÀO: niào 脲 urea
Nǔ: nǔ 钕 neodymium **(Nd)**
PÀI: pài 哌 **piperazine** and pài 蒎 **pinane**
TÈ: tè 铽 terbium **(Tb)**
WǍ: wǎzú 佤族 the Va minority in Yunnan province
XUǍN: xuǎn as alternate pron. of xiǎn 癣 'tinea' (ringworm).

Comment: Note how the list above intersects about half the time with names of chemical elements or chemical compounds (bolded). From my viewpoint, this represents an unfortunate "contamination of the prime category that I needed to rationalize away"; from the viewpoint of those who coined those new terms, this would have represented "a space where we have some breathing room to create the new characters we need"!

Reality check: My tally of 263 primes is in close agreement with Chao's 261 syllables that "have no homophones" (Chao [1968a] p. 185). Given the oddball cases such as cào plus the judgement calls listed above, one can only be surprised that the disparity in our respective tallies isn't larger. (I.e., my whole scheme was independently developed, with a deliberate "outsider" slant and only limited knowledge of the literature, as detailed in **Appendix I: Technical Notes and Others**.)

APPENDIX I: TECHNICAL NOTES AND OTHERS

This appendix contains technical notes about the tools used to create the `PīnYīn` and `Hànzì`.

Authoring system: Adobe FrameMaker version 6.0, running on Windows XP.

For more control, in preference to "SaveAs PDF" I used a two-step distillation process: [1] from FrameMaker, write *.bk to *.ps; [2] from Adobe Distiller, convert *.ps to *.pdf.

For the PinYin, I used the James Dew *Easy Tones Font* called "4KeyCourier" — downloaded 01/06/09 as freeware from the Fool's Workshop. (Along with Dew's "4KeyCourier" font comes his "4-KeyTimesRoman" font as well, but I settled on the Courier version because it looks so much crisper when printed out.)

Font directory for Windows: C:\Windows\fonts\

Font directory for FrameMaker: C:\ProgramFiles\Adobe\FrameMaker6.0\.fminit\fonts\

Here is how the creator(s) of the fonts explain their use:

For First Tone, type '<' followed by letter 'a' or 'e', etc. For Second Tone, use '>' + the desired letter; for Third Tone use '[' + the letter; for Fourth Tone use ']' + the letter.

However, it appears to me that there is an even more convenient way to use these fonts (something I discovered by accident): First type the word without tone diacritics, then go back into the word and add diacritics as follows: To add a first tone diacritic over 'a', place the cursor before 'a' then press '<'. To add the second tone diacritic over 'a', place the cursor before 'a' then press '>'. And so forth. Either way, it all boils down to this wonderfully simple scheme: `< > [ ]`

Easy Tones Font anomalies: Sometimes, after a fast editing session, you might be left with a misplaced or misbehaving "ghost" of the above scheme that seems impossible to delete. In that case, highlight one or more letters in the vicinity of the "ghost" diacritic and redefine that snippet as Times Roman font. The ghost goes away. Now resume editing in whichever font is appropriate to the passage. Also, if you want to apply BOLD to one of these special PinYin

fonts, do it via Format > Characters > Designer > Weight > Bold, not via the "Bold" pick on FrameMaker's floating "f Catalog"; using the latter will strip out the tonal diacritics. Finally, there is a definite (small) price to pay for using Dew's package, at least in my particular environment: The mere presence of the 4KeyCourier and/or 4-KeyTimesRoman fonts on the system causes CTRL+SPACE (the key combination for creating a forced space) to be disabled for all of the *regular* fonts, oddly enough. (See workaround below.)

Note in passing that a *.ttf font file requires *true* extraction, e.g., via WinZip, before you copy/paste it somewhere, not the phony Compressed File "extraction" option, which gives you nothing.

To display Chinese characters, I used the so-called **SimSung** font. This font moves itself down to the very end of the Windows XP font list, having renamed itself as 宋体 (sòngtǐ), after you turn on Chinese (completely) as follows: Control Panel > Regional & Language Options **> Advanced**, select "Chinese (PRC)" on the pulldown. Repeat for Control Panel > Regional & Language Options **> Details**, and reboot. (In Windows 7, simply apply **SimSun** font, spelled without the final –g; in this environment, it does not later rename itself as 宋体 . However, in moving between machines, beware the hideous muddy font that Windows 7 will sometimes silently apply *in lieu* of SimSun until SimSun is explicitly [re-]applied.)

The Hànzì themselves I pasted in from **InputKing Online Chinese**. I used that tool for obtaining the *BoPoMoFo* symbols as well, available there under **Symbols > Phonetic**.
The special cases: To get the characters for gǎ 乬 , ǒ 嚯 , cào 肏 , and chěn 硶 (the latter in a footnote), I used http://www.newconcept.com/Reference/chinese_pinyin_input.html.
(Note: They list 嚯 only under its primary reading huò, not for its secondary reading ǒ.)

The rare cases of ü (u-UMLAUT) in Chinese can be troublesome:
1. How to handle 'ü' in romanizations, e.g., in the word for donkey: When using FrameMaker along with the 4Key fonts described earlier, it appears that ALT-0252 is disabled. However, there is a workaround: I find that I am able to copy ü *(*u-umlaut) into the document from other FrameMaker documents that were created before I installed the 4Key fonts. (The same workaround can be used for hard spaces: Import one, then copy/paste to proliferate it.)
2. How to obtain a Chinese character for a word with the 'ü' *sound* in it: deliberately misspell your **PinYin**, using 'v' in lieu of 'u'. (E.g., to bring up the character for donkey, don't type 'lu' in the on-line interface; rather type the nonsense syllable 'lv'. At least for Input King that's the trick.)

Details about the English language font and layout: I've used 12-point Garamond, with line-spacing set to *15 points*. Note that so-called single-spacing actually means 14 points of separation for a 12-point font. Having realized that, we will be less surprised to find that a *single* extra point (the difference between 15 and 14 points in this case) is all it takes to produce a *subjective* effect of 'double-spacing' even though literal double-spacing means 28 points (while '1.5 spacing' means 21 points, in the context of this Adobe FrameMaker authoring system, at least). I belabor this odd effect because of its implications for dealing with drafts, editors and the (self-)publication process: In my opinion, the trick of using single-spacing-points-plus-one (or plus-two) makes the dichotomy of 'single-spaced OR double-spaced' anachronistic and obsolete. In the finished product, 15-point line-spacing is decidedly more pleasant to look at than single line-spacing; moreover, during the markup phase, 15 points is already "open enough" to prevent claustrophobia. Thus, the best of both worlds, and no more ping-ponging back and forth between different modes of line-spacing, which in turn keeps the pagination and the cross-references in turmoil.

Footnote to the problematic word tiánhá 田蛤 (discussed in connection with **Figure 4** on page **170**): On a hunch, I did a Baidu.com search on < 閩語田蛤 > which led to a website called wenin.tw which contains this snippet (accessed on 06/29/09):

巴黎朋友問我台灣青蛙怎麼說？
我說了四種．青蛙，水雞，田蛤，四腳啊 . . . 把朋友逗得笑翻了．我說，還很多呢，要不要繼續聽蕃茄的各種說法？甘啊蜜 臭柿 堪魔朵 . . .（這些都是一個台南朋友教的）

This helped confirm my long-held belief that 田蛤 for 'edible frog' in Fenn (1926) page 145 is not a misprint for 田雞 , only an old or rare form: Apparently it had once been borrowed into Mandarin, but now exists only in Southern Min.

Sources/credentials/anachronisms

In certain passages in the main text, the reader may have been bothered by a sense of anachronism or dissonance. Perhaps one feels, "This writer seems often to adopt the stance of a naive outsider. And yet, at the same time there are signs that he may once have been housebroken as an academic long ago. Apparently he is not a total heathen?" All true, guilty as charged. A brief chronology will help fit the pieces together: I started listening[48] to the

---

48. The voices were "imprinted" on me. One of them belonged to Rulan Chao Pian, although this connection dawned on me only after she had been my dissertation advisor at Harvard for several years, during the period 1970–1975.

Mandarin Primer tapes at age 14 in Berkeley in 1957. Bachelor of Arts in Oriental Languages, UCLA, 1969. Lived in Taiwan 1973–1976. Harvard Ph.D. 1975 (unpublished dissertation: *Rhythm and Meter of Tsyr in Recitation*). Taught Chinese from 1976–1979. Published three articles in the early 1980s. Then followed nearly 30 years of complete isolation from sinology. However, I maintained some contact with Chinese characters via Japanese kanji, as I studied Japanese morphology and syntax on my own circa 2000–2001. Not until 2007 did I develop a renewed interest in Chinese *per se*, probably because one could sense that the giant was beginning to stretch and yawn and throw its weight around after a 200-year nap (or period of "humiliation"). This dictionary project I conceived while still in the "isolated" mode, however, in 2008; hence, the tone and perspective of an "outsider." Belatedly, in 2009, I began the long process of trying to tie *some* of its ideas back into the existing body of scholarship, e.g., by revisiting Y.R. Chao's *A Grammar of Spoken Chinese*. (I had excitedly read that large volume cover-to-cover the year after it was published. Some of the content stuck, but much of it had drifted into rather remote corners of my mind during the ensuing 40-year interval between 1969 and 2009.) To some extent, I still value my "outsider" approach, and for that reason I have deliberately not attempted to place *everything* into the context of the existing scholarship, only enough to suggest that I am not a complete "heathen" when it comes to academic protocol and niceties. I believe there is much to be said both for painting the ceiling inside an existing castle — as an "insider" — and for starting a new wing of the castle, so to say — as an "outsider." But the final result is that there are indeed some rough edges, a sense of anachronism here and there.

File created January 5, 2009, St. Paul, Minnesota; update date: **November 2, 2010 11:48 pm**

APPENDIX J: FINALS IN BOPOMOFO, à *LA* GWOYEU TSYRDEAN, WITH PINYIN UPDATES

By far the best presentation of the Hànyǔ finals is the one that occurs incidentally to the exposition of *Gwoyeu Romatzyh*,[49] in one of the prefaces to the *Gwoyeu Tsyrdean* 國語詞典 (i.e., the several pages of the 國語羅馬字聲調拼法表 that begins on page 5 of the 13-page preface). I consider that presentation of the finals so worthwhile that I've prepared a special "updated" version of it below. In this version, I keep the original *BoPoMoFo* (see page **166** below), but use PīnYīn (on page **163**) in lieu of the original GR spelling. My emphasis is on the Table of Finals, but for context I precede it with the Table of Initials, likewise in *BoPoMoFo*.

Indirectly, this appendix is also a response to the table of finals that appears in Chao (1948) p. 22 ("Memorize!") and again in Chao (1968a) p. 24. In that table, he summarily drops -o (and hence -io) out of the picture, by a peculiar logic that is not explained at all in the earlier work, and still only hinted at in the later work, thus:

"*e* and *o* can count as one"　　— Chao (1968a) p. 23

Pray, what does that mean? Answer: Apparently, the subtext to be supplied by the hapless reader is this: "In the syllabary '*BoPoMoFo DeTeNeLe* ...' we observe that *-o* and *-e* are mutually exclusive. I.e., there is Bo but no Be; there is De but no Do. Therefore, thinking somewhat mathematically, we collapse these two into one, arbitrarily choosing the abstract symbol *e* to efficiently represent them both. Or, saying it another way, '*e* and *o* can count as one'."

Usually I find Y.R. Chao's wit and cunning nothing but delightful,[50] but here he is being "too clever by half" and creates a mess, with the result that he counts "37 finals" (a figure that does not hold up under scrutiny). There is exactly one *right* way to wrestle these finals to the ground, and it is the way done so nicely in the *Gwoyeu Tsyrdean* as cited above (possibly,

---

49. Alias 'GR' alias Tonal Spelling system.
50. One can literally open *A Grammar of Spoken Chinese* to any page at random, and while away a pleasant half-hour. Then to another page, and another ... It is a book that one could "live in" like a Bible, but ...

ironically, with the participation, or even sole authorship, of Y.R. Chao himself, for all we know, since he *was* one of its chief editors and proponents!) In any event, as we can see quite clearly in the ensuing table(s), where I take my lead from the *Gwoyeu Tsyrdean*, there are 39 finals, not 37. (Whether approached my way or his, the derivation process will sound convoluted. But eventually, can't one simply do a brute force count of the items in the table? Yes, *except* there is one double-counting pitfall, as noted below.)

First, for context, here is the Table of Initials, using PinYin in lieu of Tonal Spelling …

BO	PO	MO	FO
DE	TE	NE	LE
GE	KE	HE	
JI	QI	XI	
ZHI	CHI	SHI	RI
ZI	CI	SI	

…followed by the Table of Finals:

			1	2	3
			I (YI)	U (WU)	Ü (YU)
n/a	–i [****]		<== unproductive in this context, hence numbered 'n/a' on first pass		
1	–A		–IA (YA)	–UA (WA)	
2	–O		–IO (YO)	–UO (WO)	
3	–E				
4	–Xè		–IE (YE)		–UE (YUE)
5	–AI		–IAI (YAI)	–UAI (WAI)	
6	–EI			–UEI (WEI)	
7	–AO		–IAO (YAO)		
8	–OU		–IOU (YOU)		
9	–AN		–IAN (YAN)	–UAN (WAN)	–UAN (YUAN)
10	–EN		–IN (YIN)	(WEN)	–UN (YUN)
11	–ANG		–IANG (YANG)	–UANG (WANG)	
12	–ENG		–ING (YING)	–ONG (WENG)	
13	–ONG		–IONG (YONG)		
n/a	–ER		<== unproductive in this context, hence numbered 'n/a' on first pass		

The cell marked **[****]** is home to a very important vowel sound, the one in words such as zhǐ 只 chī 吃 shì 是 rì 日 zì 字 cí 词 sī 思 .

But viewed from the standpoint of the phonological system, it occupies an anomalous little corner of the grid (if we dare say such a thing). As such, it presents a challenge to whatever system is being built or used to represent it symbolically.

As it happens, PⅰnYⅰn handles it rather poorly, allowing this sound and the vastly different vowel sound in lì 力 to be represented by the same letter, 'ⅰ'.[51] (But I've tried to mitigate that problem on the fly by using lower/upper case to distinguish the two in this incarnation of the table at least.) Note that even *BoPoMoFo* turns a bit strange as it tries to process this corner of the grid. It offers a special symbol (业 flipped on its head), by way of *acknowledging* the vowel in question...,

> In principle, this is the final vowel in zhǐ 只, chī 吃 etc., but in practice, this symbol is left implicit, never spelled out: 帀

...yet *BoPoMoFo* never actually *uses* that symbol for anything, outside of a purely *analytical* table such as this one. (To see it in context, please refer to page **166**.)

Rather, this particular final sound is left implicit in a certain family of initials; consequently, this is how you would spell the seven words listed above...

业 彳 ㄕ 日 ㄗ ㄘ ㄙ

...with initials only, so to say. No vowels allowed! (Really it is quite elegant, once you become accustomed to it, the only drawback being that one can easily forget the implied vowel even exists — as an entity in its own right, I mean.)

Back to the romanization question: Is there any system that *does* handle this vowel well?

Yes, GR (Tonal Spelling) reserves the letter 'y' for this sound, spelling our seven illustrative words as *jyy, chy, shyh, ryh, tzyh, tsyr, sy*, while 'i' is reserved for the sound in 力 , spelled *lih*, thus avoiding any possible confusion for the student. (In passing we should note that while the *Gwoyeu Tsyrdean* is a Chinese-Chinese dictionary, it does contain romanized headings, side by side with the *BoPoMoFo* headings, and GR is the system used for those headings. But here we are not interested in GR; rather, we're trying to "modernize" the table by redoing it in PinYin.) Onward ...

A brute force Cartesian product of the 3 column headings (I, U, Ü) with the 13 row headings (–A, –O, –E ... through –ONG) that look plausibly "productive" would give us 3 x 13 = 39

---

51. Psychologically, this would appear to be the exact opposite of the circumstance regarding Chao's claim that "*e* and *o* can count as one." Whereas Chao, gratuitously, collapses two perfectly good symbols into one, when the authors of the PinYin system allow 'ⅰ' to represent two sounds, I would assume it is because they find themselves in a dilemma, boxed in by the dual constraints of the English alphabet and the peculiarity of the phonological system itself that I am describing here.

possible "compound finals" (let's call them); but as you can see, only 22 such "compound finals" appear in the table (i.e., the inner rectangle in black font that contains finals –IA, –UA, –IO, –UO, etc.) That's because many potential combinations, such as U married to –AO in row 7 for a 'WOW' sound, don't exist in this particular sound system. Next, we take the superset of the column headings, plus ALL of the row headings, plus the populated cells, to arrive at an unadjusted total:

$3 + 13 + 2 + 22 = 40$. Here is the required adjustment:[52] Subtract one to correct the double-counting of -ONG (in row 13) and WENG (in row 12). The total is then 39. From here, the narrative continues back on page .

Not that it has any earthshaking significance, but as you read the whole "story," you may note a pattern that repeats on three tiers in fractal fashion, using that term in a loose, figurative sense: At three successive levels, we multiply rows by columns, and prune away 25% to 50% of the whole to arrive ultimately at the subset that is pertinent to the actual Hànyǔ language.

---

52. Or, when counting the finals by brute force, this is the pitfall mentioned earlier, resulting from an anomaly in the otherwise elegant *BoPoMoFo* system.

Now, for a more pleasant effect, here is how the *BoPoMoFo* array looks in its native form, as an alphabet in its own right:

ㄅ	ㄆ	ㄇ	ㄈ
ㄉ	ㄊ	ㄋ	ㄌ
ㄍ	ㄎ	ㄏ	
ㄐ	ㄑ	ㄒ	
ㄓ	ㄔ	ㄕ	ㄖ
ㄗ	ㄘ	ㄙ	

ㄧ	ㄨ	ㄩ

ㆢ			
ㄚ	ㄧㄚ	ㄨㄚ	
ㄛ	ㄧㄛ	ㄨㄛ	
ㄜ			
ㄝ	ㄧㄝ		ㄩㄝ
ㄞ	ㄧㄞ	ㄨㄞ	
ㄟ		ㄨㄟ	
ㄠ	ㄧㄠ		
ㄡ	ㄧㄡ		
ㄢ	ㄧㄢ	ㄨㄢ	ㄩㄢ
ㄣ	ㄧㄣ	ㄨㄣ	ㄩㄣ
ㄤ	ㄧㄤ	ㄨㄤ	
ㄥ	ㄧㄥ	ㄨㄥ	
ㄨㄥ	ㄩㄥ		
ㄦ			

Now that we've left the romanized forms behind, note how the clutter of parenthesized forms, such as "–IA (YA)," goes away. Except for the next to last row, where a stubborn irregularity seems entrenched, the table now makes sense. *This* is the Chinese sound system! This is the "code" you need for breaking into the *Gwoyeu Tsyrdean*, a kind of medium-sized OED for Chinese. *BoPoMoFo* is the perfect tool (almost) "for carving Hànyǔ at its joints," to paraphrase an expression that is popular in philosophy departments.

To conclude this appendix, I'll sketch out the actual sorting sequence of the *Gwoyeu Tsyrdean*, which is strongly suggested though not quite spelled out by the *BoPoMoFo* array itself. First, I like to flip the array of finals from its official form above to the following, which consists of 4 rows and 13 columns, after we exclude the oddball case of ㄦ (which is really more of a Peking dialect mannerism than a true final, in my view):

	ㄚ	ㄛ	ㄜ	ㄝ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ
ㄧ	ㄧㄚ	ㄧㄛ		ㄧㄝ	ㄧㄞ		ㄧㄠ	ㄧㄡ	ㄧㄢ	ㄧㄣ	ㄧㄤ	ㄧㄥ
ㄨ	ㄨㄚ	ㄨㄛ			ㄨㄞ	ㄨㄟ			ㄨㄢ	ㄨㄣ	ㄨㄤ	ㄨㄥ
ㄩ				ㄩㄝ					ㄩㄢ	ㄩㄣ		ㄩㄥ

Conceptually, we start by jumping to Volume IV of the *Gwoyeu Tsyrdean*, which contains all the entries that have no 'initial', only a 'final'. That volume is organized exactly as shown above, if you read the rows in turn, left-to-right, *inclusive* of all the headings. In other words, that volume is where you would find the following, to put it in terms of characters that you might be looking up, represented here in simplified form...

<p align="center">啊哦恶诶哀欸熬欧安</p>

...and finally 元云永 .

Now we turn to Volume I and form the cartesian product of the 21 initials, in ㄅ ㄆ ㄇ ㄈ order, with "all the above finals," in theory, although in practice only about half of the

possible combinations materialize in the language itself, as already noted in general terms. Specifically, the series goes like this, starting from the first page of Volume I of the *Gwoyeu Tsyrdean*:

B	ㄅㄚ	ㄅㄛ			ㄅㄞ	ㄅㄟ	ㄅㄠ		ㄅㄢ	ㄅㄣ	ㄅㄤ	ㄅㄥ
ㄅㄧ				ㄅㄧㄝ			ㄅㄧㄠ		ㄅㄧㄢ	ㄅㄧㄣ		ㄅㄧㄥ
ㄅㄨ												
P	ㄆㄚ	ㄆㄛ			ㄆㄞ	ㄆㄟ	ㄆㄠ	… and so on, cycling through the initials again.				
M												
F												

I've tried to arrange these Volume I syllables so that they are vertically aligned with the Volume IV syllables in the previous table, to show the strict parallelism. Again, the table is read row by row, left to right. Cast in terms of PinYin, we have the complete B-series as ba, bo, bai, bei, bao, ban, ben, bang, beng, bi, bie, biao, bian, bin, bing, bu, and the start of the P-series as pa, po, pai, pei, pao… Notice how many potential combinations are "missing" already. These are represented by the shaded cells. I.e., there is no such thing as ㄅㄜ, no such thing as ㄅㄝ, etc. Most of the ㄨ row is blank for ㄅ. The entire ㄩ row is blank for ㄅ (i.e., the 13 cells are shaded, all the way across). Thus, you can see already how the total should come down quickly from 800 in theory to something like 400 in reality (or, specifically, the 396 rows of our matrix).

APPENDIX K: FREE AND BOUND MORPHEMES

Figure 4 below provides a preview of the upcoming discussion by placing the free morpheme and bound morpheme of Notation Convention #2 in the context of a wider range of word and phrase types.

MW	得 děi	打 dǎ	田 tián	远 yuǎn	近 jìn 'near'

common relationship

BW	安—静 ānjìng	瑕—疵 xiácī 'blemish'	田—蛤 tiánhá 'frog'	近—来 jìnlái 'recently'

very rare relationship / see text

BM	葡～萄 pú.táo	玻～璃 bō.lí	蛤～蟆 há.má 'frog'

this simple relationship in history

is problematic for the student

TW	红—地—毯 hóngdìtǎn	大—白—鲨 dàbáishā	汉—学—家 hànxuéjiā

IF	一窍不通 yíqiàobùtōng	莫名其妙 mòmíngqímiào	有备而战 yǒubèiérzhàn

peccadilloes   not   hide   excellence

CY	瑕不掩瑜 xiábùyǎnyú	对牛弹琴 duìniútánqín	一画胜千言 yíhuà shèng qiānyán

Legend:
MW = Monosyllabic Word, i.e., a single-syllable free morpheme
BW = Bisyllabic Word from two monosyllabic morphemes (common)
BM = bisyllabic word from Bisyllabic Morpheme (unusual)
TW = Trisyllabic Word
IF = the 'Ipso Facto' type of 4-character expression, falling in the twilight zone *between* báihuà and wényán; see text
CY = chéngyǔ: a proverb/allusion consisting of 4 (or 5 or 6) characters of Classical Chinese
The ◯⬤ contrast is meant to be suggestive of the lighter texture of báihuà, relative to wényán, which is terse
and "dense" carrying a heavier semantic payload per unit of elapsed time. Meanwhile the IF type occupies a twilight zone
in-between the two: ◎

Note about fántǐzì versus jiǎntǐzì:
Some phrases that I show elsewhere using fántǐzì, to match their source, I present here using jiǎntǐzì instead. This is only
for the sake of honoring my *general* policy of using jiǎntǐzì in this volume. Beyond this, the choice has no significance.

FIGURE 4: Word and Phrase Taxonomy

Restated in terms of **Figure 4**, the intent of my Notation Convention #2 is to distinguish the MW type from BW (or BM). The typical Chinese curriculum focuses on vocabulary of type MW and BW; it also gives adequate coverage to BM and TW vocabulary. However, it gives poor coverage of CY, and almost no coverage of the IF type (except to the extent that IF phrases might occur in passing in the selected readings). Yet IF and CY occupy a huge part of the *spoken* language. For that reason, and because of the relation of the BW type to certain IF and CY expressions, we will devote considerable space to the latter even though strictly speaking they are tangential to the topic at hand.

Given the technical, quasi-scientific flavor of the terms 'free morpheme' and 'bound morpheme' one may be lulled into an assumption that "to the linguist, at least, this is cut-and-dried; to someone out there, it all makes sense, so I can relax." But look again at the passage seen earlier in footnote **6** on page **9**:

> "... when we say a morpheme is free we mean it is *sometimes* free, whereas if a morpheme is bound it is *always* bound"   – Chao (1968a) p. 144, emphasis added.

The notion of a 'free morpheme' he immediately qualifies; the term 'bound morpheme' *should* have been qualified as well (i.e., with a word *other* than 'always'), as I'll explain below. But first, two examples pertaining to the (qualified) 'free morpheme':

From the relation between jìn 近 and jìnlái 近来 , as shown in the upper right corner of **Figure 4**, we can see why Chao includes the qualifier '*sometimes*': jìn is a free morpheme, meaning 'near', but that is only the beginning of its story: Here we see it combining easily with lái to form a brand-new compound, jìnlái, meaning 'recently'.

Or, take the case of ài, which I've presented implicitly (via its notation in the matrix) as a stand-alone syllable (free morpheme). In the sentence Wǒ ài nǐ (I love you), monosyllabic ài works as advertised. I've told no lie. However, off in the wings, there is a closely related word that is multisyllabic: liànài 恋爱 'to be in love with'. Thus, it overlaps slightly with ài. So, in presenting ài as a stand-alone syllable, didn't I offer only a half-truth?[53] Some might complain that I did. But at least I did so "with my eyes open," and with good intentions — the philosophy being that the student needs to start *somewhere*, preferably with a classification that at least *sounds* straightforward (even though it isn't, really). Another kind of problem: chèn 趁 *is* a free morpheme (as implied by the format I assign to it in the matrix),

---

53. This same grey area is discussed in Packard (2000) p. 120f. (building on examples such as má.fán in Chao [1968a] p. 142); likewise in Li and Thompson (1981), *passim*.

but, in practice, it is almost always followed by the particle `.zhe`, which has the feel of adding English *-ing*. So, again a half-truth, and another potential embarrassment for me. From these few examples of 'free morpheme' behavior, one can extrapolate innumerable others.

* * * * * * * * * * * * * * * * * * * *

As for bound morphemes, the first lesson that the teacher might wish to convey is this: Just because `xiězì` 'write characters' works as a monosyllabic V-O pair, this doesn't mean @`miáo-jī` 描机can be slammed together by analogy to mean 'describe the machine'; rather, a pair of bisyllabic lexemes is required: `miáoshù jīqì` 描述机器 . Whereas, @`miáo-jī` would come across as gibberish, at worst; or, as an attempt to say `miàojì` 庙季 'the temple fair season', at best. (As a counterexample, given the right circumstance, `rì-jī` 日機 could indeed be understood to mean `rìběn-fēijī` 'Japanese airplane [attacking]'; see Yang [1972] p. 102-107 *passim*. "Context is everything.")

So, the lesson is learned. *But,* soon enough (egged on by examples such as `rì-jī` perhaps), the student starts to backslide into the earlier (and now "forbidden") frame of mind, musing to this effect:

In the written language (and verbally, too, for *some* contexts at least), wouldn't each of the following bisyllabic words be understood as soon as the first of the two syllables was encountered...

> `yì.sī` 意思 , `yǒngyuǎn` 永远 , `má.fán` 麻烦 , `xiácī` 瑕疵

...thus making the second character in some sense *redundant*?

As if in further confirmation of such musing, the student might start to notice examples of the following nature, too:

> `xiábùyǎnyú` 瑕不掩瑜
>
> His peccadilloes do not detract from [lit. cover, conceal] his overall excellence.

Here, the character 瑕 is clearly holding its *own* inside a terse four-character expression (a `chéngyǔ`成语 or `yànyǔ`谚语), seemingly "broken off" from the compound `xiácī` 瑕疵 'flaw or blemish', which the student heretofore *thought* was composed of two *bound* morphemes that were "always bound." What happened?

The explanation this time is very different from the matter of "high context" for `rì-jī`. Here we've slipped from `báihuà` (the vernacular) into another dimension, that of `wényán` (Classical Chinese), which aspires to the density and terseness of a neutron star, let's say. This

slipping back and forth between dimensions occurs all the time in Chinese, whenever one feels the need to insert a condensed parable, fable, adage, homily, saw, cliché …

From the student's viewpoint, the appearance of xiá functioning on its own will likely register as an "exception to the bisyllabic rule" or "confirmation that the two syllables in xiácī 瑕疵 really are redundant"; but historically a different story is told: At some point, say two hundred years ago, just to pick a time, the free morpheme xiá started to fail (because it was surrounded by too many homophones), so it joined forces with cī to become a bisyllabic word, and thus avoid the ambiguity. Skimming along the surface of the language, all one might perceive today is "a pair of bound morphemes," thus setting oneself up for the confusion noted above.

Thus we see that the notion "always bound" is a half-truth, valid only within a special context where one is giving a preliminary definition of 'free morpheme' versus 'bound morpheme', from the very narrow perspective of the present, and with thousands of years of history and hundreds of chéngyǔ all artificially excluded from the picture.

The double-headed arrow connecting two instances of xiá 瑕 in **Figure 4** illustrates the problem, and shows why the student is blind-sided by such cases. I also decompose the long double-headed arrow into a short upward arrow (representing the historic path) and short downward arrow (representing the student's misperception of the relationship). No rule has been broken. No redundancy has been confirmed. Rather, one has been forcibly shifted from one archeological layer of the language to another. And yet, this all happens *within* the confines of báihuà, i.e., it is not a matter of leaving báihuà for a moment, entering the world of Classical Chinese, then returning to báihuà. Rather, such transitions are an integral part *of* the vernacular *itself*. That's bad news for the foreign student, but it must be faced.[54]

The 'loganberry' and 'cranberry' footnotes to morphology

Before leaving the BW row of **Figure 4**, let's look at the 'redundancy' assumption from a slightly different angle, now using the compound ānjìng 安静 'peaceful' as an example. We assume this compound was created long ago to avoid ambiguity among homophones, but nowadays sometimes it may appear redundant in certain contexts, especially when written not spoken, making the student wonder, "Can't I just use ān 安 , especially if I'm writing, not

---

54. On the bright side, it is fun to think of the Hànzì or monosyllables behaving not unlike atoms that form molecules, then break apart and reconfigure themselves to form other molecules; see Chao [1968a], pp. 159, 163, re: "ionization."

speaking?" To understand why the answer is "No" (for báihuà, whether spoken or written), consider the case of 'loganberry' in English. The morphologies of ānjìng and 'loganberry' are admittedly quite different, but there is a similarity in their psychologies worth exploring for a moment: By cold logic, *logan* could operate as a perfectly good word that required no suffix such as *-berry* to clarify it: It is, after all, unique in the whole English language, beyond the shadow of a doubt. It can denote only one thing: a loganberry. (Whereas the words 'dewberry' and 'raspberry' don't work this way: They desperately *need* the suffix *-berry*.)

Nevertheless, there are "sociological" rules[55] that operate upon a language, and one of them says, in effect: "Thou shalt not break off 'logan' as a free morpheme just because logic supports that path." That rule might be motivated by conservatism. Another desideratum of language is to have fun, and the sociological rules are surely tuned into that aspect as well: It is simply more fun to say 'loganberry' than 'logan', more fun to say yì.si than yì as 'idea'. In such cases, one might say that "fun trumps logic" and helps keep the components of Chinese bisyllabic words from drifting back in the direction of their long-ago monosyllabic state. (On the face of it at least, this notion must seem at odds with much of **Appendix L: Muchengxue and the Monosyllabic Mystique**, I realize. But both forces *are* present in the language. How they play out in a given context is not so easy to predict.)

In passing, it is amusing to note that Y.R. Chao mentions the word 'cranberry' in a similar context (1968a, pp. 139, 167). But he is looking at these cranberry/loganberry words from the opposite direction, searching for an analogy on the English side where a free morpheme *could* in theory be broken off, while my point is that while the potential exists, our sociological rules *forbid* its consummation. Two parts of the same elephant.

* *

The next row in **Figure 4** to consider is the one labeled BM = Bisyllabic Morpheme. Compared to BW, this type is fairly uncommon. Some examples are...

<div align="center">

pú.táo 葡萄 , bō.lí 玻璃 , méi.guì 玫瑰 , níngméng 柠檬

</div>

...meaning 'grape', 'glass', 'rose', and 'lemon'.[56] (The discussion of há.má 蛤蟆 we defer for a moment.)

---

55. I've used the term 'sociological' as a reference to Chao (1968a), p. 137; different context, similar idea.

56. *Mandarin Primer* has luó.bo 'radish', jī.gu 'grumble', pú.sa 'bodhisat', luójì 'logic' (luójí nowadays); Chao (1948), p. 39. For other examples, see Chao (1968a), pp. 139 and 190–191. See also pánghuáng on page **111** above.

Here, in the written language at least, and in many verbal contexts, too, one may make the case that the first syllable *uniquely* identifies the item in question. Which is to say that the second syllable makes a *zero* contribution, semantically. Thus, we have a case that is "still worse" than the earlier one, where sometimes the two bound morphemes might appear *redundant*: Here, the second syllable (of a bisyllabic morpheme) looks as though it might actually be *superfluous*.

Accordingly, we wonder: Does half of *this* type sometimes get broken off as well (in the manner of 瑕 above)? Answer: Yes, but only very rarely, as most bisyllabic morphemes seem to enjoy an 'untouchable' status which trumps the logic of 'superfluous'. (Again, one might compare the psychology of 'loganberry', although its morphology is not analogous.)

A possible exception to the rule is illustrated by há.má 蛤蟆[57] in **Figure 4**. This bisyllabic morpheme, thought to be hermetically sealed, so to say, has apparently been "opened up" and has had a semantic chunk broken off for reuse elsewhere, in the term tiánhá 田蛤 . The latter I stumbled on in the Fenn dictionary (p. 145), where it is glossed as 'edible frog'.[58]

This kind of relationship is so rare that one is prompted to consider alternative interpretations: Perhaps the existence of the term tiánhá only means this: Once upon a time, há 蛤 was a "real" word, i.e., a Classical Chinese monosyllable, free to do as it pleased, e.g., combine with tián to form a brand-new compound. Which is to say our modern perception of há.má 蛤蟆 as a bisyllabic morpheme would need to be abandoned as erroneous (see footnote **57**). But I will stick to my hypothesis of "breaking off" one half of a bisyllabic morpheme, the *general* idea of which is supported by the morphological minuet

---

57. The word há.má is properly written with the character for shrimp (primary reading xiā, secondary reading há), like this 蝦蟆(using fántǐzì), but usually it is written this easier way instead: 蛤蟆. (This is a traditional substitution for convenience; nothing to do with modern "simplified" characters *per se*.) In the latter compound, note that the character 蛤 is again being used in a *secondary* reading, while its primary reading is gé, as in gé.li 蛤蜊 'clam'. In the *Gwoyeu Tsyrdean*, p. 1540, há.má 蝦蟆 is defined as a "dull brown marsh frog with black spots, good at jumping, and with a song like 'xiàxià'." Because of the way this word is listed in dictionaries, only as a two-syllable entity, it seems to be a very strong candidate for a bisyllabic morpheme, although I don't see it cited by others along with the usual suspects, pú.táo, ló.bo, bō.lí, pú.sà, wū.gū, etc.

58. I cite the Fenn dictionary of 1926, specifically, because the term tiánhá 田蛤 simply doesn't appear anywhere else. And even in Fenn, I had to wonder if those two characters might just be a misprint for tiánjī 田鸡 , literally "field chicken," the latter being a relatively *well*-known synonym for qīngwā 青蛙 'frog'. But then thanks to the Baidu search engine I was able to confirm tiánhá as a term that is alive and well in Southern Min. For details, please refer to **Appendix I: Technical Notes and Others**.

between yīngwǔ 'parrot', bā.gēr 'mynah', and yīng.gēr [again] 'parrot', as described in Chao (1968a) p. 167.

Before leaving the BM type, it is worthwhile taking a slight detour into the mercurial nature of lǎo 老 , which further complicates the picture as drawn above. As used in lǎo-Wáng 'good ole Wang' or in lǎo-sān 'third daughter or third son', lǎo- still feels like a *prefix* with varying degrees of vestigial meaning. But in an animal name such as lǎohǔ 'tiger', lǎushǔ 'mouse, rat', or lǎoyīng 'eagle, hawk', lǎo loses its identity, having melted over time into a single bisyllabic morpheme, such that it carries *zero* vestigial meaning. This is my rough paraphrase of Li and Thompson (1989), pp. 37 and 83 note 4. (Another case to note in passing is lǎoyuǎn 老远 'very far'; this occurs in Yang (1972) pp. 156, 170, e.g.)

While of note to the linguist as a technical curiosity, the bisyllabic morpheme is *"invisible" from the student's perspective*, where it registers only as part of a much larger problem: the surprise of stubbornly bisyllabic words [of all kinds] in a language s/he had approached under a vaguely monosyllabic assumption (which springs back to life via exceptions of the 瑕 type noted above!). This important distinction between the student's perspective and linguist's perspective is mentioned in a separate but related context — "disagreement over the definition of *compound*" — in Li and Thompson, p. 45.

* *

As mentioned earlier, the chéngyǔ (represented by CY in **Figure 4**) is one of the most neglected aspects of the "conversational Chinese" curriculum. And even when it is not neglected, it tends to be misunderstood. One sees suggestions in Chinese teaching materials to this general effect: "Why don't you spice up your conversation by learning a few chéngyǔ; these will surprise your Chinese listeners." Surprise them indeed! The whole point of chéngyǔ, as with similar phrases in English, is to revel in *le mot juste* — not just to sling phrases about at random. Whether in English or Chinese, some expressions of this type get used exactly *once* in an entire lifetime, remaining dormant all the rest of the time. The other thing to realize about chéngyǔ is that others will be using them at will, in fact compulsively (*if* the time is ripe for that *mot juste*), not just randomly once or twice to "spice up the conversation." And if you are caught out by the chéngyǔ being batted about by your hosts, then how comfortable are you going to feel about your supposed "working knowledge of Chinese"?

It gets worse. There are innumerable four-character expressions of a type I call '*ipso facto*'. In English, there are literally dozens of Latin phrases such as *ipso facto, quid pro quo, pro bono, de facto, per se* and *semper fi*[*delis*] that "everyone knows." These are not *quite* vernacular English (since in fact *not* everyone knows them, only "many people, above a certain level of education"); but at the same time, these terms hardly count as "classical allusions" that would impress a Latin professor! Thanks to a certain ballet scene between Elmer Fudd and Bugs Bunny, "everyone" (of a certain generation or two) knows something about the Bacchanal in *Tannhäuser*. Not that this makes them all Wagner experts; some of them have no idea that scene is a direct allusion to an opera called *Tannhäuser* by Wagner. But there it is: At some level, they now *know Wagner*. Even more interesting is the case of "The Masquerade Is Over," a Broadway song immortalized by jazz greats Nancy Wilson and Cannonball Adderly. At one point Nancy Wilson sings, "I guess I'll have to play Pagliacci." By context, we inkle the meaning even if we have zero knowledge of the allusion. And after hearing the song, one is now primed to actually learn, at some future date, all about the opera *Pagliacci* and its creator Leoncavallo, long ago and far away in Italy. Isn't that a remarkable kind of bootstrapping process? (In my case, for example, the two dates were 1963 for the "inkling" via jazz and 2001 for knowledge of the opera itself.)

Obviously, similar things go on in spoken Chinese in its interplay with the classics, but in a far greater quantity and variety given the deeper historical inventory to draw upon. Thus, a body of *four-character* expressions that "everyone knows." These have *some* degree of Classical Chinese flavor (from a little to a lot), but many cross an imaginary line (in my mind at least) such that they are *not* what one would call chéngyǔ.

(Note that this is all happening on a broad sliding scale, at the far end of which a phrase may be improvised on the spot to accommodate a certain situation in the conversation. In short, these expressions occupy a twilight zone, midway between the informality of báihuà [the vernacular] and the formality of wényán [Classical Chinese]. Not only do the expressions exist on a sliding scale, but their context matters, too. Consider the expression 'sour grapes' in English. In a setting where the educational level is low, the expression 'sour grapes' may exist simply as an expression in the air. In a higher socioeconomic setting, it might be recognized as something to do with a story handed down from antiquity. In a yet "higher" cultural setting, it might be recognized as the legacy of a specific Aesop's Fable about a fox. Similarly in Chinese, depending on the educational background of the listener, a given phrase might be perceived as [a] just a common saying or súyǔ, or [b] as a pointed allusion, i.e., as a formal chéngyǔ.)

And the punch line: Since Chinese-English dictionaries are focused on the monosyllabic zì and the bisyllabic or trisyllabic word, only a few of these *four*-character expressions can be expected to make an appearance in their pages.

Take the case of zhāoQín-mùChǔ 朝秦暮楚 (literally, In the morning it's the Qin State,[59] in the evening it's the Chu State), meaning 'capricious, fickle, unreliable'. This we would expect to find in a large dictionary such as the *Gwoyeu Tsyrdean* (and yes, it appears on p. 2618), but when it comes to a medium-sized dictionary such as *Zhongda*, all bets are off. If we're *lucky*, it might be there; if it is not there, one can hardly be surprised, because these are extras, icing on the cake so far as medium-sized Chinese-English dictionary tradition is concerned. On p. 1040, the *Zhongda* has several *other* "morning-evening" expressions, but it happens not to offer a definition for this one in particular. (The *HarperCollins* likewise does not include the expression zhāoQín-mùChǔ, which is not surprising since it is a smaller dictionary.) What about zhōngyōngzhīdaò 中庸之道 ? The *Zhongda* defines this phrase as 'Doctrine of the Mean', which is fine as far as it goes, but this fails to explain its actual *use,* where its meaning is: 'to do everything in moderation.'[60] Meanwhile, the *HarperCollins* compilers do not list the four-character phrase, but they do include the word zhōngyōng 中庸 , and for this they give its *secondary* meaning only, 'mediocre', without even mentioning its *primary* role as synonym for zhōngyōngzhīdaò, one of the *Four Books*. Quite a bum steer! So one begins to see the difficulty with dictionaries. And these were just two random examples I pulled out of the air.

Below are some sample four-character phrases jotted down from a book in the public library in Maplewood, Minnesota. To my eye, only two or three of these are chéngyǔ; in general, they belong to the type I call *ipso facto* (represented by the 'IF' row in **Figure 4**).

---

59. The same Qin State that would later found the Qin Dynasty, of book-burning fame.
60. Illustrated in Yang (1972) p. 160 for instance, where she contrasts three kinds of traveler: Americans are arrogant, the Japanese too deprecatory (getting on one's nerves for just the opposite reason); only Chinese take the middle road (zhōngyōngzhīdaò), and thus, acting sensible, are welcomed universally as world travelers, she opines.

True, we could quibble at length over the exact number of chéngyǔ among them, but the salient point is that some kind of "third flavor" makes itself felt in these examples:[61]

> sìmùjiāotóu, yíqiàobùtōng, xīnmǎnyìzú, shìsèbúduì,
> bùchǒubùcǎi, yánshénsànhuàn, wúliáodǐngtoù, yǒubèiérzhàn,
> mòmíngqímiào, shuādìtōnghóng, xīnhuāngyìluàn, lěngcháorèfèng,
> wànbúyuànyì, bēixǐjiāojí, tóngbìngxiānglián, lǐsuǒdāngrán,
> yīyīshuōbù, yǐdébàoyuàn, yǔzhòngshēncháng, yìběnzhèngjīng,
> bàochóuxuěhèn, yìyánjiǔdǐng, tóubènnùhǎi.

> — Liang (2005), pp. 65–75 (block of 10 pages chosen at random)

(The list above represents all of the four-character phrases found in the indicated page range.)

Liang is a popular author of 'young adult' novels. Some of these expressions occur in his narrative; some in a character's interior monologue; and others in the dialogue itself. Above I show them first in romanization only, to approximate their effect when spoken — say when having the novel read to you aloud. (Granted, I have taken them out of their context, which is often quite helpful in giving/clarifying the meaning.)

Below, I repeat the list, using characters instead:

> 四目交投 , 一竅不通 , 心滿意足 , 勢色不對 ,
> 不瞅不睬 , 眼神散渙 , 無聊頂透 , 有備而戰 ,
> 莫名其妙 , 刷地通紅 , 心慌意亂 , 冷嘲熱諷 ,
> 萬不願意 , 悲喜交集 , 同病相憐 , 理所當然 ,
> 一一說不 , 以德報怨 , 語重深長 , 一本正經 ,
> 報仇雪恨 , 一言九鼎 , 投奔怒海　　　　— Liang (2005), pp. 65–75

Now they are easier to grasp. But in a few cases, even with the characters spelled out, one might feel inclined to consult a dictionary to be sure what a term means, e.g., shēncháng lit. 'deep long' alludes to the Hànshū 漢書 , a classic, where these two characters are glossed as "with purity and depth of meaning," per *Gwoyeu Tsyrdean* p. 3129. Finally, we should note that

---

61. Not to say that *any* symmetrical four-character expression is automatically to be approached as an item of Classical Chinese or as an instance of my '*ipso facto*' type. For example, here are two phrases that reside on the báihuà side of the line (just barely, I would say): huángjīn-shídài 黄金時代 'golden age'; shūjí-zìhuà 書籍字畫 'books and paintings' (or more literally, 'books, documents, calligraphy and paintings'). But since we are dealing with a sliding scale, there will be disagreements about where and how to draw the line.

the title itself includes an example: lányánzhījǐ 蓝颜知己 as a takeoff on hóngyánzhījǐ 红颜知己 .

Before leaving this very important sub-topic, let's take a quick look in Yang (1972) to see what kind of four-character phrases might appear in her memoir. As evidenced by the use of particles such as .ba 吧, .ne 呢 and .me 嚜 on pp. 16, 52 and 117, respectively (also the 'horse meat' double-entendre built around mǎ 馬 and .ma 嗎 on p. 143), this book is written in a self-consciously báihuà style, in harmony with the philosophy of the author's husband, the linguist Y.R. Chao. Nevertheless, the urge to employ handy, familiar four-character expressions, some of which have a decidedly classical flavor, is clearly irresistible. Sometimes the author introduces the phrase as something slightly unusual; sometimes she drops it into the mix like any other báihuà expression. (And indeed, would *you* want to be the one to find a more "friendly, báihuà way" of expressing 見利忘義 [defined in the *Zhongda* as 'to forget moral principles upon the sight of profits']? Sometimes a well-worn Classical Chinese phrase *is* the way to ease and clarity.)

Here is my random list, then, using characters only in this instance, just to give a range of flavors of the kind of expressions used by Yang Bu-wei:

朝秦暮楚　無人負責　上有天堂　下有蘇杭　生絲出口　熟絲進口
一人犯法　大家遭殃　眼觀四面　耳聽八方　臨難知好友　亂世見忠臣
高枕而臥　有始無終　軍閥當政　兩粵方言　竟成永別　追念無已
樂極生悲　分類歸齊　退休養老　東砲西奔　活靈活現　果不其然
拼死忘生　冤家路窄　以德報怨　見利忘義　果報有靈

　　– Yang (1972) pp. 39, 50, 52, 56, 58, 59, 69, 71, 76, 87, 88, 92, 94, 96, 100, 102

In this case I have not aimed to provide the exhaustive list of such phrases in the page range indicated, as I did earlier for Liang (2005), only a random sampling based on a quick perusal of the first half of the book, by way of reemphasizing the importance of this sub-topic. Some pages contain multiple examples, hence the disparity between the number of phrases given and number of pages cited.

Also, several of the expressions Yang uses happen to be 4 + 4 patterns, or in one case, 5 + 5. All such 'double' phrases I present in the top two rows, followed by the 17 independent four-character phrases in the ensuing three rows. This, too, leads to dissonance between the presentation sequence and the linear string of page numbers cited.

All the above may be described as a "footnote to Notation Convention #2." The impetus is to acknowledge, in passing at least, *all* of spoken Chinese, not just those parts that are easily fitted into the linguists' customary boxes.

APPENDIX L: MUCHENGXUE AND THE MONOSYLLABIC MYSTIQUE

Note: This appendix should be read in conjunction with **Figure 4** (Word and Phrase Taxonomy) on page **170**.

Consider the beginning student's viewpoint: Given that every written character corresponds to a morpheme in the spoken language (note the *direction* of the mapping here), and given that the teacher seems to be fixated on tones, the student will find it reasonable to assume that the three-way combination of character plus syllable plus tone *must* be definitive of one specific *word*. It will seem that learning Chinese words means committing to memory certain combinations of *character-syllable-tone*. It is almost inevitable that the student pass through this stage, at least for a brief period. This is where the myth is born, out of a half-truth on Day One.

By and by the student gains awareness that the Chinese morpheme is often *not* a word, only one half or less of a word, and by several different mechanisms (as detailed in Li and Thompson [1989] pp. 13–14 and 28–84). Moreover, moving along a different axis, the student discovers that a given tone-syllable combination *may* be associated with a dozen or more characters, although thankfully the number of such cases does not seem overwhelming. Thus, over time, the student's notion of a Rebus System Deluxe is seriously challenged on all sides. And yet, once planted in the mind, the sentimental notion of "one character, one word" dies hard.

Beyond those private sentiments, there are external forces that seem to prop up the myth, even validate it, as well. Consider the following scenario. A teacher has finally persuaded the class to approach Chinese with the linguist's notion of a 'syntactic word' — a pragmatic outlook where the word is recognized to be monosyllabic (e.g., yǎo 咬 'bite') *or* bisyllabic (pèifú 佩服 'admire') *or* trisyllabic (hóngdìtǎn 红地毯 'red carpet' [treatment]), as required by the semantics or by the history of the language. Now, for variety, she decides to give the class a poem to study over the weekend. The poem will likely contain phrases of such compactness as mùchéngxuě 暮成雪 (explained in footnote **62** below). On the instant, the

monosyllabic mystique is resurrected with a vengeance, now with the full weight of Li Po [Lǐ Bái 李白], no less, backing it up from beyond the grave, as if to say: "Chinese *is* monosyllabic, and gloriously so: It is precisely that feature, above all others, that allows me to write poetry so intense that it is not just good but *i-m-m-o-r-t-a-l*!" For context, the phrase in question comes from these opening lines of his 樂府 style poem entitled 'Bring on the Wine' 將進酒 :[62]

> 君不見　黃河之水天上來　奔流到海不復回
> 君不見　高堂明鏡悲白髮　朝如青絲暮成雪

Not to say an equally compact construction might not occur in a quoted chéngyǔ, e.g., duìniú tánqín 对牛弹琴 'play the lute [while] facing an ox', i.e., to cast pearls before swine; dàjīng xiǎoguài 大惊小怪 'to exhibit great astonishment at a small strangeness', i.e., to overreact in a silly or foolish way. Or, for that matter, the same degree of compactness occurs often enough even in báihuà conversation, e.g., wǒ pà gǒu 我怕狗 'I'm afraid of dogs'. So it needn't be a poem that triggers a new cycle of infatuation with the monosyllable, to the detriment of multisyllabic thinking. But poems are especially good at doing the job, for they intimate the potential for a *marathon* of high density, line after line, whereas we know the high syntactic density of a phrase such as 我怕狗 will occur only randomly in báihuà, usually with no particular rhetorical effect achieved or intended. Thus, come Monday morning, our hypothetical teacher might feel that she has gone one step forward and two steps backward in the effort to help her students move beyond the tenacious myth.

Coming at it from the other direction, the myth is reinforced this way, too:

> ... practically all *Chinese students of English* refer to the unit of Chinese writing as a *word* and never as a character.   –Chao (1968a) p. 137 (emphasis added)

Moreover, Chao warns us not to go too far in trying to debunk the myth because "The so-called 'monosyllabic myth' is in fact one of the truest myths in Chinese mythology"; *Ibid.* p. 139. My reaction: If we were talking only about Mandarin, I think Chao's warning would sound slightly out of place (too strong), but looking at the big picture, namely the Chinese Language *Family* (as I call it in **Appendix B: Dialects**), then yes, the presence of the venerable

---

62. Nine days out of ten, I would say the poem is simply untranslatable, but today I'll be foolhardy and try to give some impression of its opening quatrain. Note how wordy the English is, at times, compared to the Chinese: *Do you not see? How the Yellow River flows down from Heaven / Racing to the ocean never to return! / Do you not see? In the Great Hall a tall mirror gleams, only to frame its lonesome owner lamenting white hair: / Though blackest silk at dawn, it turns to snow in the evening* (mùchéngxuě, lit. 'evening become snow').

monosyllable can be felt strongly in all those non-Mandarin "dialects" that are actually the true Chinese *languages* of today. Note that the same admonition appears elsewhere, in a version where he leads into it with a qualifying statement:

> *But so far as Classical Chinese and its writing system is concerned*, the monosyllabic myth is one of the truest myths in Chinese mythology. –Chao (1968b) p. 103 (emphasis added)

Here, our viewpoints converge.

BIBLIOGRAPHY

Boyce, C.D. 'Min Sandhi in Verse Recitation', *Journal of Chinese Linguistics* 8:1 (1980) 1–14.

Chao, Yuen Ren 趙元任. *A Grammar of Spoken Chinese*. University of California Press (1968a) 847 pp.

_____. *Language and Symbolic Systems*. Cambridge University Press (1968b) 240 pp.

_____. *Mandarin Primer.* Harvard University Press (1948) 336 pp.

Chiang, Yee 蔣彝. *Chinese Calligraphy: An Introduction to its Aesthetic and Technique.* Methuen & Co. (1938) 230 pp.

du Sautoy, Marcus. *The Music of the Primes.* HarperCollins (2003) 335 pp.

Duden: Das Herkunftswörterbuch: Etymologie der deutschen Sprache. Dudenverlag (2006) 960 pp. [Volume 7, the etymological volume of the Duden dictionary project]

Fenn, Courtenay H. *The Five Thousand Dictionary [Chinese-English].* Harvard University Press (1942 [1926]) 696 pp.

Gwoyeu Tsyrdean 國語詞典. Taibei (1969 [1947]) xxxii + xix + ii + xiii + 4485 p., followed by supplements, colophon, indices and appendices, totalling an additional 361 pp.

HarperCollins Chinese Concise Dictionary (2005) 957 pp.
Comment: The Chinese-English team and English-Chinese team obviously had some serious (non)communication problems on this project. For example, the E-C half of the dictionary doesn't have entries for 'gangway', 'gargle' and 'gauze' even though these words are represented in the C-E half of the dictionary (as xiántī, shùkǒu, shābù, on pp. 356, 307, 284, respectively). Also, on the E-C side, some of the Chinese terms are only half-baked glosses, not genuine Chinese counterparts to the English words. On the plus side: To each definition in a primary entry, and to every compound listed under that primary entry, this dictionary assigns a grammatical classification: míng 名, dòng 动, xíng 形, etc. (whereas the *Zhongda* dictionary, for instance, does not).

Hsu, Ying 徐英 and J. Marvin Brown. *Speaking Chinese in China*. Yale University Press (1983) 314 pp.

Li, Charles N. and Sandra A. Thompson. *Mandarin Chinese: A Functional Reference Grammar.* University of California Press (1989 [1981]) 691 pp.

Liang Derun and Zheng Jiande, eds. *Zhongda Chinese-English Dictionary.* The Chinese University Press, The Chinese University of Hong Kong (2003 [1999]) 1141 pp.

Liang Wangfeng 梁望峰 . *Wo shi ni.de lanyan zhiji* 我是妳的藍顏知己 . Hong Kong (2005) 322 pp.

Liu Fu, Li Jiarui. *Song Yuan Yilai Suzi Pu* 刘復,李家瑞,宋元以来俗字谱. Shanghai: Academia Sinica (1930)

Mathews, R.H. *A Chinese-English Dictionary.* (1943 [1931]) xxiv + 1226 pp.

Packard, Jerome L. *The Morphology of Chinese: A Linguistic and Cognitive Approach.* Cambridge University Press (2000) 335 pp.

Yang Bu-wei 楊步偉 . *Zaji Zhaojia* 雜記趙家 . Taibei 臺北 Zhuanji Wenxue Congkan 傳記文學叢刊 24 (1972) 206 pp.

Note:

ACKNOWLEDGEMENTS AND ERRATA

To Bell Yung, a big thank-you for suggesting that I work '3D' into the title. (Before receiving that suggestion, I had nurtured a delusion that if I called the book *Chinese Soundscape Lexicon* its nature would be immediately self-evident.)

Neither is the other half of the title, 'Chinese As It Is', my own idea. The phrase 'As It Is' I lifted from the *Bhagavad-Gītā As It Is* (1972, 1986) by A.C. Bhaktivedanta Swami Prabhupāda.

(A caveat: Happy though I am with that stolen phrase, one would be remiss not to point out the following aspect of the Swami's expansion of the classic into a 924-page extravaganza. The first four layers — consisting of Sanskrit text; roman transliteration; English equivalent; and English translation — are a dream come true for those of us who like to work across cultural borders. All of that looks solid, save for the stray typo in a transliteration. But when it comes to the fifth layer, the one he identifies on page v as his "elaborate purports," the swami sometimes combines the worst of two traditions: western evangelism and eastern hocus-pocus. He is, after all, the founder of the International Society for Krishna Consciousness, a group with an agenda that goes far beyond the great *Gītopanishad* 'As It Is'. The PURPORT on page 378 exemplifies the problem.)

So much for the title; onward to the cover art and text:

Fred Wahlquist provided the cover art and book design, along with invaluable advice on the text, as well. Many thanks to Jane Mackay (at janemac.net) for her editing of the final text. In that round of editing, our focus was narrowly on the English text that precedes and follows the dictionary proper; meanwhile, given the nature of the beast, it is inevitable that a few typos or errors or points of contention remain on the Chinese side. In that connection, any corrections or suggestions would be most welcome here: conalboyce@gmail.com. My intention is to post updated PDF versions of the dictionary from time to time on conalboyce.com (where one may also find my monograph on Japanese morphology and syntax, and a manifesto on another great language, that of chemistry, in the unlikely role of spiritual redeemer).

Index