

# Chinese language sources for Chinese Pidgin English: What we know and what we need to know<sup>1</sup>

Geoff P. Smith  
University of Hong Kong

## Abstract

The existence of the contact language Chinese Pidgin English (CPE) has been copiously documented, but its exact nature is obscured by a number of factors. The time depth since its inception and the fact that the language is no longer spoken make access to reliable data difficult. Evaluation of such data that exist is made more difficult by the spurious nature of some sources. Even those European observers who were not ridiculing the language or inventing passages may have recorded what they heard through a filter which tended to hear forms in terms of their relation to standard English. A valuable source of comparison is provided by the writings of Chinese-speaking observers during the time when Chinese Pidgin English was spoken. There are two main sources of information: *The Common Language of the Red-haired Foreigners*, a short CPE phrasebook printed around 1850 in Guangzhou, and Tang Tingshu's six-volume *Chinese-English Instructor*. Some descriptive work on these works has been carried out, but a great deal of analysis remains to be done.

## What is Chinese Pidgin English?

Chinese Pidgin English (CPE) is rather special in the history of Pidgin and Creole languages. It was one of the most widely discussed of such contact varieties during its heyday, and indeed gave the name 'pidgin' to this type of language. Its origins are in the intense and lucrative trading contacts between Chinese and Western merchants beginning around the sixteenth century, and it was the commonest medium of inter-ethnic communication along the South China Coast from about the beginning of the 19th century to the early decades of the 20th century. More complete accounts of the formation of the pidgin can be found in Baker and Mühlhäusler (1990) and Bolton (2003).

However, many questions remain about the exact nature of the language. It appears that a widespread maritime variety of simplified or pidginised English may have contributed to it, as well as features of the two main

---

<sup>1</sup> This paper is a slightly modified version of a presentation at the Society for Pidgin and Creole Linguistics conference in Curaçao, August 2004.

languages in contact, English and Cantonese. It also seems likely that an earlier pidginised variety of Portuguese may have been in use in the area when Portugal dominated maritime trade from the 15th Century onwards, and this may have interacted with the incipient CPE to some extent. Links with other pidgin languages such as those arising in the south-west and central Pacific (Baker 1987, Franklin 1979, Siegel 1990) and Hawaii appear to be fairly tenuous, but have not been exhaustively explored. Some other pidgin languages emerged from multiple roots, but CPE involved only two major contributors - Cantonese and English. As a result, there was no useful role for the language once a critical mass of bilinguals emerged, unlike the situation, for example, in Melanesia, where the English-based pidgin continued to serve as a valuable medium of communication in a situation of extreme multilingualism when communication with English-speakers was no longer the main function of the language.

The major problem facing researchers attempting a comprehensive investigation of CPE is that the language is now no longer spoken, virtually no tape-recorded data of the language in use exist and evidence has to be gleaned from written historical records. One or two speakers were reported to remember the language to varying degrees until the middle or even later decades of the 20th century, but CPE effectively ceased to be a useful medium of communication in the mid-20th century as greater knowledge of English became established in the region. The main records contributing to our knowledge of CPE until recently have been accounts by European travellers, missionaries and other visitors to the Far East of interactions between speakers. More recently some Chinese language sources have come to light, and these offer a valuable point of comparison to those English-language records.

### **English Language Sources**

Numerous travellers to the China Coast in the nineteenth century commented on the language commonly spoken in trade and master-servant relationships in the area, for example, William Hunter's *The Fan Kwaie at Canton before the Treaty Days, 1825-1844* which was originally published in 1882. Since the context of use was initially trade or business, the name business English became established, and "pidgin" appears to be derived from this. The word *pidgin* for 'business' is attested as early as 1807 (Baker and Mühlhäusler, 1990:93) although Bolton (2002:4) reports that it is not until 1859 that the collocation "Pidgin English" is first found. Pidgin, variously spelled as pigeon etc. in other documents, became the name not only for this language variety, but as a generic name for other contact varieties, usually considered according to the prevailing opinion as a simplified and corrupted or

improperly acquired form of English. As with other pidgin languages, observers often commented with various degrees of negativity or even downright hostility on the variety, and such terms as “vile,” “absurd,” “bastard” and “infantile” were typically included in the descriptive vocabulary.

It is against this background that instances of language in use were recorded. Many English-speaking observers would thus have a natural bias to interpret utterances as deviant forms of English, and even those who became fluent in CPE may not have been trained linguists, and may have recorded samples in idiosyncratic spelling based on that of Standard English. The best-known work on CPE is probably Charles Leland’s *Pidgin-English Sing-song*, which consists of many poems and stories in the pidgin, together with some vocabulary notes. However, it emerges that Leland had never actually visited the Far East and wrote his works, which were intended to amuse and entertain, in London. While he does appear to have used many authentic words and structures from CPE, items from different stages of development are mixed up (Baker and Mühlhäusler 1990:88) and structures from other known pidgin languages such as the transitivising suffix *-im* or *-um* appear to have been inappropriately included. The role of his own imagination has also to be factored in. Academic study of CPE was pioneered by the linguist R. A. Hall who courageously insisted on treating Pidgin and Creole languages as serious objects of study, in direct opposition to the received academic wisdom of the time, and his grammar and texts (1944) are a landmark in the study of the language. However, some of Hall’s sources, including Leland, are problematic in that they can not be considered genuine texts from one unique period of the development of the language.

### **Chinese Language Sources**

Some Chinese language documents giving details of CPE provide an interesting and important comparison with the English-language sources. Firstly, some short phrase books of CPE were produced in the region around the beginning of the 19th century (Selby and Selby 1995). Traders and others in contact with overseas merchants had a great deal to gain by some familiarity with the language, and records indicate that these learning booklets were in common use among the Chinese population in trading contact with foreigners at that time. One of these booklets is described below. The other source is a more serious academic study of English by a Chinese writer Tong Ting-ku (唐景星). Although the major focus of this work was Standard English, it contains some records of CPE versions of some of the English forms cited. These two sources are described in greater detail below.

***Hùhng mòuh tùng yuhng fàan wá* (紅毛通用番話)**

These booklets were apparently common around the “factories” of the Pearl River delta in the middle of the 19th century. 紅毛通用番話 (*hùhng mòuh tùng yuhng fàan wá*) literally means ‘the language of the red-haired foreigners,’ i.e. Europeans. An earlier version had a slightly less complimentary title with 鬼 *gwái* ‘ghost, devil’ in place of 番 *fàan* ‘foreign.’ It appears that several versions of the booklet appeared, and that similar booklets giving guidance in speaking basic Macau Portuguese were produced as early as the 1750s (Baker 1989:3). The “Red Hair” booklet held in the British Museum and dated around 1850 is the source of the examples given here. It has 16 pages of words or phrases dealing with numerals, occupational vocabulary, etc. For each entry, the term is given in Chinese, and beneath it and slightly offset to the right, the CPE pronunciation is given as closely as possible in Chinese characters, as the following examples show:

水	老
手 (séui sáu = sailor)	婆 (lóuh pòh = wife)
些	威
利	父 (wài fuh i.e. wife)
文 (sè leih mahn i.e. sailorman)	

Only the Chinese characters, of course, appear in the original. In the above examples, the Cantonese pronunciation according to the Yale system and a gloss have been added for clarification. The booklet contained a total of nearly 400 entries, all single words or formulaic expressions. Thus, while this is a valuable source for lexical and phonological issues, it says little about the grammar of the language.

As noted by Shi (1993), learning a pidgin language through written characters is highly unusual in the context of the way Pidgin and Creole languages are normally acquired. A number of analyses of the “Red Hair” booklet have appeared. Shi (1993) gives some useful phonological correspondences, while a more thorough analysis of the content was made in Baker (1989) with a detailed translation and suggestions for the etymology of various entries as well as further discussion of phonological equivalents. The Yale version of the pronouncing characters and target English or pidgin forms for the entire 372 entries are listed as an appendix to Bolton (2003) and he produces some further suggestions regarding the source of certain items in a set of footnotes, including a few possible derivations from Swedish.

### *The Chinese-English Instructor* (英語集全 *yìng yǔ jí quán*)

A much more substantial work was produced, probably in 1862 (Selby and Selby, 1995:123) by Tang Tingshu, also confusingly known as Tong Tingku or Tong King-sing (唐景星). Tang appears to have been an accomplished English speaker with a linguistic insight far ahead of his time. He produced a monumental six-volume work, the *Chinese-English Instructor* (英語集全 *yìng yǔ jí quán*), hereafter CEI, which attempted to render Standard English comprehensible to Chinese speakers. In addition to single words and phrases, more extended sentences and chunks of dialogue are featured. Each of the six volumes deals with a different subject area. In addition to its undoubted linguistic value, the snatches of dialogue also serve as an interesting insight into the everyday life of traders in both camps during the period. The pronunciation of the English forms is indicated not only by the nearest equivalent Chinese character, but also by a series of diacritics, which appear to use conventions developed in describing minority languages within China (Selby and Selby, 1995:125). The most valuable parts for scholars of CPE appear to be the marginalia recorded especially in volume six, which give the CPE equivalent of some of the English expression.

The CEI has been mentioned in a number of works, but no comprehensive analysis has appeared. The most detailed discussions to date are Selby and Selby's (1995) description of the general nature of the work with some examples and Bolton's (2003) general description of Tang's life and work. Further analysis is currently being undertaken by Bolton and a number of other researchers such as Martino, following up on her previous (2003) work on the origin and development of CPE.

### **Future Research Directions**

As we have seen, a certain amount of descriptive work has been carried out on these two sources, but a great deal remains to be done. Shi (1993), Baker (1989) and Bolton (2003) have produced a reasonably comprehensive analysis of the "Red Hair" booklet, but there remain problems with some of the characters used. Since the booklet was hand-written, some of the characters are rather hard to read and a certain amount of guesswork is used in identifying the target character in some instances. There is also the possibility of errors in the form characters of similar shape being erroneously used by the writer. Bolton, for example, indicates (2003:277) that the 步 *bouh* in expression 孛步 *mà bouh* appears to be an error, and if the similarly shaped 沙 *sà* is substituted, the entry makes much more sense as the CPE *massa*, master. However, it is difficult to be certain about errors while there remain

uncertainties about changes which may have taken place in the language over the past 150 years. This is seen in the use of various archaic characters, and the nature of written Chinese in the Cantonese-speaking areas during the Qing Dynasty needs to be considered. A number of characters used do not show up in modern dictionaries, and there is also the possibility of sound variation and change to consider. This is illustrated in the case of the character 米 *mai* ‘uncooked rice’ which appears to have been pronounced as “mi” in the 19th century Cantonese used in Macau (Baker, 1987:20). If the writer of the “Red Hair” pamphlet was influenced by Macau pronunciation, and represented the CPE “me” with the character 米, this could have led to the widespread use of the possessive “my” for the first person singular pronoun in the CPE of other areas. The use of “my” for “me” is highly unusual in Pidgin and Creoles worldwide, and it would also be exceptional for pidgin pronunciation to be influenced by written language.

The position of a number of characters incorporating the 口 *hau* ‘mouth’ radical is also worth examining. There are many regular characters in Standard Written Chinese with this radical, for example 喊 *haam* ‘scream,’ where the mouth radical indicates some connection to vocalisation. However, adding the radical to an existing character can also be a device for creating new characters for spoken forms, as is commonly done with those Cantonese vernacular words which are not part of Standard Written Chinese. Thus to indicate the character *deih*, the Cantonese pluralising particle for pronouns, the character 地 *deih* ‘ground’ is modified with a ‘mouth’ radical to create a novel character whose pronunciation is indicated by the right hand side. In this case, the pronunciation is identical to the base character, but in other cases, the similarity may only be approximate, for example, the standard character 既 *gei* ‘already’ is used as the basis for the Cantonese possessive, which has a slightly different pronunciation *ge*. As noted by Selby and Selby (1995:124), ersatz pronouncing characters in these early Chinese sources are commonly created by adding this radical to the left of a standard character, and it can be seen that there are many in use in both the “Red Hair” booklet and the CPE entries in the CEI. It is not entirely clear why characters of this type are used when standard characters are available. It could be merely to reinforce the fact that the character is used for pronunciation and not meaning, or it could indicate some modification of the sound.

One puzzling aspect of these two sources is the large number of characters used to represent relatively few sounds. It might be expected that once the principle of representing English or CPE in Chinese characters had been established a standard syllabary would soon be established. This is especially so as the practice of representing English sounds by Chinese

characters has continued to the present, in names of people and places, as well as instructional sections in local almanacs. Smith (2004) discusses some of the problems involved with reference to the representation of the names of sporting teams and players. However, there is little evidence of a standard syllabary emerging as far as these two CPE sources are concerned. No full analysis has been made, but my own preliminary investigation shows the following in use in the two sources. Ignoring some problematic characters, a total of 574 characters were identified as representing the various English sounds. In CEI there were 402 characters used, while in the “Red Hair” booklet the total was 266. However, only 115, or 20% of the total are common to the two works, a surprisingly small percentage, which seems to indicate that the origins are quite different, and that representations were *ad hoc* rather than following a standard syllabary.

For the CEI, a full comprehensive analysis of the CPE entries remains to be carried out, and this is the most immediate need. The 5000 or so characters of the marginalia in Volume 6 alone represent a significant source of CPE in the form of relatively complete utterances, so a grammatical analysis of these would be illuminating. Most of the problems outlined above with the “Red Hair” booklet also apply to the CEI and require investigation.

## References

- Anon. (around 1850). 紅毛通用番話 (hùhng mòuh tùng yuhng fàan wah - The Language of the Red-haired Foreigners). Guangzhou
- Baker, P. (1987). “Historical developments in Chinese Pidgin English and the nature of the relationships between the various pidgin Englishes of the Pacific region”, *Journal of Pidgin and Creole Linguistics*, 2(2):163-207.
- Baker, P. (1989). A provisional assessment of huŋ mow tuŋ yʌn fan wa (a vocabulary of the common words of the language of the red-haired foreigners), a 19th century Pidgin English phrasebook written entirely in Chinese characters. Unpublished research report.
- Baker, P. and P. Mühlhäusler. (1990). “From business to pidgin”, *Journal of Asian Pacific Communication*, 1(1):87-115.
- Bolton, K. (2002). Hong Kong English: Autonomy and creativity, in K. Bolton (Ed.) *Hong Kong English: Autonomy and Creativity*, pp.1-25. Hong Kong: Hong Kong University Press, 2002
- Bolton, K. (2003). *Chinese Englishes: A Sociolinguistic History*. Cambridge: Cambridge University Press.

- DeFrancis, J. (1984). *The Chinese Language: Fact and Fantasy*. Honolulu, University of Hawaii Press.
- Franklin, K. J. (1979). "Some comparisons between Chinese Pidgin English and Melanesian Pidgin English", *Talanya*, 6:
- Hall, R.A. Jr., (1944). Chinese Pidgin English grammar and texts, *Journal of the American Oriental Society* 64:95-113.
- Hunter, William C. (1965) [1882]. *The Fan Kwae at Canton before the Treaty Days, 1825-1844*. London: Kegan Paul, Trench & Co.
- Leland, C. G. (1897). *Pidgin-English sing-song, or songs and stories in the China-English: with a vocabulary*. London, Kegan Paul, Trench, Trubner and Co. (Fourth edition).
- Martino, E. (2003). Chinese Pidgin English: Genesi ed evoluzione. MA dissertation, Università degli Studi di Roma: La Sapienza.
- Matthews, S. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Selby, Stephen and Anne (1995). China Coast Pidgin English. *Journal of the Hong Kong branch of the Royal Asiatic Society*, 35:113-41.
- Shi, D. (1993). "Learning pidgin English through Chinese characters", in Byrne, F. and J. Holm (eds.), *Atlantic meets Pacific: A Global View of Pidginization and Creolization*. pp.459-465. Amsterdam, Benjamins..
- Siegel, J. (1990). "Pidgin English in Nauru", *Journal of Pidgin and Creole Linguistics*, 5(2):157-186.
- Smith, G. (2004). A Cantonese Syllabary for English Soccer, in M. Baynham, A. Deignan and G. White (Eds.) *Applied Linguistics at the Interface. British Studies in Applied Linguistics 19*. London: Equinox
- Tang, Tingshu. (1862). Ying yu ji quan (The Chinese and English instructor.) Guangzhou. (唐景星: Also known as Tong Ting-ku or Tong King-sing )