The Phonology of Consonant Voicing in Shanghainese

Henry Li Xinyuan

Spring 2021

Contents

1 Introduction

1.1 Background and motivation

Shanghainese (上海话, 上海闲话, 沪语) is a dialect of Wu Chinese spoken in and around the metropolitan area of Shanghai, China. Traditionally considered a prestige dialect of the region, its role as the regional lingua-franca is gradually being replaced by Mandarin [Wellman 2013].

Numerous studies have also reported phonological changes in Shanghainese over the past few decades, many of which were linked to the increased influence from Mandarin [Wellman 2013]. One of the commonly reported changes was the "loss of voicing distinction" [Chen 2003, Qian 2003, Gu 2004, Gu 2007]. We found this claim particularly interesting, owing to the fact that many minimal pairs in Shanghainese exist where the consonant voicing is the only difference between them. We therefore set out to study voicing in Shanghainese, bearing in mind to look out for potential signs of it disappearing.

1.2 Basic analysis and techniques used

On the surface, the claim that Shanghainese is losing its voicing distinctions due to influence from Mandarin appears reasonable and consistent with other phonological changes that have been observed in Shanghainese. In Mandarin, there are (almost) no contrasts between voiced and unvoiced consonants, as can be seen in the table below [Duanmu 2000]:

	Labial	Denti-alveolar	Retroflex	Alveolo-palatal	Velar
Nasal	[m]	[n]			[ŋ]
Stop	[p] [p^{h}]	[t] [t^{h}]			[k] [k^{h}]
Affricate		[t͡s] [t͡s^{h}]	[t͡ʂ] [t͡ʂ^{h}]	[t͡ɕ] [t͡ɕ^{h}]	
Fricative	[f]	[s]	[ʂ] [ʐ]	[ɕ]	[x]
Approximant		[l]	[ɻ]		

The only voiced-unvoiced distinction occurs between /ʂ/ (written in pinyin as sh) and /ʐ/ (written in pinyin as r), however as the romanisation might suggest, it's very common for them to be pronounced as /ʃ/ and /ɹ̝/ respectively. In real life, the unaspirated stops are also interchangeable with their voiced counterparts.

Shanghainese on the other hand, exhibit a voiced-unvoiced distinction in all of its affricates and fricatives, as well as a three-way distinction between voiced, unvoiced and aspirated unvoiced in its stops. Below is a table of Shanghainese consonants, as transcribed by Xiaonong Zhu in Zhu 2006:

	Labial	Denti-alveolar	Retroflex	Alveolo-palatal	Velar	Glottal
Nasal	[ʔm] [m̤/m]	[ʔn] [n̤/n]		[ʔɳ] [ɲ̈/ɳ]	[ʔŋ] [ŋ̈/ŋ]	
Stop	[p] [p^{h}] [p̤/b]	[t] [t̤/t^{h}] [d]			[k] [k^{h}] [k̤/g]	[ʔ]
Affricate		[t͡s] [t͡s^{h}]		[t͡ɕ] [t͡ɕ^{h}] [t͡ɕ̤/d͡ʑ̤]		
Fricative	[f] [f̤/v]	[s] [s̤/z]		[ɕ] [ʑ̤]		[h] [../ɦ]
Approximant		[ʔl] [l̤]				

Note that in all of the three categories mentioned above, Zhu comments that the [+voice] can instead by realised as breathy voiced (transcribed with two dots underneath the sound in question). In our subsequent analysis, we will demonstrate that there is a substantial difference between the way voiced consonants are articulated in Shanghainese versus in other languages (such as English).

Past studies on phonological changes in Shanghainese have commonly been (attempted) population-wide studies, usually employing surveys as the means of determining articulatory characteristics of certain sounds and words. For the purpose of this study, we will not be relying on self-reported speaking habits as the basis of our analysis; instead, we will take recordings of a small number of native Shanghainese speakers, which we will then use for the following:

(a) Observe the waveform for each word, focusing on the the distinctive patterns that characterise voiced consonants

(b) Play the recordings to other native speakers of Shanghainese and report the accuracy with which they can distinguish between [+voice] and [-voice]

Due to the limitations of our resources, we will be running our study on a very small number of participants. As such, we will not be able to confirm or reject the conclusions made in past population-wide studies. Instead we will be making a snapshot of how voicing is realised and perceived at this point in time by a small number of speakers. We hope that some of the insights gained in this study will address the issues in some of the assumptions in past studies about the underlying phonology of Shanghainese, as well as inspire similar studies of a larger scale which can bring us a more complete picture of the topic.

1.3 Comments on some distinctive characteristics of Shanghainese phonology

It is common to describe the syllabic structure in Chinese languages as being composed of an initial (consonant or zero-onset) and a final (mandatory nucleus vowel, preceded by optional medial vowel and followed by optional ending consonant). The choice of the ending consonant is very limited: in the case of Shanghainese, it can only be one of /ŋ/ and /ʔ/, the latter of which is omitted in most day-to-day speech. Since neither has a counterpart in Shanghainese that differ with them only in voicing, we will not cover them in this study; instead we will focus exclusively on initials.

Another observation is that inflection is very rare in most Chinese languages, Shanghainese included [Quian, Nairon and Zhongwei Shen 1991]. The only type of morphological change in Shanghainese is tone sandhi, which is beyond the scope of our study; hence we will not make a distinction between underlying and surface represenations of words. For the same reason, we will be also looking only at monosyllabic words (one-character words in Chinese).

Finally, we will transcribe tone as a number at the end of each syllable. For the most part in this study, tone is a confounding variable that is controlled for. ”Homophony” refers to words with identical transcriptions including tone.

2 Analysis

2.1 Basic knowledge of voicing distinction

First we run ourselves a sanity check: are Shanghainese speakers still aware of the difference between voiced and unvoiced consonants?

We refer to the Wu Chinese dictionary compiled by the Wu Chinese Society as our ”standard” for transcribed Shanghainese. We choose a number of words which, according to that dictionary, are minimal pairs that differ only in initial voicing (an example being 肥 /vi1/vs 飞 /fi1/). We show them in writing to a number of Shanghainese speakers, asking them to report whether they regard these words as distinct. This, incidentally, is the method used in many of the previous survey-based studies that we referenced, albeit at a much larger scale both in terms of number of surveyees and size of the data collected.

The results were that all of the speakers we interviewed possessed a basic understanding of all of the voicing distinctions present in Shanghainese, as they agreed that the pairs of words presented to them were distinct.

2.2 Collection of data

2.2.1 Speakers

In total, speech data from 3 different individuals were collected. The speakers are native speakers of Shanghainese, are of similar age profile, and live in relative geographic proximity within the city of Shanghai. The speakers were not exposed to the ”sanity check” study described above.

2.2.2 Choice of words spoken

We made an effort to keep the number of words to within the suggested limit (50). In order to rule out the effect of tone sandhi, the words that were chosen were all monosyllabic. We chose words based on the following (overlapping) criteria:

(a) Words that are minimal pairs differing only in initial voicing (eg. 肥 /vi1/vs 飞 /fi1/), or in the case of stops, [+voice], [-voice, -aspiration], [-voice, +aspiration] 3-way minimal pairs (eg. 鼻 /bɪ4/ vs 笔 /pɪ4/ vs 匹/p^hɪ4/)

(b) Words that have a [+voice] initial in Mandarin and a [-voice, -aspiration] initial in Shanghainese (肥 Shanghainese: /vi1/, Mandarin: /fəi/), or vice-versa (导 Shanghainese: /dɔ3/, Mandarin: /tau/)

(c) Minimal pairs where the initials are [∅], [h] and [ɦ], respectively (eg. 挨 /a1/ vs 哈 /ha1/ vs 鞋 /ɦa1/). Note that in the table of Shanghainese consonants, [ɦ] was considered interchangeable with [..], zero-initial with murmur sound. An observation made in Zhu 2006 and reaffirmed in preliminary interviews with Shanghainese speakers is that [ɦ] is far more similar to [∅] than to [h], as a result of that interchangeability. As such, for the purpose of this study, we will treat [∅] as an unvoiced initial (technically true) and [ɦ] as its voiced counterpart.

(d) Words that were cited in previous studies as having lost their voicing in Shanghainese, coupled with their minimal-pair counterparts differing only in initial voicing according to standard Shanghainese transcriptions:

1. 腐 /vu3/, claimed to have evolved into /fu3/ [Gu 2004]. Contrasted with 富 /fu3/
2. 徐 (transcribed /ʑi1/, spoken as /ʑy1/ in collected data), claimed to have evolved into /ɕy1/ (homophonous to 虚) or /d͡ʑ̈y1/ (homophonous to 跪) [Chen 2003].
3. Loss of initial murmur [../ɦ] -> [∅] [Gu 2004]

In Shanghainese, as is the case with most other Chinese languages, there is little to no liaison between the final of the previous word and the initial of the next word [Zhu 2006]. Therefore the difference in environment that could affect the pronounciation of a consonant is the following final portion of the syllable. For each [+voice], [-voice] minimal pair, we made an effort to cover as many different finals as possible. An example is shown in the table below:

final	k	g	k^h
a	街 /ka1/	茄 /ga1/	咖 /k^ha1/
e	改 /ke2/	倚 /ge2/	凯 /k^he2/
æ	革 /kæ4/	个 /gæ4/	刻 /k^hæ4/
ɒŋ*	工 /kɒŋ1/	共 /gɒŋ1/	空 /k^hɒŋ1/
aŋ	降 /kaŋ3/	戇 /gaŋ3/	抗 /k^haŋ3/
u	古 /ku2/		课 /k^hu2/

*Transcribed ɑŋ, recorded ɒŋ

Once the list of words had been finalised, they were put in a randomised order with some of them duplicated. They were then presented to the speakers in writing. The speakers then read each word individually. The waveforms of each recording were then inspected and compared.

2.2.3 Perception study

Armed with recordings of a sizeable number of [+voice], [-voice] minimal pairs, we proceed to test whether speakers of Shanghainese are able to distinguish between such minimal pairs. We recruited the help of 2 additional native speakers of Shanghainese for this part of the study, hence we have a total of 5 sets of data. The recordings from the 3 speakers were played to each in a random order (including recordings from themselves). After each recording, the participants were asked to report the word that they heard. The results are recorded as such:

(a) Correct: the participant reported the correct word or some other homophonous word.

(b) Incorrect word, correct initial: the participant reported a word with the correct initial and a different final.

(c) Incorrect word, correct voicing: the participant reported a word with an initial that has the same voicing property as the original.

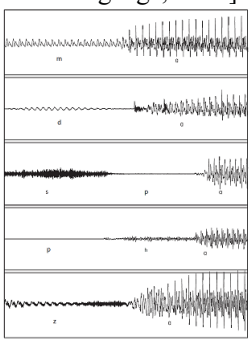
(d) Incorrect: the participant mistook the word for another with a different voicing property, or failed to make an identification.

(a), (b) and (c) would all be considered instances where [+ voice] or [-voice] is correctly perceived by the listener, although we managed to capture a number of non-voicing related phonological shifts in Shanghainese within (b) and (c).

2.3 Waveform observation

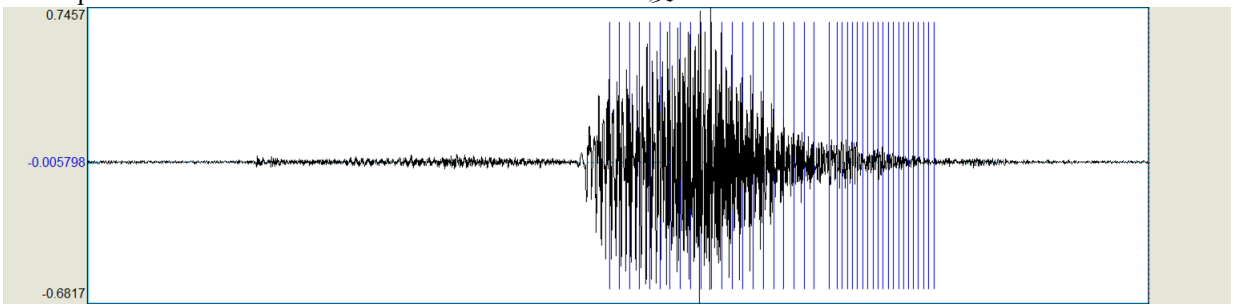
We look at the 200ms waveforms in the recordings that we took, comparing the [+ voice] [-voice] minimal pairs.

Zsiga characterised the waveform pattern of an unvoiced stop as being close to a flat line followed by a burts, that of a voiced stop as periodic with lower amplitude compared to syllabic sounds, and that of an aspirated stop as evidenced by the distinctive aspiration noise. Zsiga also pointed out that the waveform of a voiceless fricative is close to that of random noise, while those of a voiced fricative combine periodicity and noise with periodicity dying out towards the end of the consonant. Attached is a picture of a few characteristic waveform patterns [Zsiga, The Sounds of Language, 2013]:

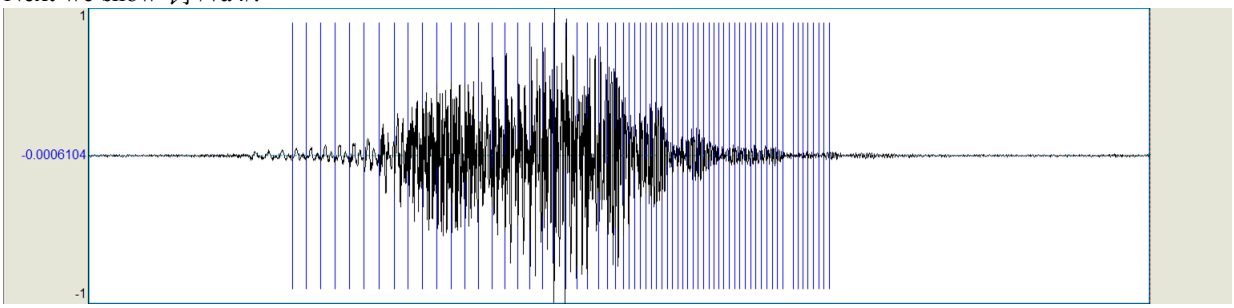


2.3.1 Fricatives

In the case of fricatives, the waveforms of the voiced initial demonstrate a clear difference from those of the unvoiced initial, consistent with the ”periodic noise” vs ”random noise” distinction remarked by Zsiga. One pair of such examples is shown here. First we show the waveform for 发 /fa4/:

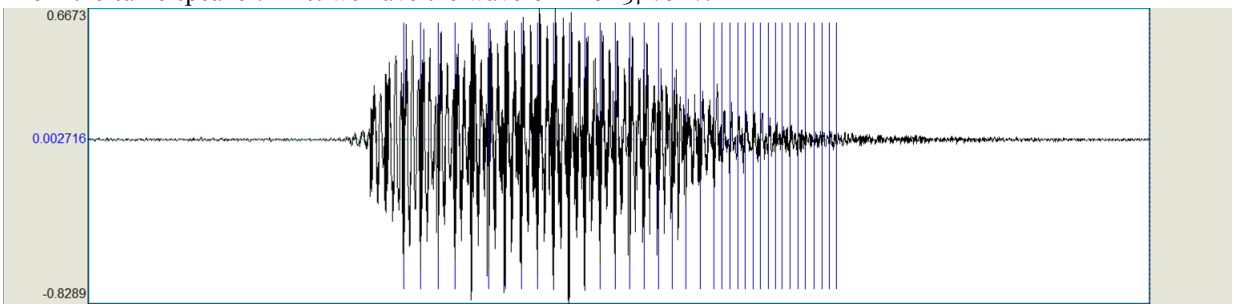


Next we show 罚 /va4/:

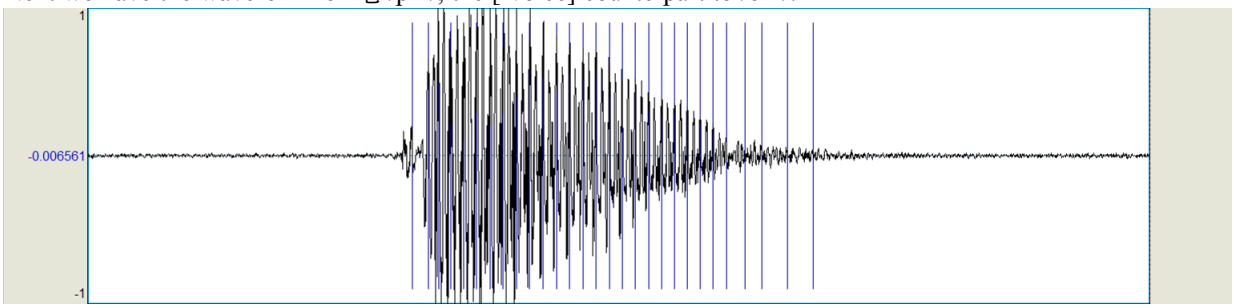


2.3.2 Stops

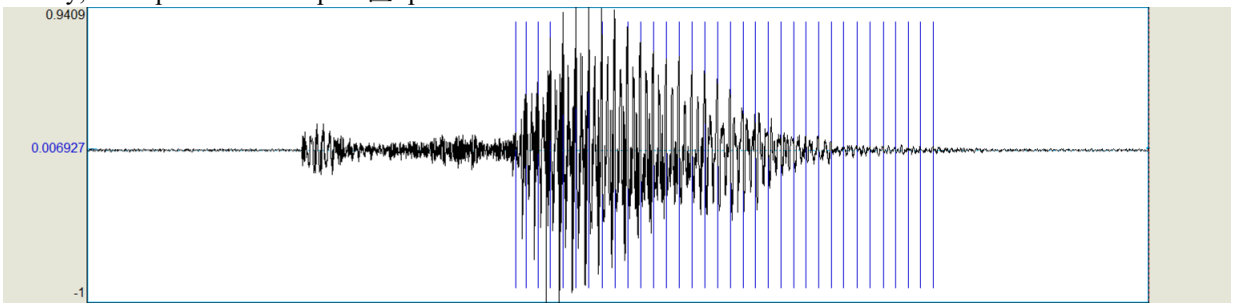
We observe that in the case of stops, aspiration can clearly be seen in the waveform across all speakers. On the other hand, the waveform of voicing sounds described and illustrated by Zsiga is not present in any of the words with voiced initials. Below are the waveforms for a [+voice], [-voice, -aspiration], [-voice, +aspiration] initial minimal pair from the same speaker. First we have the waveform for 鼻 /bɪ4/:



Next we have the waveform for 笔 /pɪ4/, the [-voice] counterpart to /bɪ4/:



Finally, the aspirated counterpart 匹 /p^hɪ4/:



The aspiration noise is both very easy to see on the waveform, and is also distinctive to the ear when the recordings are cut to just the initial. However, the voicing of the initial is far from obvious: the waveform show silence up until the stop burst. The speaker having a particularly short voice onset time is a possibility that's ruled out by observing the waveforms from recordings by other speakers, which demonstrate similar patterns.

2.4 Analysis of perception data

The participants to this study demonstrated a strong ability to distinguish between voiced and unvoiced initials. In total, 900 data points were gathered. 31 of these cases fell in the category d (outright incorrect) outlined in the earlier section on perception study design, with an additional 21 falling into category c (correct voicing, incorrect initial), adding to an error rate of 5.78%. Compare this to studies such as Miller, Nicely 1955 where participants were asked to identify English [-sonorant] consonants, in which the participants made errors 13.7% of the time in a minimal-noise condition, this error rate is relatively low (although worth noting that our study has a much smaller pool of consonants to choose from as well as a syllabic context for the consonants, hence participants should have an easier time making a correct identification).

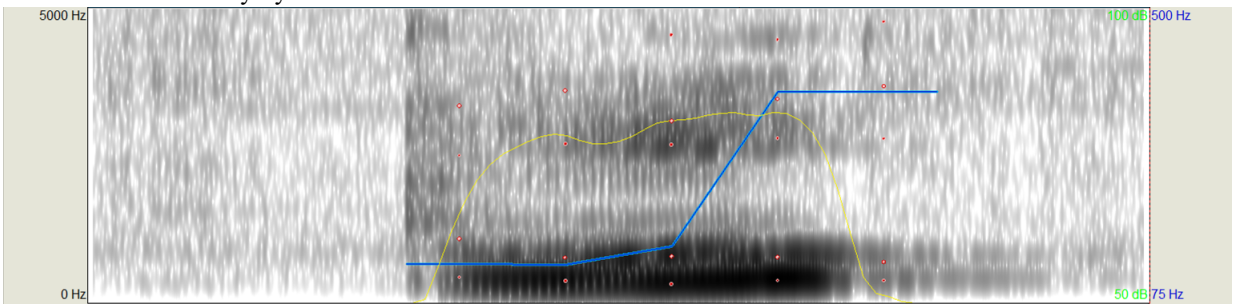
We note that among the 31 outright incorrect identifications, 23 were words with tones 2 or 3, quite significantly larger in proportion to the number of words in the dataset that are tones 2 or 3 (18/60 after random word duplication). An example is the word 降 /kɑŋ2/: the utterance of speaker 1 was mistaken as /gaŋ2/ or /gaŋ3/ by 4 listeners, with one claiming that /kɑŋ2/ and /gɑŋ2/ were identical. We therefore consider the possibility that tone is significant in perception of initial voicing in Shanghainese.

2.5 Revisit recordings: intensity and pitch contour

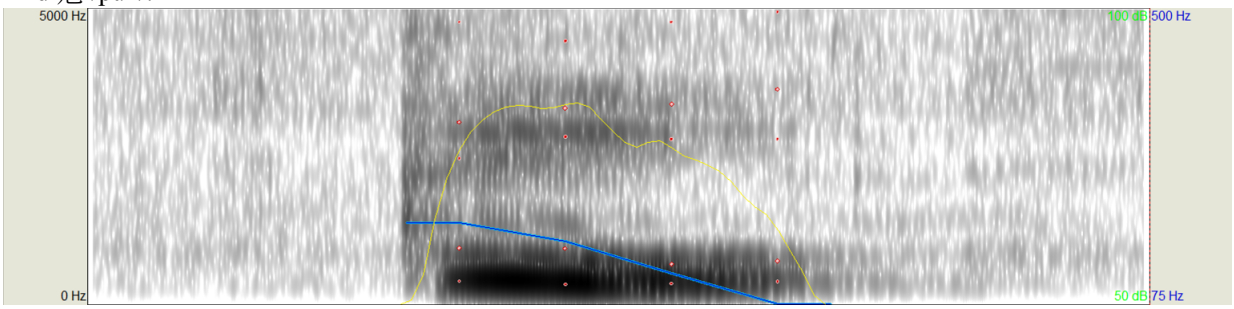
We attempt to reconcile the apparent lack of voicing shown in the waveforms with the great accuracy that listeners are able to point out voiced initials from their unvoiced counterparts, as well as address the possibility that tone plays a role in voicing detection in Shanghainese.

We refer back to the recordings, this time focusing on the spectrogram.

Below is the spectrogram for 爬 /bu1/, with the intensity (yellow line) and pitch (blue line) contours annotated, as calculated automatically by Praat:



And 疤 /pu1/:



We can observe clear patterns here:

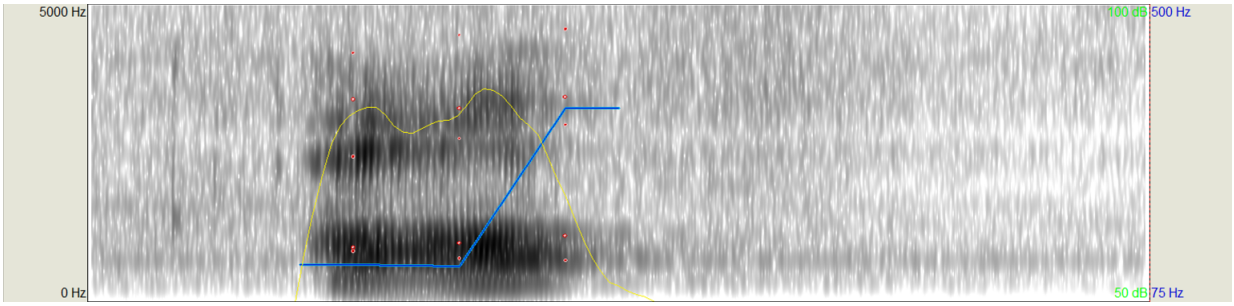
(a) The (mono-syllabic) word with the [+voice] initial has a rising pitch contour (in its only syllable), whereas the word with the [-voice] initial has a falling pitch contour.

(b) The word with the [+voice] initial has an intensity contour that rose slightly during the vowel onset until a sharp fall near the end, while the word with the [-voice] initial fell in intensity immediately after the vowel onset.

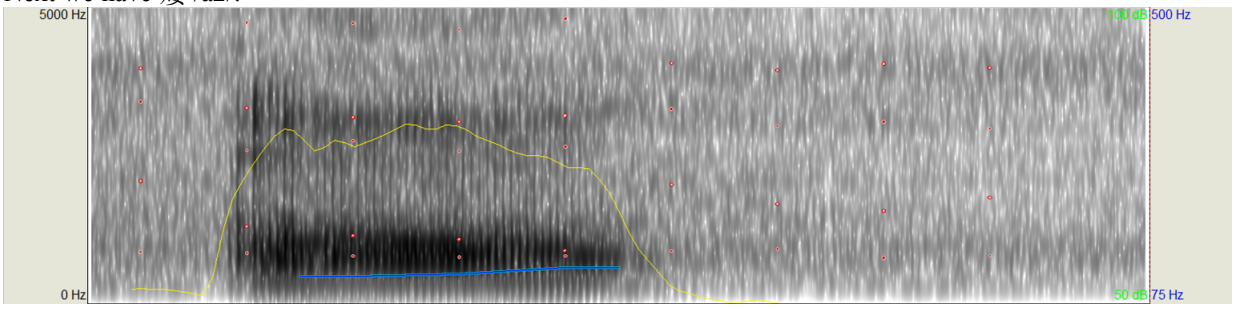
The two rules mentioned appear to be somewhat connected and might be part of a more general connection between tonal and intensity contours; that is unfortunately beyond the scope of our study.

We observe similar patterns in the spectrograms of other words in tone 1, including words that have other stops or fricatives in syllable-initial position.

We next compare two words with tone 2. First we have 也 /ɦa2/:

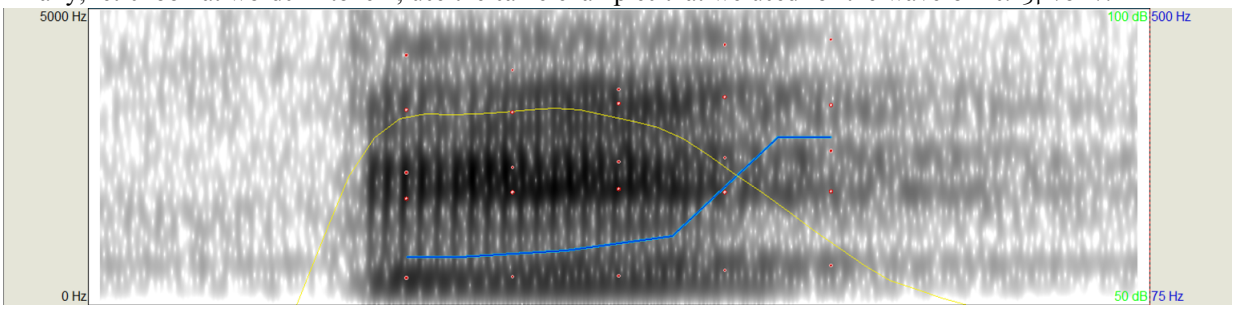


Next we have 矮 /a2/:

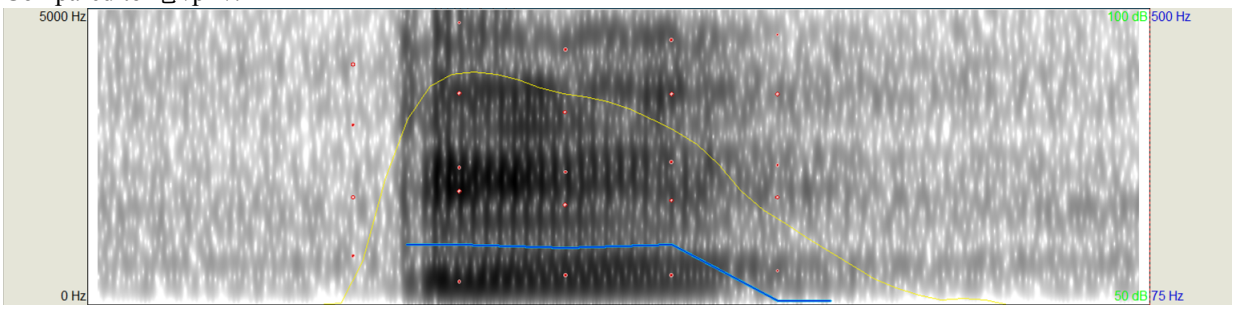


Note that in this instance, both syllables display a rising tonal contour, although the one with the [-voice] (zero-onset) initial has a much less significant rise than the one with the [+voice] initial. The same can be observed in the tonal contours of other tone 3 words.

Finally, let's look at words in tone 4, use the same examples that we used for the waveforms. 鼻 /bɪ4/:



Compared to 笔 /pɪ4/:



We still observe a sharply rising tonal contour in the word with the [+voice] initial, whereas in the case of the word with the [-voice] initial we observe a relatively constant tonal contour.

We must now revisit the way we transcribed tone in our study. We used 1, 2, 3, and 4 to denote the four tones in the Middle Chinese tone designations: Ping (平), Shang (上), Qu (去), Ru (入), respectively. Middle Chinese also has the [+Upper] feature describing whether the a syllable has upper tonal register, a feature independent of the four tones. The [+Upper] and [-Upper] feature, in addition to which of the four tones a sound falls into, decide the pitch contour of the syllable [Zhu 2006]. Specifically, in a system of tonal contour description where 1 is the

lowest and 5 is the highest, tones 1 (Ping), 2 (Shang), and 3 (Qu) are all designated as having contour 14 when in a [-Upper] environment, while tone 4 (Ru) is designated as having contour 24 when [-Upper]. By comparison, when in a [+Upper] environment, tones 1, 2, 3, and 4 have tonal contours 52, 34, 34, and 44, respectively [Zhu 2006].

While the precise tonal contour varies due to word and speaker difference, the rough outline of these contours as described by Zhu match up with the following:

(a) The subjective observation of the listeners post study (by asking them questions such as "Do you feel like the word 也 /ɦa2/ has a rising tone?").

(b) The observed tonal contour as calculated by Praat.

(c) The perceived error distribution by the listeners: errors in voicing detection occur mostly in words that are tones 2/3 [+Upper].

(Note that while the tonal contour of 4 [-Upper] is closer to that of 2/3 [+Upper] than 1/2/3 [-Upper], tone 4 are exclusive for short vowels and thus easily distinguished from the other tones)

We would thus propose the following rules regarding Shanghainese consonant voicing and the tone of the associated syllable:

(a) All syllables are [+Upper] by default.

(b) [+Upper] $\rightarrow$ [-Upper]/[+Voice, +Consonontal]_$\emptyset$: Words assume a sharply rising contour when following a [+Voice] initial, or in Zhu's terminology, take on a [-Upper] tone register

(c) [+Voice] $\rightarrow$ [-Voice]/$\emptyset$_[-Upper]: Voicing is not pronounced properly in initials when the following syllable has been converted to [-Upper].

or [+Voice] $\rightarrow$ [-Voice, +Murmur]/$\emptyset$_[-Upper]: Voicing is converted to voiced murmur when the following syllable has been converted to [-Upper]. As we lack a way to definitively point out the voiced murmur sound, we regard this formulation with much less certainty than the previous formulation. All we have shown is that voicing is pronounced substantially differently from how it's pronounced in languages such as English.

2.6 Revisit perception study

In order to test the hypothesis we just proposed, we design a simple revised perception study: we cut away the initial portion of each word and play the modified recordings to listeners. Conveniently zero-onset is in one of our [+voice] [-voice] minimal pairs. If our hypothesis holds up, we should see listeners report words with [$\emptyset$] initial when the original recording is was [-voice] initial (regardless of aspiration), and we should see them report words with [../ɦ] initial when the original recording has a [+voice] initial.

We tested the cut data on the 3 original speakers. 40 of the 60 original sounds resulted in a Shanghainese word which exists and has an initial-voicing minimal-pair counterpart, thus in total we were able to gather 360 data points. Due to the difficulties cutting the recordings, error rates were substantially higher than in the previous part of the perception study: the 207 [+Upper] tone 1/4 and [-Upper] tone 4 were almost always correctly identified (204/207), whereas the 36 [+Upper] tone 2/3 words had a hit rate of 29/36, and the 117 [-Upper] 1/2/3 words had a correct identification ratio of 102/117 (Correct here is defined as identifying the correct initial voicing). Broadly speaking, these results match up with our expectations.

In order to find out whether Shanghainese speakers primarily rely on the tonal contour to determine initial voicing, we attempted to create new recordings by appending unvoiced consonants and the finals following unvoiced initials; unfortunately we couldn't produce audio of a high enough quality to perform a similar study as before.

3 Conclusion

3.1 The phonology of Shanghainese initial voicing

We established, through observation of waveforms and spectrograms of recordings from native speakers, as well as perception studies with native speakers as listeners, the following characteristics in the voicing of initials among the speakers we sampled:

(a) The speakers did not voice their initials in a ”typical” way, or in a way similar to how English speakers would voice their consonants. Due to the difficulties determining the quality of the ”voiceless murmur” mentioned in past literature as a replacement for voicing, we claim that the Shanghainese speakers in question quite often drop their initial voicing. This phenomenon applies to both stops and fricatives.

(b) Initial voicing affects the tonal contour of the following syllable by converting it to a sharply rising tone, regardless of the original tone. In the language of past literature, it converts the syllable into [+Upper].

(c) Shanghainese speakers very likely primarily use the tonal contour of the syllable to determine initial voicing. Further study, possibly using synthetic speech sounds, is needed to test that claim.

As mentioned earlier, we do not, nor did we attempt to, confirm or contradict the claims by earlier scholars on the loss of voicing in Shanghainese: we are, afterall, only looking at a tiny number of speakers. Even in our small sample, it’s very difficult to draw conclusions on the evolution of Shanghainese: no detailed look at the spectral and waveforms patterns of historical Shanghainese words exist, and recordings are hard to come by and also often pronounced ”theatrically” or at least formally. The not-so-pronounced voicing of initials we observed raises the possibility that voicing distinctions are eroding in Shanghainese. On the other hand, both the perception tests as well as the observed impact voicing has on tonal contour suggest that losing voicing distinctions is less likely than previously assumed: it is, afterall, harder for words to lose their tonal contour than the initial consonant to lose one distinctive feature, especially in tone-sensitive Chinese languages, where such changes tend to happen systematically across all words of a particular tone [Sagart, L. 2014].

3.2 Issues

3.2.1 Issues with a non-phonetic alphabet

Languages with non-phonetic alphabets present many difficulties with phonetic study (such as excluding the possibility of wug-tests). In the context of this study, they manifested predominantly in the following ways:

(a) All words that subjects were asked to read out must exist in real life.

This severely limited our ability to construct minimal pairs covering each possible final following the initials we’re interested in.

(b) In the cases where a single written word has multiple pronounciations, a context has to be provided; likewise, listeners often struggle to report a word they heard due to not being able to come up with a context for that word, even when they’ve clearly recognised the word (demonstrated by them repeating the word accurately).

We entertained the idea of addressing this issue by teaching the listeners basic phonological concepts relevant to this study (voicing, aspiration, etc.), although concerns about the knowledge of voicing distinctions confounding the results led us to not pursue that route. This would not have been an issue if there were a romanisation of Shanghainese similar to pinyin for Mandarin.

(c) Standalone monosyllabic words are seldom said in real life; as claimed in Zhu 2006, the tonal contour of individual words is seldom discernible due to tone sandhi (a form of tone spreading). In those cases, do speakers make a bigger effort to voice their consonants, or do they rely on context to determine the initial?

This is a conundrum that merits further study.

3.2.2 Issues with reference transcription

During the design phase of this study, we ran into minor problems when it comes to choosing a reference transcription. Unlike in the case of languages such as English or Mandarin, a widely-accepted transcription of the ”correct” or ”common” pronounciation of Shanghainese does not exist. We deferred to a combination of the Wu Chinese dictionary by the Wu Chinese Society and Xiaonong Zhu's A Grammar of Shanghai Wu, although both appear inconsistent with the observed pronounciation. In one example, a listener reported the word 扛 when played the word 戇 /gaŋ3/. The word 扛 is transcribed in the Wu Chinese dictionary as /kaŋ1/, but when we approached the listener and recorded them speaking the word 扛, it exhibited the rising tonal contour characteristic of [+voice] initial words, demonstrating that the listener's recognition of voicing was accurate as per their pronounciation of Shanghainese

3.2.3 Overlap between speaker and listener

While other perception studies have come to the conclusion that base-frequency related difference in speakers cause little difficulty in speech perception [slides], it may still be possible that a speaker would be more acclimatised to their own speech and perform better than they otherwise would in a perception test. In a larger sclae study this could be averted.

4 Addendum

4.1 Observed changes in Shanghainese, unrelated to initial voicing

(a) 樱 /aŋ1/ → /iŋ1/: pronounciation of this word shifted to become more similar to its Mandarin pronounciation.

(b) The vowels /u/ and /o/ are merging: this is not only one of the most common mistakes among listeners, but also a speaking habit displayed by one of our speakers. The merge is not universal: when interviewed on this matter, one speaker claims that /pu/ and /po/ are identical whereas /ku/ and /ko/ are not. Given the lack of distinction between these two vowels in Mandarin, this is likely an actual example of phonetic change in Shanghainese due to influence from Mandarin.

(c) The finals /aŋ/ and /ɑŋ/ appear to be merging. Likewise a possible result of influence from Mandarin.

4.2 Gloss of the Chinese words mentioned in this study

We only provide a signle most common gloss for the words listed here.

Word	Gloss	Transcription	Word	Gloss	Transcription
肥	fat	/vi1/	腐	rot	/vu3/
飞	fly	/fi1/	富	rich	/fu3/
鼻	nose	/bɪ4/	徐	family name	/ʑy1/
笔	pen	/pɪ4/	虚	weak	/ɕy1/
匹	quantifier for horse	/p^hɪ4/	跪	kneel	/$\widehat{dʑ}$y1/
导	guide	/dɔ3/	街	street	/ka1/
挨	suffer a blow	/a1/	茄	aubergine	/ga1/
哈	ha (laughter)	/ha1/	咖	咖啡 coffee	/k^ha1/
鞋	shoe	/ɦa1/	改	change	/ke2/
倚	lean against	/ge2/	凯	victory	/k^he2/
革	leather	/kæ4/	个	generic quantifer	/gæ4/
刻	sculpt	/k^hæ4/	工	industry	/kɒŋ1/
共	common	/gɒŋ1/	空	empty	/k^hɒŋ1/
降	lower (v.)	/kaŋ3/	戇	silly	/gaŋ3/
抗	struggle against	/k^haŋ3/	古	ancient	/ku2/
课	lesson	/k^hu2/	发	send	/fa4/
罚	punish	/va4/	爬	crawl	/bu1/
疤	scar	/pu1/	矮	short	/a2/
也	also	/ɦa2/	扛	carry on one's back	/gaŋ3/
樱	cherry	/iŋ1/			

5 References

1. Wellman, L. (2013) Pudong and Putonghua: Sound Change and Language Shift in Shanghai. Retrieved 2021 from Yale Linguistics: `https://ling.yale.edu/sites/default/files/files/alumni%20senior%20essays/Wellman%2C%20Laura%20-%20Senio%20Essay.pdf`
2. Chen, Z. (2003) Studies on Dialects in the Shanghai Area. LINCOM: LINCOM Studies in Asian Linguistics
3. Qian, N. (2003) A History of Shanghainese. Shanghai: Shanghai People's Press
4. Gu, Q. (2004) Phonetics in New Shanghainese. Shanghai: Shanghai Normal University
5. Gu, Q. (2007) The Influence of Language Contact on Sound Changes in Shanghainese. Shanghai: Shanghai Normal University
6. Duanmu, S. (2000) The Phonology of Standard Chinese. Oxford: Oxford University Press
7. Zhu, X. (2006) A Grammar of Shanghai Wu. LINCOM: LINCOM Studies in Asian Linguistics
8. Quian, Nairon and Zhongwen Sen. (1991) The Changes in the Shanghai Dialect. Journal of Chinese Linguistics Monograph Series
9. Zsiga, E. (2013) The Sounds of Language: An Introduction to Phonetics and Phonology. Wiley: Wiley-Blackwell
10. Miller, Nicely. (1955) An Analysis of Perceptual Confusions Among Some English Speakers. The Journal of the Acoustic Society of America
11. Sagart, L. (2014) Proceedings of the Symposium/Cross-Linguistic Studies of Tonal Phenomena/Tonogenesis, Typology and Related Topics. Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies
12. Class slides