

# Phonetic Detail in the Teaching of Mandarin Pronunciation

Chris Wen-Chao Li\*

*Department of Foreign Languages and Literatures, San Francisco State University  
1600 Holloway Avenue, San Francisco, California 94132-4163, U.S.A.*

*Keywords:* Interlanguage phonology; contrastive analysis; markedness; redundancy; Modern Standard Chinese; Mandarin

---

## Abstract

This paper seeks to answer the question: should technical phonetic descriptions be introduced in the teaching of Mandarin pronunciation to beginning learners, and if so, what type of description, in how much detail, and following what order? The presenter will argue that the indiscriminate introduction of articulatory specification can impede rather than aid the acquisition of a foreign language sound system, and that if aspects of phonetic description are to be incorporated into the teaching of pronunciation, it should be done with an awareness of interlanguage phonology, distinctive feature theory, and the limitations of present-day Mandarin phonetic/phonological description. Examples from the portrayal of Mandarin vowels, voicing contrasts, tone sandhi, and consonant place of articulation in the literature will be used to illustrate the issues at hand and support the adoption of a selective and tiered approach to the use of phonetics in teaching pronunciation.

## Outline

0. Introduction
1. Contrastive Analysis
2. Markedness Theory
3. Distinctive vs Redundant Features
4. Areas of Controversy
5. Dialect Variation
6. Is pronunciation worth teaching?

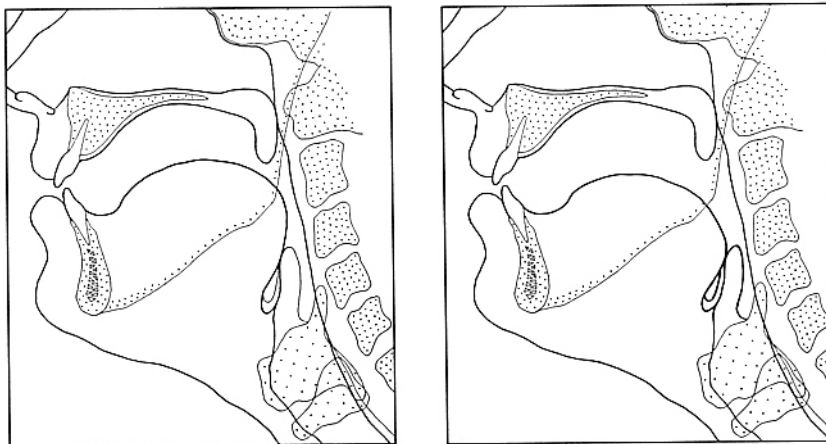
---

\* Tel.: 001 415 338 1034; E-mail: wenchao@sfsu.edu

## 0. Introduction

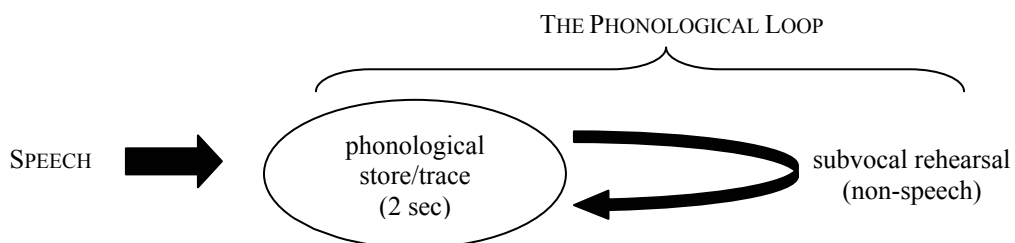
### 1. Common methods of teaching pronunciation

- Listen and repeat (no use of phonetic theory)
- Physiological diagrams (partial use of articulatory phonetics)



### 2. What is easier for the beginning student?

- Joseph Conrad (康拉德) phenomenon (Scovel 1969)
- Aptitude for oral mimicry (Kenworthy 1987; Jones 1997)
- Phonological short-term memory (PSTM)



(After Baddeley 1986; Gathercole 1998)

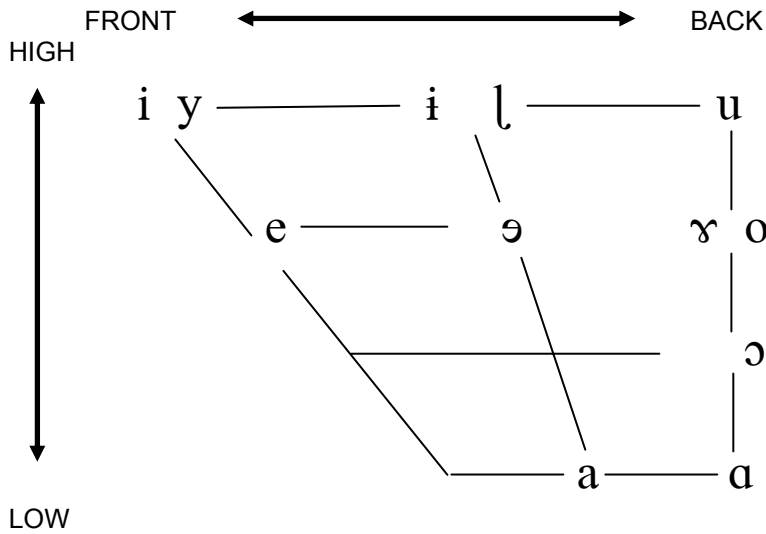
- Motor theory of speech perception (Lieberman *et al* 1967; Lieberman *et al* 1985); inappropriate acoustic image formation (Neufeld 1977)
- Interlanguage phonology (Tarone 1987): mapping of foreign sounds onto (1) native language phonemes (L1 transfer) [CONTRASTIVE ANALYSIS], and (2) universally unmarked patterns [MARKEDNESS THEORY]

## 1. Contrastive Analysis

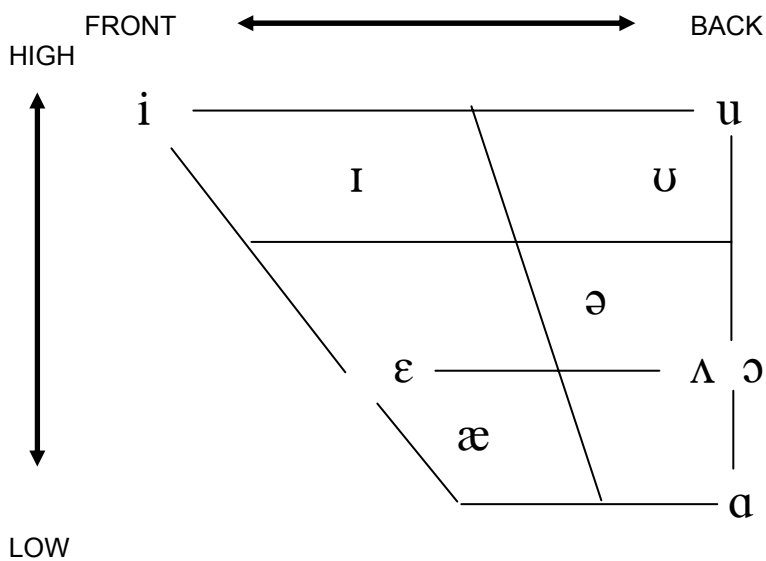
### 1. Vowels

- y (魚) → u / ju (difficulty in distinguishing 女/努/扭、旅/魯/柳)
- ɤ (餓) / ə (恩) → ə / ʌ (sounds in-between “e” and “a” to Chinese ear 德/達、跟/乾)

MODERN STANDARD CHINESE



GENERAL AMERICAN ENGLISH



## 2. Consonants

- Retroflex series (replacement with English “j”, “ch”, “sh”)
  - Alveopalatal series (replacement with English “j”, “ch”, “sh”)
- } MERGER

### MODERN STANDARD CHINESE

|             | LABIAL           | DENTAL | ALVEOLAR           | PALATO-ALVEOLAR    | ALVEO-PALATAL      | PALATAL | VELAR            | GLOTTAL |
|-------------|------------------|--------|--------------------|--------------------|--------------------|---------|------------------|---------|
| STOP        | p p <sup>h</sup> |        | t t <sup>h</sup>   |                    |                    |         | k k <sup>h</sup> |         |
| AFFRICATE   |                  |        | ts ts <sup>h</sup> | tʂ tʂ <sup>h</sup> | tɕ tɕ <sup>h</sup> |         |                  |         |
| FRICATIVE   | f                |        | s                  | ʂ ʐ                | ɕ                  |         | x                |         |
| NASAL       | m                |        | n                  |                    |                    |         | ŋ                |         |
| APPROXIMANT | w                |        | l                  |                    |                    | j       |                  |         |

### GENERAL AMERICAN ENGLISH

|             | LABIAL           | DENTAL | ALVEOLAR         | PALATO-ALVEOLAR | ALVEO-PALATAL | PALATAL | VELAR            | GLOTTAL |
|-------------|------------------|--------|------------------|-----------------|---------------|---------|------------------|---------|
| STOP        | b p <sup>h</sup> |        | d t <sup>h</sup> |                 |               |         | g k <sup>h</sup> |         |
| AFFRICATE   |                  |        | dz ts            | tʃ dʒ           |               |         |                  |         |
| FRICATIVE   | f v              | θ ð    | s z              | ʃ ʒ             |               |         |                  | h       |
| NASAL       | m                |        | n                |                 |               |         | ŋ                |         |
| APPROXIMANT | w                |        | l                | ɹ <sup>w</sup>  |               | j       |                  |         |

## 3. Phonotactics

- $\varepsilon]_{\sigma} \rightarrow ei$  (e.g., 學生 → 水聲)
- $o]_{\sigma} \rightarrow ou$  (e.g., 我 → 屋歐)

## 4. Tone

### 2. Markedness Theory

1. Marked vowels: [y], [i], [ʉ]

2. Marked consonants:

- Retroflex series “zh”, “ch”, “sh”, “r”
- Alveopalatal series “j”, “q”, “x”

3. Unmarked tonality and the description of Tone 3

- Tonal systems of the world’s languages
  - (a) 2 tone: H, L (level tones)
  - (b) 4 tone: H, L (level tones), HL, LH (contour tones)
- Traditional vs contemporary descriptions:

| TRADITIONAL DESCRIPTION: Tone 3 as <i>dipping</i> tone  | CONTEMPORARY DESCRIPTION: Tone 3 as <i>low</i> tone  |
|---|--|
| <p>“The third tone’s pitch starts at 2 (the second lowest point of your pitch range), falls to the lowest point of your pitch range, then goes upward to 4 (the second highest point of your pitch range). It is like “Mom” when you say it in a disapproving manner.” (p. 9) ...</p> <p>“When a third tone syllable is followed by a first, second, fourth or neutral tone, the pitch drops and does not rise (i.e., only the first half of the third tone is pronounced; it is called “half-third tone”. (p. 26) (<i>Interactions I: A Cognitive Approach to Beginning Chinese, 1997</i>)</p> | <p>“... in natural speech, a final Tone 3 need not be [214], but is often [21]. For example, in a study in 1998, I asked six native Standard Chinese speakers (four of them being Chinese language teachers) to read at normal speed sixteen expressions that contain a final Tone 3 ... Five of the six speakers pronounced all the sixteen final Tone 3s with [21] ... All the speakers agree that a final [214] carries some emphasis.”</p> <p style="text-align: right;">(Duanmu 2000: 221-222)</p> <p>“The third tone is a low tone with a pitch value of 214 (even though in reality the pitch value is more like 21)”</p> <p style="text-align: right;">(<i>Integrated Chinese, Level 1, Part 1: 9, 1997</i>)</p> |

### 3. Distinctive vs Redundant Features

#### 1. Vowels

##### ■ *Distinctive:*

- (1) High front rounding: [y] vs [u] vs [ju] (女/努/扭、旅/魯/柳)  
(Teach via C/A: [y] is [i] with lip-rounding.)  
(C.f. *Intergrated Chinese* I-1, p.1: “To produce this vowel, first pronounce i, then modify the shape of the mouth from unrounded to rounded”)
- (2) Secondary height:
  - (a)  $\varkappa / \varepsilon \rightarrow \varepsilon / \Lambda$  (perceived by Chinese as 餓/阿) [c.f. RP “er”]
  - (b)  $\varepsilon \eta \rightarrow \varepsilon \eta / \Lambda \eta$  (e.g., 很 [hɛn], [hΛn])
- (3) Diphthongization and fronting of “u” [努力 vs 女力]  
(C.f., sociolinguistic difference: “oo” in macho Italian vs gay accent)

##### ■ *Redundant:*

- (1) Diphthongization of “i” (e.g., 力量、弟弟)
- (2) Fronting of “ao” (e.g., 高興、好不好)

#### 2. Consonants

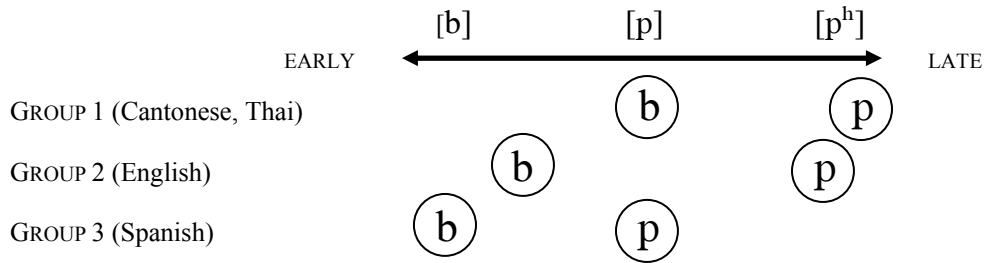
##### ■ *Distinctive:*

- (1) Lip rounding in retroflex and alveopalatal initials:  
Chinese [±round] is distinctive; English [+round] by default  
(Teach to spread lips when pronouncing “zh”, “ch”, “sh”, “r”, and “j”, “q”, “x”)

(2) Voicing and Aspiration

- Early textbook view: In English are able to perceive voicing contrasts, but not aspiration; Chinese: aspiration is distinctive, voicing is not. As a result much time is spent on teaching the distinction between [p, p<sup>h</sup>]  
 "...while /p-, t-, k-/ and /b-, d-, g-/ in Mandarin differ as to the presence or absence of aspiration, the difference between these two stop series in English is said to be that /p-, t-, k-/ are "voiceless" stops whereas /b-, d-, g-/ are "voiced" stops." (Tung 1983: 115-116)
- Phonetic literature: In English, aspiration is the main acoustic cue in the perception of stops (e.g., "b", "p"); whereas in continuants (e.g., "ch", "sh"), voicing serves as the main cue.

FIG Voicing and aspiration as continuum on VOT scale (After Catford 1977: 113)



(3) Tone: Tone Sandhi

(a) *Distinctive* (system-wide; phonemic): Tone 3 Sandhi [i.e., 洗澡 ; 養馬]

(b) *Redundant* (phonetic effects):

- T4+T4: "Chao (1968: 28-29) points out that when T4 is followed by another T4, the sequence is [53 51], instead of [51 51]. However, as Chao suggests, this may be a phonetic effect and need not be included in a phonological analysis" (Duanmu 2000: 221) [再見]
- T1/T2+T2+FT: "If in a three-syllable word or phrase the first syllable is a 1<sup>st</sup> or 2<sup>nd</sup> Tone, the second is a 2<sup>nd</sup> Tone, and the third syllable is any except the neutral tone, then the second syllable (which is in the 2<sup>nd</sup> Tone) is pronounced in the 1<sup>st</sup> Tone" (Chao 1948 :26) [三年級 ; 蔥油餅 ; 還沒完 ; 誰能飛] -- (EXPERIMENT: 蔥有柄 vs 蔥油餅 vs 聰優餅)

#### 4. Areas of Controversy

##### 1. Retroflex consonants

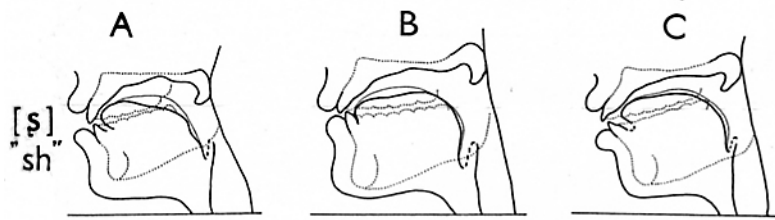
###### ■ Chinese vs Indian retroflex:

(a) “this sound does not have the tip of the tongue curled up and backwards, as it does in Indian sounds symbolized in a similar manner. All three speakers produce the constriction for this sound with the upper surface of the tip of the tongue (as compared with the under surface of the tip, as is common in the Dravidian languages of Southern India). The constriction is about the same place for all three speakers, namely at about the centre of the alveolar ridge.”

(Ladefoged and Wu 1984: 271)

(b) 「在北京话 sh 的问题上，洋人也是会犯错误的。他们只做了三个人的实验，大概都是年轻人。他们根据少数特例就下结论了。说它平，是否定卷。国内也有否定卷的强大势力。关键是，国际音标的 sh（指汉语 sh 常用的符号）已经被他们定义为舌尖下表面翻卷上去的姿态了。北京的事实并不是舌头尖倒勾发音，于是他们又走到了极端，说北京话的 sh 是顶音，不是卷舌音。那是少数北京人的偏前的一种发音。」（沈炯 2003.01.01 於北京大學中國語言文學系「學術論壇」）

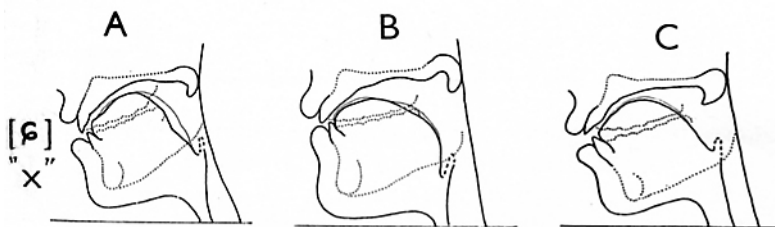
###### ■ Variant place of articulation?



(Ladefoged and Wu 1987: 269)

##### 2. Alveopalatal consonants

###### ■ Multiple place of articulation? ([ɕ] off the main IPA consonant chart)



(Ladefoged and Wu 1987: 269)

- (a) “Chinese palatals are articulated with the blade of the tongue placed against the front part of the palate: simultaneously the free part of the tongue is raised toward the alveolar ridge.” (Norman 1988: 140)
- (b) “The narrowest channel occurs near the front part of the alveolar ridge for Subject C, and notably further back for Subject B ... it is farther back than [s], but not quite as far back as [ʂ]. And, as in the previous cases, the category to which this sound is assigned has to specify not only the location of the constriction but also the shape of the whole body of the tongue.” (Ladefoged and Wu 1984: 271)

- Is IPA symbol [ç] adequate? (C.f. Polish alveopalatal)
  - (a) Ladefoged and Wu (1984: 270): Mandarin different from Polish -- Polish sound is more palatal.
  - (b) Halle and Stevens (1991: 93): Polish and Mandarin [ç] are close enough, and can be characterized by the same set of features [-anterior], [+distributed], [-back]. Different from Russian postalveolars.
- 3. 「日」母 Chinese “r”: fricative or approximant
  - Fricative [ʒ]: closer to “z”, “zh” (c.f. English “pleasure”) [Karlsgren 1926; Chao 1934; Hartman 1944; Howie 1976; Hsueh 1986; Lin 1989]
    - (a) 「[ʒ] 是唯一的一個濁擦音，必須與同部為的清擦音 [ʃ] 分立的。」 (董同龢 1970: 16)
    - (b) “The Chinese *r* differs from the English *r* in two respects. We have already noted that the [retroflex] initials are pronounced with no lip action, unless followed by a rounded vowel. The other difference is that it is shorter and has more friction.” (Chao 1948: 21)
  - Approximant [ɹ]: closer to “r” (c.f. English “run”) [Cheng 1973; Norman 1988; Kuo 1994; Duanmu 2000]
    - (a) 「外國人學漢語，把『然』讀成 [ʒan]，把『人』讀成 [zən]，是很難聽的。其所以難聽，是因為讀得太『強』了，摩擦太大了。」 (朱曉農 1982)
    - (b) 「普通話日母 [ɹ] 和英語的 [ɹ] 只有兩點不同：第一，發音部位方面，英語的 [ɹ] 是個舌尖前音，普通話日母是個舌尖後音；第二，發音方法方面，典型的英語 [ɹ] 是圓唇化的。」 (王力 1983: 21)
    - (c) “the Chinese *r* is pronounced with less friction than the comparable English fricative, and acoustically sounds much closer to the usual American pronunciation of *r*.” (Norman 1988: 140)
    - (d) 「r 聲母的發音方法差別很顯著。根據摩擦程度的不同，可以把所有變體分為強摩擦、弱摩擦和無摩擦三類 ... 不僅不同人發的 r 聲母可以有濁擦音和無擦通音之別，就是同一人發的 r 聲母，也往往有很大的差別：有人發的 r 聲母既有濁擦音又有無擦通音；也有人發的 r 聲母雖然都是濁擦音，但摩擦的強弱程度相差很大。」 (廖蓉蓉、石鋒 1994: 192)

## 5. Dialect Variation

### English

#### 1. American vs British “er”

- *Colloquial Chinese* (Kan Qian, 1995: 3): “e – like *ir* in *Sir*”
- *Colloquial Chinese* (T’ung and Pollard 1982: 9): “e – as in *her*; ê – as in *there*”

#### 2. “hw” preserving vs “hw”-less dialects

Chinese

1. 我、果、說 : [o] vs [ɔ]
2. 文、完、為 : [w] vs [v]
3. 丙、景、行 : [iŋ] vs [iəŋ]
4. 煙、前、圓 : [iæn], [yæn] vs [iɛn], [yɛn]

**7. Is pronunciation worth teaching?**

1. No, because:

- CRITICAL AGE HYPOTHESIS: No adult ever achieves perfect native pronunciation in L2 (Scovel 1969) – free dinner to anyone who can show him an individual who learned a L2 after puberty and who now speaks that L2 with perfect native pronunciation (Scovel, in Tarone 1987: 80).  
POSSIBLE CAUSES --
  - (a) *Social Pressure*: “social pressures for phonological conformity and the ways these are manifested, for instance, might be different for children than for older learners” (Tarone 1978, in Jones 1997)
  - (b) *Cognitive Development*: The close of the critical period is related to the onset of Piaget’s stage of formal operations (Krashen 1977; Rosansky 1975). In this stage of cognitive development, adolescents begin to consciously construct abstract theories of the world. (Tarone 1987: 81).
- Pronunciation is an acquired skill, and focused instruction is at best useless and at worst detrimental (Krashen 1982).

2. Yes

- ANALYTICAL ACTIVITIES: “Whereas imitation activities might be more successful with younger learners, older learners might benefit from a more descriptive or analytical approach” (Brown 1992, in Jones 1997)
- INCREASING EMPATHY LEVEL:
  - (a) “the more learners identify with native speakers of a language, the more likely they are to sound like native speakers. Conversely, learners who wish to retain identification with their own culture or social category may consciously or unconsciously retain a foreign accent as a marker of in-group affiliation” (Jones 1997)
  - (b) Increasing empathy level with alcohol: “The learners’ pronunciation of the target language improved up to a certain point and then, as subjects drank greater amounts of alcohol, rapidly deteriorated”. (Guiora *et al* 1972, in Tarone 1987)

## References

- Baddeley, A. D. 1986. *Working Memory*. Oxford: Oxford University Press.
- Brown, A. 1992. "Twenty Questions". In Brown, A., ed., *Approaches to Teaching Pronunciation*. London: MacMillan.
- Catford, John C. 1977. *Fundamental Problems in Phonetics*. Bloomington, IN: Indiana University Press.
- Chao, Yuen-Ren [趙元任]. 1948. *Mandarin Primer: An Intensive Course in Spoken Chinese*. Cambridge, MA: Harvard University Press.
- Chao, Yuen-Ren [趙元任]. 1968. *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.
- Cheng, Chin-chuan [鄭錦全]. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton de Gruyter.
- Duanmu, San [端木三]. 2000. *The Phonology of Standard Chinese*. New York: Oxford University Press.
- Gathercole, Susan E. 1998. "The Structure and Functioning of Phonological Short-Term Memory". *Actes du colloque ACFOS 2* (November 27-29, 1998, Paris).
- Guiora, A. et al. 1972. "The Effects of Experimentally Induced Changes in Ego States on Pronunciation Ability in a Second Language: An Exploratory Study". *Comprehensive Psychiatry* 13: 421-428.
- Halle, Morris and Stevens, Kenneth Noble. 1991. "The Postalveolar Fricatives of Polish". *MIT Research Laboratory of Electronics—Speech Communication Group Working Papers* 8: 77-94.
- Hartman, Lawton M. 1944. "The Segmental Phonemes of the Peiping Dialect". *Language* 20: 28-42.
- Hockett, Charles Francis. 1947. "Peiping Phonology". *Journal of the American Oriental Society* 67: 253-267.
- Howie, John Marshall. 1976. *Acoustic Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.
- Jones, Rodney H. 1997. "Beyond Listen and Repeat: Pronunciation Teaching Materials and Theories of Second Language Acquisition". *System: An International Journal of Educational Technology and Applied Linguistics* 25: 1-3-112.
- Karlgren, Bernhard. 1926. *Etude sur las Phonologie Chinoise*. Leiden: E. J. Brill.
- Kenworthy, J. 1987. *Teaching English Pronunciation*. London: Longman.
- Krashen, Stephen D. 1977. "Some Issues Relating to the Monitor Model". Paper presented at the TESOL Convention, Miami Beach, FL.
- Krashen, Stephen D. 1982. *Principles and Practice in Second Language Acquisition*. Fairview Park: Pergamon Press.
- Kuo, Feng-lan [郭鳳蘭]. 1994. "Aspects of Segmental Phonology and Chinese Syllable Structure". Ph.D. diss., University of Illinois at Urbana-Champaign.
- Ladefoged, Peter and Wu, Zongji. 1984. "Places of Articulation: An Investigation of Pekinese Fricatives and Affricates". *Journal of Phonetics* 12: 267-278.
- Li, Wen-Chao [李文肇]. 1999. *A Diachronically-Motivated Segmental Phonology of Mandarin Chinese*. New York: Peter Lang.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., Studdert-Kennedy. 1967. "Perception of the Speech Code". *Psychological Review* 74: 431-461.
- Lieberman, A. M., & Mattingly, I. G. 1985. "The Motor Theory of Speech Perception Revised". *Cognition* 21: 1-36.
- Lin, Yen-hwei [林燕慧]. 1989. "Autosegmental Treatment of Segmental Processes in Chinese Phonology". Ph.D. diss., University of Texas at Austin.
- Neufeld, G. 1977. "Language Learning Ability in Adults: A Study on the Acquisition of Prosodic and Articulatory Features". *Working Papers in Bilingualism* 12: 45-60.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Qian, Kan. 1995. *Colloquial Chinese: A Complete Language Course*. London: Routledge.
- Rosansky, E. 1975. "The Critical Period for the Acquisition of Language: Some Cognitive Developmental Considerations". Ph.D. diss., Harvard University.
- Scovel, T. 1969. "Foreign Accents, Language Acquisition, and Cerebral Dominance". *Language Learning* 19: 245-254.
- Service, Elisabet. 1992. "Phonology, Working Memory, and Foreign Language Learning". *Quarterly Journal of Experimental Psychology* 1992 45A (1): 21-50.
- Tahta, Sonia and Wood, Margaret. 1981. "Age Changes in the Ability to Replicate Foreign Pronunciation and Intonation". *Language and Speech* 24 (4): 363-372.

- Tarone, Elaine E. 1987. "The Phonology of Interlanguage". In Georgette Ioup and Steven H. Weinberger, eds., *Interlanguage Phonology: The Acquisition of a Second Language Sound System*: 70-85. Cambridge, MA: Newbury House Publishers.
- Tung, Jeffrey C [董昭輝]. 1983. *Chinese Syllables and English Syllables*. Taipei: Student Book Co.
- Wang, Jenny Zhijie. 1993. "The Geometry of Segmental Features in Beijing Mandarin". Ph.D. diss., University of Delaware.
- Wu, Yuwen. 1994. "Mandarin Segmental Phonology". Ph.D. diss., University of Toronto.
- Yan, Margaret Mian and Liu, Jennifer Li-chia. 1997. *Interactions I: A Cognitive Approach to Beginning Chinese*. Bloomington, IN: Indiana University Press.
- Yao, Tao-Chung, and Liu, Yuehua. 1997. *Intergrated Chinese*. Boston: Cheng & Tsui Co.

- 董同龢。1970。《漢語音韻學》。台北：中華文化出版事業委員會。
- 廖蓉蓉、石鋒。1994。「漢語普通話 r 聲母音質的實驗研究」。見石鋒、廖蓉蓉著，《語音論叢》：177-214。北京：北京語言學院出版社。
- 王力。1983。「再論日母的音質，兼論普通話聲母表」。《中國語文》1983 (1): 20-23。
- 薛鳳生。1986。《北京音系》。北京：北京語言學院出版社。
- 朱曉農。1982。「論普通話日母的音質」。《中國語文通訊》1982 (1)。