Scientific
Research
Publishing

Teaching and Learning the Pronunciation of Mandarin, Part II: The 12i as Backbone of the Method
有无道 *yǒu wú dào*

Leo Depuydt

Department of Egyptology and Assyriology, Brown University, Providence, RI, USA
Email: depuydt@juno.com

Abstract

Part I of this multi-part paper contains basic practical and introductory observations about difficulties faced by westerners in learning to pronounce Chinese (Depuydt, 2024). It was argued that, in learning Chinese, pronunciation deserves a certain priority or precedence over grammar, vocabulary, and writing (calligraphy). A new method was introduced for studying pronunciation. It was called 有无道 *yǒu wú dào* "The Way of *yǒu wú*," in the spirit of Laozi. And a set of twelve (12) initial consonants plus Pinyin *i* called the 12i was defined. The 12i contains most of what is difficult to pronounce for westerners in Chinese. The design of the present **Part II** is to cement the 12i as the core for learning how to pronounce Chinese. **Part III** (and possibly beyond) is where this paper assumes a truly practical dimension for teaching and learning how to pronounce Chinese. **Part III** will discuss individual sounds, first and foremost seventeen (17) out of the twenty-one (21) initial consonants, including the 12(i) plus five (5) more. Practical hints and exercises pertaining to the pronunciation of individual consonants will be presented. As for the 12i, the proposals found in this paper reflect a student perspective. They derive from efforts to develop home-baked ways of pronouncing Chinese accurately over many months and years. A first oral report on these efforts was presented, partly in English and partly in Chinese, in the fall of 2021 at the 10th International Conference of the New England Chinese Language Teachers Association (NECLTA), under the title 努力准确发音普通话: 西方初学者的一些观察 "Struggling to Pronounce Mandarin Accurately: Observations by a Beginner from the West."

1. Statement of Purpose: A New Start

The design of the entire study—of which this is Part II—is didactic and educational. The aim is to add more comprehensive and coherent structure to the learning and teaching of the pronunciation of Chinese.

This is not a research paper in the strict sense. Nor can it be, considering its aim. Hence the general absence of academic references.

Then again, some effort has been made to establish that what is written here does not contradict the current state of academic knowledge.

In assessing the current state, San Duanmu's *The Phonology of Standard Chinese* (Duanmu, 2007) has served as a sufficient guide and has all the relevant bibliography. There also relevant articles in Brill's Encyclopedia of Chinese Language and Linguistics, available online, from 2015 (Sybesma et al., 2015), including the article on "Phonetics" by Xiaonong Zhu (Zhu, 2015). But as already noted in Part I, theoretical articles on language and linguistics are not a suitable point of departure for teaching beginning students of all ages. Nor should they need to be accessed at any point in the process.

The intended audience of this study is beginning students of Chinese in the west learning to pronounce Modern Standard Chinese and their teachers.

The aim is to go beyond mere mimicking to use conscious reflection and analysis that is tailored to the needs of beginning students.

In addition to imitating teachers, students use comparison of a sound with sounds that are like it, typically sounds in their own mother tongue. The method proposed in this study is exactly the opposite. It compares sounds with what they are *not*.

Current training of the pronunciation of Chinese is fragmented. This study may be the first time that a comprehensive, coherent, and rigorous method following overarching principles and yet accessible to beginning students is attempted.

2. The 21 Initial Consonants of Mandarin

Consonants are generally more difficult to learn to pronounce than vowels. With vowels, there is no obstruction in the flow of air from the lungs through the mouth. With consonants, the flow of air is partly obstructed or even completely occluded by the lips or the tongue at various locations in the mouth from front to back.

There are **twenty-one (21)** initial consonants in Mandarin. Initial consonants occur at the beginning of a syllable.

Initial *y* and *w* are not included in the count of twenty-one (21). They are better described as located somewhere between consonant and vowel. That is why they are sometimes called semi-consonants and sometimes semi-vowels. *Y* and *w* will
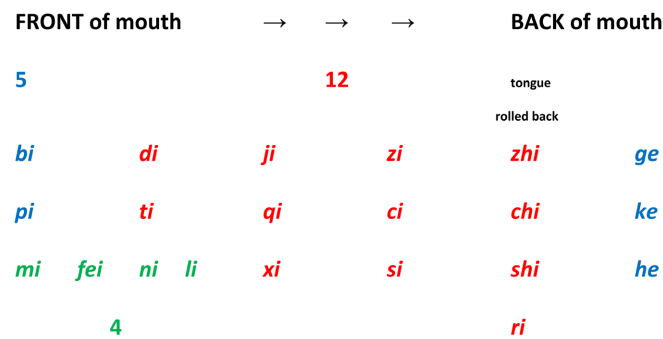
be discussed in detail in their own right later.

The only final consonants in Modern Standard Chinese ("Mandarin") are Pinyin 1) *n* and 2) *ng*, as well as 3) *r* in the syllable *er* in a handful of words including 而 *ér* "and," 耳 *ěr* "ear," and 二 *èr* "two."

A special case is final *r* in the diminutive suffix 儿 (*e*)*r*.

Final *n* and *ng* pose no serious difficulties of pronunciation. The possible tiny difference with the final *n* and *ng* of English and of many other languages is negligeable for the purpose of learning how to pronounce Chinese. Though no two consonants in two different languages are otherwise ever 100% the same.

Final *r* is generally also considered easy to pronounce. That is because it is generally associated with American *r*. But they are not quite the same. On this more later.

The 21 initial consonants are represented in the following figure. A suitable vowel has been attached to each consonant. This is also **Figure 1** of Part I of this paper, with two minor changes pertaining to the position of *fei* and *li*.
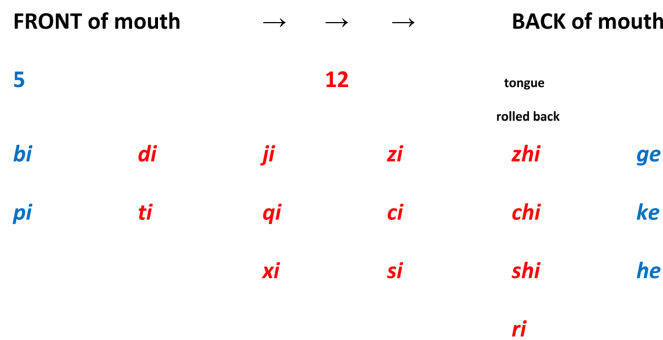
FRONT of mouth			→ → →			BACK of mouth
5				**12**		tongue
						rolled back
bi		*di*	*ji*	*zi*	*zhi*	*ge*
pi		*ti*	*qi*	*ci*	*chi*	*ke*
mi	*fei* *ni* *li*		*xi*	*si*	*shi*	*he*
	4					*ri*

Figure 1. The 21 Initial Consonants (12 + 5 + 4).

3. The 21 Initial Consonants Minus *f*, *l*, *m*, and *n*

Among the 21 initial consonants, **four (4)** will not require any discussion because they hardly pose difficulties. They are *f*, *l*, *m*, and *n*, notated in **green** in **Figure 1** above.

That leaves **seventeen (17)** initial consonants. The table in **Figure 1** may accordingly be shortened as follows.

FRONT of mouth		→ → →		BACK of mouth
5			**12**	tongue
				rolled back
bi	*di*	*ji*	*zi*	*zhi*	*ge*
pi	*ti*	*qi*	*ci*	*chi*	*ke*
		xi	*si*	*shi*	*he*
				ri	

Figure 2. The 21 Initial Consonants Minus *f*, *l*, *m*, and *n*.

4. The 17 Remaining Initial Consonants as 12(i) + 5

Of the 17 remaining initial consonants, **twelve (12)** combine most of the difficulties in pronouncing Chinese for beginners. They are Pinyin *d, t, j, q, x, z, c, s, zh, ch, sh,* and *r*. They are shown in **red** in **Figure 1** and **Figure 2**.

To understand the pronunciation of this set of 12 properly, it will be necessary to consider them in conjunction with the Pinyin vowel *i*, as follows: *di, ti, ji, qi, xi, zi, ci, si, zhi, chi, shi,* and *ri*.

This set of 12 can therefore be designated as the **12i** or the **12(i)**.

In the 有无道 *yǒu wú dào* method, the 12i is the backbone of learning the pronunciation of Chinese. It is the fundamental point of departure.

The method in question makes it possible to find with pinpoint accuracy

1) the three (3) different pronunciations of *i* in the 12i and

2) the pronunciation of all 12 consonants of 12i,

a) not only in front of Pinyin *i*,

b) but also in front of all other vowels.

The remaining **five (5)** consonants are *b, p, g, k,* and *h*. They will be interpreted below as an expansion of the 12(i).

The 12(i) exhibits perfect coherence as a set, a coherence that has not been exploited in the interest of teaching and learning the pronunciation of Chinese.

It is already common to unite nine (9) of the 12(i) into three (3) sets, as follows:

1) *j, q, x*,

2) *z, c, s*, and

3) *zh, ch, sh*.

5. The 12(i) from First Principles in Euclidian Fashion

a) Element 1: The 12(i) contains *five* (5) *simple* consonants and *seven* (7) *compound* consonants.

b) Element 2: The five (5) *simple* consonants consist of *one* (1) component. They are (Pinyin, in Latin alphabetical order):

d   r   s   sh   x

c) Element 3: The seven (7) *compound* consonants consist of *two* (2) or *three* (3) components. They are (Pinyin, in Latin alphabetical order):

c   ch   j   q   t   z   zh

d) Element 4: *Four* (4) of the seven (7) compound consonants consist of *two* (2) components. They are (Pinyin, Latin order):

j   t   z   zh

In the interest of clarity, it will be useful to identify the two components of each already now in anticipation, as follows:

j = d + x

t = d + ʰ

z = d + s

zh = d + sh

ʰ stands for aspiration (see below).

e) Element 5: *Three* (3) of the seven (7) compound consonants consist of *three* (3) components. They are (Pinyin, Latin order):

$c \quad ch \quad q$

In the interest of clarity, it will be useful to identify the three components of each already now in anticipation, as follows:

$c = d + s +{}^{h}$

$ch = d + sh +{}^{h}$

$q = d + x +{}^{h}$

f) Element 6: As can already be seen in the lists d) and e) above, *compound* consonants consist of *simple* consonants, *one* (1) or *two* (2)—with or without aspiration as second or third component.

Aspiration will be marked as ^{h}, whose pronunciation is like English *h*. Pinyin *h* is pronounced quite differently from English *h*.

g) Element 7: Only *four* (4) of the five (5) simple consonants appear in compound consonants, namely *d*, *s*, *sh*, and *x*.

R does not. The pronunciation of *r* will be clarified below.

h) Element 8: In the seven compound consonants, the first component is always Pinyin *d*.

The compound consonants can therefore be rewritten as follows, three (3) containing three (3) components and four (4) containing two (2) components:

$c = d + \ldots + \ldots$

$ch = d + \ldots + \ldots$

$j = d + \ldots$

$q = d + \ldots \quad + \ldots$

$t = d + \ldots$

$z = d + \ldots$

$zh = d + \ldots$

Ten (10) empty spots (. . .) need to be filled.

i) Element 9: In the seven (7) compound consonants, aspiration (^{h}) fills the final spot wherever it can.

That means that aspiration (^{h}) appears four (4) times: three (3) times as third component and one (1) time as second component.

It can evidently only appear once as second component; otherwise two or more consonants would be the same.

The result is as follows:

$c = d + \ldots +{}^{h}$

$ch = d + \ldots +{}^{h}$

$j = d + \ldots$

$q = d + \ldots +{}^{h}$

$t = d +{}^{h}$

$z = d + \ldots$

$zh = d + \ldots$

One consonant is now complete, namely Pinyin *t*.

Six (6) empty slots (. . .) remain to be filled—three (3) in consonants with aspiration and three (3) in consonants without aspiration.

j) Element 10: The remaining empty six (6) slots are filled by the three (3) simple consonants s, sh, and x, each filling two (2) slots.

There is in fact only one (1!) way of filling the six (6) empty slots with s, sh, and x, namely as follows:

$c = d + s +{}^h$

$ch = d + sh +{}^h$

$j = d + x$

$q = d + x +{}^h$

$t = d +{}^h$

$z = d + s$

$zh = d + sh$

Any other placement of s, sh, and x would result in two or more identical consonants.

The above list can be rearranged to bring together what belongs together.

$t = d +{}^h$

$j = d + x$

$z = d + s$

$zh = d + sh$

$q = d + x +{}^h$

$c = d + s +{}^h$

$ch = d + sh +{}^h$

Or else, in order to emphasize the cohesion:

$$t \quad = \quad d \qquad + \qquad {}^h$$

$$j \quad = \qquad\qquad + \quad x$$

$$z \quad = \qquad\qquad + \quad s$$

$$zh = \qquad\qquad + \quad sh$$

$$q \quad = \qquad\qquad + \quad x \qquad + \quad {}^h$$

$$c \quad = \qquad\qquad + \quad s \qquad + \quad {}^h$$

$$ch = \qquad\qquad + \quad sh \qquad + \quad {}^h$$

k) Element 11: The 12(i) Two-dimensionally

The lists in i) are one-dimensional. But the repetition of x, s, and sh yields the following obvious two-dimensional arrangement.

$$d + x \qquad d + x +{}^{h}$$
$$d + s \qquad d + s +{}^{h}$$
$$d + sh \qquad d + sh +{}^{h}$$

The single consonants *x*, *s*, and *sh* may be added to complete the following square:

$$x \quad d + x \qquad d + x +{}^{h}$$
$$s \quad d + s \qquad d + s +{}^{h}$$
$$sh \quad d + sh \qquad d + sh +{}^{h}$$

An alternative arrangement with a 90-degree rotation is as follows:

$$x \qquad\qquad s \qquad\qquad sh$$
$$d + x \qquad\quad d + s \qquad\quad d + sh$$
$$d + x +{}^{h} \qquad d + s +{}^{h} \qquad d + sh +{}^{h}$$

The Pinyin equivalents are as follows.

x	*s*	*sh*
j	*z*	*zh*
q	*c*	*ch*

The second arrangement will be preferred here. The reason is as follows.

In the sequence *x, s, sh*, the tongue moves horizontally from the front to the back of the mouth. The second arrangement better brings out this horizontal movement.

l) Element 12: Placing Pinyin *d* and *t*

Three consonants still need to be added to the two-dimensional arrangement in 4.11 to complete the 12i:

Pinyin *d*, *t*, and *r*.

Where to add *d* and *t* (= *d* + ^{h})? There are only two viable positions, as follows:

$$\qquad\quad x \qquad\qquad s \qquad\qquad sh$$
$$d \quad d + x \qquad d + s \qquad d + sh \qquad d$$
$$d +{}^{h} \quad d + x +{}^{h} \quad d + s +{}^{h} \quad d + sh +{}^{h} \qquad d +{}^{h}$$

Since *d* and *x* are both pronounced near the front of the mouth but *sh* near the back, the following position is more natural.

$$\qquad\quad x \qquad\qquad s \qquad\qquad sh$$
$$d \quad d + x \qquad d + s \qquad d + sh$$
$$d +{}^{h} \quad d + x +{}^{h} \quad d + s +{}^{h} \qquad d + sh +{}^{h}$$

Or also:

	x	*s*	*sh*
d	*j*	*z*	*zh*
t	*q*	*c*	*ch*

The International Phonetic Alphabet is not used here because it would just complicate the learning and teaching.

The IPA equivalents of the sounds in the above table reveal the harmony of the table above.

	ɕ	s	ʂ
t	tɕ	ts	ʈʂ
*t*ʰ	tɕʰ	tsʰ	ʈʂʰ

The cohesion in the table is such that relations between sounds are replicated both horizontally and vertically.

For example, the horizontal relation between *x* and *s* is the same as that between *j* and *z* below it vertically and the vertical relation between *x* and *j* is the same as that between *s* and *z* next to it horizontally. And so on for all pairs throughout the table.

m) Element 13: Placing Pinyin *r(i)*

Pinyin *r* is a case all by itself. Even native speakers sometimes pronounce it in slightly different ways.

Pinyin *r* is mostly associated with American *r* and therefore not given much attention by learners of Chinese because it is considered easy to pronounce. But it is not quite like American *r*.

However, the real problem is the pronunciation of the Pinyin syllable *ri*. This syllable is pronounced (and explained) in different ways.

The syllable *ri* will play a critical role in the development of the 有无道 *yǒu wú dào* method.

The 12i stands at the beginning and other pronunciations are derived from it. Pinyin *ri* is part of the 12i. *Ri* will therefore be used to determine how *r* is pronounced everywhere.

The sequence *r* + *i* is found only in one syllable and one word, namely 日 *rì* "day." There are no Pinyin syllables *rin* or *ring*.

Pinyin *r* can also be followed by the vowels *a, e, o,* and *u* in six different syllables, as in 然 *rán*, 人 *rén* "people," 让 *ràng* "let," 扔 *réng* "throw," 容 *róng* "prosperous," and 如 *rú* "like."

The 有无道 *yǒu wú dào* will be used in **Part III** of this article as follows.

Pinyin *ri* will be the point of departure. The 有无道 *yǒu wú dào* method will be used to find exactly how *r* and *i* are pronounced in their sole combination, namely the syllable and word *rì* "day."

Only then will the pronunciation of *r*—obtained as part of *ri* through the 有无道 *yǒu wú dào* method—be applied wherever else *r* occurs.

For now, the design of the present Part II is to locate *r* in relation to the rest of the 12(i).

Its place is as follows:

			r
	x	*s*	*sh*
d	*j*	*z*	*zh*
t	*q*	*c*	*ch*

The closeness of *r* to *sh* is crucial. It will be explained in **Part III** why that is and how it contributes to applying the 有无道 *yǒu wú dào* method to finding the pronunciation of both *r* and *i*.

6. The Two Dimensions of the 12i and the Mouth's Anatomy

In the tables above, there are two dimensions:

1) from left to right;

2) from top to bottom.

They correspond to two physical dimensions in the mouth:

1) from front to back;

2) from more open to more closed.

	From front			to back
As open as possible				*r*
Less open		*x*	*s*	*sh*
Hard closed	*d*	*j*	*z*	*zh*
Soft closed	*t*	*q*	*c*	*ch*

This matter has already been addressed in **Part I**.

7. The 12i and Its Extension of 5: *b*, *g*, *h*, *k*, and *p*

There are twenty-one (21) initial consonants in Modern Standard Chinese—if one excludes *y* and *w*.

The 12i stands at the center of the 有无道 *yǒu wú dào* as proposed tool of analysis.

Two reasons for the central position of the 12i are as follows.

1) The 12i encompasses most of what is difficult to pronounce for westerners.

2) The 12i makes it possible to establish

a) how Pinyin *i* is pronounced in three different ways in the 12i,

b) how the 12 consonants of the 12i are pronounced when followed by *i*, and

c) how the 12 consonants are pronounced when followed by vowels *other than i*.

Nine (9) initial consonants remain to be accounted for.

Of these nine (9), four (4) need no further attention because they pose no difficulties. They are:

f, *l*, *m*, and *n*.

They are marked in green in **Figure 1** at the outset of the article.

That leaves five (5) more initial consonants to be located. They are:

b, *g*, *h*, *k*, and *p*.

Where are they located in relation to the following layout of the 12i? Let us again contemplate the chart of the 12i.

From front			to back
			r
	x	*s*	*sh*
d	*j*	*z*	*zh*
t	*q*	*c*	*ch*

A remarkable similarity can be observed between the pair

d and *t*

to the left in the table and the two pairs

1) *b* and *p* and

2) *g* and *k*.

How so?

The three pairs can be rewritten as follows:

d and t = d and d + ^{h}

b and p = b and b + ^{h}

g and k = g and g + ^{h}

It follows that the difference between the three (3) pairs d and t, b and p, and g and k is the same as that between just d, b, and g.

How do d, b, and g differ?

The difference is simply between *lip*, *tooth*, and *throat*. In the case of b, the air coming from the lungs is *fully stopped* at the lips and then exploded from the mouth; in the case of d, at the teeth; in the case of g, at the throat. The following table can therefore be constructed.

From	front	to back
Lips	*teeth*	*throat*
b	d	g
b + ^{h}	d + ^{h}	g + ^{h}

This simple three-way division can now be expanded as follows, in full detail (Pinyin h can also be added):

FROM FRONT		TO				BACK
lips	*teeth*	*(+ palate = dome of mouth)*				*throat*
					r	
		x	s	sh		h
b	d	$d + x$ (j)	$d + s$ (z)	$d + sh$ (zh)		g
b + ^{h} (p)	d + ^{h} (t)	$d + x + ^h$ (q)	$d + s + ^h$ (c)	$d + sh + ^h$ (ch)		g + ^{h} (k)

And, in Pinyin alone, as follows:

lips	*teeth*	*(+ palate)*			*throat*
				r	
		x	s	sh	h
b	d	j	z	zh	g
p	t	q	c	ch	k

8. The Treatment of the 12i + 5 by Means of 有无道 *yǒu wú dào* in Part III (Preview)

Part III is where of 有无道 truly takes a practical turn for both students and teachers. The seventeen (17) consonants of the 12(i) + 5 will be treated in the fol-

lowing order:

1) *s* and s*h* (simple consonants of the 12i)
2) *d* and *x* (simple consonants of the 12i)
3) *r* (simple consonant of the 12i)
4) *j z zh t q c ch* (the seven compound consonants of the 12(i)
5) *b* and *g*
6) *p* and *k*
7) *h*

9. Conclusion

The foundations of the 有无道 *yǒu wú dào* have now been established. The critical and central importance of the 12i has been affirmed in this Part II.

It is now time to turn, in Part III, to the pronunciation of individual sounds.

(*To be continued*)

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Depuydt, L. (2024). Teaching and Learning the Pronunciation of Mandarin: The 有无道 yǒu wú dào "The Way of yǒu wú", Part I (Perspective: Method; The 12i). *Open Journal of Modern Linguistics, 14,* 521-542. https://doi.org/10.4236/ojml.2024.143027

Duanmu, S. (2007). *The Phonology of Standard Chinese* (2nd ed.). Oxford University Press. https://doi.org/10.1093/oso/9780199215782.001.0001

Sybesma, R. et al. (Eds.) (2015). *Encyclopedia of Chinese Language and Linguistics*. Brill. https://scholar.harvard.edu/files/ctjhuang/files/2016_brill_ecll.pdf

Zhu, X. N. (2015). Phonetics, Articulatory. *International Encyclopedia of the Social & Behavioral Sciences,* 65-74. https://doi.org/10.1016/B978-0-08-097086-8.52013-3