

Preface

Tongue Twister

[s+4 f+4 s+4, f+2 f+4 f+2;
4 is 4, 10 is 10.
f+2 s+4 f+4 f+2 s+4, s+4 f+2 f+4 s+4 f+2]
14 is 14, 40 is 40.

Story of Stone Grotto: Eating Lions by Yuanren Zhao

[f+2 f+4 f+1 f+4 f+1 f+4, f+4 f+1, f+4 f+2 f+2 f+1.
In a stone den was a poet called Shishi, a lion aficionado, sweared to eat ten lions.
f+4 f+2 f+2 f+4 f+4 f+4 f+1.
He often went to the market to look for lions.
f+2 f+2, f+4 f+2 f+1 f+4 f+4.
At 10:00, 10 lions had just arrived at the market.
f+4 f+2, f+4 f+1 f+4 f+4 f+4.
At that time, he had just arrived at the market.
f+4 f+4 f+4 f+2 f+1, f+4 f+3 f+4, f+3 f+4 f+2 f+1 f+4 f+4.
He saw those 10 lions, and using his trusty arrows, caused the ten lions to die.
f+4 f+4 f+4 f+2 f+1 f+1, f+4 f+2 f+4.
He brought the corpses of the 10 lions to the stone den.
f+2 f+4 f+1, f+4 f+3 f+4 f+4 f+2 f+4.
Since the stone den was damp, he asked his servants to wipe it.
f+2 f+4 f+4, f+4 f+3 f+4 f+4 f+2 f+4 f+1.
After the stone den was wiped, he tried to eat those ten lions.
f+2 f+2, f+3 f+2 f+4 f+2 f+1 f+1, f+2 f+2 f+2 f+1 f+1.
When he ate, he realized that these ten lions were in fact ten stone lion corpses.
f+4 f+4 f+4 f+4]
Try to explain this story.

Representations of Mandarin Sibilant Homorganic Syllables¹

Introduction

From the tongue twister and poem in the preface, we can see that a lot of words in modern Mandarin sound like [st] and [ʃt]. What are their surface and underlying representations? In the literature, there has been many proposed underlying form, but many phonologists did not give detailed reasoning, therefore creating a gap in this field. Among those who did, their only evidence was the complementary distribution of [t̪] and [i]. I contribute to the literature by using evidence from loanwords, L2 production, place licensing. And to date, I am the only who has done so.

¹ Professor Heather Goad has taught me this course and provided valuable feedback to me for this paper. For this, I am incredibly grateful. All errors are my own.

Section 1: Data

In this chapter, I present some data from Mandarin, since the reader might not be a native speaker. I strive to be theory-neutral in this section, and reserve my theory-based analyses to later chapters.

Transcription practices

- The data on this page is accurate for vast majority of, if not all, Mandarin dialects. Cantonese, on the other hand, is entirely different.
- [t̪] is used as a shorthand to indicate a sound that is similar to [t̪]. In later sections, I will reject [t̪].
- Whether Mandarin has [ŋ] is disputed, and there is likely variation between speakers, or even within 1 speaker ([ẅ] freely varies with [ŋ]). I chose not to include it in the allophone inventory, but I included it in the transcriptions. Whether Mandarin has /ŋ/ is even less known.
- Nasalization is not transcribed.
- I assume that there are no retroflex consonants, but only flat post-alveolar ones.
- I have very limited knowledge of Cantonese tone system, therefore I do not transcribe them.
- I did not transcribe English words, and please assume that the English input is General American.

Mandarin Allophone Inventory, Simple Segments

For each place of articulation, the leftmost column is voiceless aspirated, the middle column is voiceless plain, and the rightmost column is voiced plain.

	Bilabial		Labiodental		Alveolar		Post-alveolar		Palatal		Velar	
Stop	p ^h	p			t ^h	t					k ^h	k
Nasal			m				n					

	Bilabial	Labiodental	Alveolar	Post-alveolar	Palatal	Velar
Fricative	f		s	ʃ	ç	
Tap/flap		v				
Approximant	w			ɹ	j	
Lateral approximant			l			

Mandarin Allophone Inventory, Complex Segments¹

	Bilabial	Labiodental	Alveolar	Post-alveolar	Palatal	Velar
Affricate	t ^h s	ts	t ^h ʃ	tʃ	t ^h ç	tç

Attested [Ci] and [Ct] Words

	Bilabial			Labiodental			Alveolar		
Stop	p ^h {✓i ✗t}	p {✓i ✗t}					t ^h {✓i ✗t}	t {✓i ✗t}	
Nasal			m {✓i ✗t}						n {✓i ✗t}
Affricate							t ^h s {✗i ✓t}	ts {✗i ✓t}	
Fricative				f {✗i ✗t}				s {✗i ✓t}	
Tap/flap						v {✗i ✗t}			

¹ Segments that are composed of more than 1 melody

	Bilabial			Labiodental			Alveolar		
Approximant									
Lateral approximant									l {√i X+i}
Post-alveolar			Palatal			Velar			
t ^h f {X+i √i}	tʃ {X+i √i}		t ^h ç {X+i √i}	tç {X+i √i}		k ^h {√i X+i}	k {√i X+i}		
	ʃ {X+i √i}			ç {X+i √i}					
		d {X+i √i}				j {√i X+i}			

Not listed above:

w {X+i|X+i}

ɥ {X+i|X+i}

Please note that “unattested” does not entail “productively illicit”. For example, the syllable [vu] is unattested in English, but it is still productively well-formed².

Some Words in Mandarin

² I can invent a product and call it [vú.tù.pík], and it would sound natural.

The words are expressed in features.

[COR]				
	[+ant]		[-ant]	
	[+SG]	[-SG]	[+SG]	[-SG]
[+cont]		Mandarin: [sɿ4] Cantonese: [seɪ] or [si] 四 “Four”		Mandarin: [ʃɿ2] Cantonese: [sap̚] 十 “Ten”
[-cont]	Mandarin: [tɿsɿ2] Cantonese: [tɿsi] 磁 “Magnet”	Mandarin: [tsɿ3] Cantonese: [tsi] 紫 “Purple”	Mandarin: [tɿʃɿ1] Cantonese: [hɛk̚] 吃 “To eat”	Mandarin: [tɿʃɿ3] Cantonese: [tsi] 纸 “Paper”

Non-Native Words

In China, English is the most popular second language to learn. However, Chinese people usually have a lot of difficulty with it, and very few people take it seriously, which means that they do not speak it well, if at all. Therefore, all English words are rather thoroughly adapted to Mandarin phonology.

L1 Mandarin Speakers' Production of English words

English	Mandarin	
(George) Bush	[pu4.ʃɿ2]	
(Tom) Cruise	[kɿə4.lu3.sɿ1]	

English	Mandarin	
Boat	[pəʊt̪əʊt̪]	
Bushy	[bu1.çi0]	
C	[seɪ1], [çɪ1]	
Cascade	[kʰa3.sj0.kʰeɪ4.t̪əʊ]	
Children	[t̪çu.dren]	
G	[t̪çɪ4]	
Genome	[t̪çɪ1.nəʊt̪əʊmju0]	
H	[er2.t̪ʃ]	
Issue	[ji.çu]	
J	[t̪ʃeɪ4]	
Leaf	[l̪ji4.fu0]	
M	[ar2.mu0]	
Map	[maɪ4.pʰu0]	
Pepsi	[pai1.pu.çɪ]	
Q	[kʰioʊ4], [t̪çou4]	
Sheet	[çɪ4.t̪əʊt̪]	
Spring	[s̪j0.pu0.rɪŋ4]	
Stop	[s̪j0.taʊ4.pʰu0]	✗[si0.taʊ4.pʰu0]

English	Mandarin	
Truth	[tʰʃu4.si0]	✗ [tʃu4.si0]
U	[jou1]	✗ [iu1]
Z	[tseɪ4]	✗ [tɕi4]
Zebra	[tseɪ1.pu0.ra0]	✗ [tsɿ1.pu0.ra0] ✗ [tsi1.pu0.ra0]

[LAB] codas receive [u̯] as epenthesis.

[COR] codas receive [ɨ] or [ə] as epenthesis.

All other codas receive [ə] as epenthesis.

None of the epenthetical vowels receive tone. 0 indicates the complete absence of tone.

These voiceless vowels are very short in duration, compared to voiced ones in the previous syllable.

Data Authenticity

To prove that the data in this paper is not fabricated, I have attached a sample of natural speech, by a female graduate student at McGill, as a WAV file. Furthermore, you may verify any part of the data with any native speaker of Mandarin, should you have doubts about its authenticity.

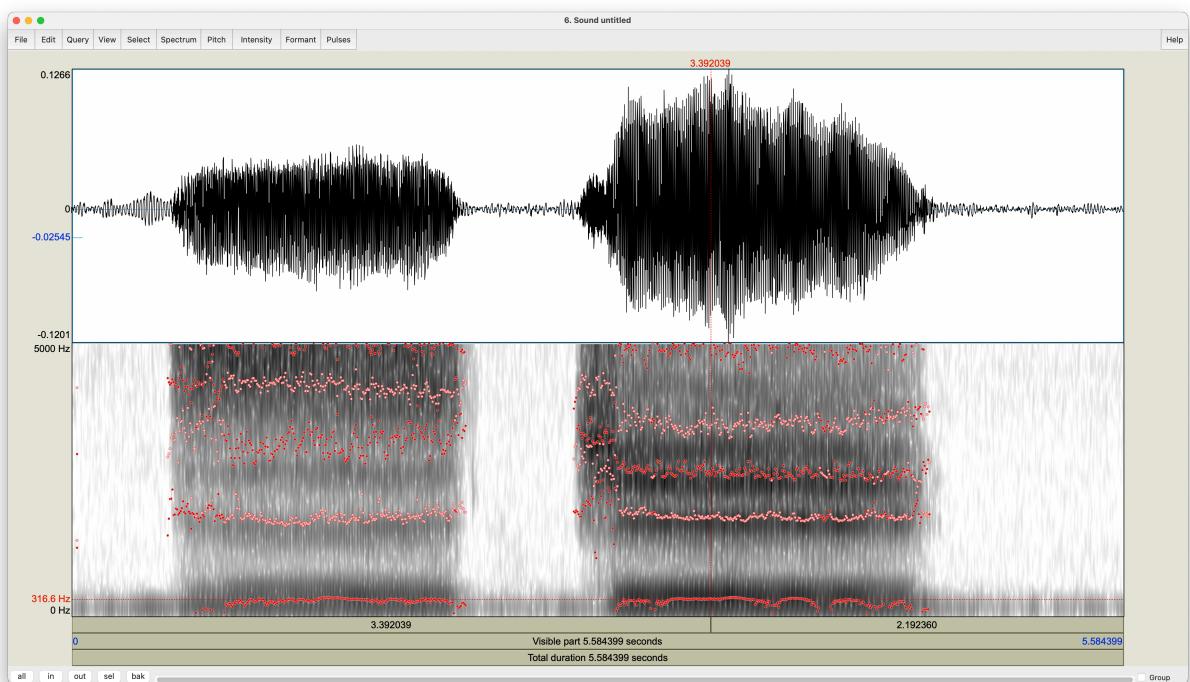
Section 2: Theoretical Assumptions

Abstract

In this section, I present my assumptions on features and feature geometry. I borrow and adapt the theory which (Clements & Hume 1995) presented.

Data

I produced the voiced and geminate counterparts of [s] and [ʃ], which are [z:] and [ʒ:]. I sustained them for 3 seconds.



Is [i] CORAL?

Many, such as (Pulleyblank 1989), have argued that front vowels are coronal. Like (Goad & Ramasimhan 1994), I argue that [i] is not coronal, with evidence from acoustics

[z:] and [ʒ:] are coronal continuants, and their F2s are about 1768 Hz. On the contrary, [i]'s F1 is much higher, between 2400 Hz and 2700 Hz, depending on the speaker and language. Therefore, they are not identical in place.

Furthermore, there is a lot of tongue movement in words that contain [zi] and [ʒi], for example, “zebra” and “genius”. On the contrary, there is barely any tongue movement for words like “yeast”, “year”, “yearn”, etc. This is another evidence that [i] is [DOR], not [COR].

Since [COR] is a feature defined on articulation, it is necessary to use an ultrasound to visualize tongue position of Mandarin speakers when they produce [i]. Evidence from distribution (phonological evidence) and acoustics (phonetic evidence) is not enough. Due to the absence of funding and an ongoing pandemic, I am not able to collect this kind of data. I leave this to future researchers.

Acoustic definition of [±front] and [±back]

To further articulate my theory about the [PLACE] node, I provide an acoustic definition of [±front] and [±back].

A vowel is [+front] if its 2nd formant is around 2600 Hz.

A vowel is [-front] and [-back] if its 2nd formant is around 1865 Hz.

A vowel is [+back] if its 2nd formant is around 1080 Hz.

The frequencies are based on a 24-year-old female participant.

[C-Place] dominates [LAB], [COR], and [DOR], which are defined by articulation. [V-Place] dominates [±front] and [±back], which are defined by acoustics.

Do [s] and [ʃ] have [HEIGHT]?

In many feature geometries, non-vocoids are said to lack height. On the contrary, I believe that any (voiced) continuant consonant may have height.

On the spectrogram, they have a clear F1, which is around 317 Hz. Its F1 is higher than [i]’s F1, but lower than [ɪ]’s F1. Later, I will argue that it is the [HEIGHT] features of [s] and [ʃ] that allow a vowel to be borne from them.

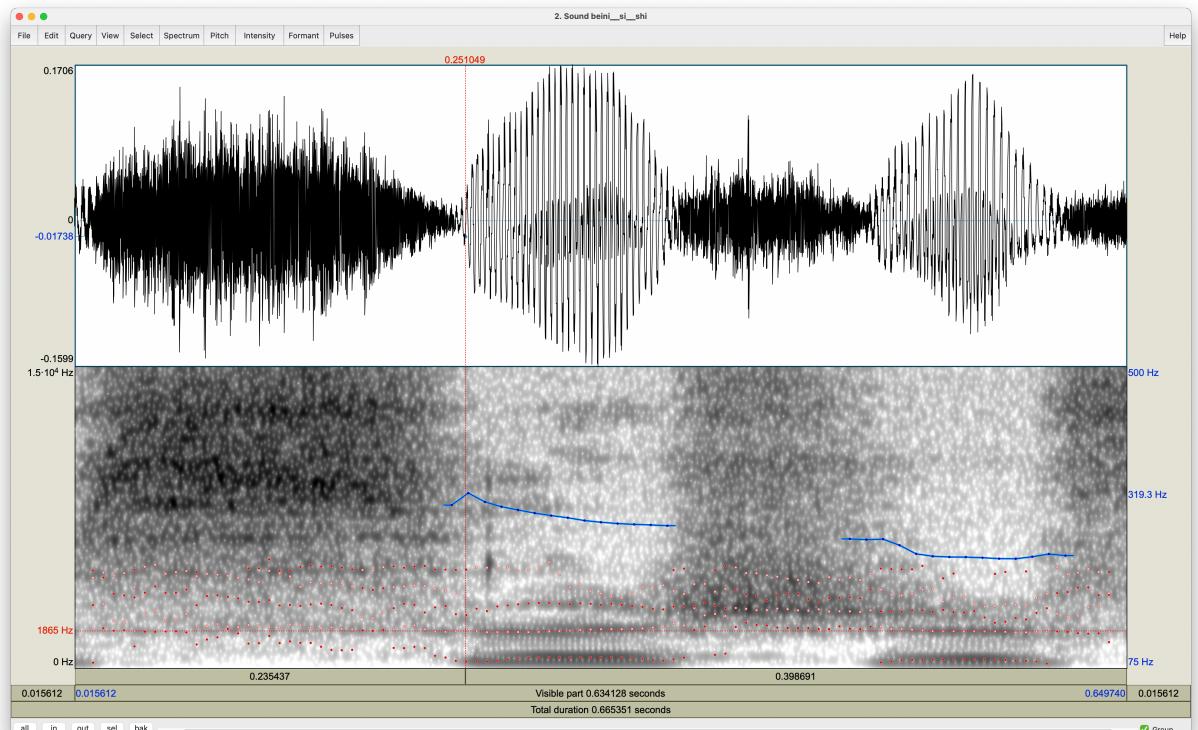
Section 3: Surface Form

In this section, I evaluate some surface transcription proposed by others, then propose my own.

All phoneticians agree that when producing the syllable, the tongue body does not move (in other words, homorganic), and the tongue tip is slightly lowered to remove constriction.

Data

Acoustic data has been collected through 3 Mandarin speakers. To ensure the most natural production, speakers were asked to translate words from English into Mandarin. Due to the high frequency of these syllables in Mandarin, I have collected significantly more data than I needed. Since the recording did not happen in sound-proof booths, 2 samples have a significant background noise. The optimal sample, recorded by a 24-year-old female graduate student was used for acoustic analysis.



Spectrogram for naturally-produced word [sì shí], which means “forty”.

[s_z, ts_z, t^hs_z, ʃ₃, tʃ₃, t^hʃ₃]

(Wiese 1997) and (Duanmu 2000) transcribed these syllables as [s_z, ts_z, t^hs_z, ʃ₃, tʃ₃, t^hʃ₃].

If the nucleus was a fricative, then we should expect to see frication darkness carried over to the formants. However, the frication ends as soon as the formants begin, and the area above the formants are clear. Therefore, I argue that the nucleus is a vowel.

I further tested this by producing words as [s_z, ts_z, t^hs_z, ʃ₃, tʃ₃, t^hʃ₃] and letting participants judge how natural they sound. Participants agreed that they sound “abnormal” and “not like Chinese”.

[s_t, ts_t, t^hs_t, ʃ_t, tʃ_t, t^hʃ_t]

(Lee & Zee 2003) and (Lee-Kim 2014) transcribed them as [s_t, ts_t, t^hs_t, ʃ_t, tʃ_t, t^hʃ_t]. (Lee & Zee 2003) specified that the [t] in [s_t, ts_t, t^hs_t] is apico-laminal or laminal denti-alveolar, and the [t] in [ʃ_t, tʃ_t, t^hʃ_t] is apical post-alveolar.

If the nucleus was indeed [t], then we should expect to see significant movement of F3, and a grey area above the formants indicating that the air is partially constricted. However, F3 is barely visible, and the area above the formants are clear, which suggests that there is no constriction whatsoever during the nucleus. Therefore, I argue that it is a vowel. (Ladefoged & Maddieson 117:1996) agrees with this view.

I measured the formants in the female student’s production. Her F1 was around 491 Hz, and F2 was around 1769 Hz. Currently, no symbol on the IPA vowel chart captures this exactly, and I do not believe that [t̪] is close enough. Therefore, I choose to create [🎄] as an ExtIPA symbol. I define [🎄] as a vowel whose F1 ≈ 491 Hz, and F2 ≈ 1769 Hz.

In this section, I have argued that these 6 syllables are best transcribed as [s^h , ts^h , t^hs^h , f^h , t^hf^h].

Section 4: Underlying Form

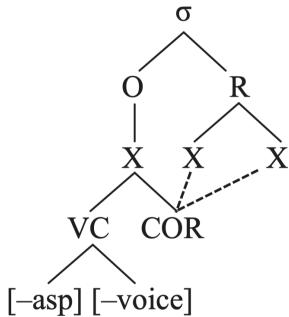
In this section, I evaluate some underlying forms proposed by others, then propose my own.

/s₁, ts₁, t^hs₁, f₁, t₁f₁, t^hf₁/

(Cheng 1973) believes that /  / is a phoneme that is independent from / i /. Under this view, Mandarin has 6 phonemes, / y /, / ə /, / a /, / u /, / i /, /  /.

/s, ts, t^hs, f, t₁f, t^hf/

(Duanmu 2000) believes that underlyingly, the vowel occupies the nucleus slot. This is possible, since / s / has been reported as a vocoid in other languages (Goad & Shimada 2014). He provided the derivation¹:



(Duanmu 2000) assumes that the surface forms are [s₁, ts₁, t^hs₁, f₁, t₁f₁, t^hf₁]. However, as I have shown in the last section, the surface forms are not syllabic fricatives.

/si, tsi, t^hsi, fi, t₁fi, t^hfi/

(Ladefoged & Maddieson 1996) and (Cheng 1966) argued that these 6 syllables are underlyingly / si, tsi, t^hsi, fi, t₁fi, t^hfi /. This argument is based on the observation that [] and [i] are in complementary distribution, and furthermore, [i] is the ONLY

¹ His theory of features geometry differs slightly from mine.

allophone that is entirely in complementary distribution with [i].² The derivation can be expressed with either one of the SPE rules³ below.

/i/ → [i] / [s, f] __

/i/ becomes [i] when it precedes coronal fricatives.

/i/ → [- VPlace] / [s, f] __

/i/ loses [VPlace] when it precedes coronal fricatives.

Under this view, Mandarin has 5 phonemes, /y/, /ə/, /a/, /u/, /i/.

Criticisms of Above Analyses

Please refer to “Attested [Ci] and [C̄i] Words” and “L1 Mandarin Speakers’ Production of English words” for evidence.

In Mandarin, [s, ts, tʰs, f, tʃ, tʰʃ] are licit onsets, and [i] is a licit nucleus. However, [si, tsi, tʰsi, fi, tʃi, tʰʃi] are not attested syllables⁴, and in fact, they are not even LICIT syllables⁵. Then, when a Mandarin speaker tries to produce English syllables [si], [zi], [fi], [ʒi], what would it sound like?

If /i/ is the underlying form of [i], then a Mandarin speaker should produce an English syllables [si], [zi], [fi], [ʒi] as [s i], [tʃ i], [f i], [tʃ ʒ i] respectively. However, Mandarin speakers produce them as [seɪ], [tseɪ], [sɪ], [tʃeɪ] instead. This vaguely suggests that /i/ may not be the underlying form of [i]. Then, for which English word do Mandarin speakers actually produce [s i, ts i, tʰs i, f i, tʃ i, tʰʃ i]?

² This is perfectly illustrated by the table “Attested [Ci] and [C̄i] Words” in Section 1.

³ No phonologist yet has tried to account for this derivation with Optimality Theory.

⁴ Evidence: refer to “Attested [Ci] and [C̄i] Words”

⁵ Evidence: refer to “L1 Mandarin Speakers’ Production of English words”

Refer to the words “(Tom) Cruise, (George) Bush, truth, cascade, spring, stop”. For [s, z, t^hs, f, ʒ, t^hf] in coda position⁶ (with no following vowel), Mandarin speakers produce [s 3, ts 3, t^hs 3, f 3, t^hf 3]⁷. This vaguely suggests that the underlying form of [3] is /Ø/, a silent vowel.

Below, I further refute /si, tsi, t^hsi, fi, t^hfi, t^hʃi/ and /s 3, ts 3, t^hs 3, f 3, t^hf 3/ with 2 arguments: place licensing and principle of Economy⁸. I also compare it to Spanish to demonstrate that, contrary to popular belief, this phenomenon in Mandarin is not peculiar, but in fact very common cross-linguistically.

Spanish	Mandarin
In Spanish, [N.P] clusters are homorganic.	In Mandarin, [COR+ 3] clusters are homorganic.
An example word is [blaŋko], “white”.	An example word is [s 3], “death”.
Is the underlying form /blaŋko/, /blanko/, or /blaNko/?	Is the underlying form /s 3/, /si/, or /s/?
How do we know which segment licenses place? If [ŋ] cannot license place, then it should ONLY occur before a [DOR] onset. If [ŋ] can license place, then it should occur word-finally. [ŋ] ONLY occurs before a [DOR] onset, therefore it does not license place.	How do we know which segment licenses place? If [i] cannot license place, then it should ONLY occur after [DOR] onsets. However, [mi] and [pi] are licit syllables, therefore [i] must be able to license place. If [3] can license place, then it should occur after [LAB] and [DOR] onsets as well. However, [LAB+ 3] and [DOR+ 3] syllables are not licit (see data sheet). Therefore, [3] cannot license place.
Therefore, /k/ is the segment that licenses place. Then, what are the features of the preceding segment?	Therefore, /s/ is the segment that licenses place. Then, what are the features of the following segment?

⁶ I choose to analyze (word-initial) sC clusters as coda + onset. This section accidentally provides evidence for the claim that Mandarin speakers parse English sC clusters as coda + onset.

⁷ Except it is devoiced.

⁸ A principle in Generative Phonology which assumes that speakers store as few items underlyingly as possible, and that the wide variety of surface forms are DERIVED.

Spanish	Mandarin
Principle of Economy: speakers store as few items underlyingly as possible, and that the wide variety of surface forms are DERIVED.	Principle of Economy: speakers store as few items underlyingly as possible, and that the wide variety of surface forms are DERIVED.
/N/ contains only 1 feature, [nas], whereas /ŋ/ contains many features, such as place features. Furthermore, if /ŋ/ was underlying, then Spanish would have 1 more phoneme. Therefore, /blaNko/ is the form which the brain stores underlyingly.	/s/ contains no vowel features, whereas /s/ and /si/ do, such as height features. Furthermore, if / / was underlying, then Mandarin would have 1 more vowel phoneme. Therefore, /s/ is the form which the brain stores underlyingly.
The brain can derive from [blaŋ.ko] from merely /blaNko/.	The brain can derive from [s] from merely /s/. An example of Optimality Theory derivation is provided below.

Morphological constraints:

MimimalWord: Every (morphological) word contains at least 1 (phonological) syllable⁹

Phonological constraints:

Markedness:

Nucleus: Every (surface form) syllable must have at least a (non-silent) nucleus

Place(O = N): Onset and Nucleus must share place¹⁰

Faithfulness:

Max(Segment): Do not add a segment

/s/, "four"	Nucleus	Place(O = N)	Max(Segment)
[s]	X !		
[si]		X !	X
👉 [s]			X

⁹ English word minimality constraint uses the unit — mora, whereas I use the unit — syllable.

¹⁰ If /i/ were [COR], then [si] would have won.

Section 5: Discussion

In this chapter, I discuss my conclusions.

In this last section, we have seen that, in order to get [s ] on the surface, one only needs to store /s/, and that storing /s / (which entails that /  / would have to be an additional vowel phoneme), is superfluous. However, this is not to say that no every Mandarin speaker has the same phoneme inventory in his/her brain. [s] and [] belong to the same morpheme, therefore there is no morpho-phonemic alternation from which speakers can draw a certain conclusion about its underlying form. In fact, in non-generativist approaches to phonology, /s / may be the most reasonable underlying form. Other phonologists reject the concept of “underlying form” entirely.

[s] in Other Languages

Other language have vowels that sound like []¹ as well, and are commonly transcribed as [i̯]. In Malayalam, [i̯] is the epenthetical vowel, which suggests that it is the underlyingly /i/ (Narasimhan 1995). It is in complementary distribution with [i]: during fronting, /i/ → [i̯]. In Russian, [i̯] is also in complementary distribution with [i], which has led many phonologists to believe that they are allophones of /i/. Nimboran presents a more surprising case: there is a minimal pair [kɪ̯] “feces” and [ki] “woman”, which means that /i̯/ and /i/ are different phonemes (Clements 1991). In this paper, I drew an analogy to Spanish, which has greatly informed my analysis. I suggest that future researchers investigate these languages thoroughly, and draw analogy to them, which may greatly inform their theory.

¹ “sound like” means F1 ≈ 491 Hz, and F2 ≈ 1769 Hz.

Bibliography

Cheng, Chinchuan. 1973. A synchronic phonology of Mandarin Chinese. The Hague: Mouton.

Cheng, Robert L. 1966. Mandarin phonological structure. *Journal of Linguistics* 2(2), 135–262.

Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. *The handbook of phonological theory*, 7, 245-317.

Clements, G.N. (1991, December 12). Place of Articulation in Consonants and Vowels: a Unified Theory. Zenodo. <http://doi.org/10.5281/zenodo.3735363>

Duanmu, San. 2000. The phonology of standard Chinese. New York: Oxford University Press.

Ladefoged, Peter, and Ian Maddieson. (1996). The sounds of the world's languages. Cambridge, MA: Blackwell.

Lee-Kim, S. (2014). Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association*, 44(3), 261-282.
doi:10.1017/S0025100314000267

Lee, W., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, 33(1), 109-112. doi:10.1017/S0025100303001208

Narasimhan, Rama. (1995). Coronals, Velars, and Front Vowels.

Pulleyblank, E. G. (1989). The role of coronal in articulator based features. In CLS. Papers from the General Session at the... Regional Meeting (Vol. 25, No. 1, pp. 379-393).

Wiese, Richard. 1997. Underspecification and the description of Chinese vowels. In Jialing Wang & Norval Smith (eds.), *Studies in Chinese phonology*, 219–249. Berlin: Mouton de Gruyter.