A counterexample to homophony avoidance*

Geoffrey Sampson
University of South Africa

1.   A hypothesis revived

The idea that avoidance of homophony is an important factor influencing which sound changes do or do not occur in languages has a long pedigree within linguistics. It was introduced almost a hundred years ago by Jules Gilliéron (1918: 14) and was made widely known by André Martinet in writings from the 1930s onwards, culminating in his 1955 book *Economie des changements phonétiques*. Martinet claimed that the extent to which a particular phonological contrast resisted elimination through sound change depended on the *rendement fonctionel* (variously translated into English as 'functional yield' or 'functional load') of the contrast, that is, roughly, the quantity of minimal pairs which the contrast served to keep distinct. Although this idea might look plausible a priori and in Martinet's version became very influential, it has often been criticized, notably by Robert King (1967). King made the hypothesis precise enough to test quantitatively, and used the histories of four Germanic languages as test-beds. He concluded "functional load, if it is a factor in sound change at all, is one of the least important" (King 1967: 848). But King's careful investigation was not accepted as settling the issue. Paul Lloyd briefly surveyed the history of the homophony-avoidance hypothesis since King's article, pointing out the various difficulties in supposing that languages systematically avoid creating homophones, but adding, "And yet it is clear that wholesale mergers of phonemes do not occur" (1987: 38). That is, while Lloyd knew that phoneme mergers are not always avoided, he took it to be self-evident that their occurrence is constrained within some kind of limits.

Recently the hypothesis has been revived. Blevins & Wedel (2009) argue for a phenomenon they call 'inhibited sound change'. They quote Lyle Campbell (1996: 77, and cf. 2004: 322) as arguing that "While scholars opposed to teleological explanations in linguistics have never been friends of the explanation of certain changes as due to the avoidance of pernicious homophony, such avoidance is … an undeniable empirical reality." (By his unexplained term 'pernicious' homophony, Campbell appears to mean homophony likely to lead to real confusion.) Matthew Baerman (2011: 3) tells us that he was converted to believing in homophony avoidance through research aimed at demonstrating the opposite. (He also quotes a Google search which confirms that the idea is widely taken for granted (2011: 2 fn. 4).) Garrett & Johnson mention other recent publications which have argued for a mechanism of homophony avoidance, though they themselves caution that "[r]esearch in this area is intriguing but not yet definitive" (2013: 82 fn. 35).

A first point to make is that a claim "language change avoids creating homophones" is too ambiguous to rank as a specific hypothesis. If it were interpreted as (i) "language change *never* creates homophones", it would be obviously false: any linguist can quote cases where a pair of words once pronounced differently have been turned by some sound change into homophones. But even if interpreted as a statistical tendency rather than an absolute constraint, it could mean either (ii) that what languages tend to avoid is adoption of a sound change which would create many homophone pairs or alternatively (iii — a weaker statement) that when a rule which would create homophones is adopted by a language, individual forms that would become homophones will tend to be exempted from application of the rule, contrary to the Neogrammarian idea that sound laws are exceptionless but in the spirit of work on lexical diffusion. (We shall see that even weaker interpretations are also available.)

Alternative interpretations of the claim about homophony avoidance have not been a salient topic in the literature cited above; and even if some of those writers have been explicit about what they were claiming, many linguists today who take homophony avoidance for an established finding have not spelled out which interpretation they have in mind. Consequently the present paper is not concerned with refuting one particular version of the homophony-avoidance idea, but to challenge those sympathetic to the idea to produce some version of it which is in principle falsifiable, which makes a strong enough claim not to be trivial and which is compatible with the facts to be presented here. The paper analyses well-documented historical data which, in my view, must refute any version of the homophony-avoidance idea that is not so weak as to be uninteresting.

Contrary to Paul Lloyd's statement, wholesale mergers of phonemes can occur. They have occurred repeatedly in the roughly 3000-year evolution of modern Mandarin Chinese from the Old Chinese of China's early period as a literate

society. At the least, linguists who continue to develop theories of homophony avoidance need to ensure that their theories are compatible with the Chinese facts to be discussed, which are very unlike any phenomena known to me from the European languages that have more commonly been considered in this connexion.

2. The Chinese background

Chinese is a language in which syllables are highly salient units with clear boundaries (there are no 'interlude' consonants like the /t/ of English *butter*, which belongs as much to the preceding as to the following syllable),[1] and with marginal exceptions morphemes are realized as single syllables. Particularly at the beginning of the period studied here, Chinese was grammatically close to the ideal type of isolating language, and there was little or no distinction to be made between morphemes and 'words'. (The linguistic usefulness of the concept 'word' is debated even with respect to the modern language.) In consequence there have been relatively few and insignificant phonological processes applying across syllable boundaries; the history of Chinese phonology consists mainly of the evolution of individual syllable shapes. Again and again, that history has involved changes which greatly reduced the number of possible distinct syllables, and hence increased the incidence of homophony between morphemes. By now, that incidence is very high.

We shall look in detailed quantitative terms at the segment of that history since the 'Middle Chinese' (Norman 1988: 23, Baxter 1992: 14–15) of about 600 C.E. (This is the period when, according to the majority scholarly consensus, Chinese began to fission into most of the present-day dialects, so that dialect comparison is one source of evidence for Middle Chinese pronunciation; others are the extensive philological work carried out at the time, including 'rhyme tables' compiled for the use of poets, and wholesale loans of vocabulary from Chinese into neighbouring languages.) Homophony-promoting changes were already happening between Old and Middle Chinese — it is sure, for instance, that Old Chinese had syllable-initial consonant clusters, which were all reduced to single consonants by the Middle Chinese period; but there is too much room for debate about Old Chinese phonology to discuss it quantitatively with confidence, whereas open issues in the reconstruction of Middle Chinese phonology are limited to matters of detail that are unlikely crucially to affect the points discussed here.

Among the modern Chinese dialects, Mandarin is a good choice to compare with its ancestor not only because it is the standard and the best-documented

---

1. For 'interlude', see Hockett (1955: 52).

variety of the language, but also because some other dialects have undergone fewer mergers than Mandarin.

3. Mandarin mergers and the response to them

Some processes which have applied to syllable-structure between Middle Chinese and current Mandarin include (sounds represented in the standard Chinese *pinyin* romanization system; rules not listed in historical sequence):

– final obstruents (*–p –t –k*) dropped, i.e. merged with zero
– final *–m* and *–n* merged as *–n*
– voiced obstruents became voiceless
– apical sibilants and velars merged as alveolo-palatals before close front vowels (*z, g > j; c, k > q; s, h > x* before *i, ü*)
– initial *ng-* dropped
– the vowel system was simplified

Like other proto-languages reconstructed largely by comparing related modern languages or dialects, Middle Chinese may be to some extent an artificial construct containing postulated features that did not all co-occur at any one historical stage of the ancestor language (cf. Baxter 1992:27). But few knowledgeable scholars would argue that this list of large-scale sound changes depends on such debatable details. That is, Mandarin surely did have an ancestor language at one period which contained three syllable-final obstruents that have all been replaced by zero, which genuinely had syllables in *–m* that have turned into syllables in *–n* and so forth — whether or not that ancestor language was identical in every respect with 'Middle Chinese' as described by Baxter or other Chinese historical linguists.

The consequence of these numerous mergers is that Mandarin has far fewer distinct syllables than morphemes; most syllables are homophonous, often multiply homophonous. Many Chinese dictionaries give an exaggerated impression of this by listing numerous obsolete morphemes from the long history of Chinese letters, but Chao & Yang (1962) is one dictionary which is careful to list only elements of the living spoken language of its day. (Chao Yuen-jen was the first Chinese to apply the techniques of twentieth-century Western synchronic linguistics to his language.) Chao & Yang show the syllable *xī*, for instance, as ambiguous between 23 different morphemes, various of which had at least six distinct pronunciations at the Middle Chinese period; some examples are (with starred forms representing

Middle Chinese as reconstructed by Baxter 1992): *xjɨj "hope"; *xje "sacrifice"; *sej "west"; *xip "inhale"; *sit "cricket"; *sek "cleave".[2]

Twenty three is one of the highest figures for numbers of morphemes sharing a pronunciation, but plenty of other syllables are equally or almost equally homophonous; only a minority of syllables unambiguously represent just one morpheme. Chinese has always been a language with considerable fluidity about using the same morpheme in different grammatical roles, so the clear noun/verb/adjective distinctions among the English glosses for the examples above give a false impression of the extent to which logical functions helped to resolve ambiguity among Chinese homophones. With this level of homophony overall, almost all examples are 'pernicious'. If the vocabulary of Old Chinese, which consisted largely of single-morpheme words, had been retained unchanged while the phoneme mergers that yielded modern Mandarin had proceeded, the result would have been a spoken language that was too full of ambiguity to be usable. (This is not a speculative judgement. Classical Chinese read aloud cannot be understood, without sight of the text, even by scholars who are very familiar with that language. In writing, on the other hand, homophony is not an issue, since Chinese script is not phonographic.)

Comprehensibility was maintained by replacing most single-morpheme words of Old Chinese with two-morpheme compounds. These compounds are of various kinds, but in particular many of them are of a type that is very unusual in European languages, comprising a pair of synonyms or near-synonyms. For instance, "friend" in Mandarin is *péngyǒu*; in Old Chinese, *péng* and *yǒu* (or rather, the forms from which those Mandarin syllables descend) were separate words each meaning "friend", but neither morpheme occurs as an independent word in modern Mandarin. According to Chao & Yang, *péng* is seven ways ambiguous as a morpheme of Mandarin and *yǒu* is three ways ambiguous, but the compound *péngyǒu* is unambiguous. Many of the 23 *xī* morphemes are marked by Chao & Yang as occurring only as 'bound forms', that is as parts of compounds rather than as independent words. The shift from predominantly monomorphemic to

---

2. Some of the 23 morphemes are marked by Chao & Yang as 'literary', meaning that in modern speech they occur only in fragments of the Classical language which survive in present-day usage. These morphemes will be familiar chiefly to more educated speakers and hence, arguably, are less relevant to the issue of homophony avoidance. But, in the first place, only six out of 23 *xī* morphemes are marked as exclusively literary. Secondly, 'literary' does not mean 'abstruse'; Chao & Yang (1962) is a dictionary of spoken language. Thus, they mark *xī* "formerly" as literary only, but a Chinese would not need much education to understand a phrase like, say, *jīn bù rú xī* "things aren't what they used to be". (This *xī* is listed by Chao & Yang as second-tone, *xí*, but other dictionaries give it as first-tone, *xī*, and that would be the expected reflex of Middle Chinese *sjek.)

predominantly bimorphemic words is not the only device which has tended to compensate for increased homophony; another, for instance, is that Mandarin has made modest moves away from the isolating extreme towards inflexion. But the adoption of bimorphemic vocabulary is by far the most important compensation device. It is difficult to be precise about how far this shift had proceeded in the spoken language by the Middle Chinese period, since written Chinese then (and up to the beginning of the twentieth century) was conventionally couched in the vocabulary and grammar of the Classical language, modelled on Old Chinese, which has long been obsolete as a spoken language. Available evidence suggests that compounding had got under way in Middle Chinese but was less extensive and the compounds less fixed than in modern Mandarin. (See e.g. Norman 1988: §5.1, esp. 112–113.)

4. Possible replies

Some commentators object to the suggestion that Mandarin refutes the homophony-avoidance hypothesis by saying that homophony should be measured among words rather than among morphemes, and the modern vocabulary of (mainly bimorphemic) words contains no greater level of homophony — perhaps a lower level — than the monomorphemic vocabulary of Old Chinese. If this is an objection, that implies an interpretation of the homophony-avoidance hypothesis which would be much weaker than any of the alternatives discussed in §1 and to my mind would be quite uninteresting, along the lines "A language will not allow the incidence of homophony to rise excessively, or if it does then it will adopt compensating devices". That surely says very little; we scarcely need linguistic research to tell us that languages will not change in ways that make them unusable. I have found one writer on homophony avoidance, Anatole Lyovin (1977), who has expressed essentially the view just quoted, but Lyovin does not represent this as an interesting, falsifiable claim about language change (his interest lies elsewhere). For those linguists who see the homophony avoidance hypothesis as an interesting finding, I take it that their interpretation includes no "or if it does" clause. Without such a clause, the claim could be nontrivial; but Mandarin will refute it. The very fact that this language has replaced most simple words by compounds seems to confirm that the incidence of homophony created by sound changes was indeed excessive.

Blevins & Wedel's concept of 'inhibited sound change' refers to hypothesis (iii) above: many forms affected by a rule are unambiguous both before and after the change, but one or a few forms which would be turned into homophones by the rule fail to undergo it (2009: 145–146 and n. 2). Hypothesis (iii) is weaker than (ii),

but still strong enough to amount to a worthwhile claim. But it assumes a situation in which vocabulary items are scattered sufficiently sparsely through the phonological space permitted by a language structure that a sound change which would cause a few items to collide will shift most affected items to vacant locations in the space. With European languages that is often the situation. Chinese, already in the Middle Chinese period, was not like that: phonological space was densely occupied, so that a change which created any collisions at all would normally create many. It would be for defenders of hypothesis (iii) to explain whether they are assuming that phonological space always will be sparsely occupied, or are predicting that in a language where phonological space is densely occupied few or no homophone-creating changes will occur. Either way, Chinese is a clear counterexample.

Someone aiming to defend the homophony avoidance idea against the Chinese data might suggest that, if the adoption of bimorphemic vocabulary largely preceded the operation of the sound mergers, then the latter would not have created homophony when they happened. But, in the first place, this is not normally understood to have been the historical sequence (Norman 1988: 86), though the evidence either way is admittedly poor (see above). More important, postulating such a sequence would create a large puzzle about why the shift to bimorphemic words should have occurred. This would be particularly mysterious in the case of the many synonym-compounds of the *péngyǒu* type. If a word is ambiguous one can understand why speakers might want to achieve clarity by adding a synonym (compare the one-off English parallel "funny peculiar, not funny ha-ha"). But, without the pressure of homophony, it would surely not be natural for speakers to adopt a system of repeatedly saying the same thing twice? I do not believe that this defence of homophony avoidance can succeed.

5.   Quantifying homophony and its increase

I turn now to the matter of putting figures on the incidence of homophony.

The range of possible distinct syllables in Mandarin can be defined by the rules of Figure 1 (where commas separate alternatives, and elements concatenated without a comma must both/all be taken).[3]

---

3. *Pinyin* romanization, slightly modified in order to clarify the structural patterns (for instance by omitting the *y- w-* that *pinyin* prefixes to syllables beginning with close vowels), is used in preference to phonetic symbols, since Mandarin transcribed in the IPA alphabet requires obscure and complex notation, the details of which have no relevance here. (That is specially true for the phoneme shown here as *ï*, which is 'underlyingly' a close back spread vowel, IPA [ɯ], but is realized as different velarized syllabic fricatives depending on preceding consonant.) Numerals 1 to 4 represent the four Mandarin tones.

$$\text{Syll}\begin{cases}\text{(b, p, m) (i, u, (\O, i))}\\\text{f (u,)}\\\text{(d, t) (i, u, (\O, i, u))}\\\text{(n, l) (i, ü, u, (\O, i, ü, u))}\\\text{(z, c, s, zh, ch, sh, r) (u, ï, (\O, u))}\\\text{(j, q, x) (i, ü, (\O, i, ü))}\\\text{(g, k, h) (u, (\O, u))}\\\text{\O (i, ü, u, er, (\O, i, ü, u))}\end{cases}\text{(1, 2, 3, 4)}$$

(e, a) (Ø, i, u, n, ng)
(but not i in a syllable already containing i or ü
and not u in a syllable already containing u or ü)

Figure 1. Mandarin syllable-formation rules

These rules allow a total of 2056 distinct syllables. The number of morphemes listed by Chao & Yang is 4743,[4] so the average syllable is ambiguous between 2.3 morphemes. This would surely seem a bizarrely high incidence of homophony to a linguist familiar exclusively with European languages, but the average figure actually understates the homophony problem, because homophony is distributed very unevenly across the range of syllables. Various syllables represent different numbers of morphemes, ranging from twenty-plus at the upper end down to one or none. The syllable *rï4*, for instance, represents only the morpheme for "day" (etymologically, "sun"), and *rï* does not occur with any other tone.[5] To see why

---

4. This number was derived by counting index entries on Chao & Yang's pp. 255–278, ignoring entries not fitting the Mandarin syllable-formation patterns (e.g. *dü1* "toot", *pf* "interjection meaning 'What shall I do?' " — even if such forms are counted as morphemes of the language, we will never have data on comparable forms in earlier stages of Chinese) together with letters of the Chinese 'National Phonetic Alphabet', an obsolete system now of only antiquarian linguistic interest.

5. Because some phonologically possible syllables correspond to no morphemes, the number of distinct syllables actually occurring in the modern language is substantially fewer than the figure of 2056 quoted above. It might seem better to use a count of actually occurring syllables, so as to avoid questions about whether cases such as non-fourth-tone *rï* are accidental gaps or systematically ill-formed. However, for Middle Chinese, while it is reasonably practical to count phonologically possible syllables (as in the following paragraph), it would be difficult or impossible to count just how many of those possibilities were actually used; so for the sake of comparing like with like I have worked with systematic possibilities at both language-stages. (This is a conservative decision; to use a count of syllables actually occurring in Mandarin would further exaggerate the already striking disparity between the two figures.)

this skewing matters, consider two very simple systems, each having just six morphemes realized by two different syllables. In system *A*, three morphemes share one syllable and three the other; in system *B*, five morphemes share one syllable and the other morpheme is represented unambiguously by the other syllable. In either case the average ambiguity per syllable is three; but what matters to a language user is the average ambiguity *per morpheme*. In system *A* this will also be 3; but in system *B* it will be $(5 \times 5 + 1 \times 1) \div 6 = 4.3$. If different morphemes are similar in frequency, a hearer will encounter highly ambiguous syllables much more often than unambiguous syllables.[6]

The details of phoneme distribution given by Baxter (1992: ch. 2) for his reconstruction of Middle Chinese do not allow the set of possible syllables to be distilled into such concise formulae as those shown in Figure 1 for Mandarin, and readers would find it tedious to work through the details. But, for example, the four Middle Chinese retroflex initial consonants /tr trh dr nr/[7] are listed as combining with:

– two syllable-finals /æj wæj/ that occur only with one of the Middle Chinese tones
– 17 finals ending in obstruent consonants, which are therefore outside the tone system (traditionally, syllables ending in obstruents were said to have a tone of their own)
– 31 finals compatible with each of the three Middle Chinese tones

Hence there were $4 \times (2 + 17 + 3 \times 31) = 448$ possible Middle Chinese syllables with retroflex initials. Calculating the possibilities given for the other classes of initial consonant leads to a total of 5688 distinct phonologically possible Middle Chinese syllables.[8]

It seems that the mergers discussed above have reduced the range of available shapes for morphemes by a factor of about 2.8 in the passage from Middle Chinese to Mandarin. (This agrees reasonably well with the ratio of 3:1 quoted by Wang 1969: 10 n. 3, who does not explain how he derives it.) If the morpheme vocabulary of Middle Chinese had been identical in size to that counted from Chao &

---

6. Properly, quantification of the disambiguation problem ought to take into account not just numbers of morphemes but their relative frequencies. I have no data which would enable me to incorporate that consideration into the figures I quote for Chinese.

7. These are Baxter's typographically convenient notations for single retroflex consonants; they do not contradict the point above (§2) that Middle Chinese had no consonant clusters.

8. This calculation is intended to be conservative; in particular, the so-called 'double button' final-pairs are counted once rather than twice each, in view of debate about whether they represented a genuine phonological distinction.

Yang's (1962) dictionary for Mandarin, each morpheme of the former could have had a unique pronunciation. (In fact the Middle Chinese vocabulary was not so tidily organized; Middle Chinese, and indeed Old Chinese, certainly contained some homophony — as, surely, all languages do.)

6.   Vocabulary development

To be sure just how far the incidence of homophony has increased in practice it would be necessary to produce a count of morphemes that were current in the Middle Chinese vernacular, parallel to the count based on Chao & Yang. That is impossible: nobody fourteen centuries ago was doing that type of lexicography. But notice that, if one wanted to challenge the claim that homophony has greatly increased, one would need to argue that Middle Chinese had far *more* morphemes than the modern language (so that they were crowded equally densely in the much larger space of distinct syllables). While it is hard to know just which of all the morphemes recorded in writing over 3000 years of Chinese literacy were current in the speech of that period, the idea that there were far more then than now seems surprising. Surely, when a society enjoys a high level of civilization over many centuries, it would be natural for the stock of concepts encoded as morphemes gradually to increase; some morphemes will become obsolete, but these will be more than offset by new morphemes encoding new concepts.

Commonly the latter happens through a form developing polysemous extensions to its earlier meaning which in due course lose their psychological link with that sense and come to be effectively separate morphemes, and that happened frequently in Chinese: for instance, the *xī* whose original meaning was "to cleave (wood)" came, via the idea of splitting, to mean "analyse" — someone who did not know that these two examples of *xī* are written with the same Chinese graph might be unlikely to think of them as related. In other cases, morphemes are borrowed from another language; Chinese is a language relatively resistant to borrowing, but another *xī* was coined in Mandarin to translate English *selenium* (*xī* is as close as one can get in Mandarin to English /si/). The first of these two processes is much the commoner in Chinese, and, because the script usually retains the same graph for polysemous senses of a single etymon no matter how far apart the senses grow, a morpheme count based on Chinese script will underestimate vocabulary size. My count of morphemes in Chao & Yang (1962) did not attempt to distinguish between morphemes which share both a pronunciation and a written graph (it would not have been practical to do so), so my figure for Mandarin homophony is surely an underestimate. The figure of 2.8 as the factor by which homophony has increased since Middle Chinese is much more likely to be too low than too

high. And although we have seen that homophony avoidance can only reasonably be understood as a statistical tendency rather than an absolute constraint, implying that we cannot expect the hypothesis to identify a precise ceiling above which homophony creation will not proceed, an increase of this order surely goes far beyond anything that the devisers of the idea would see as compatible with it.

7.   Provisos

We cannot say that homophony avoidance has never occurred in the history of Chinese. If it had undergone the regular sound changes, Middle Chinese *pjie "inferior" would have become not the actual modern form *bēi* but *bī*; it is easy to guess that the reason for the irregularity was to avoid homophony with *bī* "vagina". Avoiding homophony with taboo words is an uncontroversial phenomenon, but it is a very special case of homophony avoidance. Unless a language were remarkably rich in taboo words (which Chinese is not), it could never be a major factor in controlling sound changes.

If Campbell's concept of 'pernicious homophony' meant only this kind of 'embarrassing homophony', then I could not quarrel with his claim. Of the four examples quoted in Campbell (2004: 322–324) three are of this type; but his leading example is Gilliéron's French case of homophony between words for "cat" and "cock" (American "rooster"), which is not. Thus Campbell appears to be using the undoubted phenomenon of avoidance of homophony with taboo words in order to lend plausibility to Gilliéron's much broader claim about homophony avoidance. Likewise Baerman (2011: 3) sees it as uncontroversial "that we may find instances where an entire lexeme *Y* is avoided outright due to homophony with some other lexeme *X*, as occurs with taboo avoidance". I would say that this is uncontroversial when *X* is indeed a tabooed word, but Baerman's "as" suggests that this is just one case of a broader phenomenon: that is controversial.

It is also not the case that every one of the phonological developments leading to modern Mandarin which affected the number of distinct syllables shrank rather than expanding that number. Historically, there was no route leading to first-tone syllables beginning with sonorants /m n l r/, since the contrast between tones 1 and 2 derives from the earlier contrast between voiceless and voiced initial consonants, and /m n l r/ have always been voiced. However, while there are fewer sonorant-initial Mandarin morphemes in tone 1 than in the other tones, there are enough to make it clear that such syllables must now count as phonologically well-formed: this is an expansion of the syllable inventory. But it is a very minor expansion, relative to a series of major contractions.

8.   Homophony embraced

When the vocabulary of a language is packed fairly densely into its pronunciation-space with relatively few unused possibilities, as was already the case in Middle Chinese, perhaps one might think it inevitable that almost any plausible sound-change affecting large numbers of forms would create homophones. But that is not true. For instance, a linguist familiar with many languages of the world but not with Chinese, told that something dramatic happened to the final obstruents /-p -t -k/, might well guess that they became fricatives, /-f, -θ or -s, -x/ — that is a very natural sound-change on a world scale (it occurred for instance in Hebrew and in Liverpool English), and it would have created no homophones at all since Middle Chinese had no final fricatives. Yet so far as I know, developments like that are rarely or never found in the history of any of the dialects which descended from Middle Chinese. In Mandarin the final stops all merged with zero, greatly increasing homophony. Far from Mandarin avoiding them, one is almost tempted to see this language as having systematically selected homophony-creating changes.

Baerman, while believing in a mechanism of homophony avoidance, seems to argue (2011: 25) that it is language-specific: some languages would be homophony-avoiding languages, others not, and Chinese (not discussed by Baerman) would clearly be one of the latter. If this idea is admissible in principle, it would offer a way of reconciling the Chinese facts outlined here with the data used by other writers to argue for homophony avoidance. My difficulty with Baerman's suggestion is that the a priori plausibility of the homophony avoidance idea stems from the fact that people normally want communication to be clear. That is presumably true everywhere, so I do not understand how a mechanism motivated by that goal could apply to some languages but be entirely irrelevant to another language.

And if it is conceived as a language-universal mechanism (as most linguists who discuss the subject have supposed), then I see no possibility of turning avoidance of pernicious homophony into a precise, falsifiable and non-trivial hypothesis which would have a chance of standing up in face of the facts of Chinese. If such a hypothesis is available, I have yet to see it stated.

References

Baerman, Matthew. 2011. Defectiveness and homophony avoidance. *Journal of Linguistics* 47(1). 1–29.

Baxter, William H. 1992. *A Handbook of Old Chinese phonology*. Berlin & New York: Mouton de Gruyter.

Blevins, Juliette & Andrew Wedel. 2009. Inhibited sound change. *Diachronica* 26(2). 143–183.

Campbell, Lyle. 1996. On sound change and challenges to regularity. In Mark Durie & Malcolm Ross (eds.), *The comparative method reviewed*, 72–89. Oxford: Oxford University Press.

Campbell, Lyle. 2004. *Historical linguistics*, 2nd edn. Cambridge, MA: MIT Press.

Chao, Yuen Ren & Lien Sheng Yang. 1962. *Concise dictionary of spoken Chinese*. Cambridge, MA: Harvard University Press.

Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan C. L. Yu (ed.), *Origins of sound change*, 51–97. Oxford: Oxford University Press.

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

Hockett, Charles F. 1955. *A manual of phonology* (Indiana University Publications in Anthropology and Linguistics, Memoir 11). Baltimore: Waverly Press.

King, Robert D. 1967. Functional load and sound change. *Language* 43(4). 831–852.

Lloyd, Paul M. 1987. *From Latin to Spanish* (Memoirs of the American Philosophical Society 173, vol. 1). Philadelphia: American Philosophical Society.

Lyovin, Anatole. 1977. Sound change, homophony, and lexical diffusion. In William S.-Y. Wang (ed.), *The lexicon in phonological change*, 120–132. Mouton: The Hague.

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

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

Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45(1). 9–25.

Author's address

Geoffrey Sampson
Linguistics Department
PO Box 392
University of South Africa
UNISA 0003
South Africa

sampson@cantab.net