

UDC 81'34

DOI: 10.18413/2313-8912-2023-9-4-0-3

Yifang Li¹ 

Svetlana V. Androsova² 

**A perceptual study of weak forms in Chinese:
the case of phonetic homophony**

¹ Shanxi Datong University
405 Xingyun St., Pingcheng District, Datong, Shanxi Province, 037009, China
E-mail: liyifang.1993@163.com

² Amur State University
21 Ignatievskoe Rd., Blagoveshchensk, 675027, Russia
E-mail: androsova_s@mail.ru

Received 10 July 2023; accepted 14 November 2023; published 30 December 2023

Acknowledgements. We express our gratitude to Amur State University and Laboratory of Experimental Phonetic Research for providing technical support for this research.

Abstract. The paper investigates the way some Chinese weak forms in spontaneous speech are perceived when taken out of a wider context. These forms are the result of syllable initials and finals modifications as well as coalescence due to the drop of syllable constituents. The results of the previous acoustic study demonstrated that not only functional but notional words have weak forms. We hypothesized that a number of those forms appear to acoustically resemble other Chinese words or word groups giving way to potential homophony. To prove the hypothesis, a perceptual study was performed. The tokens presenting potential homophones were segmented from native female spontaneous speech of 6 subjects. Two groups of Chinese native speakers (20 males, 20 females) participated in the experiment performing a discrimination task. The obtained results prove the hypothesis as in both groups, only 16% of responses identified canonical syllables, whereas 84% of responses identified the homonyms. There was certain gender difference which, however, proved insignificant for three fourths of the responses – more male listeners preferred homophones to canonical words/word groups. As expected, the reliable preference for the homophone during the discrimination task was demonstrated in 64% of the responses classified as non-identification of the canonical word/word group. Chinese native speakers are aware of this homophony, which in some instances is further expressed in the change of hieroglyphs creating not just homophones but full homonyms that have to be considered in Chinese as L2 teaching and learning.

Keywords: Chinese spontaneous speech; Discrimination task; Token; Homophone; Canonical word; Canonical syllable; Weak forms; Reliable/unreliable perception

How to cite: Li, Y. and Androsova, S. V. (2023). A perceptual study of weak forms in Chinese: the case of phonetic homophony, *Research Result. Theoretical and Applied Linguistics*, 9 (4), 47-62. DOI: 10.18413/2313-8912-2023-9-4-0-3

УДК 81'34

DOI: 10.18413/2313-8912-2023-9-4-0-3

Ли И.¹ 

Андросова С. В.² 

Перцептивные характеристики слабых форм
в китайском языке: омофония

¹ Датунский университет
ул. Синьюнь, 405, район Пинчэн, Датун, Провинция Шаньси, 037009, КНР
E-mail: liyifang.1993@163.com

² Амурский государственный университет
Игнатьевское шоссе, 21, Благовещенск, 675027, Россия
E-mail: androsova_s@mail.ru

Статья поступила 10 июля 2023 г.; принята 14 ноября 2023 г.;
опубликована 30 декабря 2023 г.

Информация об источниках финансирования или грантах, благодарности:
Авторы выражают благодарность Амурскому государственному университету и Лаборатории экспериментально-фонетических исследований за техническую поддержку данного исследования.

Аннотация. Настоящее исследование нацелено на выявление особенностей восприятия некоторых слабых форм, сегментированных из китайской спонтанной речи и предъявленных на опознание вне широкого контекста. Эти формы возникли в результате естественных модификаций инициалей и финалей слога, а также стяжения слогов после выпадения элементов слога. В ходе ранее выполненного акустического анализа выявлено, что не только служебные, но и знаменательные китайские слова были реализованы в слабых формах. Согласно гипотезе данного исследования, некоторые из этих форм схожи с другими словами и сочетаниями слов в китайском языке, что является основанием для потенциальной омофонии. Для проверки гипотезы был проведён перцептивный эксперимент, материалом для которого послужили стимулы – потенциальные омофоны, сегментированные из спонтанной речи шести женщин, являющихся носителями языка. В эксперименте приняли участие две группы носителей китайского языка (20 мужчин и 20 женщин), перед которыми стояла задача прослушать и опознать стимулы. Полученные результаты полностью подтвердили выдвинутую гипотезу, поскольку в обеих группах лишь в 16% ответов пришлось на варианты с каноническими слогами, тогда как 84% ответов содержали варианты с омофонами. Была отмечена некоторая гендерная разница, однако для трёх четвертей данных она оказалась статистически не значимой – больше мужчин предпочли омофон каноническому слову/сочетанию слов. Как и ожидалось, надёжное опознание омофона зафиксировано в 64% случаев среди ответов, классифицированных как неопознание канонического слова/сочетания слов. Носители китайского языка осознают наличие омофонии, которая в ряде случаев развивается далее и приводит к изменению иероглифов и, следовательно, к образованию полных омонимов, которые нельзя не учитывать при обучении китайскому языку как иностранному.

Ключевые слова: Китайская спонтанная речь; Задание на распознавание; Стимул; Омофон; Каноническое слово; Канонический слог; Слабые формы; Надёжное/ненадёжное восприятие

Информация для цитирования: Ли И., Андросова С. В. Перцептивные характеристики слабых форм в китайском языке: омофония // Научный результат. Вопросы теоретической и прикладной лингвистики. 2023. Т. 9. № 4. С. 47-62. DOI: 10.18413/2313-8912-2023-9-4-0-3

1. Introduction

Syllabic morpheme is a minimal pronunciation and meaningful unit in the Chinese language. According to Kasevich (2006: 524-525), it is the basic unit of the language unlike morpheme and word that occupy the periphery. Its graphic expression is hieroglyph or character that in its modern manifestation has lost its original iconic nature to a certain extent (although it is present in general) as it includes symbols that guide pronunciation (Shen, 2018: 133). Syllable is also a minimal (monosyllabic) word, and monosyllabic words are archetypal lexical units in Chinese.

Syllabic morpheme pronunciation can certainly vary, however, the patterns of this variation and variability extent are still far from being thoroughly investigated, although there are a few very successful studies in this field (see e.g. Tseng, 2005; Cheng and Xu, 2009).

Very few researchers have focused on the way naturally modified syllables (including contracted ones) are perceived. Thus, Tseng (2005: 159) reported that about 32% of 123,320 syllables studied were contracted – the result of work of professional human annotators whose task was to mark contractions and provide acoustic transcription. No perceptual studies of syllable phonemics natural modifications with non-linguist listeners have been performed.

At the same time, Chinese linguists admit (see, e.g., Fu and Li, 2018) that studying the ways language units are perceived is a vital experimental goal, however, mostly non-Chinese colleagues have succeeded in this realm so far. At present, more and more researchers make perceptual features the primary focus of their attention,

and auditory experiment has become a crucial part of new technologies of linguistic research in China. There is a marked shift from studying canonical forms or full forms as a result of careful pronunciation to non-canonical or weak forms (including reduced or contracted forms) as a result of natural modifications in the speech flow where we can observe a combination of careful and careless pronunciation types whose proportion depends on many factors. In Russian, the theory of those types was developed by Liya Bondarko and her colleagues (Bondarko et al., 1974) based on Lev Shcherba's earlier ideas first reported in 1915 and later published in the book of his Selected works (Shcherba, 1957).

Based on different languages, it was shown that in non-prominent positions in the utterance where coarticulatory pressures are very high, canonical phonemic patterns of words are subject to various consonant and vowel context-dependent quality changes and omissions giving way to weak forms, and by this, to co-existence of different phonetic and phonemic patterns of one and the same word in our mental lexicon (see e.g. (Jones, 1969; Zinder, 1981; Laver, 1994; Port, 2007; Babushkina, 2012; Androsova, 2015; Lobacheva and Androsova, 2016; Karavaeva, 2018; Priva and Strand, 2023)). Robert Port (2007) called it a very concrete, detailed auditory code (see also Borisova (2018) who mentioned native speakers using lexical coding that will definitely challenge Chinese as L2 learners' perception by the slips of the tongue and wrong word use). However, the real challenge is not in wrong word use or the slips of the tongue but in connected speech processes that are part of the code. Mücke et al. (2017) analyzed multiple works studying

connected speech processes as a result of regulation mechanisms during speech production and perception in a variety of languages including Arabic, English, French, German, Russian and some others.

For native speakers, weak forms are identifiable in a wider context in most instances. But when the context is minimal, weak forms (a word or its part) will be perceived in a different way (see Karavaeva (2018) experimenting with the perception of VCV tokens segmented from British speech flow).

As far as Chinese, there is a whole war between those who consider that the Chinese syllable does not change its phonemics and boundaries (or if it does, the changes are minor enough for the canonical syllable to be perceptually identified) and those who provide overwhelming evidence of the changes. Thus, Cao Jianfen (1992), based on the analysis of 200 Chinese neutral-tone syllables in the 2nd position in bisyllabic words, found that despite weakening of their intensity, shortening of their duration and neutralization of sound quality (consonant voicing, vowel neutralization to schwa etc), the original quality of a normal syllable is still perceptually identifiable due to relational invariance. However, high context sensitivity of the Chinese language from both perceptual and grammatical viewpoints determining the openness, flexibility and adaptability in speech (Wenzel, 2009: 124) casts doubt on the idea of identifying the syllable original quality, that has been naturally weakened, without a wider context, particularly if they resemble other Chinese syllables triggering homophony.

Chinese is a language where homophony is very extensive contrary to, e.g., English having little homophony (Huff, 2017). The reason, according to Huff, is in the number of unique syllables which in Chinese is about 7 times less than in English (around 1,200 vs around 9000, respectively). Huff identified three types of homophony: (1) identical phonemics + identical etymological tone, (2) identical phonemics

but different etymological tone, and (3) identical phonemics + unspecified tone (for a number of reasons, one of them being the case of neutral/zero tone). These types can be united under the term “inherent homophony” or phonological homophony. In addition to inherent homophony, there is homophony that results from natural modifications in the continuous speech flow giving way to syllable reduction of various degrees. We will refer to such homophony as phonetic homophony. It is crucial to understand that homophony is a perceptual phenomenon that a research can prove during perceptual studies. It means that if during acoustic study we find certain naturally modified syllables that look like other syllables, we can interpret them as potential homophones, and only when perceptual evidence is obtained, we can classify them as phonetic homophones – the ones that resulted from syllable reduction.

Instances of Chinese syllable reduction are abundant. According to a number of researches (Tseng, 2005; Cheng and Xu, 2009; Burchfield, Bradlow, 2014), the amount of reduced syllables in Chinese is greater than in English, and the reason for this is that, unlike in English, most syllables in Chinese are open (73–75%), and open syllables are more subject to reduction than closed syllables. It means that the openness of the Chinese syllable increases its chances for reduction.

In one of our previous studies based on spontaneous female speech (Li, Androsova, 2022), we obtained some statistics on modifications of Chinese syllables (syllable = minimal word in Chinese) that contextually functioned as monosyllabic or united in disyllabic and trisyllabic words or formed word combinations that functioned as single phonetic complexes (all of them further referred to as items or tokens). First, we found that 481 unique items or 882 tokens including repetitions were involved in modification process resulting in weak forms. 105 items out of 481 had more than 1 weak form. Some of the weak forms were a result of a single

modification (e.g., 1 instance of monophthonging or 1 instance of voicing, etc.) while others were a result of a series of modifications (e.g., omission + vowel quality reduction, etc.). Second, we obtained the proportion of word grammatical classes for the items in weak forms. It should be noted that, unlike in Indo-European languages, a word in Chinese does not have a permanent word-class label, instead, a particular context determines belonging to a particular word class. Among the 481 unique items having weak forms, only 148 were functional words while the majority of 333 words were notional words. However, if to consider repetitions, the proportion of functional and notional words was almost equal (470 and 412 tokens, respectively). The third finding was that contextually, among 148 unique functional words the majority (73) were adverbs (unlike in most Indo-European languages, adverbs in Chinese are considered functional words). Contextually, among 333 unique notional words the majority were nouns (149) and verbs (107) while adjectives, numerals and count words were the least frequent to be represented by weak forms – ten times less often (15, 13 and 12, respectively). Third, out of 882 instances of weak form use, 548 tokens preserved the canonical lexical tone while the remaining 334 tokens exhibited the change of canonical tone unrelated to tone sandhi but being a part of connected spontaneous speech processes. Mostly, it was the change to the neutral tone (in 302 tokens) but sometimes (in 32 tokens) Tone 1 was involved. Forth, as far as the number of syllables in the items, only 5 of them were monosyllabic words, the majority of 323 were disyllabic words, 121 items were trisyllabic words, and 32 were word combinations of 2 words (22 items), 3 words (8 items) and 4 words (2 items). The priority of disyllabic and trisyllabic words and word combinations over other patterns reflects the preference of disyllables and trisyllables for Chinese vocabulary (Arcodia, 2007; Xiao et al., 2009: 19, 40). We also found (Li, Androsova, 2022) that acoustically

weak forms of some of the items resembled other Chinese words and word combinations.

All this brought us to hypothesize the following:

(i) single Chinese syllables and syllable sequences that undergo natural modifications in non-prominent position in the utterance in the natural speech flow might resemble other single syllables and syllable sequences giving way to potential homophony;

(ii) these potential homophones segmented from the speech flow would hardly be identified as canonical syllables or syllable sequences; without a wider context, listeners might perceive many of them as other syllables / syllable sequences.

To test the hypothesis, we carried out a perceptual study whose results will be presented in the following paragraphs.

2. Materials and Methods

To perform this research, we used a combination of 3 methods from different scientific fields – linguistics, acoustics and psychology. A set of 44 tokens, or stimuli, (1 monosyllabic word, 31 disyllabic sequences and 12 sequences of three and four syllables) was constructed as testing material. The tokens were segmented from spontaneous monologues of 6 female Mandarin Chinese speakers aged 20-30 who volunteered to participate in the experiment for free. None of the subjects reported of any speech or hearing disorders, nor did any of them have any acoustic or phonetic knowledge. The recordings were made by the authors of this paper in the period of 2018–2019, in a quiet room at the frequency rate of 44000 Hz, 16-bit sampling rate, mono. At the first stage of the experiment, an acoustic study was carried out using PRAAT, version 6.0.05¹. The segmentation was performed using audiovisual control on the dynamic spectrogram. During the acoustic study, the above-mentioned tokens were segmented. The reason for

¹ Boersma, P. and Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program], (version 6.0.05), available at: <http://www.fon.hum.uva.nl/praat/> (Accessed 16 November 2015) (*In English*)

including the 44 tokens into the set was their naturally induced modification due to non-prominence in the speech flow (see more on prominence due to highlighting and chunking in (Cangemi, Baumann, 2020)) that resulted in potential homophony with other Chinese syllables / syllable sequences. The modifications were classified into three types taking into account the specifics of syllable status and syllable structure in Chinese:

Type 1 – initials (syllable-initial consonants) substitutions due to e.g. voicing and some other processes;

Type 2 – finals (vowel or vowel + n/ŋ part of the syllable) with vowel quality change as well as nasal locus change;

Type 3 – initials and finals omissions including the instances that lead to severe syllable reduction (=contraction and further coalescence).

It is necessary to make a few comments on the procedure of identifying the instances of consonant (Type 1) and vowel (Type 2) quality change. Acoustically, this change is reflected in (i) noise configuration and noise location change for consonants and (ii) formant frequency change for vowels. Some of these changes are slight variations while others are big enough to cause a phonemic change in languages like English (Cruttenden, 2001) and a syllable constituent (consonant or vowel) change for Chinese being aware of the fact that there are no phonemes in Chinese in the sense they are viewed in Indo-European languages.

At this point of Chinese phonetic studies, almost no parameter-value-based generalized results have been reported so far in the linguistic literature on acoustic cues of consonant changes and vowel formant changes in non-prominent parts of an utterance compared to its prominent parts. To the best of our knowledge, the only study in the field of consonants is the one of Rao and Shaw (2021) focusing on the acoustic cues of velar-labial merger of [x]-[f] in Zhongjiang Chinese (Southwestern Mandarin). As far as Chinese vowels, even the data on the vowel formant values are rare: Meng et al. (2006)

provided the averages for 5 formants (F1-F5) and standard deviations from those (a statistical study based of 212 female and 126 male subjects trained for broadcasting), Wang et al. (2020) targeted at building the vowel quality acoustic space in terms of F1 (vowel raising-lowering) and F2 (vowel advancement-retraction), Gu et al. (2003) measured F1 and F2 for 5 vowels in a carrier sentence each in 2 positions – focused and neutral. All the three studies involved only clearly pronounced monophthongs but obtained no averaged data for F1 and F2 in the flow of speech. The only work addressing the issue of modification acoustics is the one of (Tseng, 2005) where he pointed to F1-F2 raising/lowering, the absence of standard cues in cases of vowel and consonant deletion and the energy redistribution on dynamic spectrograms for some cases of consonant modifications.

Therefore, our decisions on identifying a consonant and vowel constituent change dwelt upon comparing (i) a naturally modified consonant spectrum with canonical spectral energy distribution of a consonant in terms of frequency, duration and intensity, (ii) F1-F2 averages measured in prosodically prominent utterance parts for clearly pronounced vowels with F1-F2 averages measured in prosodically non-prominent parts with fast speech rate and no sentence stress where the vowels were naturally modified.

Types 1-2 change the so-called phonemics of the canonical syllables while Type 3, often accompanied by Types 1-2, can change the number syllables in the sequences (Li, Androsova, 2019). We will interpret the syllables undergoing any of the modification types as weak forms.

The tokens were 14 functional words (e.g. 但是 /tanʃi/ 'but', 虽然 /sueizən/ 'although' etc) and 30 notional words including word combinations of notional words with functional words (e.g. 喜欢 /eixuan/ 'like', 阴暗 /jin'an/ 'dark', 考完 /kauwan/ 'after exam', 如果你有 /zukuonijou/ 'if you have' etc). The notional words

represented different word classes (nouns, verbs, adjectives etc).

The stimuli were randomized and given to 2 groups of volunteer listeners who were native Chinese speakers (20 males, 20 females, aged 20–23, 3d-year students of Musical Department of Shanxi Datong University, all of them having either 2nd or 3d level certificate of *Putonghua – Mandarin Chinese Standard*). We informed the listeners that the stimuli had been segmented from Chinese speech.

Each listener received a questionnaire with 44 questions (according to the number of tokens). To each question, 3 options were offered as answers: target token (canonical syllable / syllable sequence) and 2 potential homophones: first homophone represented actual realization of the syllable / syllable sequence, second was a variant acoustically close to the first potential homophone. We limited the number of listenings to 3 for each token. During the experiment, there were 3 failures to perform the task, and in 2 instances the listeners ignored the options

and offered their own variant. Therefore, the real total number of reactions was taken as 100%.

3. Results

As a result, 1758 listeners' reactions were obtained. The data was further processed and the reactions were distributed into 2 groups: 1) more than 50% identification of canonical syllable sequence (target token) including reliable canonical syllable / syllable sequence identification by 70% of listeners or more; 2) 50 or more percent of homophones identification including reliable homophones identification by 70% of listeners of more. 50% value was considered as absolute impossibility to discriminate between canonical syllable / syllable sequence and homophone syllable / syllable sequence. Reactions belonging to the second group were interpreted as non-identification of the canonical syllable / syllable sequence.

Table 1² shows the distribution of the reactions. Group 1 comprised less than 16% of all reactions, the remainder formed group 2 – 84%.

Table 1³. Potential homophones perception by Chinese listeners

Таблица 1. Восприятие потенциальных омофонов носителями китайского языка

	Stimuli given for discrimination				Perceived			
	Target token		Realized token		C (%)		H (%)	
	hieroglyphs + translation	CT	AT	homophone+translation	m	f	m	f
1	但是 'but'	tançi	tanşu	单数 'singular' / 大树 'big tree'	5	0	95	100
2	考完 'after exam'	kauwan	kaŋ	抗 'anti' / 口岸 'port'	0	5	100	95
3	时候 'when'	şixou	şou	收 'receive' / 朔 'new moon'	10	0	90	100
4	虽然 'although'	sueizān	suan	酸 'sour' / 三 'three'	10	10	90	90
5	黑龙江省 'Heilongjiang province'	xeilunŋeşəŋ	xeilunŋeţəŋ	黑龙江人 'Heilongjiang people' / 洪江人 'Hong Jiang residents'	25	10	75	90

² This is the unpublished material of Li Yifang's Ph.D. dissertation prepared under the scientific supervision of Svetlana V. Androsova (Li, 2022).

³ Tables 1–3 and Figure 1 are designed by authors.

6	黑河市 'Heihe' (city)	xeixɿʃi	xɿʃi	鹤市 'He Shi city' (in Jiang Su province) / 黑河日 'Heihe city day'	35	25	65	75
7	酸 'sour'	suan	san	三 'three' / 仨 'three'	15	10	85	90
8	比较 'compare'	bitɕiau	teiau	脚 'foot' / 腰 'waist'	0	5	100	95
9	这种 'this kind of'	tʃɿtʃuŋ	tʃuŋ	中 'among' / 阵容 'lineup'	15	10	85	90
10	喜欢 'like'	ɕixuan	ɕiŋ	行 'alright' / 宣 'declare'	0	5	100	95
11	东西 'thing'	tunɕei	tunɕji	冬衣 'winter clothes' / 懂你 'understood you'	0	0	100	100
12	女巫 'witch'	nywu	niu	牛 'cow' / 你好 'hello'	45	45	55	55
13	就是 'just/like'	teiuʃi	teiu	揪 'drag' / 交 'pay'	30	35	70	65
14	那种 'such kind'	nAɿtʃuŋ	neizɿuŋ	内容 'content' / 弄 'do'	0	5	100	95
15	其他 'other'	teitAɿ	teiAɿ	掐 'pinch' / 起亚 'Kia' (car brand)	45	55	55	45
16	主要 'main'	tʃɿjau	tʃijau	只要 'as long as' / 教 'teach'	45	30	55	70
17	小时候 'when I was young'	ɕiauʃixou	ɕiauʃou	小勺 'small spoon' / 销售 'sales'	80	85	20	15
18	然后 'afterward'	zənɿxou	zou	肉 'meat' / 绕 'wrap around'	100	100	0	0
19	这样 'such'	tʃɿjaŋ	teiaŋ	酱 'sauce' / 检验 'examin'	80	100	20	0
20	黑河市 'Heihe city'	xeixɿʃi	kɿʃi	格式 'scheme' / 鹤市 'He Shi city' (in Jiang Su province)	20	25	80	75
21	代购 'trafficker'	taikou	touɿzɿ	兜儿 'pocket' / 代号 'codename'	20	35	80	65
22	军事化 'militarization'	teuənʃixuAɿ	teuənɿxau	军号 'military bugle' / 一句话 'one phrase'	25	20	75	80
23	如果你有 'if you have'	zukuonijou	tʃuwoniou	煮蜗牛 'boil a snail' / 乳蜗牛 'milk snail'	15	30	85	70
24	现在 'right now'	ɕiantsai	ɕi'an	西安 'Xian city' / 鲜 'fresh'	15	20	85	80
25	想不起来 'can't remember'	ɕiaŋbuteilai	ɕiaŋbueiaŋ	想不想 'do you want' / 相片 'photo'	35	45	65	55
26	没有 'no'	meijou	miu	谬 'absurd' / 妙 'wonderful'	25	40	75	60
27	阴暗 'dark'	jin'an	jan	烟 'smoke' / 呀 'Yeah'	5	5	95	95

28	出来 'come out'	tʂulai	tʂuai	揣 'put' / 穿 'wear'	50	60	50	40
29	都有 '(everybody) has'	toujou	tou	逗 'to joke' / 等我 'wait for me'	20	15	80	85
30	但是 'but'	tanʂi	tanʂən	单人 'individual' / 大人 'adult'	15	35	85	65
31	得有 'must have'	teijou	tiau	迪奥 'Dior' / 丢下 'to leave'	45	35	55	65
32	东西 'thing'	tunʂei	nuŋjɿ	农业 'agriculture' / 农民 'farmer'	35	50	65	50
33	西安 'Xian city'	ei'an	ɛian	鲜 'fresh' / 天 'heaven'	75	90	25	10
34	或者 'or'	xuotʂɿ	wotʂɿ	窝着 'to be a stay-at-home guy' / 过着 'to live'	5	5	95	95
35	人家 'other people'	zənteia	lianteie	连接 'connection' / 横切 'crossing'	65	45	35	55
36	虽然 'although'	sueizən	ɛian	香 'fragrant' / 酸 'sour'	0	5	100	95
37	现在 'right now'	ɛiantsai	ɛin'an	心安 'calmness' / 寻思 'think'	25	25	75	75
38	你要 'you need'	nijau	niau	鸟 'bird' / 你好 'hello'	40	40	60	60
39	这个 'this one'	tʂɿkɿ	zɿʌ	热啊 'hot' / 热 'heat'	15	0	85	100
40	所以 'and so'	suoji	suojou	所有 'все' / 左右 'left and right'	5	5	95	95
41	大家 'everyone'	tʌteia	teiateia	家家 'every family' / 大牙 'large teeth'	65	55	35	45
42	所以 'and so'	suoji	suei	虽 'although' / 所有 'all'	75	75	25	25
43	学校 'school'	ɛyɛiaɯ	ɛiaɯ	笑 'smile' / 桥 'bridge'	15	15	85	85
44	就要 'is going to'	teiujaɯ	teauju	教育 'education' / 叫 'call'	10	10	90	90

Note to Table 1: C – canonical word/word group, H – homophones, CT – Canonical transcription (IPA), AT – acoustic transcription. Pinyin transfer in to IPA symbols is performed in accordance with the table of matches⁴.

⁴ Baidu venku (百度) [Chinese online encyclopedia], available at: <http://www.baidu.com/> (Accessed 20 December 2019) (*In Chinese*)

The proportion of canonical reactions vs homophone 1 + homophone 2 displayed certain gender differences (3 reactions or more, 7-15% in fact). In particular, the differences were, first, in the cases where homophones prevailed – the degree of prevalence was different for tokens 5, 16, 21, 23, 26, 30, 39. For token 33 with the prevalence of canonical reactions, the degree of prevalence was also gender-specific. Third, for token 28, males demonstrated impossibility to differentiate between the

canonical token and the homophone while females preferred the canonical token, however, for token 32 it turned out vice versa; males interpreted token 35 as the canonical one and females – as the homophone. In addition, males gave more canonical reactions to tokens 5, 16, 35, 39 while females gave more canonical reactions to tokens 13, 19, 21, 23, 30, 32, 33 (the difference of 3 reactions or more).

Table 2⁵ shows details on each of the 2 homophones perception in absolute values.

Table 2. Details of homophones perception (absolute values)

Таблица 2. Детализация по восприятию омофонов (в абсолютных единицах)

	Target token		Homophone 1		Homophone 2	
	m	f	m	f	m	f
1	1	0	19	19	0	1
2	0	1	19	19	1	0
3	2	0	14	11	4	9
4	2	2	18	18	0	0
5	5	2	14	18	1	0
6	7	5	12	15	1	0
7	3	2	17	18	0	0
8	0	1	13	11	7	8
9	3	2	17	18	0	0
10	0	2	15	15	5	3
11	0	0	9	12	11	8
12	9	9	11	11	0	0
13	2	6	14	8	4	6
14	0	1	18	19	2	0
15	9	11	7	7	4	2
16	9	6	6	9	5	5
17	16	17	4	2	0	1
18	20	20	0	0	0	0
19	16	20	3	0	1	0
19	16	20	3	0	1	0

⁵ This is the unpublished material of Li Yifang's Ph.D. dissertation prepared under the scientific supervision of Svetlana V. Andosova (Li, 2022).

	Target token		Homophone 1		Homophone 2	
	m	f	m	f	m	f
20	4	5	5	7	11	8
21	4	7	9	11	7	2
22	4	5	15	12	1	3
23	3	6	13	9	4	5
24	3	4	12	8	5	8
25	7	9	12	10	1	1
26	6	8	10	6	4	6
27	1	1	15	16	4	3
28	10	12	7	4	3	4
29	4	3	13	16	3	1
30	3	7	12	10	5	3
31	9	7	7	9	4	4
32	7	10	12	8	1	2
33	15	18	4	2	1	0
34	1	1	18	18	1	1
35	14	9	5	9	1	2
36	0	1	20	19	0	0
37	5	6	15	13	0	1
38	8	8	10	11	2	1
39	3	0	13	14	4	6
40	1	0	18	20	1	0
41	13	11	6	7	1	2
42	15	15	5	5	0	0
43	3	5	17	15	0	0
44	2	2	17	13	1	5

It was discovered during the data analysis that there were instances when males preferred 1st homophone, while females tended to choose 1st and 2nd homophone in equal proportion (tokens 3, 13, 24 and 28 in Table 2). The contrary result was obtained for tokens 16 and 21 where females preferred 1st homophone, while males chose homophone 1 and homophone 2 in equal

proportion. Only 1 case was found when males gave preference to homophone 2 (token 20), while female preferences for homophone 1 and homophone 2 were distributed equally. For tokens 23 and 41, both males and females preferred 1st homophone to 2nd homophone, but the preference was higher with males compared to females.

The significance of these gender differences was tested using *chi-square test* and Fisher's test. About 10% of the data turned out to demonstrate very significant gender differences and about 15% – significant differences. Thus, about ¼ of the data show significant gender difference, while ¾ of the data show insignificant difference or no difference whatsoever.

Here we are presenting the descriptive statistics on Group 1 (identification of canonical syllable by the majority – more than 50%) and Group 2 (non-identification of canonical syllables – 50% or less). 16% of male responses appeared in Group 1 (7 tokens out of 44). Reliable identification, i.e. by the vast majority (70% and more), accounted for 11% out of 16%. Non-identification of canonical syllables accounted for 84% of male responses – 37 out of 44 tokens were not identified as normal canonical syllables / syllable sequences therefore they were distributed into Group 2. Within this group,

three subgroups were formed: reliable identification ($\geq 70\%$) – 64%, identification by the majority ($>50\%$ but $<70\%$) – 18% and absolute impossibility to discriminate between canonical syllable / syllable sequence and homophone syllable / syllable sequence ($=50\%$) – 2%.

As far as the female listeners, the number of responses distributed into Group 1, Group 2 and to the subgroup of reliable responses within Group 1 was the same as for the males. However, the distribution within Group 2 was different. The number of reliable responses was lower: 59% (vs 64% with male listeners) account for reliable identification, other majority responses accounted for 23% (vs 18% with male listeners). There was no difference as far as absolute impossibility to discriminate between canonical tokens and homophone tokens. Table 3⁶ and Figure 1⁷ show the distribution into the 5 subgroups. Although the details are a little different, the distribution pattern is clearly similar.

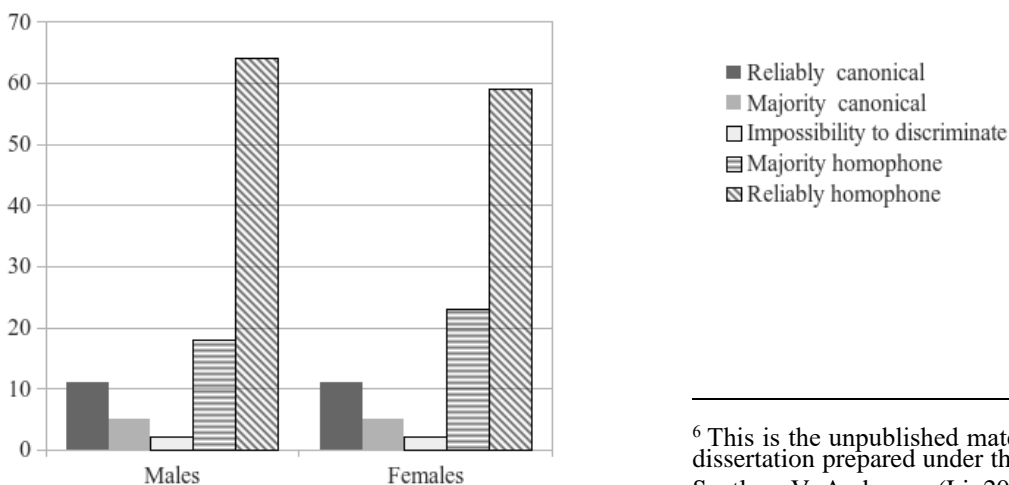
Table 3. Data grouping (%)

Таблица 3. Группировка данных (%)

	Reliably canonical	Majority canonical	Impossibility to discriminate	Majority homophone	Reliably homophone
males	11	5	2	18	64
females	11	5	2	23	59

Figure 1. Data grouping (%)

Рисунок 1. Группировка данных (%)



⁶ This is the unpublished material of Li Yifang's Ph.D. dissertation prepared under the scientific supervision of Svetlana V. Andosova (Li, 2022).

⁷ Ibid.

4. Discussion

This study showed the results contrary to Cao Jianfen (1992) as the original quality of a normal syllable in our experiment was not perceptually identified for 84% of the tokens which is the vast majority. There might be several reasons for such an outcome. Firstly, the corpus for our study included only potential homophones that emerged as a result of naturally induced modifications. No homophony was reported by Cao Jianfen. Secondly, all stimuli in the experiment of 1992 were bisyllabic sequences, but among our tokens there were also 3 and 4 syllable sequences and one monosyllabic token. Thirdly, due to the fact that our tokens occurred in spontaneous monologues and tokens for the 1992 experiment were syllables from neutral type bisyllabic words (2nd position), relational invariance that Cao Jianfen insisted upon did not work in our case.

There is nothing surprising in the obtained results of our perceptual study. All the 44 tokens are plausible candidates for homophony as they were naturally reduced to differing extents (whether it is a consonant / vowel change or syllable merger as a result of omissions and further vowel and/or consonant change) in favor of articulatory convenience triggered by context sensitivity characteristic for the Chinese syllable and speech rapidness. Below, we can see the classes of modifications exemplified by the tokens from Table 1.

1. Consonant substitutes (Table 1, tokens 2, 5, 11, 39):

– terminal [ŋ] → [ŋ] or [ŋ] → [n], e.g.:

a) [n] → [ŋ] 考完 'after exam' /kauwan/ → [kaŋ] (forelingual nasal → velar nasal);

b) [ŋ] → [n] 黑龙江省 'Heilongjiang province' /xeilunɥeʂəŋ/ → [xeilunɥeʂən] (velar nasal → forelingual nasal);

– sonantization of initials:

[ç] → [j] e.g. 东西 'thing' /tuŋci/ → [tuŋji] (fricative → glide);

– occlusion weakening and voicing:

[tʂ] → [z] e.g. 这个 'this' /tʂɤkɤ/ → [zɤʌ] (voiceless affricate → voiced fricative in initial of the 1st syllable).

2. Vowel substitutes (Table 1, tokens 1, 7, 14, 16, 40, 44):

– monophthong-to-monophthong or diphthong-to-diphthong (front to back as well as back to front vowel finals):

[i] → [u] e.g. 但是 'but' /tanʃi/ → [tanʃu] (2nd syllable, front to back);

[u] → [i] e.g. 主要 'main' /tʃujau/ → [tʃijau] (1st syllable, back to front);

– monophthonging of diphthong:

[ua] → [a] e.g. 酸 'sour' /suan/ → [san];

– diphthonging of monophthongs:

[ʌ] → [ei] e.g. 那种 'such kind' /nʌtʃuŋ/ → [neizʉŋ] (1st syllable);

[i] → [ou] e.g. 所以 'and so' /suoji/ → [suoju] (2nd syllable);

[au] → [u] e.g. 就要 'is going to' /teiuju/ → [teuju] (2nd syllable).

3. Whole syllable omission (Table 1, tokens 43, 29) e.g. 学校 'school' /ɥeɥeiau/ → [eiau] (1st syllable drop), 都有 '(everybody) has' /touju/ → [tu] (2nd syllable drop).

4. Syllable parts (consonants, vowels) omissions with further liaison: 考完 'after exam' /kauwan/ → [kaŋ], 时候 'when' /ʃixou/ → [ʃou], 虽然 'although' /sueizən/ → [suan] (Table 1, tokens 2–4).

There are instances when contracted syllables may be conventionalized or lexicalized by a fixed character in the written form (Tseng, 2005: 154). At present time, such processes start in social networks – 15 neologisms of this kind (Li, 2020) were found in Chinese online Encyclopedias Baidu (<http://www.baidu.com/>).

This naturally induced variability created homophony identified in the tokens. The same processes are abundant in spontaneous speech flow, and when they do not trigger homophony, the normal canonical syllable might be identified due to the relational invariance advocated by Cao Jianfen (1992). In any case, no matter leading to homophony or not, naturally-induced

modifications have to be identified and classified, the occurrences of a particular modification class have to be calculated, and the distribution of different classes of modifications must be revealed. The study of syllable phonemics changes initiated by (Tseng, 2005) is getting more support (Cheng and Xu, 2009; Burchfield and Bradlow, 2014) including specialists in creating models for continuous Chinese speech recognition who use context-dependent data (Hao Wu, Xihong Wu, 2007), but these issues are still far from being thoroughly studied.

5. Conclusion

Literature analysis has shown that Chinese weak forms study is still in its infancy. To the best of our knowledge, no previous research has focused on the perceptual perspective to test the hypothesis that naturally induced modifications of Chinese syllables can result in homophony or even homonymy that listeners cannot resolve in narrower context.

This multidisciplinary study was performed using a combination of acoustic, perceptual and traditional linguistic methods. The obtained results of 84% of non-identification of canonical syllables bring an important insight into the common patterns of human linguistic behavior whose integral part are weak forms. Neglecting careful articulation in those parts of utterances that lack prosodic prominence due to the lack of informativeness is a common strategy for the speaker as far as it gives an opportunity to make important parts of the utterances more prominent, more carefully articulated. It enables the listener to base on these prominent, carefully articulated parts to rebuild the utterance despite phonetic losses in non-prominent parts.

Native speakers and listeners develop this ability during their lifetime interaction in the language environment. Therefore, it is a certain implication for Chinese as L2 teaching methods as well. There is a vital need to identify and collect common weak forms and include them in practical courses of developing listening and speaking skills to

provide adequate interaction between speaker and listener characteristic for native speakers.

References

- Androsova, S. V. (2015). Nekanonicheskie fonologicheskie modeli morfem i slov v russkom i angliyskom yazykakh [Non-canonical modes of morphemes and words in Russian and English], *Teoreticheskaya i prikladnaya lingvistika* [Theoretical and Applied Linguistics], 1 (1), 5-15. (In Russian)
- Arcodia, G. F. (2007). Chinese: A Language of Compound Words?, in Montermini, F., Boyé, G. and Hathout, N. (eds.), *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, Somerville, US, 79-90. (In English)
- Babushkina, E. S. (2012). K probleme izmeneniya zvukovykh kharakteristik slov kitaiskogo yazyka v potoke rechi [Phonetic pattern changes of Chinese words in the speech flow], *Molodoi uchenyi* [Young scientist], 6 (41), 229-230. (In Russian)
- Bondarko, L. V., Verbitskaya, L. A., Gordina, M. V., Zinder, L. R. and Kasevich, V. B. (1974). Stili proiznosheniya i tipy proizneseniya [Pronunciation styles and types], *Voprosy yazykoznaniiya* [Topics in the Study of Language], 2, 64-70. (In Russian)
- Borisova, E. G. (2018). Upravlenie ponimaniem. Yazykovye edinitiy, reguliruyushchie ponimanie soobshcheniya [Understanding Management: Language Units as Indicators of Message Comprehension], *Vestnik Mosk. un-ta. Ser. 9: Filologiya* [Moscow University Bulletin, Ser. 9: Philology], 6, 34-50. (In Russian)
- Burchfield, L. A. and Bradlow, A. R. (2014). Syllabic reduction in Mandarin and English speech, *JASA*, 135 (6). <http://dx.doi.org/10.1121/1.4874357> (In English)
- Cao, J. (曹剑芬). (1992). On neutral-tone syllables in Mandarin Chinese, *Canadian Acoustics*, 3, 49-50. (In English)
- Cangemi, F. and Baumann, S. (2020). Integrating phonetics and phonology in the study of linguistic prominence, *Journal of Phonetics*, 81, 1-7. <https://doi.org/10.1016/j.wocn.2020.100993> (In English)
- Cheng, C. and Xu, Y. (2009). Extreme reductions: Contraction of disyllables into monosyllables in Taiwan Mandarin, *Proceedings*

of Interspeech, ISCA, Brighton, UK, 456-459. (In English)

Cruttenden, A. (2001). *Gimson's pronunciation of English*, Edward Arnold Limited, China. (In English)

Fu, X. and Li, L. (2018). Hebei jiaotong zhiye jishu xueyuan (2018 nian 6 yue). yuying ganzhi shiyan jishu de xintansuo [New exploration of speech recognition technology], *Journal of Shijiazhuang Institute of Railway Technology*, 2, 115-117. (In Chinese)

Gu, Z., Mori, H. and Kasuya, H. (2003). Analysis of vowel formant frequency variations between focus and neutral speech in Mandarin Chinese, *Acoustic Science and Technology*, 24 (4), 192-193. (In English)

Huff, Ch. L. (2017). A study of Mandarin homophony, Ph.D. Thesis, Hertford College, Oxford, UK. (In English)

Jones, D. (1969). *An Outline of English Phonetics*, 9th edn. (1st edn. 1918), Heffer, Cambridge, UK. (In English)

Karavaeva, V. G. (2018). Pertseptivnye kharakteristiki intervokal'nykh soglasnykh realizatsii na meste orfograficheskikh -t-, -tt- (na materiale britanskogo novostnogo analiticheskogo diskursa) [Perceptual features of intervocalic consonant realizations used for -t-, -tt- (Based on British news analytic discourse)], *Vestnik Sankt-Peterburgskogo universiteta. Yazyk i literatura* [Vestnik of Saint Petersburg University. Language and Literature], 15 (3), 424-437. <https://doi.org/10.21638/spbu09.2018.308> (In Russian)

Kasevich, V. B. (2006). Fonologicheskie problemy obshchego i vostochnogo yazykoznaniya [Phonological issues of general linguistics and Asian linguistic studies], *Trudy po yazykoznaniyu* [Essays in the study of language], Philological Faculty of St Petersburg University Press, St. Petersburg, Russia, 10-238. (In Russian)

Laver, J. (1994). *Principles of Phonetics*, Cambridge University Press, Cambridge, UK. (In English)

Li, Y. (2022). Slabye formy slogov v spontannoy rechi zhenshchin (eksperimentalno-foneticheskoe issledovanie na materiale kitayskogo yazyka) [Weak forms of syllables in female spontaneous speech (Experimental phonetic study based on Chinese)], Ph.D. Thesis, Amur State University, Blagoveshchensk, Russia. (In Russian)

Li, Y. and Androsova, S. V. (2019). Foneticheskie osobennosti slov v ikh obychnykh funktsiyakh i v kachestve slov-parazitov (na materiale kitayskogo yazyka) [Phonetic patterns of words in their ordinary functions and as parasite words (Based on Chinese)], *Teoreticheskaya i prikladnaya lingvistika* [Theoretical and Applied Linguistics], 5 (3), 99-116. https://doi.org/10.22250/2410-7190_2019_5_3_99_116 (In Russian)

Li, Y. and Androsova, S. V. (2022). Slabye formy kitayskikh slogov i ikh kombinatsiy: akusticheskiy aspekt [Weak forms of Chinese syllables and their combinations], in Fayzieva, G. V. (ed.), *Diskursivnye praktiki: zhanrovye i foneticheskie aspekty* [Discursive practices: Genre and phonetic issues], PIEF Press, Perm, Russia. (In Russian)

Lobacheva, M.V. and Androsova, S. V. (2016). K probleme statusa i modifikatsii vnutrennei struktury sloga v slogovykh yazykakh (na materiale kitayskogo yazyka) [On the status and modifications of the internal syllable structure in syllabic languages (Based on Mandarin Chinese)], *Teoreticheskaya i prikladnaya lingvistika* [Theoretical and Applied Linguistics], 2 (4), 23-40. (In Russian)

Meng, Z., Chen, Y. and Li, X. (2006). Statistical Survey of Monophthong Formants in Mandarin for Students Being Trained as Broadcasters, *Proceedings of the 20th Pacific Asia Conference on Language, Information and Computation*, Tsinghua University Press, Huazhong Normal University, Wuhan, China. 280-286. (In English)

Mücke, D., Hermes, A. and Cho, T. (2017). Mechanisms of regulation in speech: Linguistic structure and physical control system, *Journal of Phonetics*, 64, 1-7. <http://dx.doi.org/10.1016/j.wocn.2017.05.005> (In English)

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes, *New Ideas in Psychology*, 25, 143-170. (In English)

Priva, U. C. and Strand, E. (2023). Schwa's duration and acoustic position in American English, *Journal of Phonetics*, 96, 1-16. <https://doi.org/10.1016/j.wocn.2022.101198> (In English)

Rao, D. and Shaw, J. A. (2021). The role of gestural timing in non-coronal fricative mergers in Southwestern Mandarin: Acoustic evidence from a dialect island, *Journal of Phonetics*, 89.

<https://doi.org/10.1016/j.wocn.2021.101112>

(In English)

Shcherba, L. V. (1957). O raznykh stilyakh proiznosheniya i ob ideal'nom foneticheskom sostave slov [Different styles of pronunciation and canonical phonemic models of words], in Shcherba, L. V. (ed.), *Izbrannye raboty po russkomu yazyku* [Selected works in the Russian language], Uchpedgiz Press, Moscow, Russia, 21-25. (In Russian)

Shen, I. (2018). Ieroglif «chinarei» i kitaiskii ieroglif [The Traditional Chinese Hieroglyph and the Hieroglyph of the Chinari Literary Group], *Vestnik Mosk. un-ta. Ser. 9: Filologiya* [Moscow University Bulletin, Series. 9: Philology], 6, 131-140. (In Russian)

Xiao, R., Rayson, P. and McEnery, T. (2009). *A Frequency Dictionary of Mandarin Chinese: Core Vocabulary for Learners*, Routledge, London, UK. (In Chinese and English)

Wang, Z. N., Chen, Y., NG, M. L., Yao, L. and Zhang, W. (2020). Acoustic analysis of Mandarin Chinese vowels produced by young adults, *Rehabilitation Medicine*, 30 (3), 183-191.

<https://doi.org/10.3724/SP.J.1329.2020.03004>

(In English)

Zinder, L. R. (1981). Realnyi potok rechi i «rekonstruktsiya» fonemnogo sostava slov [Speech flow and word phonemic pattern reconstruction], in Avanesov, R. I., Bondarko, L. V. and Zinder, L. R. (eds.), *Teoriya yazyka. Metody ego issledovaniya i*

prepodavaniya: k 100-letiyu so dnya rozhdeniya Lva Vladimirovicha Shcherby [Language theory. Methods of language study and teaching: A commemorative volume in honor of L. V. Scherba], Nauka Press, Leningrad, Russia, 102-106. (In Russian)

Все авторы прочитали и одобрили окончательный вариант рукописи.

All authors have read and approved the final manuscript.

Конфликты интересов: у авторов нет конфликтов интересов для декларации.

Conflicts of interests: the authors have no conflicts of interest to declare.

Yifang Li, Ph.D. in Philology, Teacher, Department of Russian and Japanese Languages, Shanxi Datong University, Datong, China.

Ифан Ли, кандидат филологических наук, преподаватель факультета русского и японского языков, Датунский университет, Датун, КНР.

Svetlana V. Androsova, Doctor of Philology, Professor, Department of Foreign Languages, Amur State University, Blagoveshchensk, Russia.

Светлана Викторовна Андросова, доктор филологических наук, профессор, профессор кафедры иностранных языков, Амурский государственный университет, Благовещенск, Россия.