A Chinese phonological enigma

Geoffrey Sampson

University of South Africa

This article is to serve as the basis for a discussion section in a 2015 issue of the Journal of Chinese Linguistics*, where various scholars will be invited to comment and I shall provide a summary response to the comments.*

Abstract

Recent cross-language research has yielded strong statistical evidence in support of the idea, advocated by André Martinet and widely accepted by linguists, that languages avoid adopting sound-changes which would create many homophones. Yet we know that the history of Chinese phonology has been marked by repeated phoneme mergers and losses which led to a very high incidence of homophony, forcing the monomorphemic vocabulary of the classical language to be replaced by a largely bimorphemic modern vocabulary. This paper examines various ways in which this apparent contradiction might be resolved. None seems fully satisfactory, yet some resolution must exist.

安德烈·马蒂内（André Martinet）提出对于会产生很多同音异义词的变音形式，语言会避免引入；该观点已广为语言学家们所接受。近期的跨语言研究已提供有力的统计数据来支持这一观点。然而，我们知道汉语音韵学的发展历史特点是音素的不断合并和丢失，导致同音现象的几率很高， 使得古代语言的单音素词汇被以双音素为主的现代词汇所代替。本文考察了各种不同的方法来尝试解决这个突出的矛盾。 似乎没有一种方法完全令人满意， 然而某种解决方法肯定存在。

1 A paradox stated

This paper is about a contradiction which has emerged between, on one hand, well-established facts about the history of Chinese phonology, and on the other hand a claimed linguistic universal for which robust evidence has recently been produced.[1]

One of the most striking properties of Chinese, to people more familiar with European languages, is its very high incidence of homophony. All languages contain some homophones, for instance English /rait/ can represent any of the unrelated words *right*, *write*, or *rite*. But in English and other European languages, words which coincide in pronunciation with other words are a minority, and even in such cases it is rare for more than two or three etymologically-distinct words to share a spoken form. In Chinese, and particularly in the standard, Mandarin dialect, if for the moment we use the term "word" to refer to 字 rather than to 詞, there are very few words which

1 I thank Abby Kaplan of the University of Utah and Daniel Silverman of San José State University for discussion of the topic of this paper, and 李漢宏 (John Li) of the Open University of Hong Kong for the Chinese version of my abstract. Responsibility for the contents of the paper is mine alone.

are not homophonous with other words, and a set of homophones may contain ten or twenty members. It is difficult to be precise about which of all the words that have been used in the long recorded history of Chinese should be counted as elements of present-day spoken Mandarin, but one linguistically-sophisticated attempt to do so is Chao and Yang (1962); on average a Mandarin syllable is ambiguous between about four of the 字 listed there, with a maximum of 25-way homophony for the syllable *yù*.

Studies initiated in Qing-dynasty China have shown that this situation results from a series of sound-changes over a long period which merged phonemes that previously contrasted, or eliminated phonemes altogether (merged them with zero). Many categories of evidence, including the Tang dynasty rhyme tables, comparison among dialects, vocabulary borrowings between Chinese and other languages, and the structure of the script, make it certain that at earlier periods the phonology of Chinese was much richer than it is today. (The subject is surveyed concisely e.g. by J. Norman 1988: ch. 2; for a more detailed account of developments from Old Chinese through Middle Chinese to Mandarin, see e.g. Baxter 1992.) Scholars differ about details of reconstruction, but the large areas of consensus suggest that homophony in the Old Chinese of three thousand years ago may not have been strikingly greater than in modern European languages. Homophony has increased so much that, if the language had retained the largely monosyllabic vocabulary of the classical period, it would now be too ambiguous to be usable. (No-one can understand a passage of 文言 read aloud without sight of the script.) Consequently a vocabulary of simple roots has been replaced by one that now consists largely of compounds and derived forms of various types (cf. Li and Thompson 1987: 816ff.) – a two-syllable compound is typically unambiguous even if each of its component roots is many-ways ambiguous. The European concept "word" equated in the classical period to 字 but for modern Chinese corresponds more closely to 詞.[2] To Sinologists all these things are well known.

In general linguistics, though, there is a longstanding belief that the historical sound-changes which occur in all languages are subject to a constraint by which they avoid creating a high degree of homophony. Contrasting phonemes merge, it is claimed, only if there are not too many pairs of words distinguished by that particular contrast. This idea can be traced back to Jules Gilliéron (1918) and became central to the linguistic theories of André Martinet (especially 1955), who used the term *rendement fonctionel* (literally "functional yield", often translated "functional load" – the term was not original with Martinet) for a quantitative measure of the work done by a particular phonemic contrast in keeping words apart, and hence of the probability that future sound-changes would eliminate that contrast. Intuitively it is very plausible that changes which create more ambiguity are less likely to happen, and many linguists have taken this to be an uncontroversial truism. (See the Google count of quotations in Baerman 2011: 2 n. 4.)

Intuitively plausible or not, Martinet's idea seems to be directly refuted by the Chinese facts already cited. And even with respect to European languages the idea soon proved problematic: Robert King (1967) attempted to test Martinet's theory quantitatively, using data from Germanic languages, and concluded that functional

2 Even 文言 contained a few disyllabic loans, e.g. 駱駝 *luòtuó* "camel", 珊瑚 *shānhú* "coral", which were written with two 字 each but were undoubtedly single "words". However, that marginal phenomenon has no relevance to the substance of this paper.

yield has little or no influence on which sound-changes occur in a language. I took it that the functional yield theory, though attractive, had to be wrong, and when I became aware recently that the theory was being revived (e.g. by Campbell 1996: 77, Blevins and Wedel 2009, Baerman 2011) I pointed out that Chinese appears to refute it (Sampson 2013).

However, the literature reviving the functional yield theory has been growing, in number of papers and in strength of evidence and argumentation, so that some recent publications cannot be dismissed as lightly as the ones quoted in my 2013 paper. Martinet (1955: 58) expressed some caution about how significant a factor functional yield is in practice in determining the course of phonological evolution, but recent writers have claimed a decisive role for it. Particularly strong evidence is discussed by Wedel, Kaplan, and Jackson (2013), who examine 41 phoneme mergers in six European languages together with Korean and Cantonese, all of which occurred recently enough to allow vocabularies to be studied statistically. Comparing phoneme mergers which have actually occurred with hypothetical mergers that are equally phonetically plausible but have not occurred, Wedel et al. find support at a very high level of significance ($p < .001$) for the hypothesis that the likelihood of a merger correlates inversely with the number of homophones it creates.

(Other relevant publications which I had not seen when I wrote my 2013 paper include Silverman 2010, Kaplan 2011, Wedel, Jackson, and Kaplan 2013, Bouchard-Côté et al. 2013, Kaplan 2015.)

As things stand, then, we have arguments which seem quite cogent that language-change avoids creating excessive homophony; if I did not know about Chinese I would certainly find those arguments convincing. Yet at the same time we know that sound-changes in the history of Chinese have created a massive level of homophony. This is a real paradox. Both statements appear to be true, but they contradict each other. The aim of the rest of this paper is to explore various ways in which one might hope to resolve the paradox. None of the alternatives seems to me satisfactory. But some resolution there must be.

2 Chinese as an exceptional case

One solution might simply be to say that homophony avoidance is a valid law of sound change but Chinese is an exception. Indeed, although he does not mention Chinese by name, Matthew Baerman implies that this is reasonable when he suggests (2011: 25) that languages may differ with respect to homophony avoidance. However, I find it scientifically unacceptable. Laws by which societies control human conduct can have "special cases": there might be a law by which all ordinary residents are subject to income tax but the monarch is exempt (this was the law in Britain until the 1990s). Scientific laws are not like that: they only count as laws if they apply across the board, so an apparent exception means that the law has been inadequately formulated.

3 The relevance of script

A better approach would be to look for some special factor which might cause Chinese to behave differently from other languages with respect to sound-changes. One

distinctive property of Chinese is the logographic rather than phonographic nature of the script, which means that words which fall together in pronunciation remain distinct in writing. This was the factor which enabled 文言 to remain the standard written language until 1919, centuries after it ceased to be intelligible as a spoken language. It might be suggested that the true linguistic universal is not that languages avoid changes which make too many words *phonetically* indistinguishable, but rather that they avoid changes which make too many words indistinguishable *in all respects*. For a language with a perfectly phonemic script, the two versions of the law would have identical consequences. But for Chinese the latter version allows any amount of homophony, since written word-forms remain distinct.

This meets the requirement for laws to be universal, but it is empirically implausible. If there is indeed a constraint on words becoming indistinguishable, the obvious reason would be that users of a language need their communications to be understood. But until recent times the literacy rate in China was not high (UNESCO 2006: 192 n. 6); even in 1949 it was only about twenty per cent (*Economist* 2014). How could phoneme mergers in the speech of an entire population be affected by the presence of a minority within the population for whom the mergers did not destroy the psychological distinctiveness of words containing the phonemes? Indeed, even the literate minority did plenty of speaking and listening as well as reading and writing, so problems caused by homophony affected them too.

Linguists do not normally think of writing as capable of playing a large role in determining the historical evolution of spoken languages, and they are surely correct not to do so. (Yishaï Norman (2009) surveys various ways in which spoken language can be influenced by script, but it is fair to say that these are all marginal, relative to the issue of Chinese homophony.) It is difficult to see how our paradox could be resolved by reference to the nature of Chinese script.

4 Timing of vocabulary replacement

A further possibility would be to query the timing of the shift from monosyllabic to disyllabic vocabulary. This is normally seen as having been a response to the increasing homophony among monosyllabic roots. Thus, Li and Thompson (1987: 817–18) referred to the sound-change by which final -m merged with -n in Mandarin, so that e.g. 金 “gold” and 斤 “tael”, respectively *kim and *kin in Middle Chinese, have fallen together as Mandarin *jīn*, and they wrote:

> If [the word 金] hadn’t become the disyllabic form [金子], the ... words for “gold” and “tael” would have been homophonous. The threat of too many homophonous words has forced the language to increase dramatically the proportion of polysyllabic words ...

Likewise Jerry Morgan (1988: 112) wrote:

> Given this progressively radical reduction in the overall number of contrasting syllables, and the consequent falling together of many words once phonologically distinct, it is not surprising that the old one-word/one-syllable pattern began to weaken, and that the use of disyllabic words began to increase.

However, if it should be that the shift from monosyllabic to disyllabic words took place *before* the contrast-eliminating sound-changes, those changes would not have created much homophony between words when they occurred, so Chinese would not be an exception to the generalization about homophony avoidance.

As it happens, Li and Thompson were ill-advised in their choice of example. Instances of the particular disyllable-creating process they cited, namely suffixing 子 to a noun without any diminutive connotation, are known to have occurred early (Jerry Norman 1988: 114 quoted examples from the Tang dynasty), while on the other hand -m and -n still contrasted for the 14th-century 中原音韻. And many disyllable-creating innovations may well have occurred in speech before they showed up in the written record.

Nevertheless, it seems unlikely that in general the shift to disyllabic vocabulary could have preceded the loss of phonemic contrasts. One leading disyllable-creation process was a type of compounding which conjoins synonyms or near-synonyms. Li and Thompson (1987: 819) gave examples such as 疲乏 *pífá* "tired-tired = tired", 防守 *fángshǒu* "defend-defend = defend", 放棄 *fàngqì* "loosen-abandon = to give up". From what they wrote one might suppose that this process occurred mainly with verbal meanings, but there are also many examples with other grammatical functions, e.g. 朋友 "friend", 民族 "a race", 墳墓 "a grave", etc. etc. If synonym compounds of this type arose earlier than the phoneme mergers, that would imply that Chinese adopted a habit of saying the same thing twice even though saying it once would have been unambiguous. Is it realistic that any speech community would adopt such a pointlessly redundant habit of speech? I am not aware of any empirical evidence that the Chinese did so.[3]

Daniel Silverman suggests to me (and cf. Silverman 2006: 76–8) that the two developments may have co-evolved, so that vocabulary replacement was both triggered by phoneme mergers and enabled them to proceed further. (Relevant sound-changes are likely to have been long-drawn-out affairs in which modifications to a phoneme spread gradually from word to word across the vocabulary; on Chinese evidence for this model of sound-change as against the Neogrammarian concept of abrupt across-the-board changes, see e.g. Feng and Yip 2014.) But Silverman's suggestion seems to reduce homophony avoidance to an unfalsifiable doctrine with no predictive power. In itself the co-evolution idea appears plausible, but if *prima facie* violations of homophony avoidance can readily be explained away in that fashion, then I do not understand how findings such as Wedel, Kaplan, and Jackson's could obtain.

Li and Thompson (1987: 817) note that southern dialects of Chinese, in which fewer mergers have occurred, also retain a more monomorphemic vocabulary. That is as predicted, if it was the increase in homophony which triggered the Mandarin shift to bimorphemic words.

3 Synonym compounds are of course only one type of Chinese compound, and it may well be that they seem disproportionately salient to Western linguists because European languages contain little or nothing that is analogous. But that very fact strengthens my point. I know of no language other than Chinese which uses compounding of synonyms as a word-formation technique, so there must presumably be some special reason why Chinese uses it. I cannot think of any alternative to the pressure of homophony as an explanation.

5 Mergers versus phoneme losses

Another attempt to resolve the paradox might point to the restricted nature of the hypothesis examined in Wedel, Kaplan, and Jackson's statistical research. Their technique limited them to considering only homophones created by phoneme mergers, rather than those created when phonemes drop altogether in particular environments; they make no prediction about the latter type of sound-change. So, logically, it is possible that those Chinese sound-changes which merged phonemes did conform to their findings, and that the very high incidence of homophony in modern Mandarin was produced by other types of change. For instance, when final -m merged with -n, leaving the third final nasal -ŋ distinct, because -m was a low-frequency final consonant it is probably true, as Wedel et al. would predict, that the change created substantially fewer homophones than would have been created by a (counterfactual but equally phonetically plausible) merger of -n and -ŋ, leaving -m distinct. On the other hand the number of homophones that were or would have been created by either of these changes might well have been dwarfed by the number that were created by the loss of all final oral stops -p -t -k, but that would not falsify Wedel et al.'s claim.

But, in the first place, although the particular statistical techniques used by Wedel et al. restricted them to considering homophony resulting from phoneme mergers, they do not suggest that this is anything more than an unavoidable limitation of their research method. Other publications reviving the functional yield theory have argued that sound-changes in general, not just one category of sound-change, avoid creating homophones. And that is surely what we would expect. A universal tendency for languages to avoid becoming inefficient through excessive ambiguity is very natural and understandable, if it is indeed a reality, whereas a universal tendency to avoid generating homophones via one type of process while allowing any amount of homophony to be produced in other ways would be inexplicable and implausible.

In any case, some of the Chinese sound-changes to which Wedel et al.'s findings ought to apply do seem to refute them. Consider the Mandarin sound-change which produced the sounds spelled *j q x* in *pinyin* by merging /k k^h x/ with /ts ts^h s/ respectively before close front vowels. This affected about an eighth of the entire vocabulary, and created a huge number of homophones. When distinct wordforms which are each, say, three ways homophonous fall together, the result is nine new homophone-pairs: each word of one set is newly confusable with each word of the other. Many words affected by the *j q x* merger will have been more than three ways homophonous before it applied. The number of homophone-pairs it created must have approached ten thousand. I am not quite sure what range of hypothetical mergers Wedel et al. would count as comparable in phonetic plausibility to this one, but I would surmise that these must include mergers which would have created substantially fewer homophones.

Comparing this quantitatively with the mergers in Wedel et al.'s data is not easy, because their "online supplemental material" includes no figures and lists only mergers between individual phonemes, whereas sound-changes commonly affect classes of phonemes. For instance, the first merger they list for RP English is between /θ/ and /t/, while in reality this is one case of a sound-change which also

merges /ð/ and /d/. (I believe it applies in some Irish dialects.) But, using an English dictionary which feels comparable in scope to Chao and Yang for Chinese, I find that this sound-change yields 142 homophone-pairs.[4] In line with Wedel, Jackson, and Kaplan (2013: 410) this figure does not count inflected forms separately (e.g. *heat ~ heath*, *heats ~ heaths* count as one pair), but it does include some quite obscure pairs, e.g. *dhow ~ thou* or *bath ~ Bt.* (A dhow is a type of Arab boat; *Bt* is an abbreviation for *baronet* but can apparently be pronounced /bɑt/ as a separate word.) Whether or not this particular English merger is wholly typical, the disparity between 142 and "approaching ten thousand" for the Chinese *j q x* merger is broadly representative of the difference between homophony in Chinese and in European languages.

6 Individual sound-changes versus overall threshold

Abby Kaplan (2015) argues against my 2013 paper by pointing out that a claim that languages tend to prefer sound-changes which create fewer homophones over other sound-changes which would create more homophones does not imply that there is some absolute threshold level of overall homophony which languages cannot cross. She suggests that many of those linguists who have advocated a functional-yield theory have explicitly argued for the former but have said nothing about the latter.

I accept that the one idea does not logically entail the other, but again I would appeal to the concept of general scientific plausibility. If it were universally true that languages prefer those sound-changes which create fewer homophones, that could surely only be because an excessive level of homophony interferes with communication. And if that is so (as seems undeniable) then there must be some level of homophony which is in practice intolerable. Of course that would not be a threshold expressible as a specific number, so that *N* homophone-pairs in a language are all right but *N*+1 pairs are forbidden. But when things reach the point where a largely monomorphemic vocabulary has to be replaced by a largely bimorphemic vocabulary in order to preserve intelligibility, as happened in Mandarin, it seems certain that the language as it would have been without vocabulary replacement would have exceeded any tolerable level of ambiguity.

Kaplan also points out, correctly, that those who have discussed homophony avoidance have done so in terms of statistical tendencies rather than absolute rules, and she says that the existence of one example violating a tendency cannot refute a statistical law. This might protect the functional yield theory from the Chinese counterexample, if the move from low Old Chinese homophony to very high Mandarin homophony had resulted from a single sound-change. But in fact the present-day situation is the outcome of many separate sound-changes over thousands of years. Various different types of consonant cluster were reduced to single consonants between Old and Middle Chinese, almost certainly in a series of separate changes rather than just one. Since the Middle Chinese period, apart from the three changes already mentioned (loss of -p -t -k; merger of -m and -n; mergers yielding *j q x*) there was loss of voicing in obstruents (which created new homophones among 仄聲 words), and loss of initial ŋ-. (I do not discuss vowel changes, because these are

4 I used the computer-usable version of the *Oxford Advanced Learner's Dictionary*, available by ftp via <www.filewatcher.com/m/CUVOALD.tar.gz.816821-0.html>, ignoring proper names (the file includes many obscure names which would be unfamiliar to the average English-speaker, a phenomenon having no parallel in Chinese).

harder to individuate and may not always have affected the incidence of homophony.) Various scholars have posited further, smaller-scale changes which also reduced the range of phonetic contrasts, but the changes listed above are agreed by everyone. So we are not talking about a single exception to a universal tendency. Rather, one particular language has again and again changed in ways which increased, and often massively increased, the number of homophones. That is not consistent with a universal law of homophony avoidance even if that law is statistical rather than an absolute prohibition.

I cannot find a statement in Wedel, Kaplan, and Jackson's paper of whether they believe merger probabilities relate to minimal-pair counts linearly or by some other mathematical function. If linearly, then one might think that the number of homophone pairs produced by a Chinese merger such as the *j q x* case, being orders of magnitude larger than the numbers produced by European-language mergers, ought to be associated with a probability so minuscule that one would not expect to find a single example in the few thousand languages spoken in the world.

7 Conclusion

I did not offer to provide a solution to the paradox I have discussed, and at present I cannot suggest a solution. The paradox is real, and ought to concern anyone who is interested in relating the Chinese language to the principles of general linguistics.

References

Baerman, M. 2011. Defectiveness and homophony avoidance. *J. of Linguistics* 47.1–29.

Baxter, W.H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.

Blevins, J. and A. Wedel. 2009. Inhibited sound change. *Diachronica* 26.143–83.

Bouchard-Côté, A., D. Hall, T.L. Griffiths, and D. Klein. 2013. Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences* 110.4224–9.

Campbell, L. 1996. On sound change and challenges to regularity. In M. Durie and M. Ross, eds, *The Comparative Method Reviewed*, 72–89. Oxford: Oxford University Press.

Chao Yuen Ren and Lien Sheng Yang. 1962. *Concise Dictionary of Spoken Chinese*. Cambridge, Mass.: Harvard University Press.

Economist. 2014. Bad characters. *The Economist* 23 Aug 2014, p. 52.

Feng Shengli and V. Yip, eds. 2014. *William Labov and William S.-Y. Wang: a dialogue on sound change*. Peking: Peking University Press.

Gilliéron, J. 1918. *Généalogie des mots qui designent l'abeille d'apres l'ALF.* Paris: Champion.

Kaplan, A. 2011. How much homophony is normal? *J. of Linguistics* 47.631–71.

Kaplan, A. 2015. The evidence for homophony avoidance in language change: reply to Sampson (2013). To be in *Diachronica* vol. 32 no. 2.

King, R.D. 1967. Functional load and sound change. *Language* 43.831–52.

Li, C.N. and S.A. Thompson. 1987. Chinese. In B. Comrie, ed., *The World's Major Languages*, 811–33. London: Croom Helm.

Martinet, A. 1955. *Economie des changements phonétiques*. Bern: Francke.

Norman, J. 1988. *Chinese*. Cambridge: Cambridge University Press.

Norman, Y. 2009. *L'Influence de l'écriture sur la langue*. Doctoral thesis, Université de Paris III – Sorbonne Nouvelle.

Sampson, G.R. 2013. A counterexample to homophony avoidance. *Diachronica* 30.579–91.

Silverman, D. 2006. *A Critical Introduction to Phonology*. London: Continuum.

Silverman, D. 2010. Neutralization and anti-homophony in Korean. *J. of Linguistics* 46.453–82.

UNESCO. 2006. *Education for All Global Monitoring Report 2006*. Online at <www.unesco.org/education/GMR2006/full/chapt8_eng.pdf>, accessed 14 Aug 2014.

Wedel, A., S. Jackson, and A. Kaplan. 2013. Functional load and the lexicon: evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech* 56.395–417.

Wedel, A., A. Kaplan, and S. Jackson. 2013. High functional load inhibits phonological contrast loss: a corpus study. *Cognition* 128.179–86.