

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)

Leo Depuydt

Department of Egyptology and Assyriology, Brown University, Providence, RI, USA

Email: depuydt@juno.com

How to cite this paper: 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>

Received: April 30, 2024

Accepted: June 25, 2024

Published: June 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The aim of this paper is to propose and describe a new approach to teaching and learning the pronunciation of Modern Standard Chinese. It may be called the 有无道 *yǒu wú dào* “The Way of What’s-There and What’s-Not-There.” Both name and method are inspired by, 1) a verse in the Dao De Jing, a seminal work traditionally attributed to the sage Laozi, and 2) the dawning digital age initiated by G. Boole in which the Zeros are as important as the Ones, the absences as important as the presences. The method does not aim to replace tried and true methods. It is meant as a possible supplement. It is also designed to be 100% practical and accessible to anyone. Anyone of whatever age might use it as a guide to pronounce Chinese more accurately. The method involves learning and teaching the pronunciation of a sound by **comparing what it is (有 *yǒu*) with what it is not (无 *wú*)**, what it is not being the pronunciation of (an)other sound(s). Part of this paper was read in October 2021, half in Chinese and half in English, at the 10th International Conference of the New England Chinese Language Teachers Association (NECLTA), organized by Tufts University in Boston and held remotely due to Covid. The title was: 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 The content has evolved and expanded considerably while producing this written paper. With the rise of China and the internet in the past twenty to thirty years, there are plenty of handbooks and videos teaching the pronunciation of Modern Standard Chinese, also known as Mandarin, which is strictly speaking the language spoken by officials in the Ming and Qing Dynasties.

¹Instead of 准确发音 *zhǔnquè fāyīn* “pronounce accurately,” the title originally had 正确发音 *zhèngquè fāyīn* “pronounce correctly.” But 正确 *zhèngquè* “correctly” may have moral overtones. Moreover, it is possible, I believe, to pronounce Mandarin “accurately” with a (slight) accent. Getting rid of an accent entirely is not easy, if not impossible. Then again, pronouncing a language accurately with an accent may have its charms.

Can one more discussion add anything new? I believe it can. Ideally, an oral presentation and teaching probes would have complemented the present paper. Being restricted to writing, this presentation therefore comes with unfortunate limitations. In contemplating all the proposals presented in this paper, it may be useful for a teacher to guide students through all the steps.

Keywords

Mandarin, Pronunciation of Mandarin, New Method for Teaching and Learning the Pronunciation of Mandarin

1. Practical Approach and Student Perspective

This article presupposes no knowledge of linguistic theory on the part of the reader. There are no learned footnotes. There are few references. For the same reason, the International Phonetic Alphabet (IPA) is not used—with a couple of exceptions. Pinyin is used exclusively. Beginning students cannot learn how to pronounce Chinese from theoretical linguistic articles.

Then again, having often written on linguistics, I have acquainted myself very generally with current theoretical writings on the phonetics and phonology of Chinese to ascertain that what is written here is not in contradiction with what one finds in academic writings.

Evidently, the theoretical bibliography on the sounds of Chinese is vast. An excellent detailed bibliography is found in [Duanmu \(2007\)](#). This bibliography should lead by indirection to almost everything else that is relevant.

In addition to being practical, the article is written from the perspective of a beginning student facing the practical day-by-day struggle of voicing Chinese as accurately as possible. I believe that it was important to occasionally write about the pronunciation of Chinese from this perspective.

The empirical base of this paper consists of personal observations made while learning Chinese for seven years plus (so far) daily from more than 20 teachers (a debt that I will never be able to repay) and then through self-study, from watching some Chinese programming with a careful ear, and from consulting textbooks of Chinese. However, the line of argument follows its own trajectory.

2. Focus: The 12i

The present paper focuses on a coherent group of sounds. There are three reasons for this selective focus.

First, treating the entire Mandarin sound system would far exceed the space limits imposed on an article, even a very long article.

Second, the coherent group in question combines most of what is considered difficult by westerners in pronouncing Mandarin, aside from the tones.

And third, the present article is about introducing the method. The group in question should amply suffice to illustrate the method. Future applications may

follow.

The group consists of 12 initial consonants plus Pinyin *i* and divides into four subgroups, transcribed in Pinyin as follows:

- 1) *dí, tǐ*;
- 2) *jǐ, qǐ, xǐ*;
- 3) *sǐ, zǐ, cǐ*; and
- 4) *shǐ, zhǐ, chǐ, rǐ*.

They will be called the 12i. The proposed analysis will reveal the perfect harmony and cohesion of the difficult 12i (see especially the layout of **Figure 6** below).

When it comes to difficulty, no consonant probably poses more to students from the west than Pinyin *j*. Contrary to a widely held opinion, Pinyin *j* is not at all like English *j*, even if the two are generally in one another's vicinity. Pinyin *j* will therefore deserve special attention.²

3. The Limits of Imitation and Repetition

Native speakers learn pronunciation through many years of imitating and repeating, mostly by running after their mommy for ten years or so. However, it is difficult if not impossible to recreate such circumstances in a teaching environment. There may therefore be need for something more that may make up for what imitation and repetition cannot do by themselves. But what? The focus will be on conscious analysis (see Section 6). But first, answers are proposed to two questions.

4. How Long to Learn Mandarin?

How long does it take to attain true competence in Mandarin, certainly if learning Chinese is not the only thing that one does at the exclusion of everything else? It seems to me that a good comparison is with how long it takes to play a musical instrument on a professional level or close to it. Having gone through the experience myself, I would say 10 to 12 years of practicing almost every day. It seems to me that the same applies to the study of Chinese, and of most languages for that matter. The precise number of years is difficult to determine. But more than eight and probably less than 15 to 20, somewhere in between. That may seem like bad news. Then again, it is not necessary to learn Chinese up to a certain high level to derive many benefits from it. That is the good news. And there is lots of it.

As anyone knows, so much of the effort consists of learning Chinese characters. E. Wilkinson, in his well-known manual of Chinese history, reports that imperial courtiers thought that it “took ten to twenty years of rote learning.”³

²I began studying Chinese in September 2017. Eleven months later, in the summer of 2018, near the end of studying second-year at the Middlebury College Chinese Summer School in Middlebury, Vermont, USA, I wrote a note in my records: I finally found Pinyin *j*. After 11 months! But I recently established that I had only reached 50% of my goal. I completed the other 50% in late 2023. After six years!

³Wilkinson, 2018: 33b.

And they already spoke Chinese! Misery loves company.

The great scholar emperor Kangxi prided himself on often writing more than 1000 characters a day.⁴ And he was dismayed that, when the time came to render the final verdict on a death sentence, the trial reports were on occasion full of mistakes in matters of life and death.⁵

5. How Important Is Pronunciation?

There is an argument to be made that careful attention to pronunciation deserves precedence over 1) vocabulary, 2) grammar, and 3) writing/calligraphy, the other three of the four main components involved in learning Chinese.⁶

Yet, there is something counterintuitive about prioritizing pronunciation over vocabulary, grammar, and writing, in the following sense: 1) You want to speak Chinese and forgot how to say “kitchen” (vocabulary)? Or, 2) ...and you forget to add 吗 *ma* in a question (grammar)? Or, 3), ...and you cannot remember how to write or read the character 懂 (writing)?

The main point is this. Beginners need to overcome all kinds of deficiencies in terms of vocabulary, grammar, and writing when starting from the beginning. In that regard, it is OK to sometimes forget a vocabulary item, a rule of grammar, or a written character, or get them wrong. But if one pronounces a sound inaccurately, one *always* gets it wrong, not just *sometimes*. The return on investment in working on pronunciation is therefore quite considerable.

The metaphor of music is again useful. One can know a lot about playing an instrument. But if the instrument is out of tune, or the bow does not strike the strings in the right way, or the lips are not properly positioned on the mouthpiece, or there is no breath control in singing (all of which can be metaphorically compared to pronunciation in language), even the simplest melody sounds off. Learning new compositions (cf. vocabulary) is affected. Composing music by the rules (cf. grammar) is affected. And so is sight reading or scoring parts (cf. writing).

6. Conscious Analysis

The present focus is on conscious analysis. The contrast is with learning one's mother tongue. Mother tongues are learned unconsciously through mimicking. But later in life, it is typically not possible to recreate similar circumstances. What is more, the mother tongue serves as an obstacle in learning to pronounce a new language. That is because the reflexes of the speech organs have been set in a certain way. The need is for conscious efforts to overcome those hardened reflexes. Still, second languages are mostly produced with an accent caused by interference from the mother tongue.

⁴Spence, 1974/1988: 59.

⁵Spence, 1974/1988: 32.

⁶That is what professor Chen Wenhui 陈文慧 impressed on us at the outset of first year Chinese at Brown University in September 2017. I fully agreed with this assessment from the beginning, even though I knew zero Chinese at that point.

7. Comparison of What Something Is with What It Is Not as a Basis for Conscious Analysis

The strategy of conscious analysis that is proposed here for the purpose of pronouncing sounds is to compare them with other sounds in Chinese and in one's own language(s).

One does find efforts at comparison in the available materials. But it may be of the type “Pinyin *j* is like English *j*” or “Pinyin *j* is *not* quite like English *j*.” One can find both contradictory positions in textbooks and the internet.

In one's mother tongue, one just speaks and never gives any thought to what one is doing. The problem with trying to pronounce a sound in a second language more or less accurately is of knowing what one needs to do in order to get it right. To what can this be compared?

If one makes one's way through a city on a walk, one relies on situating oneself in relation to landmarks. These landmarks are indispensable. In my experience, the same is true for pronunciation. One needs to determine where one is in relation to outside fixed points of reference. Or otherwise, one is in danger of being lost. Or of being right without knowing why.

I cannot draw. But when I was about 12, a teacher once told me that, when learning to draw with the pencil, one should leave an incorrect line standing to serve as a reference point for the correct line.

Again, the method proposed here is not theoretical in the least. It is meant to be entirely practical. Above all, it may allow students to be close to 100% certain where the sounds are because they can locate them in relation to a fixed reference point that can easily be recognized.

8. The Way of 有无 *yǒu wú* (有无道 *yǒu wú dào*): In the Footsteps of Laozi

To use what something *is not* in order to arrive at what it *is*: that is a profound insight already found in the 道德经 *Dào Dé Jīng*, a work traditionally attributed to the Chinese sage known as 老子 *Lǎozǐ*. One of his material examples is building a house but actually using the empty space inside the house to live. There is also a spiritual dimension to the insight.

In the following verse at the end of Chapter 11 we read:

(故)有之以为利。無之以为用。

“(So) what's there (有) serves as the benefit.

What's not there (无/s 无) serves as the tool (to get the benefit)”

(translation LD, slightly free).

It seemed therefore suitable to name the method 有无(之/的)道 “The Way of 有无 *yǒu wú*,” or “The Way of What's-There vs. What's-Not-There.”

In this digital age following George Boole, the contrast between what something is and what it is not, between 1 and 0, has come to permeate all engagement with reality.

9. Personal Journeys

Every beginning student of Chinese will have to trace his or her own path. Much depends on one's own mother tongue as a point of departure. What follows below is just one path. Perhaps some may find what is said here inspiring to create their own path. The emphasis is on what I have not been able to find among what else is available. Home-baked strategies therefore needed to be developed to get to where I wanted to be. On some points, I found myself in disagreement with what I found elsewhere.

10. A Preliminary Anecdote: The Quest for Pinyin *j*

The eminent American Sinologist J. DeFrancis once referred to Americans as “notoriously incapable of pronouncing Chinese even approximately correctly.”⁷ This statement may be too severe in its generality and in its national focus.

In any event, in my experience, the sound that suffers perhaps the most from mispronunciation is Pinyin *j*. It is probably never pronounced correctly by learners of Chinese. And often not even close.

In the available teaching materials, the notion that thinking of English *j* is useful in trying to pronounce Pinyin *j* is very widespread if not almost universal. However, the fact is that the two are totally different. In English, the association with *j* may well contribute the most to the difficulty in pronouncing Pinyin *j*. In my case, it certainly did.

An analysis shows that both Pinyin *j* and English *j* can each be considered to consist of two elements. The two elements are represented in the International Phonetic Alphabet by two distinct elements. However, neither of the two elements is the same. In that regard, Pinyin *j* and English *j* could not be more different.

Then where does the notion that English *j* and Pinyin *j* are pronounced similarly come from? Two possible reasons are as follows.

The first reason is simply that they are written the same, *j*. Zhou Youguang (1906-2017), the “father of Pinyin,” was mainly inspired by English in creating Pinyin. And he maximally exploited the Latin alphabet by using 25 of the 26 letters of the Latin alphabet, all except *v* (which is sometimes used to generate Pinyin *ü* in apps).

None of the sounds typically pronounced by Latin letters is close to Pinyin *j*, but English *j* is the closest while also being quite different. So it became the obvious pick for Pinyin among the 26 available letters of the Latin alphabet.

But where does that leave German, in which *j* is the same as *y* in English “yes”? The same for Dutch and Italian. In French, *j* is again different and somewhat close to Pinyin *r* and has been used often in the past to transcribe Pinyin *r* in French publications. Then again, the use of English is common across the world. So the speakers of all these other languages should be able to relate to the choice of *j*.

⁷DeFrancis, 1984: 6.

The second reason for which English *j* is thought to be inspiring in pronouncing Pinyin *j* is that the former is still pronounced generally in the same place of the mouth as Pinyin *j*, that is towards the front just behind the teeth.

However, there are two critical differences. First, English *j* is voiced whereas Pinyin *j* is voiceless. Second, the tip of the tongue is roughly above the front teeth with English *j* but behind the bottom teeth with Pinyin *j*. These two differences are typically disregarded when speakers of English pronounce Pinyin *j*.

11. The Components of the Mandarin Sound System

The traditional distinction is between initials and finals. However, I do find the distinction between consonants and vowels useful, and I will use it instead. As is well-known, consonants involve a certain disruption or obstruction in the flow of air through the mouth whereas vowels do not.

Accordingly, the Mandarin sound system consists of four main components, the following:

1) Consonants

Initial: 21 + *y/w* = 23

Final: two (2), namely *n* and *ng* (also diminutive 儿 *er* in Beijing)

In this paper, I will deal only with a coherent group of 12 of the 23 initial consonants to illustrate the proposed method.

The three other components of the Mandarin sound system will not be treated here. They are listed and briefly described below.

2) Vowels

a) six (6) in Pinyin (*i/ü + a/o/e + u*)

b) but more than six (6) in pronunciation

c) combinations thereof

Future topics of interest:

- Full analysis vowel by vowel instead of by “finals,” including:

i) *-iao* = *i + a + u/w*, etc.;

ii) the “four a’s,” as in 看 帮 脸 他, etc.

- Practical definition of

i) the relation between *i/ü* and *y*

ii) the relation between *u* and *w*

- The pronunciation of *-e*.

- The pronunciation of *-eng* and *-ong*.

3) Tones (4 + ∅ [zero tone or fifth tone] = 5)

- On neutral tone, see now Depuydt (2022)

4) (Advanced) Prosody:

- On prosody, see recently, Yang (2016).

12. The 23 Initial Consonants of the Mandarin Sound System and the “12i”

Beginning students of Chinese will surely sense that a lot is coming at them all at

once. One of the complications is pronunciation. In that regard, students may find it useful to know with more precision and clarity the extent and the nature of the problem. Such knowledge allows for maximum focus on what is needed. It eliminates the uncertainty that something else they do not know anything about is still hanging above their head, an uncomfortable and disorienting feeling that is discouraging in the learning process.

It is a fact that most of the difficulties in pronunciation converge in what will be defined as the “12i,” twelve initial consonants followed by Pinyin *i*.

There are 23 initial consonants. Where exactly does “12i” belong in the larger group of 23 initial consonants?

Of the 23 initial consonants, the two expressed by Pinyin *y* and *w* do not seem to present any difficulty. The principal fact to know is that *y* and *w* may be weak or even absent in front of *i* and *u* respectively. So 一 *yī* “one” and 五 *wǔ* “five” can sound close to, or exactly like, just *ī* and just *ǔ*. Accordingly, *y* and *w* might be characterized as unstable, that is, they come and go.

I hope to define more clearly elsewhere the reason why these two, and exactly these two, are “unstable” and two pronunciations are possible. There is a very precise reason for this. It has everything to do with the fact that they are often called “semi-vowels” or “semi-consonants,” that is, something between vowels and consonants.

The remaining 21 “stable” initial consonants can be divided into three groups:

- 1) one group of four (4) initial consonants;
- 2) one group of five (5) initial consonants;
- 3) one group of twelve (12) initial consonants, the “12(i)”.

The first group of four (4) might poetically be called the “No Sweat Quartet,” because their pronunciation requires little explanation. They are *f*, *l*, *m*, and *n*. Perhaps the Mandarin version of these consonants differs ever so slightly from versions in western languages. But if there is any difference, it is very small.

The second group of five (5) initial consonants consists of two sets, one of two (2) consonants and one of three (3) consonants respectively. The two groups are, 1), *g*, *k*, and *h*, and, 2), *b* and *p*. I like to think of the group as “Bringing Up the Front (*b*, *p*) and the Rear (*g*, *k*, *h*),” as will be clear from the chart of the 21 initials in **Figure 1** below.

Twelve (12) initial consonants remain. I might poetically call them the “Magnificent 12.” That is because this group *combines most of what is considered difficult by westerners in pronouncing Mandarin* (aside from the four tones). For reasons that will become clear, it is necessary to consider the 12 consonants in conjunction with Pinyin *i*. So one can call them the “(Magnificent) i12 or 12i or 12(i).”

The result is the following twelve (12) syllables. There are four distinct sections, as follows:

dī, tī;
jī, qī, xī;

si, zi, ci;
shi, zhi, chi, ri.

In the chart in **Figure 1**, the 21 initial consonants are ordered along with a suitable vowel according to where in the mouth the flow of air is obstructed. The left of the chart is the front of the mouth. The right of the chart is the back of the mouth. The degree of obstruction of the air flow generally increases towards the top of the chart and decreases towards the bottom.

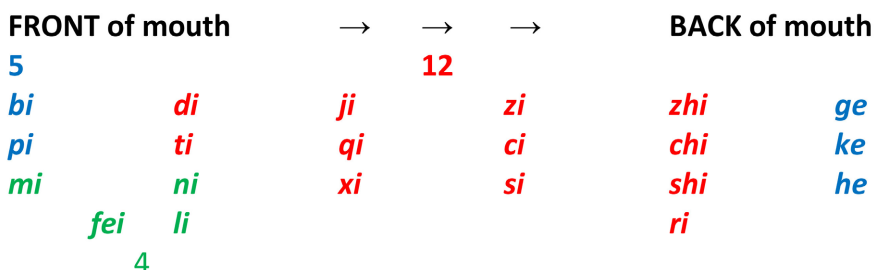


Figure 1. The 21 Initial Consonants (5 + 4 + 12).

The “No Sweat Quartet” is located together in the bottom left corner. If one eliminates those from consideration, then *b*, *p*, *g*, *k*, and *h* are located at the extremities. They are the “Frontier Five”. The “Magnificent 12” are in the middle.

Counting 21 initial consonants is entirely practical. Theoretical approaches may differ. But this is not the place to discuss the otherwise valid difference between phonetics and phonology.

For example, one prominent recent survey of phonology with copious bibliography, Duanmu’s *The Phonology of Standard Chinese* counts only 18 initial consonants.⁸ Pinyin *j*, *q*, and *x* are not counted. That is because they are said to be in “complementary distribution” with other consonants.⁹

Obviously, one cannot teach Chinese pronunciation while telling students that *j*, *q*, and *x* are in some way different from the other 18 initial consonants. The theoretical perspective cannot be the practical perspective. Beginning Chinese is not the right place or time to tell students everything about such theoretical concepts as “complementary distribution.”

I have written about language and linguistics. But when I began learning Chinese later in life, I was quite determined to keep any and all technical training in linguistics out of the picture. The aim was to acquire at least basic or intermediate skills in listening, speaking, reading, and writing. Such skills should in fact form the basis of linguistic inquiry. But often, the basis is a “structural” knowledge of a language.

13. The Two-Dimensional Map of 12i

The 12i combines almost everything considered difficult by westerners in pronouncing Mandarin, aside from the tones.

⁸Duanmu, 2007: 26.

⁹Duanmu, 2007: 31.

The degree of difficulty can already be gleaned from the following fact. Only one (1) of the 12 consonants occurs in English, namely Pinyin *t*. No consonant is probably 100% like any consonant in any other language. But English *t* and Pinyin *t* are close enough to be considered more or less identical.

It follows that proper command of the 12i should lead students a long way towards a good pronunciation of Mandarin.

The specific design of the present article is to study the pronunciation of Chinese from the perspective of 有无 *yǒu wú*.

In Laozi's words cited above, 有之以为利。无之以为用。“What's there serves as the benefit. What's not there serves as the tool (to get the benefit).”

In relation to the pronunciation of Chinese, what is the “benefit” and what is the “tool”?

The “benefit” is the correct pronunciation of a consonant. The “tool” is using the (comparison with the) pronunciation of other consonants to achieve the correct pronunciation of the original consonant.

The pronunciation of the 12i consonants is reviewed one by one below in later paragraphs from a practical perspective. But first, it will be useful to survey them all together *as a set*. Contemplating the set in its entirety may provide beginning students with a better sense of control. Guidance from a teacher may promote this process.

What is the map of 12i? Maps are about space and position and two dimensions. So is the map of 12i.

Which are the two dimensions of the map?

The first dimension is about position. It concerns the position of the tip of the tongue, which moves from the front of the mouth to the back of the mouth.

The second dimension is about space. It concerns the degree of openness of the mouth, or aperture, from wide open to less wide open to closed.

The map of 12i can be seen in **Figure 2**, as extracted from **Figure 1**.

<i>di</i>	<i>ji</i>	<i>zi</i>	<i>zhi</i>
<i>ti</i>	<i>qi</i>	<i>ci</i>	<i>chi</i>
	<i>xi</i>	<i>si</i>	<i>shi</i>
			<i>ri</i>

Figure 2. The 12i (Middle Part of **Figure 1**).

It will be useful to reorder the table partly for present purposes, turning it mostly upside down, as seen in **Figure 3**.

			<i>ri</i>
	<i>xi</i>	<i>si</i>	<i>shi</i>
<i>di</i>	<i>ji</i>	<i>zi</i>	<i>zhi</i>
<i>ti</i>	<i>qi</i>	<i>ci</i>	<i>chi</i>

Figure 3. The 12i (Rearrangement of **Figure 2**).

This map has two dimensions, horizontal and vertical.

Horizontally, the main change involves the exact place where the consonants are pronounced in the mouth. And that concerns mostly the position of the tongue. For the most part, but not exclusively, the tip of the tongue moves from the front to the back of the mouth as one moves from the left side of the table to the right side of the table.

Vertically, the difference is all about how the air flows through the mouth and therefore how wide open the mouth is. The crucial question is: To what degree is the air obstructed as it is propelled from the lungs and expelled from the lips?

There are three possibilities. The flow of air can be obstructed not at all (“unhindered”), it can be obstructed somewhat (“restricted”), or it can be obstructed entirely (“total stop”). In addition, a puff of air may follow the total stop. The stop is then said to be “aspirated.” The aspiration somewhat softens the obstruction.

These basic changes can now be inserted into **Figure 3**, as seen in **Figure 4**.

POSITION OF OF TIP OF TONGUE	→	Front of mouth	to	Back of mouth
AIRFLOW				
↓				
Unhindered				<i>ri</i>
Restricted			<i>xi</i>	<i>shi</i>
Total Stop		<i>di</i>	<i>ji</i>	<i>zhi</i>
Total Stop + Aspiration		<i>ti</i>	<i>qi</i>	<i>chi</i>

Figure 4. The 12i (Expansion of **Figure 3**).

This much for the position of the 12 consonants of 12i in relation to one another, or of each consonant in relation to all the other consonants that the consonant is not. These relations pertain mainly to *change* and *difference*.

But there is also *sameness* in the form of shared features. As consonants change from one position to the next, some features remain the same. These shared features lend a remarkable fully symmetric pattern of coherence in the map of 12i. This coherence is treated in the next section.

14. The Remarkable Coherence of 12i

In the map of 12i in **Figure 4**, changes relating to space occur from left to right and from top to bottom: the tip of the tongue moves from front to back in the mouth and the mouth’s aperture becomes smaller and then closes entirely.

That much for the *changes*. However, in the transition from one consonant to the next, there is also much that *stays the same*. In other words, the two consonants share features. These shared features create coherence between the consonants. A truly remarkable pattern of coherence emerges in the 12i.

This coherence may impress on students that the most difficult component of Chinese pronunciation is also simple in a way. I am not sure that the 12i have ever been presented in this simple way, maybe not even in theoretical work on linguistics.

What are the shared features that produce this remarkable coherence? The features derive entirely from the following fact: some of the 12i are simple consonants and some are compound consonants.

15. Simple Consonants and Compound Consonants

The 12i consists of both simple consonants and compound consonants.

There are five (5) simple consonants and there are seven (7) compound consonants.

The five (5) simple consonants are Pinyin *d*, *x*, *s*, *sh*, and *r*. The seven (7) compound consonants are Pinyin *t*, *j*, *z*, *zh*, *q*, *c*, and *ch*.

The five (5) simple consonants consist of one (1) component. The seven (7) compound consonants consist of two (2) or three (3) components.

Four (4) compound consonants contain two (2) components, namely Pinyin *t*, *j*, *z*, and *zh*. Three (3) compound consonants contain three (3) components, namely Pinyin *d*, *c*, and *ch*.

16. The Single Principle Lying at the Origin of the Remarkable Coherence of the 12i

The remarkable coherence of the 12i derives entirely from the following fundamental fact. Again, I am not sure that this fact has ever been singled out in all its simplicity, maybe not even in theoretical linguistics, but I would need to check.

All five (5) simple consonants—except Pinyin r—serve 1), not only as consonants by themselves alone, but also 2) as components of the seven (7) compound consonants. Conversely, all compound consonants contain one (1) simple consonant and one (1) or two (2) expansions, which are either themselves a simple consonant or aspiration (see below).

The unique status of Pinyin *r* will be defined below by means of the 有无 *yǒu wú* method.

It follows that the seven (7) compound consonants consist of one (1) of four (4) simple consonants plus one (1) or two (2) additional components.

These additional components can be viewed as expansions of the simple consonant. There are two (2) expansions of the simple consonants, only two (2), as follows:

- 1) *d*, which is itself a simple consonant;
- 2) aspiration, which is more or less like adding English *h* and may be represented as raised *h* (^h). Obviously, this is different from Pinyin *h*, which is expressed with friction in the throat, a sound more or less found in Arabic, Dutch, and German and other languages.

It appears that *d* is both a simple consonant and an expansion. The pronun-

ciation of *d* differs depending on what follows if anything. The position of the tongue changes in function of what follows. But it seems better to think of *d* as a unit and write it in this way.

More on the different pronunciations of *d* follows below. As regards the differences, linguists speak of allophonic behavior and the like. But again, this article is not the place for the theoretical linguistic perspective.

There are three (3) cases of expansion of the five (5) simple consonants, for a total of seven (7) instances.

First, ^h may expand a simple consonant alone by itself at the end. There is one (1) instance of this. It is Pinyin *t*.

Second, *d* may expand a simple consonant alone by itself at the front. There are three (3) instances of this.

And third, *d* and ^h may together expand a consonant at both the front and the end. There are three (3) instances of this.

The total is seven (7) instances for three (3) cases.

The seven (7) compound consonants may therefore be rewritten as expansions of four (4) simple consonants as follows:

<i>t</i>	=	<i>dth</i>	(one expansion, ^h)
<i>j</i>	=	<i>d+x</i>	(one expansion, <i>d</i>)
<i>z</i>	=	<i>d+s</i>	(one expansion, <i>d</i>)
<i>zh</i>	=	<i>d+sh</i>	(one expansion, <i>d</i>)
<i>q</i>	=	<i>d+xth</i>	(two expansions, <i>d</i> and ^h)
<i>c</i>	=	<i>d+sth</i>	(two expansions, <i>d</i> and ^h)
<i>ch</i>	=	<i>d+shth</i>	(two expansions, <i>d</i> and ^h)

17. The Perfect Harmony and Cohesion of the 12i

The relation between simple consonants and compound consonants in the 12i can now be summarized by modifying [Figure 4](#) into [Figure 5](#).

	POSITION OF →		Front of mouth	to	Back of mouth
	OF TIP OF TONGUE				
AIRFLOW					
↓					
Unhindered					<i>ri</i>
Friction			<i>xi</i>		<i>shi</i>
Total Stop		<i>di</i>	<i>d+xi</i>	<i>d+si</i>	<i>d+shi</i>
Total Stop + Aspiration		<i>dthi</i>	<i>d+xthi</i>	<i>d+sthi</i>	<i>d+shthi</i>

Figure 5. The 12i (Adaptation of [Figure 4](#)).

But it is possible to modify [Figure 5](#) even further to better bring out the perfect harmony and coherence of the 12i. [Figure 5](#) can also be presented 1) with-

out the Pinyin vowel *i* and 2) using raised *d* (^d) in compound sounds, as in the following figure: (Figure 6)

	POSITION OF → OF TIP OF TONGUE			
	Front of mouth		to	Back of mouth
AIRFLOW				
↓				
Unhindered				<i>r</i>
Friction		<i>x</i>	<i>s</i>	<i>sh</i>
Total Stop	<i>d</i>	^d <i>x</i>	^d <i>s</i>	^d <i>sh</i>
Total Stop + Aspiration	<i>d^h</i>	^d <i>x^h</i>	^d <i>s^h</i>	^d <i>sh^h</i>

Figure 6. The 12i minus *i* (Adaptation of Figure 5).

One could also write as in the following figure: (Figure 7)

Unhindered				<i>r</i>
Friction		<i>x</i>	<i>s</i>	<i>sh</i>
Total Stop	<i>d</i>	<i>d-x</i>	<i>d-s</i>	<i>d-sh</i>
Total Stop + Aspiration	<i>d^h</i>	<i>d-x^h</i>	<i>d-s^h</i>	<i>d-sh^h</i>

Figure 7. The 12i minus *i* (Adaptation of Figure 5).

When one just looks at Figure 6 in a superficial way, without even knowing any Chinese at all, it seems easy to discern the perfect harmony and cohesion of the 12(i), the set of sounds that contains almost everything that is difficult to pronounce for westerners. I do not think that this harmony and cohesion of 12i has been made fully manifest in this way before, even if it may be implied somewhere in theoretical work.

Interestingly, in the list of initial consonants in R. Dawson's popular *A New Introduction to Classical Chinese*, *sh* is absent.¹⁰ The total of initial consonants is therefore 20 instead of 21 (excluding *w* and *y*) and the perfect harmony of Figure 6 cannot be maintained.

As one can see in Figure 6,

- 1) the absence of *d* in horizontal line 2,
- 2) the presence of *d* in horizontal line 3, and
- 3) the presence of both *d* (^d) and aspiration (^h) in horizontal line 4

are spreading rightward from column 1 to the other 3 columns *whereas all else stays the same* from top to bottom.

The lone exception seems to be Pinyin *r*. How does it fit in? In Figure 6, it is closest to Pinyin *sh*. Since all the consonants in Figure 6 are explained in relation to what is closest to them in the spirit of the 有无 *yǒu wú* approach, one wonders whether what Pinyin *r* is can be explained with the help of what it is not, namely Pinyin *sh*. In fact, I believe that it can. I am not sure that it has ever

¹⁰Dawson, 1984: 8.

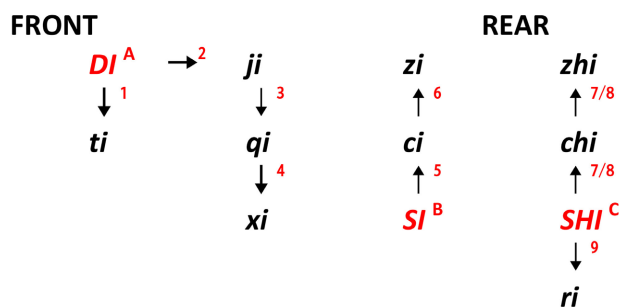
been done. More on this follows below.

Another way of looking at the cohesion of the 12(i) is as follows. Evidently, there are 12 consonants. Each of these 12 consonants consists of one (1), two (2), or three (3) components. However, there are only six (6!) *different* components in total. They are as follows: *r*, *x*, *s*, *sh*, *d* (or *d'*), and *b*. Everything in the 12i is combinations of these six (6) components.

In **Figure 5** and **Figure 6**, there are twenty-two (22) occurrences of the six (6) components in the twelve (12) consonants of the 12i. Five (5) consonants exhibit one (1) component; four (4) consonants exhibit two (2) components; and three (3) consonants exhibit three (3) components. The simple math is as follows: $(5 \times 1) + (4 \times 2) + (3 \times 3) = 5 + 8 + 9 = 22$.

It is now upon us to examine the pronunciation of the members of 12i more in detail.

Where does it go from here? In the 2021 NECLTA paper from which this article derives, I attempted to move from one sound to the next. Towards that end, I presented the following new path.



As one can see, there are three (3) points of departure, marked A, B, and C, and there are nine (9) steps.

This path will *not* be followed in the present paper. Much reflection has led me to what I believe to be a better understanding of the matter. A different approach is therefore much more suitable.

One example: Pinyin *x*. In the earlier approach, it is derived from Pinyin *q* in Step 4 (see above). The order is *q* first and *x* second. But this order needs to be reversed. The reason is not only that *x* is a simple consonant and *q* is a compound consonant. What is more, *x* is a component of *q*.

So where to begin? In the search for a suitable beginning, it may be desirable to heed the following principle.

18. Moving from the Simple to the Compound

The design of the present paper is to facilitate the pronunciation of 12i by means of the 有无 *yǒu wú* method. The essence of this method is to draw as much benefit as possible about what a consonant *is* from what it *is not*. The method is designed to be entirely practical and accessible to all students of Chinese.

But in which order should the twelve (12) consonants be discussed?

The method proposed in this paper is in many ways an effort to start from scratch. Descartes (1596-1650) was very much concerned with thinking for himself and presupposing nothing. To that end, he formulated four principles in his *Discourse on Method* aimed at “conducting reason well and seeking truth in the sciences.”

His third principle is to reason from the simple to the complex. In that regard, it has been established above that five (5) consonants of the 12i are simple and seven (7) are compound and the compound ones contain four (4) of the simple ones.

It follows that it is advised to begin with the five (5) simple consonants: *d*, *x*, *s*, *sh*, and *r*.

Pinyin *r* is a special case that will require individual attention, as has already been noted above. I believe that, in the spirit of the 有无 *yǒu wú* approach, the pronunciation of *r* can be explained in light of what it *is not*. And in this case, that which it *is not* is Pinyin *sh*, which is closest to Pinyin *r*. Details follow below.

That leaves four (4) options as a starting point: *d*, *x*, *s*, and *sh*. It would probably be possible to begin with any of these four. More on the choice made in the present article follows below.

As regards the four possible simple consonants that can serve as a beginning for this paper’s argument, there is a difference between *d* and the three others, *x*, *s*, and *sh*, as follows.

As is obvious from **Figure 6**, Pinyin *d* can be combined with the three others. But each of the three others can only be combined with *d*. In fact, *d* is part of all the seven (7) compound consonants in the 12i.

The more widespread presence of *d* may make it look like a desirable candidate to serve as a starting point. But in the end, it makes no difference whether one begins with Pinyin *d*, *x*, *s*, or *sh*. The order that will be followed in the present paper is as follows: *s*, *sh*, *x*, and *d*. It is then possible to move to the compound consonants and finally to Pinyin *r*.

But first, it is next in order to describe two powerful tools involving the 有无 *yǒu wú* method.

19. Two Powerful Didactic Tools as Applications of the 有无 *yǒu wú* Method

- 1) The Unity of Pinyin *i*
- 2) Consonant Flipping.

The first tool associated with the 有无 *yǒu wú* Method is called here “The Unity of Pinyin *i*.” Careful consideration of the pronunciation of Pinyin *i* can do much for a true and deep understanding of the proper pronunciation of the 12i, and especially of the six consonants Pinyin *z*, *c*, *s*, *zh*, *ch*, and *sh*. The unity of Pinyin *i* is described in more detail in Section 20. Its power is then brought to bear in later sections. Importantly, this first tool is devoid of all theoretical sophistication. It is designed to be accessible to all beginning learners of Chinese.

The second application will be called “The Efficacy of Cross Language Consonant Flipping,” that is, replacing English consonants in English words with their supposed “close equivalents” in Chinese (and *vice versa* in Chinese words for Chinese students of English).

In the spirit of the 有无道 *yǒu wú dào*, consonant flipping has this powerful—and hitherto never exploited, as far as I know—way of confronting western students of Chinese (and potentially also Chinese students of English and of other western languages) with what the pronunciation of a consonant *is not* to produce a better understanding of what it *is*.

The effect of the flipping can be quite startling. Just take pinyin *j*. Pronouncing English *j* in English words with Pinyin *j* has an almost comical effect. It does much to evidence how different the two are and, it is hoped, also does much to give students a better sense of how to pronounce Pinyin *j*.

Now on to the first special application of the 有无道 *yǒu wú dào*, the unity and the power of Pinyin *i*.

20. The Unity of *i* in 12i

This paper is about the twelve (12) consonants most difficult to pronounce in Chinese for Westerners. Each of these consonants can be followed by Pinyin *i*. The result is the twelve (12) syllables listed in **Figures 2-7**.

Pinyin *i* is a vowel. The present paper is about consonants. It may be somewhat counterintuitive to assign such a critical role to a vowel in a paper that deals with consonants. And yet, in my personal experience, Pinyin *i* has been crucial in understanding the nature of the consonants.

What is more, it has not only been a guide in pronouncing the consonants of 12i when followed by Pinyin *i*. It has also been a guide in pronouncing the same consonants when they are followed by other vowels. For example, the pronunciation of *s* in *sí* has provided me with an understanding of where Pinyin *s* is located in syllables such as *sa*, *se*, *so*, and *su*. It is not the same location as in English and many other languages. Close. But not quite. So where exactly? Pinyin *i* can provide critical support.

Textbooks generally present Pinyin *s* as being the same as *s* in English and other languages. But it is not quite. In other traditional transcription systems, Pinyin *s* is often represented as *ss* in front of Pinyin *i*. The doubling of *s* in *ss* quite appropriately indicates a more emphatic pronunciation than *s* in English and other languages.

But what past transcriptions typically do not reflect (as far as I know) is that—as far as my ear can discern—Pinyin *s* sounds the same in front of other vowels. I cannot readily find this view reflected in textbooks. But it may be there somewhere. And one assumes it probably is in theoretical linguistic works. Again, beginning students cannot be asked to rely on theoretical linguistics to pronounce Pinyin *s* correctly. There has to be an easily accessible way. Pinyin *i* provides this access. Details follow below.

But first about the unity of *i* mentioned in the title of this paragraph. What is meant by unity? After all, it is obvious that there are three (3) different pronunciations of Pinyin *i*. In other words, at the surface, there is no unity as far as pronunciation is concerned. So what kind of unity is meant?¹¹

The first of the three pronunciations occurs in the five (5) syllables *di*, *ti*, *ji*, *qi*, and *xi* found in columns 1 and 2 in **Figures 2-7**.

The second pronunciation occurs in the three (3) syllables *zi*, *ci*, and *si*, found in column 3 in the same Figures.

The third pronunciation occurs in the four (4) syllables *zhi*, *chi*, *shi*, and *ri*, found in column 4 in the same Figures.

Only the first pronunciation is found in English. The other two are not. In the first pronunciation, Pinyin *i* is pronounced as *ee* in English “bee.” English *ee* is written *i* in most languages. But not in English. Nor in standard Dutch/Flemish, the present writer’s mother tongue, in which it is written *ie*.

Incidentally, according to one analysis found in textbooks, there is no vowel at all in *zi*, *ci*, and *si*—as if they were just *z*, *c*, or *s*. This analysis will not be accepted here. Clearly, a vowel can be heard following *z*, *c*, and *s* in *zi*, *ci*, and *si*. It may not be a very “colorful” vowel. But it is a vowel nevertheless. How the exact color of this vowel is in fact obtained is detailed below.

Why are there three (3) different pronunciations of Pinyin *i*? This has everything to do with the consonant that precedes. In other words, difference in the preceding consonant can produce difference in the following vowel.

But how can one speak of the unity of Pinyin *i*, as was anticipated above?

It is proposed here that, in the second and third groups, as in *si* and *shi*, the pronunciation of Pinyin *i* is just as much *ee* as in *bee*!

This approach may at first sight sound contradictory because neither *i* in *si* nor *i* in *zhi* sound like *ee* in English *bee*. Then how can *i* in *si* and *zhi* be said to involve the same pronunciation as *ee* in English *bee*?

The reason is that, in pronouncing *si* (and its compounds *zi* and *ci*) as well as *shi* (and its compounds *zhi* and *chi*) the mouth engages in *exactly the same effort* as when it tries to produce English *ee*.

However, the preceding consonant “bends” the color of the vowel into something different from English *ee*. That is, I believe, the true nature of Pinyin *i* in *si* (and *zi* and *ci*) as well as in *shi* (and *zhi* and *chi*).

In sum, *trying* to pronounce English *ee* is exactly what one needs to do to pronounce *i* in *si* and *zhi* correctly.

It also follows that, in *si* and *zhi*, *the lips are spread* in exactly the same way as they are in Pinyin (*x*)*i* and English *ee*. Westerners very often round the lips. And so sometimes do, it seems to me, Chinese who learn Beijing Mandarin as a second Chinese dialect or language. More details on this matter are desirable.

To repeat, there is no rounding of lips at all in any pronunciation of Pinyin *i*. The lips are always in the same position, that of *ee* in English.

¹¹Linguists would call the three different pronunciations of *i* allophones of one and the same phoneme.

It is in a sense a fortunate circumstance that the creator(s) of Pinyin decided to use *i* for all three pronunciations. Because, on a deeper level, they are all one. They form a unity.

In my personal experience, focusing on Pinyin *i* has been critical in confidently addressing two principal difficulties of pronunciation involving the 12i. And the 有无 *yǒu wú* approach has played an essential role in the process.

What are the two principal difficulties?

The first difficulty is that, in two of its three (3) possible pronunciations, Pinyin *i* is not pronounced like English *ee* (written *i* in many other languages). This is the case in the seven (7) syllables Pinyin *zi, ci, si, zhi, chi, shi,* and *ri*.

The second difficulty is the pronunciation of the consonants in *ji, qi, xi, zhi, chi, si,* and *ri*. None of these consonants is pronounced quite like any consonant in any other language that I am aware of.

In wondering about how a Pinyin consonant is pronounced, the natural instinct and in fact common practice is to ask: What is this like? Can I compare it to something that I already know as a point of reference? A prominent example of this approach is the common advice to think of English *j* when pronouncing Pinyin *j*.

In other words, students of Chinese typically locate the beginning of their efforts to obtain 有 *yǒu* “what’s there” at 有 *yǒu* itself. The search is for something that is the same as the 12i of Chinese. However, the 12i of Chinese are so distinctly different.

In the course of studying Chinese, I gradually came to the realization that starting from 有 *yǒu* “what’s there” may not be the way to go.

Instead, this paper exploits 无 *wú* “what’s not there” as the means employed to obtain 有 *yǒu* “what’s there” as the benefit.

Details and applications follow below. But first, it is next in order to describe the second special application of the 有无道 *yǒu wú dào* anticipated above.

21. Consonant Flipping: Definition and Example

Consonant flipping is *to exchange a consonant in an English word with its supposed (close) equivalent in Chinese, or vice versa.*

Take English *j* and its supposed close equivalent Pinyin *j*. Consonant flipping would be either,

- 1) to replace English *j* in e.g. English “jeep” with Pinyin *j*, or,
- 2) to replace Pinyin *j* twice in e.g. Chinese 经济 *jīng jì* “economy” with English *j*.

There is little need to illustrate case 2). It is heard in classrooms wherever Chinese is taught in the English-speaking world, with varying results.

But I am not aware that the opposite is tried out anywhere in the teaching and learning of Chinese.

Take, for example, the English word “jeep.” It is possible to replace English *j* with Pinyin *j* in this word, as in “I have a green *jeep*. Is your *jeep* also green? No,

my jeep is red.” If Pinyin *j* is pronounced correctly in the Chinese way, the effect is almost comical.

But what does this have to do with the 有无道 *yǒu wú dào* method? Let us recall that the aim of the 有无道 *yǒu wú dào* is to help beginning students of Chinese to pronounce Chinese sounds with the help of what these sounds are *not* (无 *wú*).

In the present specific example, the aim is to establish the pronunciation of Pinyin *j* with the help of what it is not. But there is a problem. English *j* is generally thought of as something that Pinyin *j* is and not as what it is *not*.

It is true that textbooks generally point out that Pinyin *j* is not quite like English *j*. Consider, for example, the following definition of Pinyin *j* in the popular textbook *Integrated Chinese*¹².

“To make the *j* sound, first raise the flat center of the tongue to the roof of the mouth and position the tip of the tongue against the back of the bottom teeth, and then loosen the tongue and let the air squeeze out through the channel thus made. It is unaspirated and the vocal cords do not vibrate. Chinese *j* is similar to the English *j* as in “jeep,” but it is unvoiced and articulate with the tip of the tongue resting behind the lower incisors. You also need to pull the corners of your mouth straight back to pronounce *j*.”

This is accurate. But it is a lot for a second-week student of Chinese to absorb, in addition to the grammar, the vocabulary, and the writing. In fact, anecdotal evidence suggests that students of Chinese are often advised to pronounce Pinyin *j* like English *j*. And indeed, that is how Pinyin *j* is almost always pronounced by anyone studying Chinese in an English-speaking environment.

By contrast, it is assumed here that Pinyin *j* is *not* English *j*. But because almost everyone in the West equates the two, it is necessary to show in a way that is easily accessible to beginning students that the two are *not* the same.

The consonant flipping illustrated above therefore serves two purposes.

First, it confirms that Pinyin *j* is *not* English *j*. That is in a way an expression of 无 *wú*.

Second, once it has been established that English *j* is a 无 *wú* of Pinyin *j*, the need is to establish how the two are different. The need is to bring out the contrast between the two.

In that regard, one expects to hear English *j* in English “jeep.” When that expectation is not met and Pinyin *j* is heard, the contrast is quite sharp. The reason is that the expectation is overturned.

If one just compares Pinyin *j* and English *j*, there is the temptation to find what is similar between the two. And there are indeed similarities. However, replacing one by the other sharply focuses the mind on the differences. Any expectation of hearing something similar is obliterated by the sharp contrast between the two. The reason is that flipping the consonants brings out the contrast in the sense of not hearing what one expects to hear.

¹²Liu et al., 2015: 3.

The contrast just described provides some guidance as to how to pronounce Pinyin *j* differently from English *j*.

But how is Pinyin *j* exactly pronounced? That will require another application of the 有无道 *yǒu wú dào* method. This application is described further below.

22. Consonant Flipping and the 12i

What follows is a list of applications of consonant flipping as it relates to the 12i. Consider the table below. There are four columns. Column 1 is the Pinyin sound belonging to 12i. Column 2 is the English sound that commonly serves as inspiration for the pronunciation of the Pinyin sound. Column 3 is an English word that contains the English sound in question. In Column 4, the English sound is replaced by the Pinyin sound.

1 Pinyin Symbol	2 Associated English Sound	3 English Word Containing the Sound	4 Same Word with Consonant Flipping
<i>d</i>	d	deer	<i>deer</i>
<i>t</i>	t	tea	[English <i>t</i> same as Pinyin <i>t</i>]
<i>j</i>	j	jeep	<i>jeep</i>
<i>q</i>	ch	cheese	<i>qeese</i>
<i>x</i>	sh	sheep	<i>xeeep</i>
<i>z</i>	[no direct equivalent; part two of Pinyin <i>z</i> is Pinyin <i>z</i>]		
<i>c</i>	[no direct equivalent; just add aspiration to Pinyin <i>s</i>]		
<i>s</i>	s	song	<i>song</i>
<i>zh</i>	j	jerk	<i>jeerk</i>
<i>ch</i>	ch	choose	<i>choose</i>
<i>sh</i>	sh	show	<i>show</i>
<i>r</i>	r	road	<i>road</i>

23. The Pronunciation of Pinyin *s(i)* and *sh(i)*

(To be continued.)



Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Dawson, R. (1984). *A New Introduction to Classical Chinese* (2nd ed.). Clarendon Press.
- DeFrancis, J. (1984). *The Chinese Language: Fact and Fantasy*. University of Hawaii. <https://doi.org/10.1515/9780824840303>
- Depuydt, L. (2022). Neutral Tone in Chinese: A Comprehensive Theory Bridging East and West. *Open Journal of Modern Linguistics*, 12, 768-789. <https://doi.org/10.4236/ojml.2022.126053>
- Duanmu, S. (2007). *The Phonology of Standard Chinese* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780199215782.001.0001>
- Liu, Y. et al. (2015). *Integrated Chinese 中文听说读写, Textbook: Simplified Characters* (3rd ed.). Cheng & Tsui Company.
- Spence, J. D. (1974/1988). *Emperor of China: Self-Portrait of K'ang-hsi*. Vintage Books, A Division of Random House. (Originally Alfred Knopf (1974))
- Wilkinson, E. (2018). *Chinese History: A Manual* (5th ed.). Self-published by author.
- Yang, C. (2016). *The Acquisition of L2 Mandarin Prosody: From Experimental Studies to Pedagogical Practice*. John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.1>