Conditions and Effects of Homophony

In Chinese

Patrick Runkle
Finished Draft
December 10, 1997
Advisor: Donna Jo Napoli
Introduction

A billion people can't be wrong, and so, I say, Chinese language and Chinese linguistics have much to offer. From its typology to its orthography, Chinese is often severely misunderstood and misrepresented in the West. Many Chinese linguists feel that modern linguistics, even though it has made large advances, still looks at everything through a Western filter. It is crucial, therefore, to put each argument in perspective, and attempt to separate our pre-conceived linguistic notions about Western languages from the matter at hand.

In this thesis, I will look at one particular phenomenon that not only deeply piqued my interest as a student of Chinese, but also illuminates phonological issues not found in Western linguistics: Homophony. My study of Chinese homophony will look at two key areas: First, I will look at the historical record of Chinese, in order to shed some light on how the phenomenon arose. Second, I will look at the effects in the language of homophony, including a trial that I hope to run in order to clarify some important questions.

Before the meat of my thesis can begin, however, some serious background on the Chinese language is necessary.

Background
Chinese is the largest member of the Sino-Tibetan language family, which also includes Tai, Burmese, and a few other Southeastern Asian languages. Chinese is spoken by the 汉族 han4zu2 'Han Tribe,' the race to which the people living in China belong.

Chinese has an enormous number of dialects, many of which are mutually unintelligible. Many linguists refer to Chinese dialects as separate languages because of how far apart they are in terms of intelligibility. Most Chinese linguists, however, refer to the Chinese dialects as one language with many varieties, because of the 3,000-year written record, which encompasses all dialects, and because all speakers are still united under the general political and racial term 'Chinese.' Also, the syntax of most Chinese dialects, except for several small word order issues, is entirely similar. Make no mistake, though: Many varieties of Chinese are as distinct (and mutually unintelligible) as German and English.

The standard dialect of Chinese is known as Mandarin; it is the official language of the Chinese government, and is taught in schools. According to the most recent data (Li 1989), 71.5% of Chinese speakers have some form of Mandarin (in Chinese, 官话, Guanhua) as their native dialect, although the regional variation can be quite strong, as would be expected given the land area and population.
While Mandarin is primarily a Northern dialect, central China has several dialect families (Wu 吳, Gan 赣, and Xiang 湘) accounting for 13.9% of Chinese speakers. Shanghaisese is the most recognizable dialect in this group, and my trial with native speakers hopefully will include some work with Shanghaisese.

Southern Chinese dialects are the native tongue of 13.2% of Chinese language speakers. These include the Min 闽, Yue 粤, and Kejia 客家 dialect families. Cantonese (廣州話 guang3zhou1hua4 'Canton speech'), which is spoken in Hong Kong and the Canton province, is a Yue dialect. This dialect family is very well-represented in America and other overseas communities, as many of the first Chinese to emigrate came from Hong Kong and southeastern China.

Chinese is a tone language, and its tones are both distinctive and lexical, which means that the same phonetic syllable pronounced with a different pitch or pitch contour is a different lexical item. Each dialect has a certain number of set tones and tone contours that can be intoned on each syllable. Mandarin has 4 tones: a constant high tone (tone 1), a high rising tone (tone 2), a low rising (tone 3), and a falling tone (tone 4). The classic example is the syllable ma.

In Mandarin, ma1 is written as 媽 and means 'mother.' ma2 is written as 麻 and means 'hemp.' ma3 is written as 馬, which
means 'horse,' and 马 is a verb meaning 'to scold.' These words are all quite semantically distinct, and the characters are also distinct. It is important to note that every character which contains 马 as a part of it will not necessarily be phonetically related to ma, even though most Chinese characters have a phonetic 'element' like 马 in them. (More on this later...)

Chinese tones are special because they are lexical. In the African tonal languages, very few tones are actually lexical in nature, that is, they are predictable based on word class, position in the sentence, etc. But in Chinese, every word comes with an embedded tone. When put in combinations, the tones can undergo many changes that frustrate linguists. In Mandarin, for instance, the most famous tone sandhi occurs when a 3rd tone word occurs before a 2nd or 3rd tone syllable. In this case, the first 3rd tone syllable changes to a 2nd tone. In other dialects, however, there can be as many as 8 or 9 tones. Tones did not necessarily exist in Old Chinese, but they did exist in Middle Chinese.

Research on the origins of tones (tonogenesis, a term coined by James Matisoff) has indicated that they most often come from pitch distinctions that, at the outset, are predictable. All words that began with one kind of consonant had some kind of pitch associated with them, or all words
ending with the same vowel cluster had some kind of pitch movement. Over time, even though the phonetic distinctions that separated the tonal groups were lost, the tonal distinctions remained, and that's what is left in Chinese today. (Baxter 1992)

Romanization

Representing Chinese in a comprehensible way in alphabetic script is always a perplexing problem. For this paper, in the Mandarin examples I will be using 罗马拼音, which is luo2ma3pinlyin1, most oftenly called just Pin-Yin in English. This system is also called National Romanization, and it is the most comprehensible and direct (as well as accepted by the PRC) system for representing Chinese with our alphabet. For the non-Mandarin examples, and for the Old and Middle Chinese samples, I will use IPA as used in the various sources I have consulted.

The following is a good guide for translating the Pin-Yin romanization standardly used for Mandarin into IPA:

\[\begin{align*}
\text{b} &= [p] & \text{d} &= [t] & \text{g} &= [k] & \text{j} &= [ʧ] & \text{z} &= [ts] & \text{zh} &= [ts] & \text{y} &= [j] \\
\text{p} &= [p'] & \text{t} &= [t'] & \text{k} &= [k'] & \text{q} &= [tʃ] & \text{c} &= [ts'] & \text{ch} &= [ts'] & \text{w} &= [w] \\
\text{m} &= [m] & \text{n} &= [n] & \text{h} &= [x] & \text{x} &= [χ] & \text{s} &= [s] & \text{sh} &= [s] \\
\text{f} &= [f] & \text{l} &= [l] & \text{r} &= [z]
\end{align*}\]
a = [a]  o = [o]  e = [æ]  i = [i]  u = [u]  y = [y]
-i = [ɨ] after z,c,s; [ŋ] after zh, ch, sh

Homophony in Mandarin

Morphologically, Chinese is an isolating language, which means that it isolates the syllable as the smallest unit in the language, and prefers to make new words by compounding. By and large, Chinese doesn't use inflection as a tool. There is no verbal inflection, no case agreement, or anything of the sort found in Indo-European languages. Chinese does form some words by processes other than compounding, but these are not terribly common, and usually limited to reduplication and cliticization.

Unfortunately, this has led to the widely-held belief that Chinese is a barren, vague language filled with monosyllabic words. Because every character is represented one-to-one by a syllabic word, the tendency to categorize Mandarin as monosyllabic is very high.

Early work on modern Chinese linguistics (most often conducted by Europeans) is entirely embarrassing. Brenhard Karlgren, a Swiss who attempted the first dictionary of Old Chinese, made many important contributions to the field, but had so many misperceptions about the nature of the Chinese language as to render some of his work laughable. He was the first to bring the Monosyllabic Myth, as it is called, to the
West. His book is full of broad generalizations like "When we look at the sounds of the language we find that modern Mandarin Chinese is extraordinarily simple and that it is poor in its resources" and "It [Chinese] is monosyllabic, i.e., every single word consists of one single syllable." (Karlsgren 1949)

This is totally, totally wrong. There are a whole bunch of bisyllabic roots, and many of them contain characters that are completely bound and only used in that word, like 蝴蝶 hu2die2 'butterfly' and 葡萄 pu2tao 'grape'.

John DeFrancis, the world-renowned Chinese linguist, wrote an essay called "The Monosyllabic Myth," in which he debunks all of these misconceptions. He counts only a small percentage of words in Mandarin that actually occur outside of compounds, somewhere between 30-40%, which leaves 60-70% of the entirety of the language taken up by characters only found in compounds as well as the totally bound characters that can only appear in one specific word (like 'grape' and 'butterfly') or compound (like both characters of 踢踏 chou2chu2 'diddle-daddle').

I only discuss the monosyllabic debate in such depth because it is important to keep in mind while discussing Chinese homophones: If you look in a Chinese dictionary that includes romanized pronunciation, you will come upon a startling (to most Western folks, anyway) discovery, which is that a stunning number of Chinese characters are homophones.
In the common dictionary I use for Mandarin class, there are 38 words pronounced y14. The word processor that I'm using right now has the following characters listed under shì4: 是, 事, 世, 士, 尋, 室, 示, 试, 視, 式, 氏, 週, 釋, 裝, 侍, 譽, 逝, 嘱, 仕, 柿, 壩, 殭, 悟, 拆, 拆, 拆, 拆, 搞, 搞, 搞, 瞇, 瞇, 搞, 搞, 搞, 搞, 瞇

All of these have distinct meanings and usages, but the exact same pronunciation. This is pretty stunning, but in this case, two plus two does not equal four, so to speak. Most of these characters are not free to be used independently, and still more are only used in writing, and are never, never found in everyday speech.

Chinese people have no problem understanding each other, and nothing about any of this can be considered vague or difficult; it's just the way the language is. Our Western minds have trained us that a word having a homophone is a marked condition, but, in a language with only 1,277 possible phonetic words (DeFrancis 1984), homophony is a way of life.

That being said, however, homophony is still an interesting phenomenon in the language that has many fascinating effects, and that's what I'm going to be looking at in this thesis.

Conditions Of Homophony in Chinese
Over the past several years, I have discussed with numerous Chinese speakers the topic of homophony, and sometimes an interesting thing happens: American children of Chinese parents, who learned the spoken language for use at home but don't know how to write and don't have an enormous vocabulary, don't know that the homophonous pair of words are actually distinct.

I was talking to a 17-year-old informant who told me about writing a poem. The poem was in English, but she made a reference to how the word for 'clock' and the word for 'death' in Mandarin are the same. She was somewhat startled when I told her that they aren't actually the same word, even though they are completely homophonous. (鍾 zhong1 'clock' and 終 zhong1 'end, death') She had a blank expression on her face, like, "How was I supposed to know that?"

Well, when I thought about it, there really wouldn't be much of a way for her to know that they aren't the same word, except that they have different usage rules. They both can be used as nouns, though. (Word class is the thing that most often will disambiguate homophones in other languages, like the English adjective 'bare,' noun 'bear,' and verb 'bear.' [Actually, in some varieties of English, as in Chinese, these words aren't totally homophonous.])

One of the first things that entered my mind when
attempting to explain the rampant homophony in Chinese was polysemy. That is, perhaps the 38 characters pronounced yi4 actually all came from several words pronounced yi4, and over time different characters have been invented to represent the polysemous meanings of these few words. Perhaps my friend was right and 钟 zhong1 'clock' and 終 'end, death' are the same word, with different characters for its varied meanings.

It turns out that this is not the case, and that all these words are quite distinct. The clearest way to show that 終 zhong1 'end, death' and 钟 zhong1 'clock' are different words, other than through the characters (which can certainly be misleading), is from the historical record.

Chinese has an unbroken written record dating back more than 3,000 years, and reconstructionists have worked on Old Chinese rhymes and dictionaries to get an accurate picture of what Chinese must have sounded like many years ago. Lo and behold, the Chinese syllable used to be a whole lot more complex, and various processes that have acted upon it have altered the vowel cluster, reduced the onsets, limited the phonemic inventory of the whole language, and reduced the number of available syllable-final consonants. Thanks to the complex record-keeping of the ancient Chinese, lots of the sounds of the ancient language have been reconstructed in a fascinating way.
So, in Old and even in Middle Chinese, 鐘 zhong1 'clock' and 終 zhong1 'end, death' wouldn't have been pronounced the same way. In Baxter's *Handbook of Old Chinese Phonology*, the most recent, authoritative look at Old Chinese, 終 zhong1 'death' is rendered as tsyuuwng and 鐘 zhong1 'clock' is tsyowng.

Examples like this are comforting to our sense of how language is supposed to work. Saying that the Chinese language developed 38 words with precisely the same pronunciation seems entirely capricious. From an historical perspective, however, we can see the phonological change that has nudged the language in this direction.

So, Old and Middle Chinese data will be very important when looking at homophony conditions, because it shows the phonological processes that have created so many similar-sounding words in the language. The reconstruction of Old and Middle Chinese is a tedious and wholly complex affair, and I have neither the interest nor the know-how to present a complete overview of exactly what phonological changes have happened over the past 3,000 years. I can, however, mention the areas of Old and Middle Chinese reconstruction that say something about homophony. But first, an little bit of background on the way Chinese has been reconstructed is needed.

Since Chinese does not use an alphabetic writing system, reconstructing the phonology of Old Chinese is much more
difficult than, say, its Western counterparts, Latin and Greek.

The primary source for Old Chinese reconstruction is poetry, where of course rhymes give clues to pronunciation, as well as other ancient texts that have some kind of rhyme structure. Of course, the current pronunciation and the phonological information in characters also give many clues.

Middle Chinese is well-documented and there is actually very little disagreement on its phonetic structure. Of course, even more poetry from the Middle Chinese period exists, but the Middle Chinese period is also when the first dictionaries appeared, and they have been invaluable in this regard.

There is an enormous problem when it comes to making dictionaries of Chinese characters, and that is how to organize them in an ordered way. There are conventions that have developed, but until less than 3 centuries ago there wasn't an accepted way of doing it. But the books that attempted some sort of system are what interest reconstructionists.

The first known dictionaries (of Middle Chinese) worked on the 反切 (fan3qie4 'turn and cut') system, which is somewhat complex and confusing. It took characters and 'defined' them by parts of the sounds of other characters. Each character was represented by two other characters, the first character having the same onset (what the dictionaries call 'initial') as the 'target' character, and the second having the same nucleus
and coda (what the dictionaries call 'final'). So, characters
with one nucleus and coda were organized with other characters
having the same nucleus and coda. Finals of different tones
(tones did exist in Middle Chinese) were treated differently.
For example, in the dongl 東 'east' group, there would be words
such as 酥 songl 'pine' and 通 tongl 'through', whose finals are
'spelled' as 東.

There are many, many problems with this system, including
the fact that there's no way to organize on a domain smaller
than the initial, meaning that once they got to the list of 38
characters pronounced yi4, they were free to put them in any
order they chose. Differing methods of organization at the
homophone level leave a state of confusion even today. But
what these dictionaries do accomplish is to give a very
accurate portrait of what Middle Chinese must have sounded
like.

I have several examples that show some of the typical
processes that have resulted in the loss of many consonantal
distinctions, and in so doing have given rise to homophones in
the language:

Here, we see how three syllables pronounced rong2 have
been changed from yong2 (they are still pronounced yong in
Cantonese), but originated as three distinct phonetic
realizations in Old Chinese. Note that they became homophonous
in Middle Chinese:

modern < middle < old

榮 rong2 < yong < hjwaeng 'glory, honor'
融 rong2 < yong < yuwng 'melt'
容 rong2 < yong < yowng 'contain; contents'

Rong2 demonstrates how some vowels merged in the Old Chinese forms, but here is an example of how vowel distinctions, initial consonant distinctions, and final consonants have been lost.

Here, we have not only the loss of the final voiceless stop, but also an initial Ɂ consonant that got changed to a d:

毒 du2 < dowk < duk 'poison, drug'
獨 du2 < duwk < dok 'only, alone'
讀 du2 < duwk < lok 'read, study'

One final example will illustrate how several distinctive finals have been lost (the 'j' is an on- and off- glide):

樺 li4 < lek < rewk 'oak'
力 li4 < lik < rjik 'power, strength'
栗 li4 < lit < rjit 'chestnut; to shudder'
泉 li4 < lej < rets 'crime, sin'

Importantly, certain dialects preserve many of the phonological features of Middle Chinese. Even though these other dialects have changed in many of their own ways, to be
sure, the theory that they contain many of the phonetic
distinctions lost in Mandarin seems to be supported by the
available evidence, and it is a widely-held belief among
Chinese linguists. Just from my own conversations with
speakers of some Chinese dialects, I know some of the
pronunciations found in Baxter's Old Chinese phonology book
still persist.

The important effect of this is that there are fewer
homophonous words in these dialects, and perhaps fewer effects
of homophony. This is an area I would love to investigate in
detail, but this, unfortunately, is very difficult to
accomplish while in America.

**Effects of Homophony**

Now that we understand how Chinese ended up with 80 words
all pronounced *xii*, we can look at what effect (if any) this
has on the language, and on speakers of Chinese in general.
It's important to note that my point is not that homophony
makes Chinese hard to understand, or poor in its construction,
or devoid of invention. To me, homophony in Chinese is one of
those special and fascinating phenomena that makes linguistics
such a special field of study.

The first and largest tangible effect of homophones on the
Chinese language is the richness of, for lack of a better term,
punning. It's easy to imagine how a very homophonous language would be filled with puns; in Chinese, many traditions and taboos are inextricably linked to the sounds of words.

The second interesting effect of homophony is how it relates to language change. It appears that two homophonous words with similar meanings will merge into one. Puns have caused language change as well.

The last, and perhaps least clear-cut, aspect of homophony in Chinese is how it relates to the issue of monosyllabism.

a. **Puns and Fortune**

As in any language, the number of puns in Chinese is dependent on the imagination of the speaker! There are probably thousands upon thousands of possible puns on the various meanings of various words. What I'm looking at here, however, are the ways that certain puns have become ingrained in Chinese culture and thought, simply because the words are homophones. As you will see, many of them have to do with holiday traditions or other auspicious occasions, like the birth of a child.

This list has been compiled from a number of sources, and it doesn't pretend to be complete. After the list, I will discuss all of the issues raised by partial homophones.

Here goes:
1) The word for 'fish' is 魚, which is pronounced yu2. The word for 'surplus' is 餘, which is also pronounced yu2. So, there's a tradition among Mandarin speakers that, at New Year's celebration, no one is allowed to finish eating the fish, because, if there's no 魚 left over for the new year, there's not going to be any 餘, either.

2) The word for 'bat (animal)' 翼 (fu2) and the word for 'fortune' 福 (fu2) happen to be the same, and so bats are a sign of good fortune. 'Bat' in Mandarin can be either 翼蝠 bian1fu2 'bat' or sometimes just 翼 fu2 'bat,' but the homophone reading, I believe, comes from Middle and Old Chinese, when bats were exclusively 翼 fu2 'bat.' Bats appear extensively on coins and other charms in Chinese history.

3) This one is another New Year's tradition. On New Year's a traditional thing to eat is called a 年糕 nian2gao1 'year cake.' There's another word 黏 nian2, which means 'sticky' or 'glutinous' (黏米 nian2mi3 is glutinous rice), and so these 年糕 cakes are all made out of sticky rice. The folk tale surrounding this is that all families need to eat these cakes around New Year's Day in order to 'stick together' in the coming year.

4) Tangerines are often considered lucky. The word for tangerine is 橘子 ju2zi, which is sometimes written as 桔子. Now, one word for 'luck' in Mandarin is 吉 ji2. In Cantonese
（as well as other dialects), however, 吉 ji2 and 橘 ju2 are homophones, both pronounced kē. Apparently, this tradition is more recent and was 'borrowed' from the Canton region.

5) Similarly, the word for orange is 柑 gan1 (kām), homophonous in Cantonese with 金 jin1 (kām) 'gold.' Oranges are another lucky food.

6) A person's zodiac sign is important in Chinese culture, and I've heard that while having a discussion on one's Zodiac sign in Taiwan, it's unlucky to call yourself a snake. The word for snake is 蛇 she2. It could be punned here with an homophonous verb 揍 she2 which has two principal meanings, one is 'to snap or break,' and the other is 'to lose money in business.'

The most likely scenario, in my opinion, is a phonological change which happens in Taiwanese varieties of Mandarin where the initial sh, zh, and ch of standard Mandarin lose their sibilance and become s, z, and c. In that case, 'snake' would be pronounced se2, which is extremely close to the aforementioned word for death, si3.

7) 髮菜 fa3cai4 is an edible fungus that looks like hair. (hence 髮 fa3 'hair') 發財 falcai2 is a verb compound that means 'to get rich.' As you can probably imagine by now, Chinese people eat lots of 髮菜, especially around the New Year.

8) 分 fen1 as a verb has a ton of different meanings, all of them having to do with separation or allocation. In one sense,
it can mean 'to share,' and indeed, there's a verb compound 分攤 fen1tan1 'to share.' In another sense it can be 'to divide,' and so the expression 分裂 fen1li2 where 裂 is 'pear' means to 'share a pear.' This is homophonous with the expression 分離 fen1li2, which is a verb compound meaning 'to severed,' or 'to separate.' The taboo that follows from this is that it's bad to share pears because it means that you want to sever that relationship with that person.

9) Another verb of interest is 送 song4, which can be translated as 'to give as a gift,' or (rather obtusely in English), 'to see someone off,' or 'to attend to.' There is the expression 送鐘 song4zhong1, where 鐘 is 'clock,' which means to give someone a clock as a present. There is also the homophonous expression 送終, where 終 means 'death.' This expression means 'to attend to a dying relative,' or 'to bury a relative.' Naturally, it's a very bad thing to give someone a clock as a present.

10) The verb 散 san3 means 'to scatter,' or 'to fall apart.' It's homophonous with the word 傘 san3, which means 'umbrella.' Umbrellas aren't good gift ideas in China, either.

In Cantonese, because of the unlucky meaning of 散, another word is preferred for umbrella. The new word, 遮 zhe1 'cover,' is free from homophonous double readings. As far as I have been able to determine, however, 傘 san3 is still the preferred Mandarin word for 'umbrella.'
11) The word 书 shū 'book' is homophonous with the verb 败 shuǐ 'to lose (a game).'</p>
<p>Although there aren't any taboos that I'm aware of surrounding the giving of books in normal situations, it is considered quite unlucky to use this word in any kind of game environment, especially during the playing of Majong.</p>

12) Lotus seeds are considered very lucky in Chinese culture, and are commonly eaten at the New Year. The lotus flower and plant are also very important, but the seeds have a homophonous meaning that explains their particular fortune. 蓮子 liánzi3 is the word for lotus seeds. 子 zi3 has a wide variety of meanings, encompassing both 'seed' and 'son,' as well as the plurals of those words, of course. So, a homophonous phrase, 蓮子 liánzi3, means 'successive sons,' which is an enormous blessing for a Chinese family.

13) The verb 到 dao4 meaning 'to arrive' and another verb 倒 dao4 meaning 'to fall over, to turn upside down' are homophonous. Before the coming of Spring, speakers of Chinese will often have some kind of banner with the character 春 chūn 'Spring' written on it. As Spring approaches, they will turn the banner upside down to herald the arrival of Spring, and good fortune with it.

14) In Cantonese, it is unlucky to use the word 空 kong1 (hok) 'empty' because it is homophonous with both 兇 xiong1 (hok)
'fierce' and 虛 xiong1 (hoŋ) 'inauspicious, unlucky.'

Apparently, when talking about an empty room or other area, a Cantonese speaker (or real estate agent) will say that the room is 吉 ji2 (k 북한) 'lucky.' This is a fascinating choice of words, because 吉 ji2 (k 북한) carries with it no meaning of 'empty,' and so the speaker is admitting the homophony of 空 and 虛, and then saying the opposite of the homophone in order to get at the meaning desired.

15) 生菜 sheng1cai4 means 'romaine lettuce,' and it also a lucky food, as it is partially homophonous with 生財 sheng1cai 'give birth to wealth.' This compound is related to 發財 falcai2 'get rich' (above) in that both 發 fal and 生 sheng1 carry with them the idea of something coming from nothing.

16) Lucky foods abound in Chinese, and another one (from Cantonese) is 蟹 hao2chi3 'dried oyster.' In Cantonese, this is pronounced how si, and is homophonous with 好事 hao3shi4 (how sǐ) 'good affairs.' I don't believe that dried oysters are lucky to Mandarin speakers, as the pronunciation in Mandarin of the second character has a different initial and tone. No one I talked to was able to give any more information on this custom.

17) 食 shi2 (sí:t in Cantonese) 'to lose' is identical in Cantonese with 舌 she2 (sí:t) 'tongue.' Because of this, Cantonese speakers use the word 食 [no Mandarin] (lej) where
Mandarin speakers would use 舌. Interestingly enough, this word 舌 (lej) is homophonous with a very auspicious word, 利 li4 (lej) 'profits.' There's a folk tale that goes along with this which says that people in historical times, noticing that the word for 'to lose' and 'tongue' were homophonous, made sure not to say either word. Therefore, there was something of a gap in the language, because it lacked an appropriate word for 'tongue,' and scholars searched desperately through the language for a suitable replacement. They arrived at 利 li4 (lej) 'profits,' added a 'flesh' radical on the side to produce 利舌, and then thankfully had a favorable homophone instead of an inauspicious one. (An interesting side note: Because the alternate Cantonese word for 'tongue' is homophonous with 'profit,' it is also a New Year tradition for Cantonese speakers to eat duck tongue.)

Of course, this story is a tad silly, but it illuminates my point perfectly. Although some may regard it as superstitious not to say a word because it is homophonous with something that the speaker would like to avoid, I would say that it's simply more perceptive of the nature of the word. A language with a large number of homophones will naturally activate certain homophonous words in the brain when a word is said. If the homophonous word is inauspicious, and it is activated along with the appropriate word, then it simply
follows that the speaker may not want to use the inauspicious homophone. The reverse is also true: speakers should then like to activate words with auspicious homophones. It may be superstitious, but I think it makes perfect sense.

Numbers are extremely important in any language simply because of what they denote, but in Chinese numbers are found with astonishing frequency in sayings, syntactic structures, and other places where they're not totally expected. As I found, the superstitions surrounding numbers very often come from homophones:

1) The number four is written as 四 si4 and it is partially homophonous with 死 si3, which means 'death.' Hence, the number four is extremely unlucky, and there are many, many stories, myths, and the like involving the number 4.

2) Eight is written as 八 and pronounced bā1. There is a character 发 fa1 that has an enormous amount of meanings, most associated with discovery, transmission, and opening. 发 has been described to me to mean 'get rich,' and the word definitely has very positive connotations. However, the pronunciation of the two words is not the same, even though most speakers relate these two words very closely, in several dialects. Eight is a very lucky number.

3) Six is 六 liu4 and it sounds like 留 liu2, which means 'to accept.'
4) 九 jiu3 'nine' has a lucky meaning attached to it as well. Longevity is a very important concept in Chinese culture, and nine happens to be homophonous with 九 jiu3, a word meaning 'for a long time.'

5) The number ten is 十, pronounced shi2. This is homophonous with the word 死 shi2, which has the concrete meaning 'solid,' as well as the more common abstract meanings 'actually,' 'really,' or 'surely.' 十 shi2 'ten' isn't punned by itself, as far as I'm aware, but instead takes on its pun meaning when put in other numbers. 四十 si4shi2 'forty' is punned with 死 si3shi2 'die surely.' 十四 shi2si4 is punned with 死死 shi2si3 'surely die.' Many Chinese people avoid living or working on the 4th, 14th, or 40th floor of a building.

6) Cantonese only: 三 san3 (sa:m in Cantonese) 'three' is lucky because of how close it is to 生 sheng1 (sa: in Cantonese) 'be born'. Apparently, in Hong Kong, license plates with multiple 3's or 8's on them fetch a very high price.

7) It happens that the pronunciation of the string of numbers 五一八 wu3yilba1 (jé t pa:t in Cantonese) '5-1-8' is extremely close to the Classical Chinese sentence 吾一(定)發 wu2yil(ding4)fal (jé t fa:t in Cantonese). It means 'I definitely get rich,' or 'I will definitely get rich,' and so, based solely on the pronunciation of these words, May 18th is the most popular day for marriages and other ceremonies. This
custom is very pervasive in China and Hong Kong with many
dialect groups.

From these examples we see that, out of the numbers from
one to ten, 7 of them have pun meanings, which gives a portrait
of how linguistically remarkable this phenomenon is. In
Western languages, we have similar ideas about lucky and
unlucky numbers, but in no instance is the tradition based on
the sound of the number.

There's one more big area of punning--sayings and
aphorisms. Sayings in Chinese are all terribly interesting.
Most of them come from a literary tradition, and are in
Classical-style grammar and meaning. A special kind of saying
is the 成語 cheng2yu3, or 'established language,' which is a
4-character phrase, usually coming from a classical story, that
has a certain set meaning.

A good example of a 成語 phrase would be 畫蛇添足
huai4she2tian1zu2, literally 'draw snake add feet.' This saying
describes a story about a contest to see who can draw a snake
in the shortest amount of time. One man draws the snake much
faster than the others, and he gets so excited about winning
that he decides to add feet to his snake, after which, of
course, the other men point out that snakes don't have feet,
and he loses the contest. The closest English correlary to
this would probably be the story of the tortoise and the hare.
In Chinese, one only has to say this four character phrase, and the message is clear.

Punning in this area is most like punning in the western sense; it's productive and free, and most speakers make them up as they like. Still, even this form of punning is fascinating because of Chinese's isolating typology and relatively free word class distinctions. There are too many examples of this to present, but I have one that typifies the creativity and availability of puns:

A big slogan for the Communists earlier this century was wang3qian2kan4, which literally is 'turn forward look,' and means 'look to the future.' But modern Chinese have punned the central character, and so now it's wang3qian2kan4, but with the new character it means 'turn money look,' or 'look towards money,' so it's a neat commentary on Chinese society.

The Phonology of Chinese Puns

Chinese call homophones 同音詞 tong2yin1ci2 'same sound word' or 同音字 tong2yin1zi4 'same sound character.' The more technical term for homophony is 諧音 xie2yin1 'harmonious sound'. In my research I've seen both genuine homophones as well as words that are partial homophones, i.e. the tone isn't identical, or the initial differs slightly (perhaps by only one feature). In Chinese these words are said to 音近 yin1jin4
The punning data I presented above raise many questions. The 八 bal 'eight' and 發 fal 'develop' pair is particularly troubling if we just look at Mandarin; even though bal and fal are close in pronunciation, why are other similar pairs not punned? Also, the words that are homophonous except for tone also present something of a problem; if 死 si3 'death' and 四 si4 'four' can be punned, why aren't there any similar taboos about 壞 huai4 'bad, broken' and 懷 huai2 'be pregnant', and other pairs like it?

The answer, I believe, is that the partial homophones can tell us something about how these punned words are perceived in Chinese. This is where an extensive look into Old Chinese phonology and modern dialectology is extremely helpful and illuminating.

The pun words with different consonants can mostly be explained by looking at Old Chinese. The fal/bal pair, certainly one of the most pervasive lucky homophones, becomes transparent when we see that the change to an f- initial from a b- initial is a rather recent phonological alteration (Baxter 1992), and these words were pronounced identically not too long ago.

Only two other pun examples show actual consonantal differences between the pun words: 吉/柑 ji2/ju2
luck' / 'tangerine' and 蛇/死 (she2/si3)
'snake' / 'death' ('break'?). In the case of ji2/ju2, it's a clear case of a cross-dialect loan; because they are similar in some dialects, the other dialects pick up the tradition, and relate the two words, even though they aren't pronounced as homophones in that particular dialect.

I already described my view on why 蛇 she2 'snake' is related to 死 si3 'death' in Taiwan: the sh- initial in Taiwan is often said just as s-. (This is because the indigenous southern Chinese dialect spoken there lacks the j-, ʃ-, and ʂ- initials.) Then, she2 'snake' would be said as se2 'snake.' But se2 doesn't exist as a syllable in Chinese, and I think in this case it comes out sounding almost exactly like si3 'death.'

A few of the examples have divergent tones, and I think it's best in these cases to look at the Chinese tone as a feature like English stress; it certainly is distinctive, but speakers play with it in music and poetry, as well as standard speech. Tones can change over time, and a word with one tone in one dialect can have a completely different tone in another. It's not that tones are random or capricious, but there is a certain flexibility there, and I think that's what we're seeing with the punned words of different tones. Also, some of the taboo words may have been homophonous before the language even had tones. 死 si3 'death' and 四 si4 'four' is such a pervasive
taboo that it could be quite ancient. I do believe, however, that partial homophones differentiated by tone are less likely to be punned.

b. Homophones in Language Change

Language change is a fascinating, complex topic. Exactly what motivates speakers of a natural language to alter the way they speak and the processes behind the language change will probably never be fully understood. One thing I can say with certainly, however, is that homophones play a part in some aspects of language change.

Consider the English taboo word for 'penis,' 'cock,' and its non-taboo meaning 'rooster.' Using the word 'cock' to actually mean 'rooster' is fairly rare in English these days, and the taboo meaning seems to have taken over almost entirely. Taboo words, it seem, have this tendency over time in many languages. Once a taboo is associated with a word, using it in any other context (including, perhaps, its original meaning) becomes less and less favored. We see this with other words in English such as 'bitch' for 'female dog,' 'ass' for 'donkey,' and 'dick' as a nickname for 'Richard.' (But not, interestingly, with 'damn' and 'dam,' probably because, as their spelling indicates, their homophony is a more recent phenomenon, and perhaps English hasn't had enough time for an
appropriate homonym to take over.)

Chinese has an ample supply of vulgarities like 'dick' that act as taboo words, and speakers are extremely sensitive to their pronunciation. I have been unable to find a good discussion of taboo words in Mandarin, but Bauer and Benedict (1997) report that, in Cantonese, of all the most common vulgarities, only one has an exact homophone. Chinese speakers also show very much sensitivity to words which are partial homophones with vulgarities, as we do in English. (The best English example I can think of here is the song "I Wanna Funk With You Tonight," a wry pun on "I Wanna 'fuck' With You Tonight.")

Keeping all of this in mind, these unlucky puns in Chinese function as taboo words, and effect some kinds of language change. We have already seen the example of 'tongue' in Cantonese: the word found in other Chinese dialects for 'tongue' is homophonous in Cantonese with 'to lose,' and so another word is favored. There are many other examples of this in Chinese: the words for 'liver' and 'dry' are homophonous in Mandarin and Cantonese, and so, when talking about eating liver, in order to avoid the idea of it being dry and tough, speakers will use another word.

Also along these lines is the Cantonese avoidance of the word 空 kong1 'empty,' as mentioned above, because of its
homophony with 因 xiong1 'unlucky' and 六 xiong1 'fierce.' 空 kong1 'empty' is a very common word in Mandarin, and occurs most commonly in the compound 有空 you2kong4 'to have free time.' Because of the inauspicious homophone, however, Cantonese speakers don't use this compound.

It should be noted that the Chinese unlucky taboo language change differs from its Western correlary in that the Western taboo words most often come from the original meaning of the taboo--we call a person an 'ass' because of the animal--but the Chinese taboo homophones are actually different words, related only by sound. That's the key issue of interest.

I have come upon two examples of another phenomenon in Chinese language change that unfortunately is rare and hard to pin down. It seems as though, if there are two words that have a similar meaning, and the same phonetic realization, but of course different characters, one word will 'win out,' so to speak. This word will subsume the meaning of the other word, and the other word's character goes out of style. Two examples follow:

The verb 做 zuo4 means 'to do, make, or manufacture (as well as some other things),' and it appears in a good number of compounds. The verb 作 zuo4 means 'to do' more in the sense of the English 'do.' Since the meanings are somewhat similar, and they sound the same, the words are constantly interchanged,
and, in writing, it seems as though 做 is going out of style in favor of 作. I assure you that they are distinct words, and were in the past phonetically distinct, but their merger into the same phonetic realization and their closeness in meaning have caused this shift.

The next one is a compound. 厉害 li4hai4 is a common stative verb that means 'terrible, extreme.' 利害 li4hai4 is something of a noun compound that means 'advantages and disadvantages.' But the use of 厉 is diminishing, even though, in Middle Chinese, 厉 was pronounced as 'ljej' and 利 was 'lij.' In Chinese dictionaries, 利害 li4hai will have two meaning listed: 'advantages and disadvantages' and 'terrible,' as well as a listing with 厉害 li4hai4 meaning 'terrible.' The words are definitely different, but because 利 is the more common character and word, it seems to be winning out over 厉.

In order for a word to 'die,' so to speak, I believe that it has to be a perfect homophone and have a large semantic overlap, and this, of course, is a very rare phenomenon. It would be odd for a language to have two words that are said the same, mean almost the same thing, but are written distinctly and actually are two distinct words. It seems that by having one word or phrase take on both meanings, Chinese is simplifying itself in a very natural way. (It should be noted that there is a verb 受 shou4 'to receive' and another verb 收
shoul 'to receive,' and this effect is not observed. I think that's very strong evidence that this effect is only seen with perfect homophones.)

These are the two most important ways I've seen that homophones actually effect language change. Neither of them really make the language less homophonous by increasing the phonemic inventory. Indeed, a word substituted for a taboo word could be homophonous with 20 other words itself! So it isn't fair to describe this language change as a broken language trying to fix itself and become less ambiguous; instead, it's just a natural language undergoing interesting change because of one of its features.

c. **Disyllabification**

A delicate nuance of Chinese compounding is the way that some words which can be handled by just one character can also handled by two. For instance, a verb X might mean 'transmit,' but instead of just saying X, speakers use the compound XY, which also means 'transmit.'

Consider the following examples:

(a) 我知道。

wo3 zh1 dao4

"I know (something)."

(b) 我知。
"I know (something)."

Example (a) is grammatical in both Mandarin and Cantonese, but example (b), common in Cantonese, is not favored in Mandarin, where the compound is preferred.

Linguists have posited that this may be because of homophony. It may be a mechanism to both disambiguate words with lots of meanings, and to distinguish certain words from others.

This all goes back to the monosyllabic debate, and dialectology. If one of the effects of homophony over time has been the expansion of the number of compounds in the language, then, perhaps, dialects that are not as peppered with homophones will prefer to use fewer compounds than Mandarin does.

DeFrancis doesn't feel that there is a one-to-one relationship between polysyllabic words and homophony in Mandarin, and I'm tempted to agree with him. He argues that there were many polysyllabic compounds in Classical Chinese, and that multi-syllabic words and compounds are nothing new in Chinese. So, thinking that Mandarin speakers somehow had to 'fix' their language because of homophony is probably not the right idea. He doesn't deny, however, that there are more bisyllabic roots and compounds than before. I feel that as
syllables got shorter and simpler (and more homophonous with other words), perhaps Chinese speakers preferred using more compounds and adding syllables onto roots.

To this end, I think I should run a trial. I'm considering taking some speech of Mandarin speakers and of Cantonese and other dialect speakers, and looking at their use of bisyllabic compounds as opposed to monosyllabic words which mean the same thing. It would be absolutely fascinating to discover that Cantonese, Shanghaiese, and Fukanese speakers use less bisyllabic compounds because more of the phonetic distinctions of the language are preserved in their dialect. This was suggested by Thomas Lee, and it's a great idea, so I think I should run with it.

My trial will not just seek to get clues on disyllabification. Indeed, I think that I may not even find much support for that hypothesis, because of the problems of translation and oral vs. written language. My trial will also seek to take a look at the real degree of homophony in the spoken forms of these dialects. It's easy to look through a Mandarin dictionary and count words. So the actual homophony conditions in the language can't really be gauged by the sheer numbers of homophonous words, and that's what I'll be looking at with this trial.

The trial starts with the following story, which was
written by me in English. I'm going to be refining the story
to control for possible translation problems into Chinese:
"Yesterday my younger sister made me take her shopping at the
mall. She wanted to buy clothes. I didn't have any money, so
I followed her around. She bought a pair of blue pants. I
thought they didn't fit her, but she likes her clothes to be
tight. Then she bought a new pair of $100 brown shoes. I
thought they were too expensive, but she didn't care. I was
very tired and hungry, and then it was already 5:30. We went
to a restaurant and ate dinner. My sister is a vegetarian, so
she ate tofu. I had chicken. We both thought the restaurant
was a little too expensive, but the food was pretty good. It
was very crowded there. On the way home the police gave me a
ticket. When we got home, mother saw the clothes that my
sister bought, and she scolded us for spending so much money.
Then she discovered that we had already eaten, and she got very
mad. She made my favorite dish, and she went to the market to
buy ingredients for it. But by then I wasn't hungry anymore.
It was a bad day."

I tried to write the story in unidiomatic English, so
there shouldn't be enormous translation problems. The
participants will write the story the way they would say it in
their own dialects, using as much 口語 kouyu3 'oral language'
as possible, and then I'll be able to interpret the results.
accordingly, hopefully by having a session with one speaker of each dialect to go through what words exactly are homophonous.

As for the degree of homophony in the spoken language, I'll use these samples to get a numerical idea of exactly how many words sound the same now, as opposed to their phonetic realizations in the Old Chinese data.

Another possible alternative to this is to examine transcribed samples of various Chinese dialects and count compounds, then talk to speakers of each dialect and ask them to read the text and identify homophones.

I also hope to find correlaries in other languages to some of the effects I've described. This research has also turned me on to investigating homophony in language on a much broader scale: Why and how do languages in general develop homophones, and how are homophones related to language change?
REFERENCES

語音學. New York: Morton de Gruyter.

Gruyter.

曾子凡 (Ceng Zi-Fan). 1995. 廣州話, 普通話: 語詞對比研究 (Cantonese, Mandarin:
Comparitive Language Studies.). Hong Kong: 香港普通話研究社 (Hong Kong
Mandarin Press.)

Press.

Press.

Hawaii Press.

Ronald Press Company.


Press).


Commercial Press.

THANKS

I'd like to thank my Chinese professors who listen to my butchered attempts at the Chinese
language, Professors Alan Berkowitx, Karin Myhre, and Sujane Wu. They gave me lots of
ideas and feedback. Special thanks also to all the students who speak Chinese who had to
listen to my ramblings and random, odd questions, especially Sandy Lin, Min Lee, and
Elizabeth Cho. Thomas H.T. Lee, of the University of Hong Kong, gave me some great
materials and ideas too. Finally, a big hug to Donna Jo Napoli, who gives us all
encouragement and hope.