



Denveloping a Corpus-Based Online Pronunciation Learning System for Cantonese Learners of English and Mandarin

Chen Hsueh Chu, Rebecca

hsuehchu@ied.edu.hk



Presented at E-Learning in Higher Education Emerging Practices in Faculties (FHM)



The Hong Kong Institute of Education ranked 2nd in Asia and 12th in the world in Education (QS World University Rankings by Subject 2015)



Pronunciation Learning

Spoken Corpora of Chinese Learners in Hong Kong



A Corpus-Based Online English Pronunciation Learning Website

Useful resources for pronunciation learning

The HKIEd spoken corpus of China and Hong Kong English A Corpus-Based Online Pronunciation Learning System

Pronunciation practice for speakers in Hong Kong and Mainland China

The Education University of Hong Kong Library For private study or research only. Not for publication or further reproduction. Praat beginner's manual

Our spoken corpus

A spoken corpus of the English of Hong Kong and Mainland Chinese learners has been established and analyzed in order to

(a)identify Hong Kong and Mainland university students' recurrent difficulties in English pronunciation learning

(b) suggest possible remedies that may reduce or eliminate such difficulties

For private study or research only. Not for publication or further reproduction.

Our spoken corpus

This corpus contains roughly 12-hour recording data from two speech tasks

4 hours of reading-aloud
8 hours of interview
- 32 Mainland students
- 50 Hong Kong students

For private study or research only. Not for publication or further reproduction. The spoken corpus is to identify Hong Kong and mainland university students' pronunciation features in...

Three areas of segmental features
 Vowels, consonants and syllable structures

Four areas of suprasegmental features
word-level stress, pausing, linking and intonation



Users can

 browse the recordings by speakers' biographic information
 (e.g., birthplace, gender, age)

search the recordings by the pronunciation features they are interested in

For private study or research only. Not for publication or further reproduction.

A Corpus-Based Pronunciation Learning Website



Please use the Mozilla Firefox Web Browser to view the website. For private study or research only. (http://ec-concord.ied.edu.hk/phonetics_and_phonology/wordpress/

Main page of corpus website



Browse Raw Records

Hometown:	[ANY]	•	Gender:	[Either] ·	Age: [ANY]		Filter Speake
HK21	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK22	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK23	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK24	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK25	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK26	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK27	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK28	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK29	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK30	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK31	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK32	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK33	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK34	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK35	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK36	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK37	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer
HK38	Passage One	Passage	Two	Interview Q	uestion	Intervie	w Answer



-

HOW TO BROWSE SEARCH RESOURCES ABOUT

Search Passage

Suprosegmental (Passage One):

Step 1: [NONE SELECTION] ·

Step 1: Consonants Step 2: Fricatives Step 3: [NONE SELECTION]

Labiodental fricatives /v/&/f/

Post-alveolar fricatives / 3 / & / 5 /

.

.

Alveolar fricatives / s / & / z /

Glottal fricative / h /

Step 4:

* [NONE SELECTION]

©/θ/->/f/

© / θ /->/ s /

⊕ /θ/->/ t /

0/ð/->/n/

ult]

Passage Two



The Boy Who Cried Wolf

English Level: (N/A) The First Language: Cantonese The Second Language: Mandarin The Third Language: English Year of Learning English: 13 year(s)

Education Level: Bachelor Major: Liberal Studies Education

Notes: "HK" and "ML" in the speaker's name stand for "Hong Kong" and "Mainland" repectively. There was once a poor shepherd boy who used to watch his flocks in the $(/\delta/ + / d/)$ fields next to a dark forest near the $(/\delta/ + / d/)$ foot of a mountain. One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun. Raising his fist in the air, he ran down to the village shouting 'Wolf, Wolf.' As soon as they $(/\delta/ + / z/)$ heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with $(/\theta/ + / f/)$ him for a short while. This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful. However not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck. So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep. Racing down to the village, the boy of course cried out even louder than before. Unfortunately, as all the $(/\delta/ + / d/)$ villagers were convinced that $(/\delta/ + / z/)$ he was trying to fool them $(/\delta/ + / d/)$ a third $(/\theta/ + / f/)$ time, they told him, 'Go away and don't bother $(/\delta/ + / d/)$ us again.' And so the wolf had a feast.

Notes

1. The words marked in red are those specific features spotted.

The words marked in blue are those mispronounced words which are different from the featured ones.

Feature List

Syllable structure change	Consonants	Vowels
Omission of consonant(s)	Dental fricatives / ⊕ / & / ∂ / √	101
Omission of vowel(s)	Labiodental fricatives / v / & / f /	/1/
Insertion of consonant(s)	Post-alveolar fricatives / 3 / & / 5 /	/ 5:/
Show Above	Alveolar fricatives / s / & / z /	/aɪ/



Hometown: Hong Kong Gender: Male Age: 18

English Level: (N/A) The First Language: Cantonese The Second Language: English The Third Language: (N/A) Year of Learning English: 10 year(s)

Education Level: Bachelor Major: Liberal Studies Education

Notes: "HK" and "ML" in the speaker's name stand for "Hong Kong" and "Mainland" repectively.



1. The blue contour in pictures is pitch contour to indicate the intonation (pitch).

2. The <u>yellow contour</u> in pictures is intensity contour to indicate the intensity of the speech.

- 3 3. The suprasegmental features that have been spotted
 - Inappropriate pausing (IP)
 - ☑ Absence of CV linking (AL)
 - Inappropriate intonation (II)

Full Passage Recording: 🛶



Sentence 1:



Why do PEOple Usually have∩a<u>n A</u>Ccent (AL) when they spea<u>k a</u> (AL) Second LANguage ∕(II)?



Sentence 2:

Interview Answer

Speaker: HK2

Remarks:

1. The sentences in black are from interviewee, the sentences in blue are from the interviewer.

Hometown: Hong Kong Gender: Female Age: 20

2. Use "..." to indicate the fragmented sentences.

3. Use "(. . .)" to signify a word or a phrase that you cannot understand after trying several times.

English Level: (N/A) The First Language: Cantonese The Second Language: Mandarin The Third Language: English Year of Learning English: 13 year(s)

Education Level: Bachelor Major: Liberal Studies Education

Notes: "HK" and "ML" in the speaker's name stand for "Hong Kong"TandE'Mainland"repectively. of Hong Kong Library For private study or research only. Not for publication or further reproduction.

4. We put the words that are guessed from the context in the brackets ().

Speaker's Record:



[1.Interviewer: What do you think of Singapore?]

HK2: Singapore is like... I think some...it's like Hong Kong, but it has...it has a lot of the rules, the forest, and people need to...need to followed, and so for example, we cannot eat the gums in the Singapore, I think it is not very good for me, because that I...I am the one, I love freedom, and I don't want others controls me and manage me. I think that the Singapore's government do a lot to controls the people's life, but I think that the Singapore is a good place, because that he is a place to Universal (Studio), the place to watch the... I don't know how to explain. Yes, Universal Studio, that place is very beautiful and awesome. I love the Universal Studio very much. Even though I didn't go to Singapores because ..., but my friends had went to Singapore, they posted a lot of the photos about the Universal Studio, then I really really want to go there, just take a lot of the photos, yeah.

[2.Interviewer: First, tell me something about your family?]

HK2: My family is a big family; I have one sister...two...I have three sisters, and one brothers and my mother and my father. And in total my family has 7 people in there and that yes



Institute of Education A Corpus-Based Pronunciation Learning Website

Home A

About the Project Our Spoken Corpus

Learning Resources

For Chinese Learner

Home > For Chinese Learners

For Chinese Learners

- Part 1 Background of China English and Hong Kong English
- Part 2 Pronunciation of Hong Kong English
- Part 3 Pronunciation of China English

Part 4 Pronunciation Practice for Chinese Learners of English

http://ec-

concord.ied.edu.hk/phonetics_and_phonology/ wordpress/exerciseIndex.html

The Education University of Hong Kong Library For private study or research only.

Not for publication or further reproduction.

earners	Our Praat Manual Contact Us
	search here Go
	Search here Go
	Our Pages
	About the Droject
	About the Project Background of China English and
	Hong Kong English
	Comparison of English and
	Cantonese (Segmentals)
	Comparison of English and
	Chinese (Suprasegmentals)
	Comparison of English and
	Mandarin (Segmentals)
	Contact Us
	English Pronunciation of
	Cantonese speakers
	English Pronunciation of
	Mandarin speakers
	For Chinese Learners
	Learning Resources
	Online Dictionaries
	Our Spoken Corpus
	Our Praat Manual
	 Phonetics and Phonology
	Phonics



Institute of Education A Corpus-Based Pronunciation Learning Website

Home

About the Project Our Spoken Corpus

Learning Resources

urces For Chin

For Chinese Learners

Our Praat Manual

search here

Contact Us

Go

Home > Learning Resources

Learning Resources



- · Phonics full story...
- Phonetics and Phonology full story...
- Pronunciation: Useful Websites full story...
- Pronunciation: Online Videos full story..
- Typing IPA Symbols full story...
- Online Dictionaries full story...
- IPA Font Download full story...



The Education University of Hong Kong Library

For private study or research only. Not for publication or further reproduction.

Our	Pages	

About the Project

- Background of China English and Hong Kong English
- Comparison of English and Cantonese (Segmentals)
- Comparison of English and Chinese (Suprasegmentals)
- Comparison of English and Mandarin (Segmentals)
- Contact Us
- English Pronunciation of Cantonese speakers
- English Pronunciation of Mandarin speakers
- For Chinese Learners
- Learning Resources
- Online Dictionaries
- Our Spoken Corpus
- Our Praat Manual
- Phonetics and Phonology
- Phonics

Praat Beginners' Manual



- Chapter 1. Introduction: Acoustic Phonetics and Praat
- Chapter 2. Acoustic Analysis of Vowels
- Chapter 3. Acoustic Analysis of Consonants
- Chapter 4. Prosodic Analysis with Praat
- Chapter 5. Annotating Sound Files with Praat

Suggested Answers

Online Resources

1. Other recommended Praat manuals 2. Extra exercises, tutorials, and self-tests The Education University resources: charts&fonts

For private study or research only. Not for publication or further reproduction.

of Hong Kong Library

For Undergraduates:

Students of HKIEd are encouraged to use this system during or after class for self-access to practice.

In my Year One phonology course, for example, students were encouraged to use the online pronunciation website that our team created to conduct their 10-week self-access learning activities and prepare reflective reports with their pronunciation learning portfolios.

For private study or research only. Not for publication or further reproduction.

For teaching training program:

In PGDE phonology courses, students were required to collect sound samples, research the first language phonology and its impact on second language, and share their knowledge of specific speech communities with the class. The recommended spoken corpus was one of the resources for them to collect to analyse the

For private study or research only. Not for publication or further reproduction.

For Graduate students:

A workshop on Praat applications was held at HKIEd campus to help students learn some hands-on techniques in Praat to analyze their speech data.



For Frontline Teachers:

This website provides abundant pronunciation teaching resources and pronunciation practice for teachers to use in their classes.

The Education University of Hong Kong Library For private study or research only. Not for publication or further reproduction.

A Corpus-Based Online Mandarin Pronunciation Learning Website



Not for publication or further reparties://ec-concord.ied.edu.hk/Mandarin Pronunciation/

The Education University of Hong Kong Library For private study or research only.

Not for publication or further reproduction. http://ec-concord.ied.edu.hk/Mandarin Pronunciation/

Our Spoken Corpus

The spoken corpus contains:

The Hong Kong Institute of Education 乔港教育學院



Online Mandarin Pronunciation Learning A Spoken Corpus of the Mandarin of Hong Kong learners

- 8 hours recording data of Mandarin spoken tasks (reading of monosyllabic words, reading of multisyllabic words, reading of passage, and free speech) from Hong Kong university students
- phonological annotations of the recordings mainly focused on segmental features (vowels & consonants)
 The Education University and Suprase gmental features (tone & retroflex finals)

Our Spoken Corpus - Browse



Home How To About Browse Search Resources Login



Online Mandarin Pronunciation Learning

A Spoken Corpus of the Mandarin of Hong Kong learners



In The blood during memory or Encoded on Electronic 2019



Our Spoken Corpus - Search



Home How To About Browse Search Resources Login

2



Online Mandarin Pronunciation Learning

A Spoken Corpus of the Mandarin of Hong Kong learners



© The Hong Kong Institute of Education | December 2



Our Spoken Corpus –

Task 1: Reading of monosyllabic words

Tasks 1&2 Word Reading

Remarks : Please read the below marking note for details. Speaker: 23 Gender : Female 0:57 🌒 🕳 Speaker's Record: Download Age: 21 1:33 🌒 🕳 Beijing Sample: Download Education Level : 0:57 🌒 🕳 Taiwan Sample: Download Bachelor Major : English Marking Scheme of Reading Task 1 First Language : 稀 佳 车 垮 勢 恆 Cantonese Second Language : English 吼 憂 翁 爾 哲 氧 Third Language : Mandarin Year of Mandarin 謊 罿 帥 徐 蹲 卓 Learning: 10 軒 狼 阻 瀉 矯 吝 Feature Color : Markings in Red: Vowel 淙 胚 恩 惨 凶 屏 Features Markings in Orange: **Consonant Features** 쁿 寺 竣 宰 嶽 綴 Markings in Purple: Suprasegmental Features 勃 The Markingsin Giern University 淹 倍 喚 of Hispron Micistions Library For private study or research only. Not for publication or further reproduction:

Our Spoken Corpus

Task 1: Reading of monosyllabic words



Our Spoken Corpus – Spoken Tasks

Task 2: Reading of multisyllabic words



Marking Scheme of Reading Task 2

不用	包括	徘徊	狹窄	狹 窄	反省
踐踏	汛期	轉臉	瓦房	兒歌	耕耘
炎 熱	沉著	提 防	提 防	圓圈兒	謙 譲
遵 守	遵守	挑戰	舒服	電影兒	性質
研究	等 於	農村	概念	分水 嶺	視網膜



Our Spoken Corpus

Task 2: Reading of multisyllabic words



Marking Scheme of Reading Task 2

	不用	包括	徘徊	狹窄 xiá→xié	狹窄 zhǎi→zǎi	反省 Incomplete Tone 3
	踐踏	汛期	轉臉 Incomplete Tone 3	瓦房	兒歌	耕耘
	炎熱 rè→rì	沉著 <mark>zhuó→zuó</mark>	提防 dī→tí	提防 fang→fáng	圓圈兒	謙 讓 ràng→yàng
	遵守 zūn→zhūn	遵守 <mark>shǒu→xǒu</mark>	挑戰	舒服	電影兒	性質
	研究	等 於 děng→děn	農村	概念	分水嶺 Incomplete Tone 3	視網膜
rivat	The E海京蜃樓Uni of F Shèi、ochUil y r e study or research on iblication or further re	ly.				

Our Spoken Corpus

Task 3: Reading of passage



Marking Scheme of Reading Task 3





Marking Scheme of Reading Task 3

有一年的春天,我因病被	迫在家裏休 息 數周。我注 視 著我的女兒們所養的蠶正在結 繭,這 使
	地、勤奮地工作,我感到我和它 們 非常相似。像它們一樣,我總是
	固目標上。我之以如此, 或 許是因為有某種力量在鞭策 著 我——正
private study or research only. t for pu如靈磁鞭策審密結繭du般an.	

Our Spoken Corpus – Spoken Tasks

Task 3: Reading of passage

生活對於(yú→yū)任何人都非易事。我們必須相信,我們對每一件事情都具有天賦(fù→fú) 的才(cái→chái)能,並且(Incomplete Tone 3)無論付出任何代價(jià→jiè),都要把這件事完 成。當事情結束(shù→sù)的時候(hou→hòu),你要能問心無愧地說:"我已經盡我所 (suŏ→shuŏ)能了。"

有一年的春天,我因病被迫在家裏休息(xi→xī)數周。我注視(shì→shī)著我的女兒們所養的 蠶(cán→chán)正在結繭(Incomplete Tone 3),這(zhe→zhe'r)使我很感興趣(qu→qù)。望著這 些蠶(cán→jiǎn)執著地、勤奮地工作,我感到我和它們(men→mēn)非常相似。像它們一樣,我 總是耐心(xīn→xīng)地把自己的努力集中(zhōng→zōng)在一個目標上(shang→shàng)。我之 (zhī→zī)以如此,或(huò→huà)許是因為有某種力量在鞭策著(zhe→ze)我—正如蠶

(cán→chán)被鞭策著去結繭一般。



For private study or research only. Not for publication or further reproduction.

Our Spoken Corpus

► Task 4: Free speech

Task 4 Free Speech

Speaker : 23 Gender : Female	Speaker's Record: 🕨 🗨 1:32 📢 🚥 Download
Age : 21	Topic: 我的學習生活
Education Level : Bachelor	
Major : English	我的學習生活在小學的時候是非常的煩悶,因為小學的生活很簡單,老師都會為我們安排了一
First Language : Cantonese	切。所以小時候的生活主要就是唸書,然後應付考試、測驗,等等。父母對我的要求也并不是很
Second Language : English	高,不像是現在那樣很多的怪獸父母都會要求自己的孩子去上什麽什麽補習班。我當時是非常自
Third Language : Mandarin	由,總是隨自己意願去做的。所以雖然煩悶,但是那些測驗、考試都是能應付得了的。最後就很
Year of Mandarin	成功地上了一所挺不錯的第一等級的中學。在中學上面,中學的學習生活比小學的步伐快很多,
Learning : 10	測驗的次數的還有那個頻率都比小學來的高好幾倍,所以我在中學的學習生活都是在非常繁忙
LI	的考驗和測試中度過。學習生活之中我有遇到過幾個恩師,他們對我的學科上面,讓某些學科
	上面的知識增長了不少。尤其是高中最後的一年,我遇上發現在我身邊的老師其實都是個寶
The Education Univers	
of Hong Kong Library private study or research only. for publication or further reprod	從我在那個高中那個公開考試裡面就更好地掌握了。

Pronunciation Practice



Not for publication or further reproduction.

Pronunciation Practice



Not for publication or further reproduction.

Pronunciation Practice – Tongue Twisters

standard Hong Kong Library

	ABOUT 關於 SPOKEN CORPUS 語料庫	i FOR CANTONESE LEARNERS 學習	LEARNING RESOURCES 資源
TONGUE TWISTERS	HOME主頁		
Short Vowels 1. 坡上立著一只鵝(é),坡下就是一條河(hé)。〕 河,還是河渡鵝?	寬寬的河,肥肥的鵝,鵝要過河,河	Search 要渡鵝,不知是鵝過 OUR PAG	ES
Sample Reading		Our Spoker	
Slow: • 00:00		00:00	arning Resources ese Learners
Standard:		Our Praat N	Ianual
▶ 00:00		00:00	Learning Websites
			iation Learning Website
2. 山上五棵樹(shù),架上五壺醋(cù),林中五 醋,射死林中的鹿,取出箱中的褲。	只鹿(lù),箱裏五條褲(kù)。伐了山	上樹,搬下架上的	504
Sample Reading			Visits
Slow:		00:00 🛋	

00:00 🛋

Pronunciation Practice - Songs



HK&Overseas: https://www.youtube.com/watch?v=7CKOG-qQKJY Mainland: http://www.letv.com/ptv/vplay/21820202.html



Pronunciation Practice - Jokes

		ABOUT 關於	SPOKEN CORPUS 語料庫	FOR CANTONESE LEARNERS 學習	LEARNING RESOURCES 資源
JOKES		HOME主頁	Jokes		

1. 有一天,小明在家中看電視,媽媽說:"等一下你去沖個涼吧,用<mark>發(fā)</mark>灑,小明不明白,問:"什麼是發 灑?",媽媽想了想說:"哦,我搞錯了,應該是<mark>花(huā)</mark>灑!



2. 小明約了朋友見面,遲到了,他和朋友說:"不好意思,剛剛和媽媽吃了辦!"朋友不解,問:"吃花<mark>辦</mark> (bàn)?"小明意識到自己說錯了:"哦不是,是吃**飯(fàn)**!不要笑話我了哈哈!



Search ...

OUR PAGES

Our Spoken Corpus

General Learning Resources

For Cantonese Learners

Our Praat Manual

Recommended Learning Websites

ciation Learning Website



Pronunciation Practice - Jokes

Online Exercise

Your Score : \$-





Useful Resources



Learning Resources



- Chinese phonology <u>full story...</u>
- Recommended Learning Websites <u>full story...</u>



Search ...

OUR PAGES

Our Spoken Corpus

General Learning Resources

For Cantonese Learners

Our Praat Manual

commended Learning Websites

English Pronunciation Learning Website



Useful Resources

ABOUT 關於 SPOKEN CORPUS 語料庫 FOR CANTONESE LEARNERS 學習

RNERS 學習 LEARNING RESOURCES 資源

HOME 主頁 RECOMMENDED LEARNING WEBSITES



Websites

• 拼音導修室網站[Pinyin Learning Website]

http://www.cbs.polyu.edu.hk/VTP/mandarin/main.htm

The website provides systematic introduction of Pinyin system, including syllable structure, correcpondence to IPA system and so on. Learners can access a series of exercises whilst learning Pinyin.

• 普通話學習網[Putonghua Learning Website]

http://pth.linqi.org/yuyin/index.html

This website compiles useful learning materials of Mandarin, including Pinyin system, place of articulation, segmental and suprasegmental features of Mandrin. Moreover, vocabularies, correspondencee of Cantonese and Mandarin, learning exercises and useful Mandarin learning tools are introduced.

T]e普通話世界[Butonshua World]

of Hong Kong Library

For private study or research only.

Not for pub Ratatighaayofiditia platferm bulltlatin expert who has taught Mandarin for years. Mandarin teaching experience,

Search ...

OUR PAGES

Our Spoken Corpus

General Learning Resources

For Cantonese Learners

Our Praat Manual

Er

Recommended Learning Websites

Website



Useful Resources



Videos in Chinese

1. Articulation of tones, j, q, x, n, l

https://www.youtube.com/user/zeq160/videos

These videos teach in detail how tones, and difficult sounds such as j, q, x, n, and l in Mandarin should be produced with help of model reading.



2. 50 common mistakes in speaking Mandarin

https://www.youtube.com/watch?v=uKV8ta06PI4

The video lists 50 common mistakes for Cantonese speakers in speaking Mandarin, with model reading.



2	38		
0	ar	з.	

OUR PAGES

Our Spoken Corpus

General Learning Resources

For Cantonese Learners

Our Praat Manual

Recommended Learning Websites

n Learning Website



Rebecca CHEN

Thank you!

hsuehchu@ied.edu.hk

For private study or research only. Not for publication or further reproduction.

Education University