

Geoffrey Sampson

Does Simple Imply Creole?

Abstract

A number of linguists argue that (contrary to earlier linguistic axioms) some languages are structurally simpler than others, and that creoles are among the simplest languages. John McWhorter has gone further and claimed that creoles alone are specially simple. I accept the general point about a gradient of simplicity, but I show that Old Chinese is a non-creole which has all the hallmarks, including structural simplicity, claimed by McWhorter to co-occur only in creoles.

1. Introduction

I owe largely to Fred Karlsson my introduction to a recent trend in linguistic thinking, according to which certain human languages are structurally far simpler than others. Notably, John McWhorter (e.g. 2001a) has claimed that creoles are simpler than “regular languages” (McWhorter’s term for non-creole languages, i.e. languages with long histories);¹ and David Gil (e.g. 1994) has shown that Riau Indonesian, a colloquial (“basilectal”) local dialect of Malay,² lacks much of the structural apparatus, such as word classes, which Eurocentric linguists wrongly suppose to be universal.

I take these claims to be correct—and important, since they contradict views which for much of the twentieth century were treated as axiomatic. (McWhorter’s view of the simplicity of creoles has been vigorously

¹ I am grateful to John McWhorter for discussion of our different points of view on the topic of this paper.

² Malay and Indonesian (or “Bahasa Indonesia”) are separate national names for the same spoken language, written using slightly different spelling conventions. I shall refer to it as Malay.

challenged, but I have the impression that the challenges are founded more on political commitments than on empirical scientific disagreements.) However, McWhorter goes further, and argues in effect not merely that “if creole then simple”, but also that “if simple then creole”. In response to Gil’s argument that Riau Indonesian has very much the kind of simplicity which McWhorter sees as characteristic of creoles, McWhorter replies (2001b: 405–408; 2005: 69–71) that Malay in general should be classified as a creole; he discusses historical considerations tending to make this perhaps surprising proposition plausible.

I am not qualified to debate the history of Malay with McWhorter, but I do not accept the general point: the “if” of “if creole then simple” is not an “if and only if”. This paper seeks to demonstrate that, by reference to another language which is not a creole and which has the kinds of structural simplicity that Gil and McWhorter discuss.

2. A counterexample: Old Chinese

My example language is Old Chinese, defined by William Baxter (1992: 1) as the Chinese language of roughly the eleventh to seventh centuries B.C. This was the period of the earliest monuments of Chinese literature, notably the “Book of Odes” (or “Classic of Songs”), from which my own examples will be drawn. Old Chinese is ancestral to all present-day Chinese “dialects” or “languages”, such as Mandarin, Cantonese, Hakka, etc., and many characteristics which made Old Chinese structurally simple are inherited by modern varieties of Chinese. However, Old Chinese is a clearer counterexample: partly because it differs from all modern Chinese varieties with respect to one specific feature, tone, that McWhorter sees as important in this context, and partly because recent varieties of Chinese have tended to embody rather more explicit structural apparatus than their ancestor. (The latter fact may perhaps be partly a consequence of the increasing complexity of Chinese civilization—see Deutscher (2000) for the relationship between cultural sophistication and grammatical complexity; over the last 150 years it has also been encouraged through the influence of European languages.)

I begin by making the point that Old Chinese has never to my knowledge been taken to be a creole, and there seems to be no reason whatever for supposing that it was one. It is a member of the Sino-Tibetan language family (on which see Thurgood and LaPolla 2003), and its predominantly monosyllabic, isolating character is shared with other

members of that family (hence cannot have resulted from a creolization event). So far as I know, creoles are usually or always created by members of a subordinate ethnic group in contact with a more powerful group who speak a different language; but the speakers of Old Chinese were by far the most culturally advanced group in the East Asia of their time—historically, other East Asian nations borrowed the elements of their cultures from the Chinese, far more than the other way round. No language is ever wholly cut off from the influence of neighbouring languages, but Old Chinese gives the impression of being as “pure” in that respect as any language I have encountered. The most obvious non-Chinese element in premodern Chinese is the subset of vocabulary consisting of polysyllabic foreign loan words, such as *luo4tuo0* “camel”,³ but according to Pulleyblank (1995: 9) no clear cases of such loans are found before the Han dynasty—i.e. several centuries after the Old Chinese period. (On contacts between Chinese and other languages see Norman 1988: 16–22.)

Thus, there is no evidence that Old Chinese was anything other than a “regular language”, descended by normal processes of language change from an unwritten Proto-Sino-Tibetan ancestor. I shall show that it had the hallmarks which McWhorter sees as typifying creoles. But if one were led by that to postulate some hypothetical creolization episode in Chinese prehistory, this would turn “if simple then creole” into a circular truism, rather than an interesting empirical claim. Experts on early Chinese history would, I believe, be incredulous if it were suggested that Old Chinese was acquired “more often by adults outside of a school setting than by children” (the feature that McWhorter identifies (2001b: 408; 2005: 71) as crucial for the empirical status of his thesis).

3. Examples of structural vagueness

I move on now to examine the specific properties which Gil sees as refuting Eurocentric assumptions about language universals and which McWhorter sees as distinctively creole, and how far Old Chinese shares those properties.

The overall picture Gil gives of Riau Indonesian is one of a language where the logic of utterances is strikingly vague and/or ambiguous, relative

³ Here and below I represent Chinese words via the *pinyin* transliteration of their modern Mandarin pronunciations, except that for typographic simplicity diacritic tone marks are replaced by numerals. The Old Chinese pronunciations were very different; where they are relevant, they will be shown in Baxter’s (1992) reconstruction.

to any European language. This has long been recognized as a characteristic of Malay in general, in writings about language stemming from outside the world of modern academic linguistics. For instance, a Malay self-instruction manual (Lewis 1947) comments:

Malay is a flexible, eel-like language (p. xii)

[a word] which, in the list below, is given an adjectival meaning (...), will sometimes be found as a noun (...) at other times it will be a verb. Moreover, when it is a verb it may correspond to any of the verb forms found in the paradigm of e.g. a French verb; it is singular or plural, it is past, present or future, it is active or passive, according to the context (p. 51)

Subordinating Conjunctions are less common in Malay than in English. (...) A speaker will usually balance a pair of short simple sentences, and expect his hearer to infer the relationship between them (pp. 134–135)

To show how similar Old Chinese is in these respects, I begin by quoting a few sample lines from various poems in the *Book of Odes*. I have chosen lines which support my point, but such lines do not need to be searched for; they are entirely typical, even though there are also lines which would illustrate the point less vividly.

For each example I show:

- the Chinese original
- word-by-word English glosses of the Chinese words, with slashes showing where the Chinese word is ambiguous between more specific English translations. (Where a Chinese graph is ambiguous between semantically-unrelated homophones, I gloss only the meaning relevant to the given context, provided it is clear which this is. Stops link multiple English words used to express a single meaning.)
- Bernhard Karlgren's (1950) translation of the line, accompanied in some cases by my own explanatory material in square brackets. Karlgren's book gives scholarly renderings which aim to make the original meaning clear rather than to achieve poetic effects in English.

- (1) *gui1 yan4 Wei4 hou2* (54)
Return.home condole Wei marquis
“I went home to condole with the prince of Wei”
- (2) *fei3 wo3 si1 cun2* (93)
not I/me/my think dwell.on/be.among
[these girls:] “they are not those in whom my thoughts rest [I love another]”
- (3) *zhong1 feng1 qie3 bao4* (30)
everlasting/indeed wind and violent/sudden.rain
[An opening, scene-setting line:] “There is wind indeed and violent weather”
- (4) *jia4 yu2 yu3 xing2* (88)
yoke I/me together.with go
[asking her menfolk to arrange a marriage:] “let yoke the carriage for me to go with him”
- (5) *yong3 shi3 fu2 gu4* (56)
eternal swear will.not/cannot announce
[a woman is having a secret affair with an important man:] “forever, he swears, he will not tell (of our love)”
- (6) *nü3 yue1 guan1 hu1,*
girl open.inverted.commas look interrogative.particle,
shi4 yue1 ji4 ju1 (95)
gentleman open.inverted.commas perfect.particle final.particle
“a girl says: ‘have you been and looked?’
The knight says: ‘I have’”

4. Lack of word classes

A central feature of Gil’s contrast between Riau Indonesian and the “Eurocentric” model of language is that the former has no distinctions among open word classes—the categories noun, verb, adjective are not differentiated, there is simply one general “open” class of words. This is true of Old Chinese. Some words have meanings which map more naturally into English words of one part of speech than others: the first word of (1), *gui1* “to return, go back where one belongs”, is most likely in any context to be rendered by an English verb. But many words stand for concepts which in English are associated equally closely with more than one part of speech, and in those cases one will not normally be able to say that one translation is more correct than another.

In (2), *sil* is glossed “think” in Karlgren’s dictionary of Old Chinese (1957) but is translated by him in this passage as “thoughts”; either part of speech would do equally well to render the Chinese (an alternative translation might be “they are not those I think about and dwell on”). In (3), *bao4* is glossed by Karlgren (1957) as either “violent”, an adjective, or “sudden rain”, a noun phrase (these are not accidental homophones, a sudden cloudburst is seen as “violent” weather). In another context, *bao4* would more naturally be translated by the noun “violence”; for instance, in the line *zhi4 yu2 bao4 yi3* “arrive at *bao4* oh” in Ode 58, where a wronged wife is complaining about her husband, the prepositional nature of *yu2* “at” makes a nominal translation of *bao4* more natural—Karlgrén’s version is “I have (come to:) met with maltreatment”. For that matter, only the fact that the English noun “wind” has no closely-related verb inhibits one from translating *bao4* in (3) as a verb; if one accepts “blow” as an adequate verbal equivalent to “wind”, then a rendering such as “Everlastingly it blows and storms” would be equally justifiable grammatically.⁴

5. Logical indeterminacy

The two features of Riau Indonesian which Gil (2005) sees as most extremely different from European languages are “indeterminacy of thematic roles”, and “indetermina[cy] with respect to ontological types”: that is, the language is commonly inexplicit both about case relationships (elements are not clearly identified as agent, patient, benefactive, etc. with respect to actions mentioned) and about the logical status of propositions—the same utterance can be vague as between describing an action, ascribing a property to a thing, asserting the existence of a thing with a certain property, and so forth.

The same is true of Old Chinese. In (1), neither the quoted line itself nor neighbouring lines mention, via a pronoun or otherwise, the person who is returning home and planning to condole; and this is usual—probably more often than not, when an Old Chinese word is used in a verbal sense it lacks an explicit subject. Likewise, the relationship between “return home” and “condole” is not spelled out. Karlgren translates it as

⁴ The choice between the glosses “everlasting” and “indeed” for *zhong1* is a case where one Chinese graph stands for apparently unrelated homophones and the context does not clearly disambiguate; Karlgren chooses the translation “indeed”, I would have thought “everlasting” slightly more plausible. “Everlasting” versus “everlastingly”, on the other hand, is a purely English distinction with no parallel in Old Chinese.

“to”, i.e. “in order to”, which is appropriate because, as we learn from a later passage, the poet was prevented from reaching home; if we did not know that, “I went home and condoled ...” would be an equally justifiable translation.

In (1), the poem as a whole allows us to infer that the agent of *gui1* and *yan4*, “return home” and “condole”, must be the poet. But such things are not always clear. In (5), Karlgren chooses to translate on the assumption that the man is promising to the woman poet that he will not kiss and tell (“he swears that he will not ...”); but one could equally well take the wording the other way round (“I swear that I will not ...”). The man might be thought to have more to lose from publicity; elsewhere the poem suggests that the woman is unmarried (there was little or no social stigma attached to fornication at this period). It is routine for such issues to be left vague in Old Chinese.

Even when an overt pronoun occurs, its “thematic role” is frequently debatable. In (4), Karlgren translates the first-person *yu2* as agent of *xing2* “go”, and treats *jia4* “yoke” as having an object understood (“yoke the carriage”). But on the face of it, since the speaker is begging her family elders to marry her off to the man under discussion, one might equally take *yu2* as object of *jia4*: “yoke me (to him) ...”. Perhaps it could be argued that this is too metaphorical a use of *jia4* “yoke” to be plausible in this language; but nothing in the *grammar* of the line forbids that interpretation.

Example (6) reports an entire question-and-answer exchange containing only a single content word, *guan1* “look”. The man’s answer literally says that somebody has done something, but who and what are left wholly inexplicit.

As for “ontological types”: example (4), for instance, is very comparable to examples (9–12) in Gil (2005), being interpretable as asserting the existence of things (“there is wind and violent weather”), occurrence of processes (“it eternally blows and storms”), ascription of a property (“everlasting are the wind and cloudbursts”), etc. Karlgren translates (2) as negating an equation (“they are not those who ...”), but another translator might treat it as negating an action (“it is not that my thoughts are dwelling on them”). We have already seen that, out of context, (1) might be understood as asserting either a single action (“I am returning in order to ...”) or two separate actions (“I returned, and I condoled ...”); a similar remark could be made about (4).

Let me repeat that there is nothing unusual or unrepresentative about the handful of examples I have used to illustrate the structural nature of Old Chinese. This language regularly departs from the Eurocentric model in

very much the same ways that Gil says Riau Indonesian departs from it. What is more, whereas Riau Indonesian is a “basilectal”, colloquial variety of a language which has an “acrolectal” official variety that is more explicit in various respects, there is no reason to think that Old Chinese as recorded in the documents that have come down to us existed alongside a language-variety which was more explicit but went unrecorded. Old Chinese as illustrated here appears to have been the standard language of an early stage of what is now the world’s oldest and most populous civilization.

6. McWhorter’s diagnostics

John McWhorter (1998: 792–799; 2001a: 126 n. 1; 2005: 10–11) gives diagnostics for classifying a language as creole. According to McWhorter, all *and only* languages which emerged historically through the creolization process have all three of the following properties:

- no inflexional affixation
- no lexical or grammatical tone
- no semantically-unpredictable derivation

The first of these properties in particular relates to the distinctive features Gil ascribes to Riau Indonesian, since languages which make thematic roles and ontological types explicit commonly use inflexion for the purpose. It is because McWhorter believes that Riau Indonesian has all three of these diagnostic properties that he wants to classify Malay as a creole.

7. No tone in Old Chinese

A first point is that it seems surprising to find tone in this list with the other two properties. It is easy to see languages as simpler if they lack inflexion and semantically-unpredictable derivation, but it is not obvious (to me at least) that tone makes a language complex. One might regard a language with a lexical contrast between two or three tones, and a very restricted inventory of segmental phonemes, as just as simple as a non-tone language with a richer range of consonants, vowels, and diphthongs. McWhorter is

aware that Chinese has some of the characteristics he ascribes to creoles; it is possible that his diagnostic list includes tone specifically in order to exclude Chinese from the class of creole languages.

If so, this does not succeed for Old Chinese. All modern Chinese dialects are tonal, but it has been recognized for some time that Chinese tones developed within the historical period. Old Chinese was not a tone language. (The tone contrasts reflect earlier contrasts among consonants; for instance, Mandarin tone 1 versus tone 2 reflects an earlier contrast between voiceless and voiced initial consonants, and tone 4 commonly reflects a lost final *-s*.)

8. No inflexion in Old Chinese

Old Chinese wholly lacks inflexional affixation; it is commonly seen as an extreme case of the “isolating” or “analytic” language type. The only serious challenge to this generalization that I am aware of is that it has sometimes been suggested that the language had a contrast between nominative and oblique cases in pronouns, because (taking first-person pronouns as an example) the forms *wu2* and *wo3*, whose Old Chinese pronunciations are reconstructed by Baxter as *nga* and *ngaj?*, are typically found in subject and non-subject contexts respectively. But it is, I think, now generally agreed that this is not a case of true grammatical inflexion, but of phonetically reduced versus full forms, somewhat akin to French *me* versus *moi* (Pulleyblank 1995: 76–77); non-subject pronouns tend to occur in positions where phonetic reduction is less likely, but there are cases where *nga* means “me” rather than “I”.

9. Old Chinese resembles Riau Indonesian with respect to derivation

The one of McWhorter’s creole diagnostics which might exclude Old Chinese is the third, lack of semantically-unpredictable derivation. Most Old Chinese words were simple roots, but it is clear that there was some derivational affixation. For instance, there was a suffix *-s* which in some cases had a verbalizing and in other cases a nominalizing force; examples (quoting reconstructed Old Chinese forms, with modern Mandarin reflexes in brackets) are:

ak “bad”, *aks* “to hate” (*e4*, *wu4*)

lak “to measure”, *laks* “a measure, degree” (*duo2*, *du4*)

And there was a prefix which Baxter reconstructs as [ɦ] (a voiced h) that sometimes converted transitive into intransitive senses:

kens “see”, *ɦkens* “appear” (*jian4*, *xian4*)

Although the full facts about Old Chinese derivational morphology are far from clear, there is little doubt that the meanings of derived forms were sometimes not wholly predictable from the meaning of the roots and the nature of the affix (one might expect a verbalizing suffix added to a root for “bad, wicked” to give the sense “behave wickedly”, or perhaps “make bad, spoil”, rather than “hate”).⁵

If creoles never have any semantically-unpredictable derivation at all, then we must accept that Old Chinese is different in kind from creoles. But McWhorter argues that Riau Indonesian should be regarded as a creole; that seems to imply that he cannot be committed to the strong claim that creoles entirely lack semantically-unpredictable derivation, because Riau Indonesian does appear to have some.

Here I am hampered by having limited knowledge of Malay, and no personal knowledge of Riau dialect. But I have some acquaintance with the former; and Gil’s own examples permit deductions about Riau Indonesian.

General Malay certainly has semantically-unpredictable derivational affixes, such as *ber-*, *ter-*, *-kan*, *-i*. But use of these is often associated with high (“acrolectal”) style rather than the basilectal Riau dialect Gil discusses. In Gil’s own example (Gil 2001: (23–24)) of the difference between how Malay students translate English formally, and how Malay is actually used in everyday life, the acrolectal translation of English “That match was great!” involves Malay *permainan*, in which affixes *per-* and *-an* are added to the root *main* “play”, while the basilectal version uses the unaffixed root.

Nevertheless, even basilectal Riau Indonesian does seem to contain some semantically-unpredictable derivation. For instance, Malay has the words *gigi* “tooth” and *gigit* “bite”: both forms evidently occur in Riau

⁵ One Sinologist, Laurent Sagart (1999), believes that the derivational morphology of Old Chinese was far richer than suggested here, holding that any consonant cluster is evidence of derivational complexity. However, I am not convinced by Sagart’s arguments. He urges that the claim that roots contained only single consonants is a strong hypothesis which in terms of Popperian scientific method should be believed unless it can be refuted, but this seems unsound: it is not clear what kinds of evidence would be accepted by Sagart as a definitive refutation.

Indonesian (see Gil 1994: (4), (14)). And again, Malay has a word *mainan* “toy” derived from the root *main* “play” just mentioned; we have seen that Riau Indonesian uses *main*, and elsewhere (1994: (25)) Gil quotes a use of *mainan*. I do not know enough about Malay to say what the broader derivation patterns are which these individual cases exemplify (I do not know, for instance, whether there are other cases where adding *-t* to a word for a body-part gives a word for an action characteristic of that body-part); but it seems undeniable that these are cases of derivational morphology (they are treated as such by an online dictionary I have consulted), and it would be hard to deny also that derivational processes which use affixation to give “bite” from “tooth”, or “toy” from “play”, must contain an element of semantic unpredictability. Actions performed by the teeth include chewing and gnashing as well as biting; in colloquial English, teeth are sometimes called “gnashers”, but “biters” always refers colloquially to animals or misbehaving children, not specifically to teeth.

On Gil’s evidence, Riau Indonesian does contain *some* semantically-unpredictable derivation. So, if McWhorter classifies this language as a creole, he must accept that that classification is compatible with a limited amount of such derivation. In a corpus of Old Chinese similar in size to the small collection of Riau Indonesian examples I have seen in Gil’s writings, it is very doubtful whether there would be more evidence than this of semantically-unpredictable derivation in Chinese. Impressionistically, I would guess that there might probably be even less evidence, and there could easily be none at all. The case for seeing Old Chinese as typologically creole-like seems fully as strong as the case for seeing Riau Indonesian as creole-like.

10. An accident of history

There is a further point here. Even if my treatment has understated the incidence of semantically-unpredictable derivational morphology in Old Chinese, a time came in the history of Chinese when that morphology was no longer visible; in Old Chinese the derivation of *fikens* “appear” from *kens* “see” was transparent, but in modern Mandarin the historical relationship between their reflexes *xian4* and *jian4* is entirely opaque. (It is not regularly the case that *x-* words are related to *j-* words.) By some point after the Old Chinese period, *xian4* was for all practical purposes just another simple root word.

It would be very difficult to say just when that change was complete, and whether it happened before or after Chinese began to develop tones; but these two processes seem quite unconnected with one another. McWhorter seems to require that tones must have arisen before relationships of the *kens/fikens* type became opaque—because, if it happened the other way round, the intervening language-stage would have all three properties which McWhorter claims to co-occur only in creoles. But even if the Chinese language developments did as a matter of fact occur in the order McWhorter requires, as they may well have, it is very hard to see why they *had* to occur in that order. Surely there can be nothing in the nature of human language which would have prevented a language like Old Chinese losing transparent relationships of the *kens/fikens* type and only afterwards going on to evolve tones—since there is no apparent way in which one of these features impinges on the other. So even if Chinese never did fully exemplify all three of McWhorter's diagnostics, another non-creole language could do so.

11. Conclusion

I conclude that structural simplicity does not imply a history of creolization. Creoles may all be simple, but not all simple languages are creoles.

References

- Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*. Trends in linguistics; Studies and Monographs 64. Berlin & New York, NY: Mouton de Gruyter.
- Deutscher, Guy (2000) *Syntactic Change in Akkadian: The evolution of sentential complementation*. Oxford: Oxford University Press.
- Gil, David (1994) The structure of Riau Indonesian. *Nordic Journal of Linguistics* 17: 179–200.
- (2001) Escaping Eurocentrism: Fieldwork as a process of unlearning. In Paul Newman and Martha Ratliff (eds.) *Linguistic Fieldwork*, pp. 102–132. Cambridge: Cambridge University Press.
- (2005) Word order without syntactic categories: How Riau Indonesian does it. In Andrew Carnie, Heidi Harley & Sheila Ann Dooley (eds.) *Verb First: On the syntax of verb-initial languages*. *Linguistik Aktuell/Linguistics Today* 73. Amsterdam & Philadelphia, PA: John Benjamins.
- Karlgren, Bernhard (1950) *The Book of Odes: Chinese text, transcription and translation*. Stockholm: Museum of Far Eastern Antiquities.

- (1957) *Grammata Serica Recensa*. *Bulletin of the Museum of Far Eastern Antiquities* 29: 1–332.
- Lewis, M. Blanche (1947) *Teach Yourself Malay*. London: English Universities Press.
- McWhorter, John H. (1998) Identifying the creole prototype: Vindicating a typological class. *Language* 74: 788–818.
- (2001a) The world's simplest grammars are creole grammars. *Linguistic Typology* 6: 125–166; a version reprinted in McWhorter (2005), chapter 2.
- (2001b) What people ask David Gil and why: Rejoinder to the replies. *Linguistic Typology* 6: 388–412; reprinted in McWhorter (2005), pp. 68–71.
- (2005) *Defining Creole*. Oxford: Oxford University Press.
- Norman, Jerry (1988) *Chinese*. Cambridge: Cambridge University Press.
- Pulleyblank, Edwin G. (1995) *Outline of Classical Chinese Grammar*. Vancouver: University of British Columbia Press.
- Sagart, Laurent (1999) *The Roots of Old Chinese*. Current Issues in Linguistic Theory 184. Amsterdam & Philadelphia, PA: John Benjamins.
- Thurgood, Graham & Randy J. LaPolla (eds.) (2003) *The Sino-Tibetan Languages*. London & New York, NY: Routledge.

Contact information:

Geoffrey Sampson
Department of Informatics
Sussex University
Falmer
Brighton, England BN1 9QH
www.grsampson.net