

# Teaching Chinese Phonetics: A contrastive study of English and Chinese language acquisition and language instruction

Sanyi Xu

I was invited by Professor Shiao-ling Yu of the Department of Foreign Languages and Literature at Oregon State University to observe the teaching of both Chinese and Japanese on that and other campuses in the western United States during most of 2012. Here is an overview of my experiences and conclusions.

## 1. The research process

Although my research focused on both Chinese and Japanese phonetic acquisition by Americans, this report mainly concerns my observations in the Chinese arena.

### 1.1 The main agenda

- (1) In the spring term of 2012 I sat in on the following Chinese classes: Level 113, taught by Mark Ren, and Level 213, taught by Chenhui Ho. I also observed these Japanese classes: Level 113 by Luke Yamaguchi and Level 213 by Fumiko Nakamura. In the autumn term, I observed Mark Ren's Level 111 and Shiao-ling Yu's Level 311 Chinese as well as Setsuko Nakajima's Level 111 Japanese classes.
- (2) In May 2012 I attended two types of Chinese classes at Portland State University: first year class taught by Meiru Liu conducted under the auspices of the Confucius Institute at PSU, and the third year class taught by Jonathan Pease via the Department of World Languages and Literatures in the Faculty of Liberal Arts & Sciences. I was also allowed to sit in on the first year Japanese classes of Emiko Konomi, Kenta Nishimaki and Nathaniel Bond as well as the second year classes taught by Suwako Watanabe.
- (3) In January and February of 2013 at the University of Oregon, I attended the first level Chinese courses taught by Fengjun Mao and Denise Huang Gigliotti at the Department of East Asian Languages and Literature. Moreover, I observed the master's course for future teachers of Chinese by Zhuo Jing-Schmidt, as well as the three first-year Japanese classes taught by Reiko Hashimoto, and the second year Japanese class taught by Naoko Nakadate. I also interviewed Kaori Idemaru, a professor of linguistics, about

pronunciation instruction in her upper level Japanese classes.

- (4) In June 2012, I attended an Alliance for Language Learning and Educational Exchange (ALLEX) training conference at PSU for American and international students intending to teach Chinese and Japanese in their respective countries. One workshop there was conducted by Haohsiang Liao, who demonstrated how to teach a beginning-level Chinese class to a small class of American students while the audience watched. I also observed a demonstration Japanese lesson by Masayuki Itomitsu. I gave an overview to all the ALLEX students about the difficulties speakers of Japanese or Chinese experience when faced with mastering the complexities of English pronunciation, and the similar frustrations experienced by American speakers of English as they navigate the pitfalls while attempting to master Japanese or Chinese pronunciation.
- (5) In November 2012, I went to Redwood City, California to participate in the 3rd Annual Chinese Language Education Forum organized by the California Chinese Education Research Center. I attended presentations by researchers from Asia, Europe, and the Americas on topics such as the historical development of Chinese grammar, pronunciation and vocabulary and how they impact Chinese instruction around the world.
- (6) In October 2012 at a Confucius Institute training event held at PSU, I attended several lectures about innovations in new educational Chinese texts by the editors of several famous Chinese textbooks. Key features of those textbooks were described by some famous Chinese teachers in the U.S.A. Following those presentations, I had the chance to speak with the many of the participants extensively.

## **1.2 research method**

- (1) Primarily, I sought to be present in language classes during instruction and observe specific teaching methods while taking notes.
- (2) To some degree I was also a “visiting expert assistant” who helped students as directed by their teacher – answering their questions in class, conversing with them using the targeted language components they were learning, and moderating small group practice sessions, and finally helping them grapple with the errors they encountered.
- (3) After class, I attended the students’ learning groups, interacting with them, conducting individual tutorials, guiding students to help them recognize their mistake patterns, and also learning from them.
- (4) With the consent of the college, the teachers and the students themselves, I recorded some classroom sessions, collecting relevant data.
- (5) I also obtained permission from some students to record their voices individually or in group outside of class.

- (6) When conducting my research away from the OSU campus, I obtained permission to attend classes and to record student activities in class.
- (7) I also interviewed several teachers, collecting relevant information about their teaching behaviors and beliefs.
- (8) I devoted a lot of time to studying English during this sabbatical in order to assess my own experiences of trying to master its phonological system and to improve the reliability of my speech analyses.
- (9) I often sought out relevant information published in English, Japanese and Chinese, using the OSU university library resources or via the Internet, in an effort to absorb a variety of different information sources.
- (10) While this report is a brief analysis of the results of my audio recordings and class notes, I will continue to evaluate the body of information that I have collected over this year.

## **2. Summary of relevant research achievements**

### **2.1 Research overview**

Using Oregon State University as my primary source, I visited Portland State University and the University of Oregon to listen and learn from over 20 professors. I also interviewed some of them, participated in language teaching conferences, and observed their classes, and exchanged ideas about the difficulties of teaching Chinese or Japanese pronunciation.

As mentioned above, my research focuses on Chinese and Japanese phonetic acquisition by students in the United States. In this paper uses the terms “English” and “Chinese” refer to American English and Mandarin Chinese respectively.

### **2.2 Research production**

This research focuses on pronunciation acquisition and its implementation. This usually involves nothing more than the sound of vowels, consonants, tones and intonations. However in this report I also address syllabification.

#### **2.2.1 Syllables in English and Chinese**

The “sound” of a language is in its syllables, sentences and paragraphs. But initially it is in the syllable. So whether it is sound teaching or pronunciation, research should focus firstly on syllables.

To discuss syllables, we need to define them. The *Oxford Modern English-English Dictionary* defines a syllable as:

Any of the units into which a word is divided, containing a vowel sound and usually one or more consonants.

In simple terms, a syllable is a structure containing one or more vowel sounds, and usually one or more consonants, with the vowel as the center. But differences in the construction of Chinese and English syllables are a source of initial pronunciation difficulties among many North American Chinese learners. The Chinese syllable structure is a very “closed”, but the English syllable is relatively “loose”.

In Chinese syllables, consonants are used only one at a time—one at the onset of a syllable, or one at the coda (n/ng) of a syllable, or one at each of the onset and coda of the syllable. Unlike English, Chinese lacks consonant clusters such as “truths”.

The sound of the vowel in each Chinese syllable is fixed, and doesn’t change even though other syllables may be added to it; whereas in English, the sound of the vowel in the syllable often changes when a suffix change is made (“economy” [ɪkɑːnəmi] – “economic” [ekəˈnɑːmɪk]) or an additional syllable is added to the word (“geography” [dʒiːɑːgrəfi] – “geographical” [dʒiːəˈgræfɪk]).

In Chinese, the position of the consonant sounds in each syllable is fixed. Sounds cannot “migrate” or be “borrowed” when pronounced and / is used at the beginning or end of adjacent syllables. In English however, the consonant sounds shift depending on the words being used (the cats are lazy → the ca tsar lazy).

The Chinese syllable structure is closed—a syllable, or a series of syllables is pronounced using any of the four tones of Chinese, and words maintain their stress / accent patterns. However, in English this is not the case: related words such as “e-co-no-my” and “e-co-no-mic” may have differing syllables and be pronounced with differing stress patterns, such as “em-pha-sis” vs. “empha-tic.”

English has a “loose” syllable structure in which individual word stress patterns vary according to their grammatical environment. For instance, when nouns become adjectives, as in “Oregon”, “state”, “university” → “Oregon State University” a new primary stress pattern emerges. This would not happen in Chinese.

## 2.2.2 The vowels of Chinese and English

Now let us consider how the vowels of Chinese and English differ. Since this study is concerned only with pronunciation acquisition, I shall expound on those elements that appear to cause trouble for students.

### 2.2.2.1 The monophthongs of Chinese and English

It is generally accepted that there are 7 monophthongs in Chinese: /a/, /o/, /ɤ(e)/, /i/, /u/, /y(ü)/, /ɤ(er)/, as opposed to the 12 monophthongs of English: /i:/, /ɪ/, /e/, /æ/, /ɜ:/, /ə/, /ʌ/, /ɑ:/, /ɒ/, /ɔ:/, /ʊ/, /u:/.

I propose that the 4 syllables: en[en], eng[ɛŋ], and an[æn], ang[ɑŋ], should be credited to the aforementioned ranks of Chinese monophthongs, since the vowels in these 4 monophthong syllables differ from other syllables. Additionally, O is not a Chinese vowel, thus Chinese monophthongs are 10 in number. Just as there is no vowel O in Chinese, there is no y(ü) in English, although for several other Chinese vowels there have similarly parallel sounds in English.

Because English syllables often differ from Chinese syllables, the vowels function differently in many syllables. Therefore, when teaching it is necessary to correctly explain Chinese vowel sounds. The best way to do this, of course, is to pronounce the vowels in their syllable settings. For the A found in the English “hot” [hɑ:t] and “bar” [bɑ:(r)], the vowels are relatively close to the Chinese A, but neither of them is an exact replica of the Chinese vowel A. In the former, the lips do not relax; in the latter, because of the influence of the coda consonant R basic attack, the lips became slightly rounded. As for the vowel E, in English the pronunciation [ə] of “about” [əˈbaʊt] is similar to the Chinese E, but the Chinese tongue position is higher, and the lip opening larger. For the I in Chinese, the tongue position is higher than its two English compatriots found in “pig” and “pea”, the friction is stronger, and additionally, the I in Chinese has two (non-English) allophones. As for the Chinese U, because the lips are long rounded, the sound is very different from that of the two similar sounds in English, (“put” or “boot”) since the lips are extended so far forward. For example, in the sound “du” of “*shoudu*”, if the lips are not extended far enough, an American student’s pronunciation will sound wrong. The monophthongs of Chinese and English are contrasted in the following table.

**Table 1. The Monophthongs of Chinese and English**

	Chinese			English					
	Front	Central	Back	Front		Central		Back	
				Short	Long	Short	Long	Short	Long
Close	i y ɿ ʅ		u	ɪ	i:			ʊ	u:
Close mid	e	ə ɤ	o	e					
Open mid		ə				ə ʌ	ɜ:	ɒ	ɔ:
Open	æ	A	ɑ	æ					ɑ:

\* According to Celce-Murcia, Briton & Goodwin (1996), long sounds such as [i:], [u:] should be written with two symbols like [iy], [uw]. However I have continued to use the notations of the long vowels since they do not constitute a consideration that affects the conclusions put forth in this study.

#### 2.2.2.2 Diphthongs in Chinese and English

Both Chinese and English have diphthongs, but the actual meaning of “diphthongs” differs between these

languages. The English diphthongs are the sounds of two vowels produced in sequence, i.e. one after the other. They take more time to pronounce than single vowels. By contrast, Chinese diphthongs consist of two vowels that fuse temporally, producing a single pure vowel sound that is short in duration. In Chinese, there are two kinds of diphthongs: closing diphthongs: ao[ɑu], ai[æi], ei[eɪ], ou[əu] (the mouth closes when finishing the sound); and opening diphthongs: ia[iA], ie[iɛ], ua[uA], uo[uɤ], üe[yɛ] (the mouth opens when finishing the sound). In the Chinese opening diphthongs, in fact, the first element of the diphthong basically degenerates into a semivowel, because of the movement from the initial position of the lips as they shift to the open position to pronounce the final vowel sound. Chinese closing diphthongs are the fusion of two vowels, where pronouncing the second vowel requires a closing motion, and thus the monophthong characteristic of the first vowel is lost as the mouth shifts to produce the second vowel.

There are 8 diphthongs in the English vowel system. One group of diphthongs, known as centering diphthongs, are produced by the motion of the tongue as it moves from the side to the center of the mouth, or from the back or front of the mouth to the center of the mouth to produce diphthongs such as /ɪə/, /εə/and/ʊə/. Another kind of diphthongs, called closing diphthongs, are created when the tongue moves upward from a lower to a higher position in the mouth, and one's mouth changes from being relatively open to relatively closed. The sounds /eɪ/, /oʊ/, /aɪ/, /aʊ/ and /ɔɪ/ are examples of such diphthongs. In English, with both these diphthongs, generally the initial vowels are more clearly pronounced, while the second vowels are fuzzy. But, even so, with English diphthongs, each vowel blend is clearly different from the fusion of two vowels in Chinese. However, the pronunciation of words such as yard[jɑ:rd], yes[jes], cure[kjʊr], watch[wɑ:tʃ], wet[wet], work[wɜ:rk], weed[wi:d], which are clearly noted as long- or short-vowel monophthongs in general textbooks, that is to say, semivowels, closely resemble the first sounds of Chinese opening diphthongs, and thus these common American sounds could be incorporated as objects of beneficial comparison when teaching Chinese diphthongs.

Although these English semivowel sounds can appear as the initial sound of a syllable, their Chinese counterparts are stronger since they are located in the head of a syllable, when pronounced as a semivowel the features are more obvious, as in Chinese. For example in the word work [wɜ:rk], we exclude retroflex follow-up part. This should be pronounced as /ʊə/, with the uo in Chinese as basically the same.

There are two groups of vowel combinations in Chinese that are not usually addressed by any textbooks or phonetic books. These are the nasal vowels, which are ignored by most researchers, but important for correct pronunciation. However, since they are divided into two groups—those of “tongue front” and “tongue back”—they each should be addressed separately during instruction. Two groups of diphthongs exist separately in the following nasal vowels: “ian[iæn]” and “iang[iɑŋ]” , “in[iɛn]” and “ing[iɛŋ]” (according to

the draft, the two vowels spelling should be “ien” and “ieng”); “uan[uæn]” and “uang[uɑŋ]”, “uen[uen]” and “ueng[uəŋ]” (the Chinese Pinyin / Chinese Phonetic Alphabets, “CPA” for short, regards “ueng” and “ong” as two kinds of spelling, but in fact there is no difference in how they are pronounced); “üan[yæn]” and “üen[yen]” (which is often romanized as ün and un) and “üeng[yəŋ]” (which is often romanized as iong by the “CPA”). The diphthongs of Chinese and English are contrasted in the following table.

**Table 2. The Diphthongs of Chinese and English**

	Chinese				English			
	Front	Closing	Opening	Back	Centering	Closing	j	w
Close	ie ye ue			iə yə uə	ɪə ʊə		ju ju:	wɪ wi:
Close mid		eɪ əʊ	uɿ			eɪ ʊʊ		
Open mid			iɛ yɛ		ɛə	ɔɪ	je jɜ: jɔ:	we wɜ: wɔ:
Open	iæ yæ uæ	æɪ ɑʊ	iA uA	iɑ uɑ		aɪ ɑʊ	jæ jʌ ja:	wæ wʌ wɑ:

\* According to Celce-Murcia, Brinton & Goodwin (1996), diphthongs such as [eɪ], [oʊ] should be written as [ey], [ow] and is now considered to consist of “a simple vowel plus a glide. However, I have continued to use the notations of the Diphthong since they do not constitute a consideration that affects the conclusions put forth in this study.

### 2.2.2.3 The triphthongs in Chinese and English

Most general textbooks suggest that there are no triphthongs in English. However, if we add a semivowel to a diphthong, it can be said that triphthongs exist in English. There are four triphthongs in Chinese: iao, iou, uai, uei. Accordingly, these also exist in English: why[wai], way[wei], yea[jei].

English and Chinese triphthongs share the same basic principle: a vowel fusion occurs. Because Chinese has three vowel triphthongs, the transition between vowels is more compact than in English. Moreover, Chinese triphthongs share similarities with English triphthongs

The principle for pronouncing Chinese and English triphthongs is very similar: English starts with the initial semivowel, adds the monophthong and finishes with the terminal vowel; whereas in Chinese, one starts with the initial diphthong and then adds the final vowel. However, since in Chinese the three vowels fuse into one, the transition between them is more compact than in English. In addition, the triphthongs of Chinese have further similarities with the triphthongs we have listed here in English, since in both languages the first sounds could be considered semivowels.

### 2.2.2.4 Combinations of vowel + nasal consonants in Chinese and English

There are syllables that end with “vowel(s) + nasal consonants” in both English and Chinese, a combination

which Chinese describe as “nasal vowel consonants” (*biyunmu* 鼻韵母). However, there are significant differences in their behaviors. In Chinese, the sound of each syllable remains separate from adjacent ones. In other words they make no sound changes because of the vowels or consonants in the words that immediately precede or follow the syllable. This is in direct contrast to English where the pronunciation of a vowel within a syllable can change when the syllabification of the word changes. As a case in point, note how the “a” of “human” and the “a” of “hu-ma-ni-ty” are pronounced differently.

The pronunciation of Chinese nasal vowel syllables, (whether monophthong vowel + “n” or “ng”, or diphthong vowel + “n” or “ng”) is strictly governed by the constraints of the syllable. This holds true for the combination of “vowel + vowel”, as well as for the combination of “vowel(s) + nasal consonants”. In the Chinese language, vowel(s) + nasal consonants are divided into two obvious groups: “vowel(s) + n”, in which all the vowels become “front vowels”; and “vowel(s) + ng”, where all the vowels become “back vowels”.

In teaching Chinese phonetics, there is an incorrect view regarding nasal vowel syllables. Some mistakenly believe that a “nasal vowel syllable” is created simply by adding “n” or “ng” to a vowel, or to a diphthong vowel. However, for example, simply recognizing “an” and “ang” as “nasal vowel syllables”, does not adequately address the qualitative differences between the pronunciation of the vowel “a” in both its settings, in which, as dictated by the Chinese phonetic syllable system, the “vowel(s) + n” creates a “front vowel syllable” [æ]; while the “vowel(s) + ng” creates a “back vowel syllable” [ɑ].

To correctly begin the pronunciation of a nasal vowel syllable, one must begin with the tongue already in position: forward for “vowel(s) + n” or back for “vowel(s) + ng”; followed immediately by the sounding of the vowel. If one begins with the placement of the “A” (a “central vowel”) and then moves the tongue forward or back to close with the pronouncing of the “n” or “ng”, the pronunciation will be incorrect.

I would also like to clarify some of the errors in pronunciation that arise because of the simplified spellings of some of the nasal vowel syllable pronunciations presented in the “CPA”. A few of the spellings have a certain disparity with the actual pronunciation. There are many engaged in teaching Chinese who do not understand this point, causing errors in teaching. This point should be clear: the following five nasal vowel syllables are written one way, but should be pronounced as shown with an additional vowel—

in=ieng ing=ieng ün=üen ong=ueng iong=üeng

In the “CPA”, the first three of the above syllables are correctly listed in the diphthong “i+en”, “i+eng” and “ü+en” groups, but their spellings do not correctly represent their true pronunciation. Here the vowels are not pure vowels, but diphthongs. Careful reading of this part of the “CPA”, should clarify this point. However, the last two nasal vowel syllables present a different challenge. The “ong” is always pronounced as if it belonged in the “u+eng” group. The “iong” should always be pronounced as if it were in “ü+eng” group.



There are 4 monophthong nasal vowel syllables in Chinese. As mentioned above, in these 4 syllables, “an” / “ang” and “en” / “eng”, although the “a” and “e” on their own are considered “central vowel” monophthongs, their pronunciation in conjunction with the nasal consonant is significantly altered—“an” / “en” become “front vowels”, and “ang” / “eng” become back vowels.

As is seen below, there are two series (front or back vowel pronunciation) of nasal vowel diphthong syllables: “ian” and “iang”,

“in” and “ing” (spelling and pronunciation referenced above),

“uan” and “uang”,

“uen” and “ueng” (spelling and pronunciation referenced above),

“üan”, …(which has no corresponding “ng” version)

“üen” and “üeng” (spelling and pronunciation referenced above).

The pronunciations of the “a” and “e” in the pairs of nasal diphthong vowel syllables, that is to say, the vowels that are the “primary sonorous component of the diphthong”, (shown above) will differ slightly from the pronunciation of their non-nasal monophthong / diphthong / triphthong syllables, and their monophthong nasal vowel syllable siblings. The differences will arise from the placement of the vowel in the mouth due to the absence or presence of other vowels (in diphthongs / triphthongs), or the absence or presence of one of the nasal elements. Understanding this should facilitate correct pronunciation.

### 2.2.3 Consonants of Chinese and English

Both Chinese and English can have consonants at the onset and coda of a syllable, but while English can have one or a cluster of consonants to begin and / or end a syllable, Chinese limits this to one, whether at the onset or coda of a syllable.

Additionally, English can have a variety of coda consonants, while the only consonant codas permitted in Chinese are the two nasal consonants. In this “loose” English syllable structure, the sound of the consonants can change, even within identical syllables (as in decease s=c / disease s=z), but in Chinese, no matter what other syllables are around it—the sound of a syllable never changes. In English, with the so-called liaison effect, coda consonants can be transformed into onset consonants of a following syllable (human—humanity), but in Chinese this never happens—the onset and coda sounds of a Chinese syllable are sacrosanct. In the English system, consonants are either voiceless or voiced sounds, but in the Chinese system, consonants are aspirated and un-aspirated sounds.

And while English can present its consonants in clusters of two or more, this is not the case in Chinese. In addition, in English there are frequently consonants present in the syllable which are not pronounced,

(campaign, thorough, knife), while Chinese consonants are always pronounced.

Some English consonants, by turns, are voiceless in one setting, but voiced in another, (butte—butter); and colloquially, in the flow of speech, with the release of a terminal voiced consonant one can sometimes hear the echo of its voiceless counterpart, (good, bad, bed).

In a colloquial setting of Chinese, there may be occasions when one hears voiceless consonants in the place of the true un-aspirated Chinese consonant, but this is not always obvious.

Since Chinese uses the same letters as English (when not using characters) to present the pronunciation of the language, it is very important to remember that the two language systems have very different consonant pronunciations, as Tables 3 and 4 make clear.

**Table 3. The Consonants of Chinese**

Pronunciation Place	Plosive		Affricate		Fricative		Nasal	Lateral approximant
	Voiceless		Voiceless		Voiceless	Voiced	Voiced	Voiced
	Unaspirated	Aspirated	Unaspirated	Aspirated				
Bilabial	b[p]	p[pʰ]					m[m]	
Labio-dental					f[f]			
Alveolar	d[t]	t[tʰ]					n[n]	l[l]
Dental			z[ts]	c[tsʰ]	s[s]			
Alveolo-palatal			j[tɕ]	q[tɕʰ]	x[ɕ]	y[j]	n[ɲ]	
Retroflex			zh[tʂ]	ch[tʂʰ]	sh[ʂ]	r[ʐ]		
Velar	g[k]	k[kʰ]			h[x]			

**Table 4. The Consonants of English**

Pronunciation Place	Plosive		Affricate		Fricative		Nasal	Lateral approximant
	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiced
Bilabial	b[b]	p[p]					m[m]	
Labio-dental					v[v]	f[f]		
Dental					th[ð]	th[θ]		
Apical /Alveolar	d[d]	t[t]					n[n]	l[l]
Dental			z[dz]	c[ts]		s[s]		
Alveolo-palatal			j[dʒ]	ch[tʃ]	z[ʒ]	sh[ʃ]	n[ɲ]	
Retroflex								r[r]
Velar	g[g]	k[k]					[ŋ]	
Glottal						h[h]		

### 2.2.3.1 The onset consonants

The head consonant of a syllable in Chinese is generally called the initial consonant. The Chinese initial consonants, as presented in the “CPA” presents consonants classified according to their six different types of pronunciation groups, giving learners a very good reference point for understanding and mastering the Chinese consonant pronunciation system.

Since the production of these consonants within their groups have the same or similar articulation points, the presentation found in the “CPA” facilitates the mastery of Chinese consonants. In this section, when I write of consonants, I will follow the table of the “CPA”:

b, p, m, f;  
d, t, n, l;  
g, k, h;  
j, q, x;  
zh, ch, sh, r;  
z, c, s.

Because English has a system of pairs of voiced and voiceless consonants, the production of a consonant is determined by how or if the vocal cords vibrate during the articulation. On the other hand, in the Chinese system of pairs of aspirated and un-aspirated sounds, the focus of pronunciation is on the movement of air as it goes through the vocal tract.

While the English focus of pronunciation is on where and how the vocal chords may or may not vibrate, the Chinese focus is on how and where to produce or prevent the puff or burst of air that typifies the aspirated or un-aspirated consonant, which is crucial to correct pronunciation.

In Chinese, when pronouncing un-aspirated sounds there must not be any accompanying puff or burst of air in the production of the sound. But when pronouncing aspirated sounds, the following two considerations are important:

- (1) There must be a strong accompanying puff or burst of air at the place of articulation.
- (2) Speakers need to keep the tension of the part(s) of the mouth in focus –such as the lips, teeth and tongue.

—For the bilabial consonants “b”, “p”, “m”, there must be an emphatic engagement of the lips; for the labiodentals “f”, there must be an emphatic engagement of the teeth and tongue. (Although the “m” does not constitute much of a problem, there must be an emphatic engagement of the lips.)

—The apical(alveolar) “d”, “t”, “n” and “l”, sounds are more demanding: the articulation of “d” and “t” are both produced higher in the mouth in English than in Chinese; the Chinese “d” and “t” belong to the back-of-

the-teeth sound production, while the English “d” and “t” are alveolar. Getting students to place the tongue in a correct position is crucial to good pronunciation.

—For the velar “g”, “k”, “h”, the main problem is that of pronouncing the “h”. The English “h” is a glottal fricative, while the Chinese “h” is the fricative sound. English “h” sounds “smooth”, but “h” in Chinese has a strongly “crunching” fricative sound.

—The palatal Chinese “j”, “q”, “x”, for English speakers, are difficult sounds. The key problem relating to producing these Chinese sounds is that the tip of the tongue does not make contact with the back of the upper teeth or gums—the tip of the tongue should “roll down”, and be placed against the back of the lower teeth (which leaves the “center” of the surface of the tongue somewhat raised, close to the hard palate). Because of the influence of English jeep [dʒi:p], cheese [tʃi:z], sheet [ʃi:t], American students tend to pronounce the Chinese “j”, “q”, “x” as though they were apical consonants. In other words, students often use their tip of the tongue to touch the back of the upper teeth or adjacent gums. This is a leading factor of incorrect pronunciation of these three consonants.

—The four retroflex Chinese consonants “zh”, “ch”, “sh”, “r”, might not pose great difficulties for American students since English has similar sounds. Some instructors use [dr], [chr], [shr], [r] to remind American students to pay close attention to these sounds. However, when beginning the study of these four consonants, it is important to get students to raise their tongue higher.

—The Chinese coronal consonants “z”, “c”, “s”, when blended with the vowel “a” do not appear to be problematic for American students. However, when these same students reference familiar English words such as “birds” [bɜ:rdz], “tsetse fly” [ts], or “see” [s] in an attempt to correctly pronounce the Chinese “z”, “c”, “s”, their over-riding tendency to use a close-but-not-correct pronunciation. The problem is this: the two English sounds ([dz], [ts]) are created with the tip of the tongue touching the back of the upper teeth, but in Chinese, this placement of the tongue does not produce the desired results. Pronouncing these three Chinese consonants correctly requires that the tip of the tongue touch the back of the lower teeth. Because of this, as noted below, there may be some irregularities in the pronunciation of the vowels that appear with them in syllables such as “zi” [tsɿ], “ci” [tsʻɿ], “si” [sɿ] etc. In these cases, the vowel has a unique pronunciation because of the requirements of the consonants that precede it.

### 2.2.3.2 Coda consonants in Chinese and English

Because there are only two Chinese coda consonants, the “n” and “ng” the system is fairly simple. But, as mentioned above, although the Chinese syllables are very compact, these coda consonants can present difficulties within the syllable, because any syllable that ends with either of these two consonants becomes a

“nasal vowel syllable”, whose vowels undergo significant pronunciation changes. If one keeps these changes in mind, Chinese coda consonants will not cause difficulties.

#### 2.2.4 The accent / tone in the Chinese and English

The English and Chinese accent / tone systems are different. While English uses a stress accent / tone system, Chinese uses a pitch accent / tone system. In English monosyllabic words there is no occurrence of stress accent / tone distinctions; but in the presence of a diphthong vowel or a long vowel, there are stress accent / tone distinctions. And in this later case, the stress accent / tone is mainly a falling one, as in “I” [aɪ]. However in Chinese, each syllable has one unchangeable tone. Regardless of the structure of the syllable (be it composed of a monophthong, diphthong or triphthong vowel), the tones have their own patterns that do not change as vowels increase.

The stress accent / tone element in English is mainly heard in polysyllabic words. The stress accent / tone of a word depends on the syllabic structure of the word, the etymological source of the word and / or the part of speech of the word. In English polysyllabic words, the strong stress can be on any syllable of a word.

On the other hand, in Chinese polysyllabic words, even though there may be some regional changes in the structure (as in the Pekinese “hua” + “er” = “huar”), the general rule states that each syllable must maintain its original tone type. (And within the Chinese syllable structure, it is generally accepted that the two patterns of the Chinese “third tone” allophones do not contradict the general rule as stated above.)

Since the English stress accent / tone and the Chinese pitch accent / tone occasionally have similarities, American students sometimes mistakenly think that by using “their” English stress accent / tone when pronouncing a Chinese word is equal to the correct pronunciation of the Chinese pitch accent / tone. Therefore, in the pronunciation of Chinese, careful attention must be paid to the correct production of the proper tones. More especially, the first and fourth tones must remain relatively high, so that they can be differentiated from the lower second and third tones, thus producing more accurate Chinese tones.

##### 2.2.4.1 Tones of monosyllabic words

As mentioned above, in English monosyllabic words the accent / tone system is not engaged, unless the monosyllabic vowel is a diphthong or a long vowel when the stress accent / tone distinctions do come into play. Thus, in pronouncing the diphthong “I” [aɪ], although the dictionary does not indicate a differentiation of stress between the two vowels in the diphthong, in reality, the first vowel of the diphthong is pronounced higher (or stronger) than the second vowel. But in Chinese, each character is a syllable, each syllable has its own tone, and the tone is innate, not changing due to any increase of vowels.

In addition, the rising, falling or “continuous” tones of Chinese can be represented using “1 2 3 4 5”, which correspond to the “do re mi fa sol” of a musical scale:

the “first tone” can be written as 55 / sol sol;

the “second tone” can be written as 35 / mi sol;

the “third tone” can be written as 214 / le do fa;

the “fourth tone” can be written as 51 / sol do;

In general, textbooks have examples such as the following:

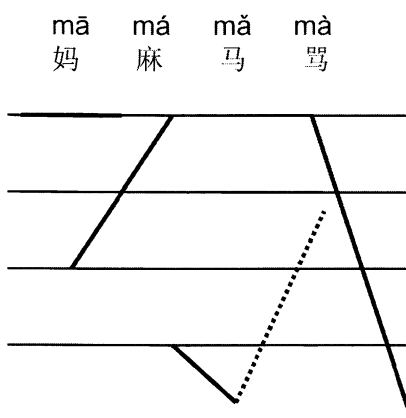


Figure 1. The Four Tones of Chinese

#### Four tones in Chinese

The “**first tone**”, “55”, although it does not exist in English, is not difficult to produce: it is the high-pitched continuous sound that does not “rise” or “fall” at any time during its pronunciation.

The “**fourth tone**”, “51”, starts at point slightly higher than the high-pitched “first tone” and descends rapidly to a pitch that corresponds roughly to the change in pitch singing “sol do” in the musical scale. Starting with the correct higher pitch is very important.

The “**third tone**”, “214”, exists in English when one says “yes” as the answer to a request s/he is not entirely willing to do: “Would you be willing to take out the trash?” “Y-eeh-sss / 214”.

More difficult tone is the “**second tone**”, “35”. It can be heard in English “tag- questions” such as “You are a student, aren’t you?” / “35”. In the final part, the “you?” sounds like the sound of the “second tone” of Chinese. Using this pattern lets students get such a sense of the Chinese second tone, that they should be able to get twice the result with half the effort.

## 2.2.4.2 Tones in dissyllabic words

In the following schematic, all possible tone combinations of dissyllabic words are represented, including the “light tone” which is an unstressed syllable. Because each syllable has its own tone, plus the light tone, the dissyllabic words in Chinese have total of 20 types, as indicated in Table 5.

Table 5. Examples of the Tone Patterns in Chinese

	First tone	Second tone	Third tone	Fourth tone	Light tone
First tone	jīntiān 今天	gōngyuán 公园	hēibǎn 黑板	ānjìng 安静	māmā 妈妈
Second tone	shíjiān 时间	xuéxí 学习	méiyǒu 没有	xuéxiào 学校	yéye 爷爷
Third tone	lǎoshī 老师	lǎngdú 朗读	shǒubiǎo 手表	mǐfàn 米饭	jiějie 姐姐
Fourth tone	miànbāo 面包	wàiguó 外国	diànnǎo 电脑	zàijiàn 再见	bàba 爸爸

Because in the English accent / tone system, dissyllabic words have only two patterns: strong-weak (such as “order” [ˈɔ:rdə(r)],) or weak- strong (such as “prefer” [prɪˈfɜ:(r)]), an English speaker naturally expects the Chinese dissyllabic words to function in the same manner: that is to say, when the first syllable is a stressed syllable, the second syllable must be an unstressed syllable, and vice versa. This expectation leads students to misunderstand the Chinese system, in which dissyllabic words do not rely on the stressing of one of the syllables, but upon the use of one of the possible combinations (seen in the chart above) of the distinctively pronounced 4 tones and light tone to shape the pronunciation of the word.

Having said this, there are two Chinese dissyllabic words whose pronunciation, closely mimicking the “strong-weak” pattern (pattern 1) and the “weak-strong” pattern (pattern 2) of English, can be considered as correct pronunciation .

Pattern 1: first tone + light tone

Pattern 2: third tone + fourth tone

But American students, trying to frame Chinese as a “stress pattern” system, tend to produce the following errors:

Pattern 1: Rather than relying on the pure pronunciation of the tones within the dissyllabic word to establish the correct presentation of the syllables, the student will attempt to pronounce the first syllable as the stressed syllable. And because the student has stressed the first syllable, s/he will very naturally pronounce the second syllable as an unstressed sound. As mentioned above, the “first tone + light tone” dissyllabic words fall into this category, yet are considered to be correctly pronounced. But for all other dissyllabic words, these

anglicized patterns of pronunciation would be erroneous.

Pattern 2: In the event that the student feels it appropriate to stress the second syllable, s/he will very naturally pronounce the first syllable as an unstressed sound. As mentioned above, the “third tone + fourth tone” dissyllable words fall into this category, yet are considered to be correctly pronounced. For all other dissyllable words, these anglicized patterns of pronunciation would be erroneous.

#### 2.2.4.3 The tones of tri-syllabic words or words of more than three syllables

For Chinese trisyllabic words, and for those words of more than 3 syllables, there exist a plethora of pronunciation patterns. Chinese trisyllabic words are often composed of these combinations: “a monosyllabic word + a dissyllabic word”; or “a dissyllabic word + a monosyllabic word”. And for many of the three-syllable words, the pronunciation of the second syllable may be a little weaker than the other two syllables.

The Chinese quadra-syllable is often composed of combinations of words: “a pair of dissyllabic words” ; “a monosyllabic word + trisyllabic word”; or “a trisyllabic word + monosyllabic word”.

The pronunciation challenges, predictably, are related to the students’ predilection for applying “their system” to the new system they are learning. The stress patterns of English words of three or more syllables, falling into 3 basic types, (the last syllable is stressed; the penultimate is stressed; or the antepenultimate is stressed), American students tend to apply one of these three English “3 or more syllabic” stress patterns to the Chinese words they encounter in this particular polysyllabic context.

For example, the pronunciation of “中国人Zhongguoren” (the word means “the Chinese people”), should be “55 35 35” (first tone + second tone + second tone). However, the middle syllable, (normally pronounced as the “second tone”) becomes the “light tone”, pronounced “55 light 35”.

When American students pronounce “Zhongguoren” , they commit all three possible errors:

**Zhongguoren**

Zhong**guo**ren

Zhongguo**ren**.

In an amusing corollary to this specific error, many students will often say “Zhong-**guo**-ren”, mimicking the rising-falling pronunciation stress pattern of its actual English translation— “Chinese” (even though the word “Chinese” is a dissyllabic word !); and they will do the same thing when pronouncing “**Ying**guoren”, using the rising-falling pattern of its actual translation – “**English**”!

In other cases, faced with a tri-syllable word whose first syllable is a “first tone syllable”, students mistakenly reason that this combination equates to “stress syllable” status, and consequently assign “un-stressed status” to the other syllables of the word, resulting in erroneous “high-low-low / strong-weak-weak” patterns, typical



of the American English pronunciation system.

### 2.2.5 Intonation of Chinese and English

Both Chinese and English both allow multiple shifts of intonation in the delivery of the same sentence in order to reflect the primary focus of the sentence, and they both allow a drawing out of the stressed syllable and a louder pronunciation of the stressed syllable. (For instance, the sentence “Yes.” is said to have five kinds of intonation.) However, in this article, I will only mention two kinds — the rising and the falling intonations, which are a primary source of intonation errors for American students.

Chinese clearly exhibits rising intonation when the question marker “*ma*” is found at the end of some interrogative sentences. Without “*ma*”, Chinese interrogative sentences do not exhibit any rising intonation.

In English, rising intonation is clearly recognizable. The rising intonation generally appears before a short pause, as in “I bought apples, pears, **bananas**...and mangoes”. / “I bought pears, bananas, **apples** and mangoes.” And as in the examples given, the rise in intonation occurs on the penultimate syllable of the word stressed.

At the end of direct “yes or no” questions, the rising intonation is also obvious as in the example, “Are you a **student**?” It can also be seen in tag questions, using a sentence + when there is an expression of doubt, hesitation, or an uncertainty (“You are a student, aren’t you?”). This intonation typically rises on the last stressed syllable of the sentence.

Chinese intonation errors made by American students generally concern the overuse of rising intonation. In a Chinese sentence representing a feeling of uncertainty or doubt, students tend to use rising intonation, mimicking what happens in similar English sentences.

你是学生? *Nǐ shì xuésheng?* (correct, with no rising intonation)

*Nǐ shì xuésheng?* (**xué** indicates erroneous rising intonation)

Correctly using falling intonation does not present much of a problem for American students, since it is used in Chinese declarative sentences, as it is in English:

I am a student. (intonation rises on “*stu*”, falls on “*dent*”)

我是学生。 / *Wǒ shì xuésheng.* (intonation falls on “*xué*”)

However, American students do frequently use rising intonation in places where they should use falling intonation. But the reason for the error is simple—students are using the rising intonation to mimic the English rising intonation to express uncertainty, doubt. (see above)

When attempting to correctly reproduce Chinese intonation, American students often make the following errors, due to their anglicization of Chinese:

In the Chinese declarative sentence,

我叫玛丽。 / Wǒ jiào Mǎlì. (third-fourth-third-fourth tones / 214-51-214-51) “I am Mary.”

since each syllable has its own tone, and since it is a declarative sentence, there is no rising or falling intonation.

But the American student, subjecting these 4 Chinese syllables to the English,

“I am Mary” “strong-weak-substrong-weak” pronunciation pattern will say:

“wo” (“strong / high” — the equivalent of the first tone or fourth tone)

“jiao” (“weak”)

“Ma-li” (“substrong” + “weak”)

and thus does not correctly pronounce the Chinese sentence.

In the question,

他是谁? / Tā shì shéi? (first-fourth-second tones / 55-51-35) “Who is he?”

since each syllable has its own tone, and since there is no “ma” in this interrogative sentence, there is no rising or falling intonation.

But the American student, subjecting the 3 Chinese syllables to the English,

“Who is he?” — “weak-strong-weak” pronunciation pattern or

“Who is he?” — “weak-weak-strong” pronunciation pattern,

will produce one of the following incorrect Chinese patterns:

Ta (weak) shi (strong) shei (weak)

Ta (weak) shi (weak) shei (strong)

and thus does not correctly pronounce the Chinese sentence.

In the declarative sentence,

我去学校。 / Wǒ qù xuéxiào. (third-fourth-second-fourth / 214-51-35-51) “I go to school.”

since each syllable has its own tone, and since it is a declarative sentence, there is no rising or falling intonation.

But the American student, subjecting these 4 Chinese syllables to the English,

“I go to school” “strong-weak-weak-weak” pronunciation pattern, will say

Wo (Strong/high — the equivalent of the first tone or fourth tone)

qu (weak, substrong)

xue-xiao (subhigh — the equivalent of the first tone or fourth tone, sub weak)

and thus will not correctly pronounce the Chinese sentence.

In the following non- “ma” interrogative sentence,

你去学校? / Nǐ qù xuéxiào? (third-fourth-second-fourth tones / 214-51-35-51) “Will you go to school?” since each syllable has its own tone, and since it is a non- “ma” interrogative sentence, there is no rising or falling intonation.

But the American student, subjecting these four Chinese syllables to the English,

“Will you go to school” (substrong-weak-weak-weak-strong) pronunciation pattern, will say:

Ni (substrong) qu (weak) xue (weak) xiao (strong) .

and thus will not correctly pronounce the Chinese sentence.

In the following “ma” question,

你去学校吗? / Nǐ qù xuéxiào ma? (third-fourth-second-fourth-light / 214-51-35-51-light)

Will you go to school? / Are you going to school?

since each syllable has its own tone, and since it is a “ma” interrogative sentence, there is rising intonation.

But the American student, subjecting these five Chinese syllables to the English,

“Will you go to **school**” (substrong-weak-weak-weak-strong) pronunciation pattern or,

“**Are** you going to **school**” pronunciation pattern, will say

Ni (strong / substrong) qu (weak) xue (strong) xiao (weak) ma (strong) ?

### 3. Summary

#### **Application of research results:**

As I mentioned, with the advice and help of Professor Shiaoling Yu and support of her colleagues, my research was very fruitful this year. This paper has underscored the need for a critical re-evaluation of the “CPA”. More importantly, it has stressed the need to reexamine how pronunciation is taught. And since this examination was conducted primarily on OSU students learning Chinese, it is logical that these observations should enhance the teaching of Chinese at OSU.

This paper emphasizes that sound is only in the syllable. But because many students don't understand the structure of Chinese syllable, their pronunciation is not consistent. Clearly showcasing the fact that the Chinese syllable is a compact entity, one whose sounds are not affected by the consonants or vowels adjacent syllables, (a phenomenon particularly prevalent in spoken English) may help guide students to better produce authentic Chinese sounds.

Although there are monophthong, diphthong and triphthong vowels in both Chinese and English, the production of their sound within their syllables is noticeably different. Teaching each vowel by pronouncing it within the setting of its specific syllable seems to offer the best results.

Furthermore, no matter what language one is studying, the mechanics of producing authentic consonant

sounds depend on the proper placement the tongue in the mouth. The sound of Chinese and English consonants, such as “d, t, j, q, zh, ch, z, c”, (to mention some of the more challenging ones), differ greatly, precisely because of the differing dictates of each language governing the placement of the tongue in the mouth when producing these consonant sounds. This is a crucial detail that is often overlooked in class, but which urgently needs to be addressed.

As for the “accent / tone” discussion, the great differences between Chinese and English require that students understand and practice the dictates of the Chinese “tone” system, and not substitute the English “accent” system for “tone”. The English “accent” system has a conspicuous, if unfortunate, impact on the acquisition of Chinese “tone”—a point that teachers should be very aware of. For example, when studying Chinese disyllabic words, (in which each syllable has its “tone”—which a student may mistake for “stressed” and thus “unstressed” the other syllables), it is important to alert students “against” English disyllabic words, (in which, if one syllable is stressed, the others are necessarily unstressed). Understanding these differences between English and Chinese, students will be better able to avoid unnecessary errors when dealing with Chinese “tones”.

As for Chinese and English intonation, observation also recognizes that when students speak Chinese, they “naturally” tend to adopt English intonation. Remind students not to use English intonation to fit Chinese intonations, thereby you may perhaps get twice the results with half the effort.

In a word, this research highlights some issues to which numerous researchers have paid scant attention. I hope it will be able to provide some reference points for teachers of Chinese in the United States, thus improving the teaching of Chinese in the United States, which may contribute to an improvement in the communication between the peoples of these two wonderful countries—America and China.

### **Acknowledgements**

Heartfelt thanks to the guidance and care of the following persons: Professor Shiao-ling Yu, Mark Ren, Chenhui Ho, Setsuko Nakajima, Fumiko Nakamura and Luke Yamaguchi and other teachers whose classes I observed at the Oregon State University. I also wish to thank Professors Meiru Liu and Patricia J. Wetzel of Portland State University as well as many teachers whose classes I observed. I would like to further express my appreciation to my housemate Ms. Donna Knoy gave and Ms. Cindy Nair of the International Student and Faculty Services of the OSU as well as the staff of the Department of Foreign Languages and Literature of the OSU. For help with this manuscript, I acknowledge the help of Tim Newfields and Kaya Taguchi of Toyo University.

## Bibliography

- 桂明超.《再论美国英语语调对美国学生学习汉语声调的干扰》. 云南师范大学学报 (对外汉语教学与研究版), 2003.1.
- 林焘, 王理嘉. 语音学教程. 北京大学出版社, 1992.
- 潘文国. 英汉对比纲要. 北京语言学院出版社, 1997.
- 王理嘉. 音系学基础. 语文出版社, 1991.
- 徐通锵. 基础语言学教程. 北京大学出版社, 2001.
- 续三义. 对日汉语语音教程. 北京语言大学出版社, 2000. 2005第二次印刷.
- 续三义.《关于汉语普通话复韵母》. 人文丛刊. 学苑出版社, 2005.7.
- 续三义.《日本的汉语语音教学小议——以韵母为例》. 国际汉语教育. 2010年第4辑. 2011.4.1.
- 张凤桐. 英国英语语音学和音系学. 四川大学出版社, 1996.
- 张金生.《英汉元音对比与英语语音教学》. 解放军外国语学院学报. 第1期. 2002.1.
- 赵德梅. 英汉比较语音学. 青岛海洋大学出版社, 1995.
- 周考成. 英语语音学引论. 四川大学出版社, 1990.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press, 1996.
- Lass, R. *Phonology*. Cambridge University Press, 1984.
- Martin J. B. & Rahilly, J. *Phonetics: The Science of Speech*. Oxford University Press, 1999.
- Murray, T. E. *The Structure of English*. Allyn and Bacon, 1995.
- Oxford Advanced Learner's Dictionary, 8th edition*. Oxford University Press, 2010 ( Electronic Dictionary).