

[[Chinese Mac](#)] [[Wubihua Tutorial](#)]

五笔字型

Wubizixing

For Speakers of English

How to master the most rapid Chinese typing method for computers

By Joe Wicentowski, 魏伟

Copyright 1996

Table of Contents

- [Chapter 1: Introduction](#)
 - [Chapter 2: Rules to Know Before Typing](#)
 - [Chapter 3: Practice with the Regions](#)
 - [Chapter 4: The Isolation Rule](#)
 - [Chapter 5: Time-Saving Rules](#)
 - [Appendix A: Graphical Summary of All Rules Used to Enter Single Characters](#)
 - [Appendix B: Table of Second-Level Characters](#)
 - [Acknowledgments](#)
-

Chapter 1: Introduction

In this computer age, the written Chinese language may seem simply too complex to survive. With more than 40,000 characters lacking any phonetic information, such as letters, it is difficult to imagine a keyboard that could include every character and the skill necessary to master such a keyboard.

Necessity has produced many methods of typing, usually much more efficient than the one just proposed. Unfortunately, among these methods an inverse relationship exists between the amount of time necessary to master the method and the efficiency of the method.

For example, the popular Pinyin method requires little training; one need only transliterate the character being entered into English letters. However, since multiple characters correspond to a single pronunciation, isolating the desired character from a list of many, does slow and interrupt the process. In addition, phonetic input methods that use one Chinese dialect are unusable by speakers of another dialect.

五笔字型 (Wu3bi3zi4xing2) solves both the dialect problem and the speed problem. Wubi, as this text will refer to the method, has these advantages:

- Any Chinese character can be entered by using four or less keystrokes.

- Based on the written character, rather than pronunciation, multiple methods for different dialects are not necessary; therefore, Wubi is universal.
- A skilled operator can input 160 characters per minute, a rate exceeding any Western language.

The only drawback to Wubi is that extensive memorization is required to become proficient. For anyone familiar with the complexity of Chinese characters or the speed limitations of the Pinyin input method, Wubi is clearly worth the work. For English speakers, however, interpreting the Chinese documentation for Wubi could prove difficult and frustrating. This document, the first of its kind, clearly presents both the basics and intricacies of the Wubi input method for the benefit of speakers of English.

1.1: Fundamentals

Every Chinese character can be broken down into its root characters, or 字根 (zi4gen1). Conversely, characters can be built by combining these roots. For example, observe the composition of 李 (li3), a common family name:

木 + 子 = 李

In addition, observe the composition of 明 (ming2), which means bright:

日 + 月 = 明

This idea -- that characters can be built by combining roots -- is the fundamental principle behind the Wubi input method. The keyboard contains five regions, organized by the direction of the root's first stroke. Each region contains five keys (making twenty-five keys in total), and each key contains a number of character roots. Diagram 1 shows the regions of the Wubi keyboard:

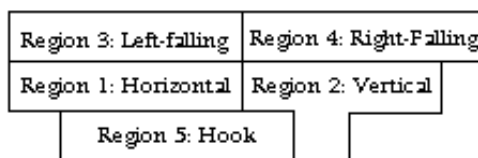


Diagram 1: Keyboard Region/Direction

To continue the 李 and 明 examples, observe in which regions (either 1, 2, 3, 4, of 5) the roots of these characters would fall.

李:

- 木 begins with a horizontal stroke. Therefore, the first root is found in Region 1.
- 子 begins with a hook stroke. Therefore, the second root is found in Region 5.

明:

- 日 begins with a vertical stroke. Therefore, the first root is found in Region 2.
- 月 begins with a left-falling stroke. Therefore, the second root is found in Region 3.

In addition, an example of a root character from Region 4 is 之. Notice that it begins with a right-falling stroke, which often appears as a dot.

1.2: Arrangement of the Wubi keys

Just as we refer to each region by its number, we refer to each key as a combination of both its region and its key number within the region. For example, Diagram 1 shows that G is in the first region and that it is the first key in that region. Therefore, G is 11. Moving leftward, the second key, F, is 12. The last key in the first region, A, is 15. Each region proceeds in the same manner, from the inner-most key on the keyboard to the outer-most.

We will refer to the individual keys by the key number and letter: neither 12 nor F, but 12F. Thinking of the keys in this notation will help your conceptual understanding of Wubi. However, we will refer to the Wubi codes, composed of multiple keys, by only the letters: JQVO, for example.

The Basic Keyboard Diagram, below, shows the essential though incomplete Wubi keyboard: the keys (like 11G and 15A) and the representative root character for each key. The representative root is the root which is most similar to the other roots on its key, and therefore learning the representative roots simplifies memorization of the complex Wubi keyboard, shown on Diagram 3.

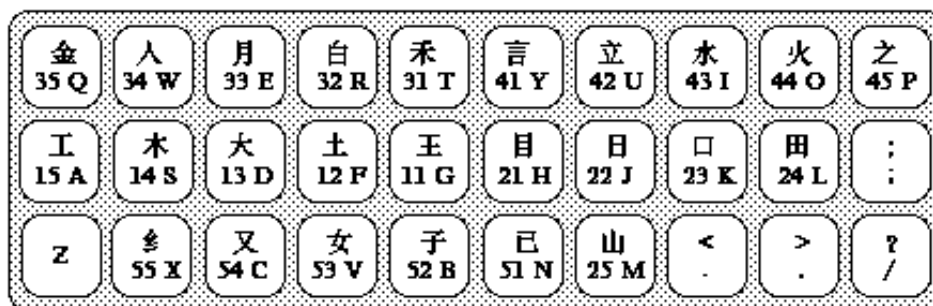


Diagram 2: Basic Keyboard Diagram

Diagram 2 shows that the representative root of 41Y is 言, whose first stroke is one dot, or one right-falling stroke. Diagram 3 shows that the other roots on 41 also contain just one right-falling stroke: 文, 方, and 广. The representative root for 42U is 立, which contains two dots (the dots between the horizontal lines), and the other roots on 42U also contain two dots: 六, 冫, and 辛. The more-detailed Note the similarity between the representative root -- also shown in Diagram 3 as the top-left root on each key -- and the other roots on each key.

In general, the roots are placed logically so as to assist the user's memory. 11G, in the horizontal region, includes the root 一 (yi1); 12F includes 二 (er4); and 13D includes 三 (san1). Be careful, however, because 四 (si4) begins with a vertical stroke; so it fits on 24L, not 14S.

"Leftover" roots -- the roots that do not fit into this pattern -- often appear on the fifth key of each region. These keys are the most difficult to memorize.



Diagram 3: The Complete Wubi Keyboard

Click [here](#) for a larger, higher-quality keyboard diagram. At first, the arrangement of the keys may seem arbitrary and overwhelming. With consistent practice and experience, though, the knowledge will come. Keep reading.

[\[TOP\]](#)

Chapter 2: Rules to Know Before Typing

2.1: Characters with more than four roots

In the introduction, it was mentioned that, with Wubi, any character can be entered with four or less keystrokes. But what about characters that contain more than four roots? The rule is to enter the first three root characters as usual, but to enter the final root as the fourth key. For example, the character 输 contains five roots: 车 + 人 + 一 + 月 + 亠. But, according to this rule, only the first three and the last are entered. Thus, the code is 24L + 34W + 11G + 22J. Check this with Diagram 3, above, if necessary.

2.2: Typing representative roots

A representative root can be typed by entering its key four times. The Wubi code for 言, the representative root of 41Y, is YYYYY.

2.3: Typing the other characters that appear on the complete keyboard diagram

When a character appears on a key but is not a representative root: First, type the key that the character lies on. Second, build the character by entering the key that contains its first stroke. Third, enter the key that contains its second stroke. Finally, enter the key that contains its final stroke. For example, 五 (wu3) is on 11G, but it is not the representative character. The first stroke, 一, is on 11G. The second stroke, 丨, is on 21H. The final stroke, 丿, is on 11G. Thus, the code for 五 is GGHG. These building strokes are always 一, 丨, 丿, 丶, or 乙; they are always found on 11G, 21H, 31T, 41Y, or 51N.

If the character has only one or two strokes, fill in the remainder of the code with L. For example, 丿 is on 31T, is not a representative root, and only contains one stroke: Thus its code is TTL.

2.4: Shortcuts

Wubi provides many typing shortcuts, often allowing the user to enter only a portion of a character's complete code to produce it. A character has a reduced code under three circumstances: First, when Wubi logic has eliminated all other characters based on the code already entered; and/or second, when the character is a commonly-used character (and therefore was given a shorter code before other, less-commonly-used characters); and/or third, when the character is one of the 25 first level characters (whose codes consist of a single key).

An example of the first case is 魏 (wei4). In its verbose form, the code for 魏 is TVRC. However, no other characters have the code TVR followed by another key (such as TVRK or TVRA). Therefore, TVR will produce 魏.

An example of the second case is 笔 (bi3). In its verbose form, the code for 笔 is TTFN. Whereas other characters exist with the codes containing TT, such as 禾 (TTTT), 笑 (TTDU), 竹 (TTGH), and 竿 (TTHF), 笔 is perhaps the most commonly-used of the group. Therefore, rather than giving a less-commonly-used character the reduced code, Wubi has assigned the code TT to 笔.

An example of the third case is 和 (he2). Meaning "and," it is an often-used character. In its verbose form, the code for 和 is TKG; however, it has a shortened code: T. Similarly, the verbose code for 是 (shi4), which means "is," is JGH; its shortened code is J. With a few exceptions, the shortened code is, logically, always derived from the verbose code. The first level characters will be listed before every section in the next chapter.

Both the verbose and shortened code will produce the character, but clearly the former is more efficient. This is just another way that Wubi logic speeds up typing. We will return to the shortcuts while presenting more examples of typing.

2.5: Keeping an open mind

It is especially important for students of Chinese to keep an open mind about stroke order and the roots that make up a character. Although most characters can be broken down into the roots that appear on the Wubi keyboard easily, some are tricky. Therefore, when all else fails, try to think of the character in a different way; think of the roots in a different configuration.

The remaining rules require actual Wubi experience to understand; therefore, we shall begin with practicing with the regions right away!

[\[TOP\]](#)

Chapter 3: Practice with the Regions

3.1: Region 1

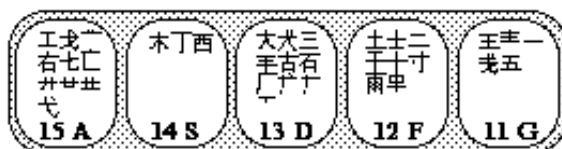


Diagram 4: Region 1

Glance over region 1 and get a general sense of the roots on each key. Be ready to start forming associations in your mind. Remember that every root begins with a horizontal stroke.

First, which characters have single-key codes? What are, in other words, the first level characters of region 1? The first level characters are as follows:

G: 一 F: 地 D: 在 S: 要 A: 工

Note that the root of each character lies on its corresponding key.

We will now analyze some sample characters, following the rules of Wubi typing that have been outlined.

林:	The first root character, 木, begins with a horizontal stroke, so we look to region 1. Glancing over the region 1 keys, we see that 木 is in 14S. The second root is the same, so we type 14S for it, too. Thus, the code for 林 is SS.
开:	The first root, 一, is in region 1 and resides on key 11G. The second root, 升, is also in region 1 but resides on 15A. Thus, the code for 开 is GA.
夺:	The first root, 大, is in region 1 and resides on key 13D. The second root, 寸, is on 12F, so the code for 夺 is DF.
西:	This character is on key 14S, but it is not the representative root. Therefore, we first enter the key that it resides on, 14S. Next, we enter the strokes: 1, 2, and 4. The first stroke, 一, is on 11G. The second stroke, , is on 21H. The last stroke, 一, is on 11G. Therefore, the code for 西 is SGHG.

3.2: Region 2

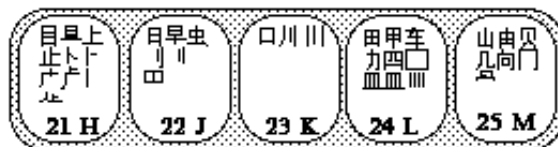


Diagram 5: Region 2

Glance over region 2 and get a general sense of the roots on each key. Remember that every root begins with a vertical stroke.

The first-level characters are as follows:

H: 上 J: 是 K: 中 L: 国 M: 同

Note that the root of each character lies on its corresponding key.

We will now analyze some sample characters, following the rules of Wubi typing that have been outlined, as well as incorporating Region 1 into the examples.

另:	The first root, 冂, is in Region 2 and resides on key 23K. The second root, 力, is also in region 2 but resides on 24L. Thus, the code for 另 is KL.
	The first root, 卜, is in Region 2 and resides on key 21H. The second root, 贝, is also in region 2

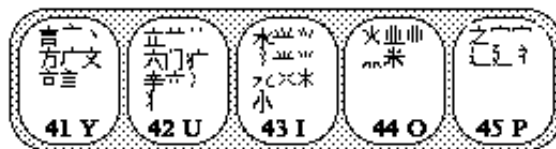


Diagram 7: Region 4

Glance over region 4 and get a general sense of the roots on each key. Remember that every root begins with a right-falling stroke (or a dot). The first-level characters are as follows:

Y: 主 U: 产 I: 不 O: 为 P: 这

Note that the root of each character lies on its corresponding key.

We will now analyze some sample characters, following the rules of Wubi typing that have been outlined, as well as incorporating Regions 1 through 3 in the example characters.

就:	A commonly-used character, 就 is only a second-level Wubi code, made up of its first two roots (the dot, line, and box and 小. Its code is YI.
样:	Another simple character, the code for 样 is SUD.
提:	The code for this character is RJGH.

3.5: Region 5

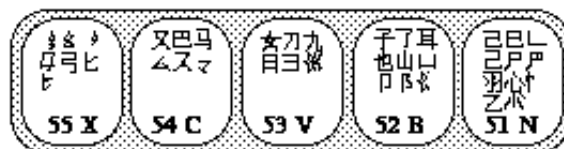


Diagram 8: Region 5

Glance over region 5 and get a general sense of the roots on each key. Remember that every root begins with a hook stroke.

The first-level characters are as follows:

N: 民 B: 了 V: 发 C: 以 X: 经

Note that the root of each character lies on its corresponding key.

We will now analyze some sample characters, following the rules of Wubi typing that have been outlined.

帮:	The roots, in order, are 三, 丿, 冫, 冂, 丨. Thus, the code for 帮 is DTBH.
魏:	The first two roots are 禾 and 女. The third root, in Wubi, is neither 丿 nor 田. Rather, it is 白. As noted in Section 2.4, no other characters' codes are TVR followed by another character. Thus, the last key for 厶 is not necessary; TVRC is reduced to TVR.
伟:	The roots are 亻, 土, the hook, and 丨. Thus, its code is WFNH.

[\[TOP\]](#)

Chapter 4: The Isolation Rule

4.1: Why?

The magic of Wubi is its ability to isolate the desired roots of a character from four keys, which contain between three and ten roots each. However, Wubi logic cannot always isolate characters based on root information. In these cases, of course, Wubi conjures up more magic.

First, a few examples: What distinguishes 叭 from 只, since the roots are identical? What distinguishes 沐 from 洒 and 汀, since the second roots of each character appears on the same key? From what we have learned so far about Wubi coding, nothing does. The Isolation Rule, however, dictates how.

4.2: When?

It turns out that the Isolation Rule is necessary when two conditions are met: First, that the character consist of less than four roots, and second, that these, alone, do not produce the desired character.

4.3: How?

Clearly, the difference between 叭 and 只 is the configuration of the character as a whole. 叭 has a left-right configuration, while 只 has an up-down configuration. The only configuration in addition to left-right and up-down is the miscellaneous configuration. It includes roots such as 里, whose roots intersect, and 国, whose roots are contained in another root. Diagram 9 describes these regions:

Configuration	Sample Configurations
Left-right	
Up-Down	
Miscellaneous	

Diagram 9: Overall Character Configurations

In the second case -- 沐, 洒, and 汀 -- not only the second roots of each character distinguish the characters, but also the last stroke of the entire character distinguishes them. To continue the example, the last stroke of 沐 is 丶, of 洒 is 一, of 汀 is 丨.

In an effort to isolate any character in a rule that covers both the configuration and the last stroke circumstances, the Wubi Isolation Rule states the following:

If the character contains less than four roots and if entering just these roots does not produce the character, then the last key -- the isolating key -- can be determined by using Diagram 10.

		Configuration		
		Left-Right	Up-Down	Miscellaneous
Last stroke	Horizontal	11G	12F	13D
	Vertical	21H	22J	23K
	Left-falling	31T	32R	33E
	Right-falling	41Y	42U	43I
	Hook	51N	52B	53V

Diagram 10: The Isolation Rule

To put the first examples to rest, their codes are as follows:

叭:	The code, based only on the roots, is KW. Since the character has a left-right configuration, and since the last stroke is 丶, the isolating key is 41Y. Thus, the code for 叭 is KWY.
只:	The code, based only on the roots, is KW. Since the character has an up-down configuration, and since the last stroke is 丶, the isolating key is 42U. Thus, the code for 只 is KWU.
沐:	The code, based only on the roots, is IS. Since the character has a left-right configuration, and since the last stroke is 丶, the isolating key is 41Y. Thus, the code for 沐 is ISY.
洒:	The code, based only on the roots, is IS. Since the character has a left-right configuration, and since the last stroke is 一, the isolating key is 11G. Thus, the code for 洒 is ISG.
汀:	The code, based only on the roots, is IS. Since the character has a left-right configuration, and since the last stroke is , the isolating key is 21H. Thus, the code for 汀 is ISH.
里:	The code, based only on the roots, is JF. Since the character has a miscellaneous configuration (日 and 土 overlap), and since the last stroke is 一, the isolating key is 13D. Thus, the code for 里 is JFD.

[\[TOP\]](#)

Chapter 5: Multi-character combinations

Until now, this manual has only dealt with entering single characters. However, Wubi logic does incorporate rules for commonly-used, multi-character combinations, and the codes for these combinations remain within the four keystroke limit. The rules are surprisingly simple. However, one rule is extremely important to remember: When using multi-character combinations, the shortened codes as described in Section 2.4 may not be used. Only the verbose codes may be used.

5.1: Two-character combinations

For commonly-used, two-character combinations, the rule is to enter the first two roots of each character. For example, the code for 经 (jing1) is XCA, and the code for 常 (chang2) is IPKH. But, together, the code for the combination 经常, which means day-to-day, is XCIP. Other examples follow:

机器:	The roots of this combination are 木, 几, 口, and 口. Thus, the code is SMKK.
汉字:	The roots of this combination are 讠, 又, 宀, and 子. Thus, the code is ICPB.

5.2: Three-character combinations

The rule for commonly-used, three-character combinations is to enter only the first roots of the first two characters and to enter the first two roots of the third character. For example, the codes for 计算机 (ji4 suan4 ji1) would be YF, THA, and SM, respectively. Together, the code for the combination is simply YTSM. Other examples follow:

可能性:	The roots of this combination are 丁, 厶, 卜, and 丿. Thus, the code is SCNT.
小朋友:	The roots of this combination are 小, 月, 厂, and 又. Thus, the code is IEDC.

5.3: Combinations of more than three characters

The rule for commonly-used combinations of more than three characters is to enter the first root of the first three and of the last character in the combination. For example, 五笔字型 (Wu3bi3zi4xing2) consists of four characters, and the first roots of each character are 一, 竹, 宀, and 一, respectively. Therefore, the code for 五笔字型 is GTPG. Other examples follow:

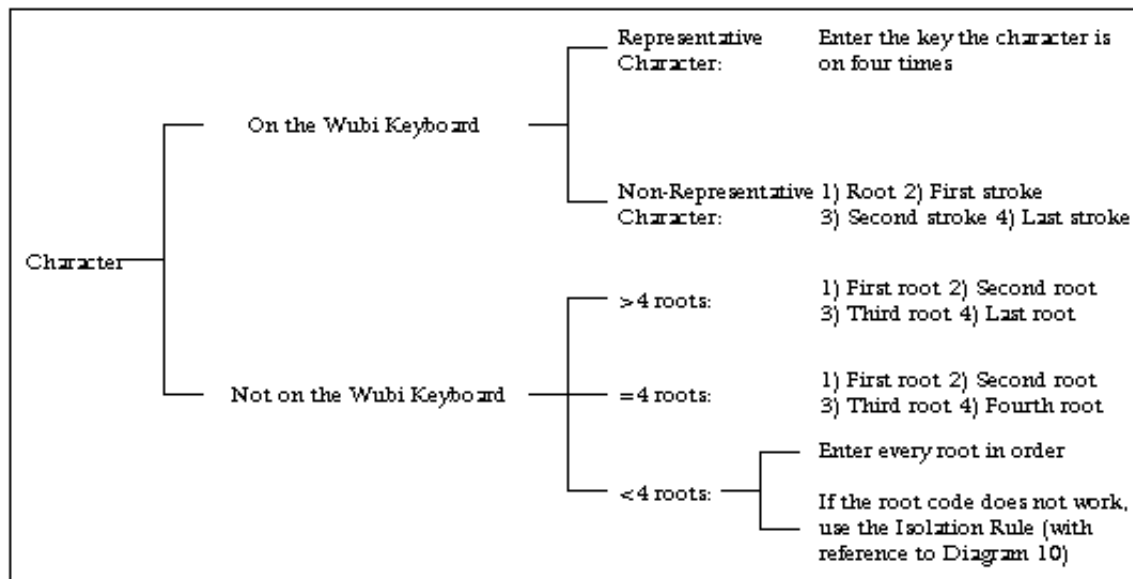
一分为二:	The roots of this four-character combination are 一, 八, 丶, and 二. Thus, the code for this combination is GWYF.
集体所有制:	The code for this five-character combination is WWRR.
人民大会堂:	The code for this five-character combination is WNDI.
中华人民共和国:	The roots of this combination, which means the People's Republic of China, are 口, 亻, 人, and 口. Thus, the code for this seven-character combination is KWWL.

5.4: The Z key

Z is the only letter not included in the Wubi root keyboard; it is reserved for a special function: the wildcard. If a user is unsure of one or more roots in a code, he or she may replace that letter with Z, and the program will list all characters or combinations whose code contain the given keys. For example, if a user knows that the code of 情 begins NG but cannot find the key that contains 月, he or she can enter NGZ. The list shown will be similar to this: 1 刁 2 恒 3 怀 4 情 5 司 6 屋 7 怔 8 怦 9 慷. Using the mouse, cursor keys, or numeric keypad, select the desired character. The Z can be entered anywhere and as many times as is desired. Note that not all programs that allow Wubi input use this function.

[\[TOP\]](#)

Appendix A: Graphical Summary of All Rules Used to Enter a Single Character



[\[TOP\]](#)

Appendix B: Table of Second-Level Characters

Second-level characters are formed by entering a key on the vertical axis followed by a key on the horizontal axis, as in the table below.

	G	F	D	S	A	H	J	K	L	M	T	R	E	W	Q	Y	U	I	O	P	N	B	V	C	X
G	五	于	天	未	开	下	理	事	画	现	玫	珠	表	珍	列	玉	平	不	来	-	与	屯	妻	到	互
F	二	寺	城	霜	载	直	进	吉	协	南	才	垢	圾	夫	无	坟	增	示	赤	过	志	地	雪	支	-
D	三	夺	大	厅	左	丰	百	右	历	面	帮	原	胡	春	克	太	磋	砂	灰	达	成	顾	肆	友	龙
S	本	村	枯	林	械	相	查	可	楞	机	格	析	极	检	构	术	样	档	杰	棕	杨	李	要	权	楷
A	七	革	基	苛	式	牙	划	或	功	贡	攻	匠	菜	共	区	芳	燕	东	-	芝	世	节	切	芭	药
H	睛	睦	-	盯	虎	止	旧	占	卤	贞	睡	-	肯	具	餐	眩	瞳	步	眯	瞎	卢	-	眼	皮	此
J	量	时	晨	果	虹	早	昌	蝇	曙	遇	昨	蝗	明	蛤	晚	景	暗	晃	显	旱	电	最	归	紧	昆
K	呈	叶	顺	呆	呀	中	虽	吕	另	呼	员	听	吸	只	史	嘛	啼	吵	-	喧	叫	啊	哪	吧	哟
L	车	轩	因	困	-	四	辍	加	男	轴	力	斩	胃	办	罗	罚	较	-	-	边	思	-	轨	轻	累
M	同	财	央	朵	曲	由	则	-	崭	册	几	贩	骨	内	风	凡	赠	峭	-	迪	岂	邮	-	凤	-
T	生	行	知	条	长	处	得	各	务	向	笔	物	秀	答	称	入	科	秒	秋	管	秘	季	委	么	第
R	后	持	拓	打	找	年	提	扣	押	抽	手	折	扔	失	换	扩	拉	朱	接	近	所	报	扫	反	批
E	且	肝	-	采	肛	-	胆	肿	肋	肌	用	遥	朋	脸	胸	及	胶	膛	-	爱	甩	服	妥	肥	脂
W	全	会	估	休	代	个	介	保	佃	仙	作	伯	仍	从	你	信	们	伙	-	亿	他	分	公	化	
Q	钱	针	然	钉	氏	外	旬	名	甸	负	儿	铁	角	欠	多	久	匀	乐	炙	锭	包	凶	争	色	-
Y	主	计	庆	订	度	让	刘	训	为	高	放	诉	衣	认	义	方	说	就	变	这	记	离	良	充	率
U	闰	半	关	亲	并	站	间	部	曾	商	产	瓣	前	闪	交	六	立	冰	普	帝	决	闻	妆	冯	北
I	汪	法	尖	洒	江	小	浊	澡	渐	没	少	泊	肖	兴	光	注	洋	水	淡	学	沁	池	当	汉	涨
O	业	灶	类	灯	煤	粘	烛	炽	烟	灿	烽	煌	粗	粉	炮	米	料	炒	炎	迷	断	籽	姜	炆	-
P	定	守	害	宁	宽	寂	审	官	军	宙	客	宾	家	空	宛	社	实	宵	灾	-	官	宇	安	-	它
N	怀	导	居	-	民	收	慢	避	惭	届	必	怕	-	愉	懈	心	习	情	屡	忧	忆	敢	恨	怪	尼
B	卫	际	承	阿	陈	耻	阳	职	阵	出	降	孤	阴	队	隐	防	联	孙	耿	辽	也	子	限	取	陛
V	姨	寻	姑	杂	毁	-	旭	如	舅	-	九	-	奶	-	婚	妨	嫌	录	灵	巡	刀	好	妇	妈	姆
C	-	对	参	-	戏	-	-	台	劝	观	矣	牟	能	难	允	驻	-	-	-	驼	马	邓	艰	双	-
X	线	结	顷	-	红	引	旨	强	细	纲	张	绵	级	给	约	纺	弱	纱	继	综	纪	弛	绿	经	比

[\[TOP\]](#)

[\[Wubihua Tutorial\]](#) [\[Chinese Mac\]](#)