



**HAL**  
open science

## Gan, Hakka and the formation of Chinese dialects

Laurent Sagart

► **To cite this version:**

Laurent Sagart. Gan, Hakka and the formation of Chinese dialects. Dah-an Ho. Dialect Variations in Chinese, Academia Sinica, pp.129-154, 2002. ⟨halshs-00095941⟩

**HAL Id: halshs-00095941**

**<https://shs.hal.science/halshs-00095941v1>**

Submitted on 18 Sep 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## **Gan, Hakka and the formation of Chinese dialects<sup>1</sup>**

L. Sagart

Centre de Recherches Linguistiques sur l'Asie Orientale

Ecole des Hautes Etudes en Sciences Sociales (UMR 8563 du CNRS),

Paris, France

**Abstract** : the author argues that Hakka and southern Gan are sister dialects, as they share several innovations not found elsewhere; that they arose out of the Chinese dialect spoken in central Jiangxi in Song times, a stratified dialect which included a non-Chinese substratum, probably Miao-Yao; an archaic layer; and a more recent layer with an important Late Middle Chinese component. It is claimed that the linguistic boundary between southern Gan and Hakka arose secondarily due to the effect of an old administrative and geographical boundary. It is also argued that Hakka devoicing took place in the south, when Hakka was in contact with the Miao-Yao language She, and that the old dialect of the city of Ganzhou may have played an important role in the formation of Hakka.

---

<sup>1</sup> This paper incorporates elements of an unpublished conference paper presented at ICCL-6 (Leiden University, 19-21 June, 1997) and entitled “preservation and innovation in Gan, and the relationship of Gan and Hakka”.

The relationship of Gan and Hakka is one of the most hotly disputed questions in Chinese dialectology. Two important issues are: (a) do Hakka and Gan subgroup together, or does Hakka subgroup with Yue and Min ? and (b) did Hakka and/or Gan arise in the south, or are they northern dialects transplanted into the south ?

The idea that Hakka is a northern dialect transplanted into the south has its roots in Lo Hsiang-lin's 1933 book *Kejia Yanjiu Daolun*. Lo's intention was to disprove the notion, then prevalent among Cantonese intellectual circles, that the Hakkas were a non-Chinese people. Based on a survey of Hakka clan genealogies, Lo claimed that the ancestors of the Hakkas had migrated in several waves from Northern China to their present area of concentration. He pointed out areas of agreement in phonology and lexicon between Hakka and early forms of Chinese. To him, this appeared to support the idea of a northern, and ancient, origin of the Hakkas. Lo's theory soon became influential among students of Chinese dialects: it was adopted in important works, most notably by Luo Changpei (1940), Yuan Jiahua et al. (1962), Hashimoto (1973: 3; 436) and Lu Guoyao (1988).

The view that Hakka and Gan 'subgroup' together, or, in lay speech, that they go back to a common ancestral Chinese dialect which is unique to them, has its origins in the dialect survey of 1936, when it was found that Hakka and Gan agree in treating the old voiced stops and affricates of Middle Chinese as voiceless aspirates<sup>2</sup> (I call this process 'aspirated devoicing', for short). Li Fang-kuei (1938) made the treatment of these Middle Chinese initials the cornerstone of his classification of Chinese dialects, and naturally, in his scheme, Gan and Hakka formed one dialect group. Shortly

---

<sup>2</sup> The idea was foreshadowed by E. H. Parker in 1884, see Sagart (1988).

afterwards Luo Changpei, under the influence of Lo Hsiang-lin's book, investigated Linchuan, a Gan dialect in central-east Jiangxi. He cited a number of phonological features<sup>3</sup> shared by Linchuan and Meixian Hakka, and concluded that "part of the inhabitants of Jiangxi are the Hakkas who have not migrated". He added: "these two dialects probably represent two divisions of one group in the larger groupings in the Chinese dialects" (Luo 1940 [1958]: 240). More recently my teacher Prof. Lu Guoyao (1988) has claimed that the Tong-Tai 通泰 dialects of northern Jiangsu share a number of features with Gan and Hakka (including aspirated devoicing) and that all three dialects must be descended from northern Chinese of the 4<sup>th</sup> century. Li Rulong and Xin Shibiao (1999) have extended Lu Guoyao's proposal to *all* those Chinese dialects with aspirated devoicing (Gan, Hakka, Tong-Tai, Shanxi-Shaanxi), claiming that they go back to one northern dialect spoken more than 1500 years ago.

J. Norman (1986; 1988) has maintained a very different theory: that Hakka is a southern dialect, not closely related to Gan and even less so to any northern dialect. Claiming that the shared characteristics of Gan and Hakka listed by Luo Changpei are superficial (Norman 1988: 222), he pointed out a completely different set of mostly conservative features shared by Hakka, Min and Yue<sup>4</sup>. In his view, Hakka, Min and Yue, are descended from a common historical source, present in south China since Han and Sanguo times (first to third centuries CE), which he calls 'Old southern Chinese'.

---

<sup>3</sup> "Development of ancient voiced stops and affricates into voiceless aspirates; development of ancient  $\gamma$ - and  $x$ - into  $f$ - in  $he_2 kou_3$ ; preservation of ancient  $-m$  and  $-p$ ; distinction between 1<sup>st</sup> and 2<sup>nd</sup> division of rhyme groups 咸, 山 and 蟹; development of rimes 魚 and 虞 into  $-i$  after ancient  $k$ - and  $ts$ - series; development of rime 侯 into  $-\epsilon:u$ ; colloquial pronunciation of group 梗 as  $-aŋ$  or  $-iaŋ$ ; etc."

<sup>4</sup> Retention of bilabials where other dialects have changed them to labiodentals; strong resistance to palatalization of velars, treatment of the old Chinese rhyme group  $ge_1$  as  $-ai$  or  $-oi$ ; also several lexical features: a special term for the verb 'to poison', derived from the corresponding noun by tone derivation; a word for 'cockroach' derived from a form reconstructible as  $*dzat$ ; similar terms for 'leech' and 'louse'.

Although the two ideas of a northern origin of Hakka and a common origin for Gan-Hakka are historically linked, there is no necessary link between them. For some years (Sagart 1988, 1993) I have been arguing that Hakka and southern Gan are closely related dialects which evolved in the south, out of a common ancestor (Proto-Southern-Gan) spoken in central Jiangxi in Song times by a population in part recently migrated from northern Jiangxi and from northern China and speaking a southern variety of Late Middle Chinese influenced by northern Gan. In support of my view, I observed that southern Gan and Hakka are to some extent mutually intelligible. I discussed the demography of the settling of Jiangxi by Chinese speakers; I also pointed out some inconsistencies in Lo's theory, arguing that clan genealogies are not always entirely reliable.

In the past ten years or so many new and interesting developments in Gan and Hakka dialect studies have taken place. A large body of linguistic data has been published (Li and Zhang 1992); also, S.T. Leong's posthumous book on Hakka migrations in the period 1550-1850 and the formation of Hakka ethnicity has considerably clarified the recent history of the Hakkas (Leong 1997). Notice must also be taken of recent and significant progress in Hmong-Mien historical linguistics (Wang and Mao 1995). In this paper I would like to discuss the contribution of these recent developments to an old debate.

## A Hmong-Mien substratum in Hakka ?

It has long been suspected that a Miao-Yao (a.k.a. Hmong-Mien) substratum could be found in Hakka. A.-G. Haudricourt once suggested that Cantonese and Hakka might be the results of the evolution of Chinese on a Tai and Miao-Yao substratum, respectively. Deng Xiaohua (Deng 1999)

presents a list of words shared by Hakka and Hmong-Mien languages, including She 𪗇 . At least one of these words gives substance to the hypothesis of a Hmong-Mien substratum in Hakka: Proto-Hmong-Mien \*ntəŋ<sub>C</sub> ‘to wear on the head’ (reconstruction by Wang and Mao 1995), She təŋ<sub>5</sub>. That word is definitely Hmong-Mien in origin, it cannot be a Chinese loan to She. Yet it is also found in Hakka: Meixian tuŋ<sub>1</sub>. One can imagine how a She-speaking population becoming bilingual in Chinese would maintain a She word for the act of wearing a head garment, if this garment were typically and emblematically She<sup>5</sup>. One can also picture how this word might have been introduced in the target Chinese dialect once that population became monolingual in Chinese. This is not to say that the Hakkas as a whole are descendants of She speakers who shifted to Chinese: rather, the ancestral Hakka population was presumably formed of Chinese-speaking immigrants who were joined by She language shifters. This scenario is a classical one: one is reminded of the wholesale shift of Gaulish speakers to Latin in France, in the first part of the first millennium CE, through a stage of Gaulish-Latin bilingualism. Gaulish speakers shifting to Latin following the military conquest by Julius Caesar in 57 BCE joined an original population of Latin speakers who were settlers from Italy. The modern population of French speakers arose out of the merger of these two communities, immigrants from Italy speaking Latin and Gaulish language shifters, with later influx from other sources, notably Germanic. Solid evidence of a Hmong-Mien substratum in Gan still eludes us, but the dissymmetry of treatment between –k and the other final stops, –p and –t, in most varieties of Gan, is reminiscent of Hmongic<sup>6</sup>.

---

<sup>5</sup> She women in Fujian still wear a pretty red cap, which is the most obvious outward sign of their national identity.

<sup>6</sup> Hmongic assigns Chinese loanwords ending in –p and –t to its entering-tone (tone D), but Chinese loanwords ending in –k to its Departing-tone (tone C).

## ‘Archaic’ features of Gan

Following the conquests of Qin Shi Huang and Han Wu Di in the 3<sup>rd</sup> and 2<sup>nd</sup> centuries CE, Chinese-speaking populations penetrated Jiangxi from the north, settling in the plains around Poyang lake and along the Gan River valley. While these populations (as reflected in a series of censuses beginning in 2 CE) were numerically weak compared with Chinese speakers in the rest of China (including the lower Yangzi Valley), two important cities were established: in the north, Yuzhang 豫章 near the site of the modern Nanchang ; in the south, Nanye 南野 near the modern city of Nankang 南康, the latter with a garrison protecting the vital Dayu pass leading across the Dayu range to the Guangdong area. There were smaller settlements between these two, along the Gan River valley. Demography appears to have peaked in the late 2<sup>nd</sup> century CE, before undergoing a catastrophic decrease in the course of the 3<sup>rd</sup> century CE, followed by stagnation until the beginning of the eighth century (Sagart 1988:148ff; 1993 :11). The archaic features of Hakka cited by Norman in support of a link with Min and Yue are relics of that period. This is an important observation which must be integrated into our understanding of the stratification of Hakka. Now, although this is not well-known, such conservative features can be found in Gan too. To be sure, they are less common there than in Hakka, the obvious reason being the greater influence of northern Chinese on Gan. I give some details below.

## Gan exceptions to labiodentalization

There were in Old Chinese and in Early Middle Chinese words with bilabial stop initials (p-, ph- and b-). In Mandarin the evolution of these initials is conditioned by the following vowel. Preceding a certain set of vowels, they change to f (they ‘labiodentalize’); preceding all other vowels, they remain bilabial stops (p- and ph-). It is well-known that in the colloquial layer of Min, labiodentalization has not taken place. That is, words which in Mandarin have initial f- have a bilabial,

p- or ph-, in Min. It is known that Hakka and Yue also show examples of this kind. The same is true of some dialects of southern Hunan. In Gan, these examples may be cited:

- **to float.** 浮 Early Middle Chinese bjuw ‘to float’ results in Mandarin *fu*<sub>2</sub>. This is a colloquial word which often shows a bilabial initial in the south (Hakka, Min, Cantonese, Nanling, Changsha, Shuangfeng, and even in Wuhan Mandarin). In Gan (data from Li and Zhang 1992: 73): Yongxin 永新 p<sup>h</sup>ɒ<sub>2</sub> (col.); Jishui 吉水 p<sup>h</sup>au<sub>2</sub>; Liling 醴陵 p<sup>h</sup>au<sub>2</sub> (col.); Yifeng 宜豐 p<sup>h</sup>au<sub>2</sub> (col.); Anyi 安義 p<sup>h</sup>au<sub>2</sub> (col.). The *Hanyu Fangyin Zihui* adds Nanchang p<sup>h</sup>au<sub>2</sub> (col.). For all these locations except Jishui, a literary reading with f- is also given.
- **to divide.** In Yiyang 弋陽 the word *fen*<sub>1</sub> 分 Early Middle Chinese pjun ‘to divide’ has a literary pronunciation fɛn<sub>1</sub> and a colloquial pronunciation pɛn<sub>1</sub> (Li and Zhang 1992).
- **to help, to support.** In Yongxin (Liu Yongguang et al. 1992: 712) the word *fu*<sub>3</sub> 輔 Early Middle Chinese bjuX ‘to help, to support’ has a lit. reading fu and a coll. reading p’u (in the compound 輔導 ‘give guidance in study or training’); tones are not given. Bilabial readings for this character are extremely common in Gan (see for instance Li and Zhang 1992: 34), but details are seldom given.
- **hibiscus.** In Yongxin (Liu Yongguang et al. 1992: 712) the colloquial name of the hibiscus flower 芙蓉花 has a bilabial reading p’u<sub>2</sub> for the first character *fu*<sub>2</sub> 芙 Early Middle Chinese bju. The corresponding literary reading is fu (no tone given).

Other examples occur here and there, but are given as mere character readings, no details being given on actual usage. One instance is 蓬, Early Middle Chinese bjuwng, Nanfeng 南豐 p<sup>h</sup>uŋ<sub>2B</sub> (Oshima 1995: 125).

## the Middle Chinese retroflex stops in south-western Gan

In my 1993 book (p. 250) I noted the existence of a few dialects in SW Jiangxi like Ji'an 吉安 where the Middle Chinese retroflex stop initials tr-, trh-, dr-, but not the Middle Chinese palatal affricates tsy-, tsyh-, dzy-, are sometimes reflected as alveolar stops in colloquial words, as in Min and Hakka. The same is true of Nanwen Cun 南汶村, a Gan dialect spoken in Yongxin 永新 district, halfway between the towns of Yongxin and Lianhua 蓮花 in the Wugong Shan 武功山 area of SW Jiangxi. In 1988-1989 I investigated this dialect in the speech of Mrs Li Lunjin 李綸錦, aged 88-89, a native of Nanwen Cun who was then living in Hsinchu 新竹, Taiwan. Her pronunciation is representative of this dialect in the final years of the Qing dynasty. [LS1] Examples:

- 豬 ty<sub>1</sub> 'pig'
- 竹 tjo<sub>1</sub> 'bamboo', as in 竹竿 tjo<sub>1</sub> kō<sub>1</sub> 'bamboo pole' (but tʃo<sub>1</sub> in 竹筍 'edible bamboo shoots')
- 帳 tjã<sub>5</sub> '(mosquito) net', (literary reading tʃã<sub>5</sub>);
- 長 tjã<sub>3</sub> 'to grow', as in 長肉 tjã<sub>3</sub> jo<sub>1</sub> 'to grow fat'
- 著 tjo<sub>1</sub> as in 著衫 tjo<sub>1</sub> sã<sub>1</sub> 'to wear clothes'

I have found no cases of words with Early Middle Chinese palatal *affricates* (i.e. the Zhang 章 initials) having t- and th- reflexes in Nanwen Cun. These initials are always represented by affricates (tʃ/tç, tʃh/tçh). The number of examples in my data is as follows:

	Nanwen Cun affricate reflexes: tʃ-, tʃh-, tç-, tçh-	Nanwen Cun stop reflexes: t- and th-
EMC retroflex stops (知 etc.)	27	5
EMC palatal sibilants (章 etc.)	35	0

**Table 1: modern reflexes of the Early Middle Chinese retroflex stops and palatal affricates in Nanwen Cun, Jiangxi.**

The unsophisticated Chi<sup>2</sup> test (Yates corrected) yields a statistically significant difference ( $p < 0.04$ ) between the figures for the Early Middle Chinese retroflex stop and palatal sibilant series. We must conclude that there exists a layer in Nanwen Cun in which the two series have different reflexes, t-, th- for the Early Middle Chinese retroflex stops, and tʃ-, tʃh-, tɕ-, tɕh- for the Early Middle Chinese palatal sibilants.

The same situation is observed in Anren 安仁, a dialect of the Gan-Xiang transition in SE Hunan<sup>7</sup>, about 100 kilometers SW of Yongxin. The Middle Chinese retroflex stop initials in Anren are described by Chen Manhua as having two layers (Chen 1995: 71-73): a colloquial layer with dental stop reflexes t and th-, and a literary layer represented by tʃ- and tʃh- or ts- and tsh-. In that dialect, the Early Middle Chinese palatal affricates are represented by tʃ- and tʃh-, without any examples of t- or th-. Examples:

	col. layer	lit. layer
竹 *trjuwk ‘bamboo’	tiu <sub>7</sub>	
筑 *trjuwk ‘to build’	tiu <sub>7</sub> (~ 壩)	tʃiu <sub>7</sub> (建~)
中 *trjuwng ‘middle’	tiŋ <sub>1</sub> (~ 間)	tʃuen <sub>1</sub> (~ 華)
長 *trjangX ‘to grow’	tioŋ <sub>3</sub> (~ 大)	tʃoŋ <sub>3</sub> (生~)
砧 *trim ‘to hack’	tien <sub>1</sub> (~ 板)	tsã <sub>1</sub> (~ 板)

**Table 2: colloquial and literary reflexes to the retroflex stop initials in Anren (Hunan)**

<sup>7</sup> For a review of scholarly opinion on the classification of Anren dialect, see Chen Manhua (1995: 2-3). Anren is variously classified as Gan, or as close to Gan within Xiang.

The hilly area between Hunan and Jiangxi, where both Nanwen Cun and Anren are located, is an area where preservation of early features may be expected to reach higher levels than in the Poyang plains or Gan River Valley<sup>8</sup>.

The existence of an old layer of forms in Hakka and in Gan was recognized in my previous work on Gan and Hakka (Sagart 1988: 149), but somewhat downplayed and not adequately documented. In consequence I would like to modify and supplement my theory (Sagart 1988; 1993:26) in the following way: the most immediate common ancestor of southern Gan and Hakka, a language which I believe was spoken in central Jiangxi in Song times, and which I call ‘Proto-Southern-Gan’, was an already stratified dialect with an ‘archaic’ layer and a thicker Late Middle Chinese layer with northern Gan admixture. To the archaic layer belong the features identified by Norman for Hakka and by myself, for Gan. This stratification is the result of the superposition of two Chinese populations in central and south Jiangxi: a population of early Chinese settlers and a larger population of northern migrants in Tang times.

A superposition of two originally distinct populations in the modern Hakka heartland in northern Guangdong is explicitly mentioned by the Song geographer Wang Xiangzhi 王象之 in his *Yu Di Ji Sheng* 輿地一勝: according to Wang, in Mei-zhou [Meixian] in Guangdong,

“There is plenty of land in the prefecture, but the people are idle and few of them work at agriculture. They rely entirely on migrants from T’ing-chou and Kan-chou to do the farming”  
(source: Shiba 1968 [1970])

---

<sup>8</sup> Note that in many southern Gan dialects the alveopalatal affricates have been changed to t- and th- (Sagart 1993). Since the *she shang yin* initials normally evolve to alveopalatal affricates in Gan, in these dialects the normal reflexes of the *sheshang yin* are alveolar stops! we will never know if among those mainly innovative t’s and th’s there are not a few *conservative* ones...

It is tempting to equate the two populations in Wang's description with the two main waves of settlers which contributed to the Chinese population of south China.

## subgrouping should be based on innovations

The idea that subgrouping should be based on uniquely shared innovations (rather than on shared preservations, or retentions) is well established in historical linguistics. It is usually attributed to Brugmann. Evolution specialists use exactly the same concepts, under different names, to determine the shape of the evolutionary trees that represent the relationships between animal species.

Discussing the genealogy of Burgess Shale animals, Gould (1989:214-215) uses the terms 'shared-and-derived traits' and 'shared-but-primitive traits' for notions which are the exact analogs of the 'uniquely shared innovations' and 'shared retentions' of historical linguistics. He writes:

Rats and people share both hair and a vertebral column. Both are homologies, structures which are inherited from common ancestors. If we are searching for a criterion which will properly unite rats and people into the genealogical group of mammals, we can use hair, but the shared vertebral column will not help us at all. Why the difference ? Hair works because it is a *shared-but-derived* character, confined to mammals among the vertebrates. A vertebral column is no help because it is a *shared-but-primitive* character, present in the common ancestor of all terrestrial vertebrates –not just mammals – and most fish.

Later on (p. 215) he talks about "the fallacy of basing groups on shared-but-primitive traits".

Let us now see how the principle of subgrouping on innovations applies to the Chinese case. It is universally agreed that all varieties of Chinese go back to a unique ancestor language, homologous with Gould's vertebrates: let us call that language 'Common Chinese'. The precise time and place of Common Chinese is open to discussion –my own feeling is that it was probably the standard language in northern China at the time of the Chinese conquest of the south, Early Han-. We now wish to discover the genealogy of Chinese dialects which are descended from Common Chinese: in

Gould's terms, we are looking for criteria which will allow us to determine subgroups of vertebrates like fishes, reptiles, mammals, birds. There are competing theories of what the subgroups are. Luo Changpei's proposal was that Gan and Hakka are derived from a common ancestor which is itself a descendant of Common Chinese. This is, in effect, a hypothesis about subgrouping. Likewise, the proposals by Lu Guoyao (1988), by Li Rulong and Xin Shibiao (1999) are subgrouping hypotheses: they set up hypothetical proto-languages which are conceived as descendants of Common Chinese. The same is also true of Norman's 'Old Southern Chinese' hypothesis. Under Brugmannian assumptions, all these subgrouping hypotheses should rely on uniquely shared innovations. Indeed, the hypothesis of Li Rulong and Xin Shibiao relies on the observation that aspirated devoicing – definitely an innovation–, is uniquely shared by Hakka, Gan, Huizhou, etc. However, I will show below that this particular innovation is of a stereotyped kind, and that it may well have occurred independently in all these dialects. As to Norman's theory, it appears to be based principally on preservations, rather than on innovations: the observation that Hakka, Min and Yue share a number of phonologically conservative features (such as preserving p's where other dialects have f's) does tell us that they are related languages, but this we already know: they are descendants of Common Chinese (vertebrates, in Gould's terms). What is needed to demonstrate that Hakka, Yue and Min are more closely related together than each of them is to any other Chinese dialect (this is in essence Norman's proposal), is a set of innovations which are uniquely shared by Hakka, Min, and Yue. Such Min-Yue-Hakka innovations, if they exist, will provide evidence that a language ancestral to Min, Yue and Hakka once existed in which these innovations could be made. Very well; but do such innovations really exist? I examine the question in the next section.

## are there uniquely shared innovations of Hakka-Min-Yue ?

In fact, surprisingly few of the uniquely shared characteristics of Min, Hakka and Yue are innovations. In support of his proposal that Min, Yue and Hakka all derive from his ‘Old Southern Chinese’, Prof. Norman cited mostly shared retentions in his article of 1986. In his book of 1988 (pp. 213-214), he added a few propositions for shared innovations. While he did not specifically claim that each of these innovations were unique to Hakka, Min and Yue, I think that such a claim has to be made if ‘Old Southern Chinese’ is to stand. In fact, as shown below, most of these innovations are also shared by at least some Gan dialects.

- **to poison.** Norman (1988: 213) claims that a uniquely shared innovation of the southern dialects consists of having derived a verb ‘to poison’, in tone *qù*, out of the noun 毒 ‘poison’. The corresponding Middle Chinese form would be dawH. Thus Fuzhou 福州 (Min) *thau<sub>5</sub>*, Hakka *theu<sub>5</sub>*, standard Cantonese *tou<sub>6</sub>*, all ‘to poison’. However this is not a *uniquely* shared innovation of Hakka-Min-Yue: the same form is found in two southern Gan dialects: Nancheng 南城 *hou<sub>6</sub>* and Jianning 建寧 *hau<sub>6</sub>* ‘to poison, esp. fish’<sup>9</sup> (Li and Zhang 1992:372). Outside of Gan, note also Jiangyong 江永 *tau<sub>6</sub>* ‘to poison fish’ (Huang 1993:175). Jiangyong is an unclassified dialect of south Hunan. There are also in several south Chinese dialects including Hakka, Gan and Xiang, *qusheng* words for ‘to poison’ going back to a proto-form \*nawH, which seem related, perhaps via a nasal prefix (I have in mind a post-archaic derivation something like \*m-dawH > ndawH > nawH, where m- is the voluntary action prefix proposed in Sagart 1999).

---

<sup>9</sup> change of th- to h- is regular in these two dialects, see Sagart 1993.

- **cockroach.** Norman (1988: 214) argued that the name of the cockroach: Amoy ka-tsuat<sub>8</sub>, Cantonese kǎ-tsaat<sub>8</sub>, Meixian vɔŋ<sub>2</sub> ts<sup>h</sup>at<sub>8</sub>, derives from an Old Southern Chinese proto-form \*dzat. He pointed out that this form is not attested in the standard historical lexica, implying that it is an innovation of Old South Chinese. Even if this form is an innovation<sup>10</sup>, it is, again, not *uniquely* shared by Min, Hakka and Yue: most of the words for ‘cockroach’ in Gan dialects are also derivable from Early Middle Chinese \*dzat, \*dzret, \*dzræt, for instance Anyi 安義 ts<sup>h</sup>at<sub>8</sub> tɕi<sub>0</sub> p<sup>h</sup>ɔ<sub>2</sub> li<sub>0</sub>, Pingjiang 平江 dzat<sub>78</sub> ke<sub>1</sub> bɔ<sub>2</sub> etc. (for a longer list, cf. Li and Zhang 1992 p. 254).
- **louse.** Norman (1988: 214) claimed that adding a female animal suffix to the name of the louse is a Southern Old Chinese characteristic. He cited Amoy sat<sub>7</sub> bu<sub>3</sub>, Hakka set<sub>7</sub> ma<sub>2</sub>, Cantonese sat<sub>7</sub> na<sub>3</sub>. In Gan similarly one finds Yiyang 弋陽 sɛ<sub>7</sub> mo, Xinyu 新余 sɛ<sub>78</sub> p<sup>h</sup>o<sub>2</sub> (for more forms, see Li and Zhang 1992 p. 254).

I conclude that a set of innovations shared uniquely by Hakka, Yue and Min, has yet to be presented. Failing this, the idea of a unified ‘Old Southern Chinese’, at least as defined in Norman (1988), cannot easily be maintained. What seems more likely on present evidence is that, in the first centuries CE, there existed, scattered in various parts of southern China, Chinese settlements where varieties of Chinese were spoken; that the speakers in each of these settlements had few contacts with the Chinese speakers in the other settlements, because population levels were low, and because Chinese was surrounded by other languages; that, as a result, each of these areas of concentration of

---

<sup>10</sup> It is possible that the Min, Hakka and Cantonese forms reflect ~~𧈧~~ Early Middle Chinese (*Guang Yun*) tsret ‘a small cicada’ (*Er Ya, Fangyan, Guang Ya*). Dauzat (1922: 142) points out that there is frequent interchange in French dialects between the names of the cicada, the cockroach and the cockchafer. The voiced initial in southern

Chinese speakers developed its own linguistic innovations. This will account for (a) the existence in modern southern Chinese dialects of linguistic materials with archaic-looking features, and (b) the lack of uniquely shared innovations between Min and the other southern dialects. In my opinion, in the first half of the first millennium CE, there were distinct Chinese dialects in coastal Jiangsu-Zhejiang, in the Gan River valley, and in the Xiang River valley. The old coastal Jiangsu-Zhejiang dialects formed the nucleus out of which modern Min and Wu (perhaps also Huizhou) arose (one of their uniquely shared innovations is a word \*nowng for ‘human being’<sub>[LS2]</sub>); while the old Chinese dialects of the Gan River Valley form the earliest layer in Gan and Hakka. Uniquely shared innovations of Gan and Hakka are presented in the next section.

### some possibly uniquely shared innovations of Gan-Hakka

The classical Gan-Hakka theory as expounded by Luo Changpei is based on a number of shared features (footnote 3), some of which are innovations. However, few, if any, of these innovations, are *uniquely* shared by Gan and Hakka, and we must agree with Norman’s criticism that Luo’s features are superficial: they cannot demonstrate a close relationship between Gan and Hakka. At most, they indicate that Hakka and Gan have participated in a number of areal developments. And yet, even if Luo did not cite any, uniquely shared innovations of Gan and Hakka do exist. I list below a few lexical innovations shared by some Hakka and some Gan dialects, and, to my knowledge, by no other dialects. Of course, we are far from a complete knowledge of the vocabulary of Chinese dialects, so that cognate forms might yet turn up in other dialects. Note that in our search for uniquely shared innovations of Gan and Hakka, we do not require the features in question to be found in all

---

dialects may be due to the effect of the prefix. Note that Cantonese has a variant kǎ-tsaat, reflecting a voiceless initial.

Hakka and all Gan dialects. We allow for the possibility that these once innovative features may have been overlaid by more recent northern features in many Gan and/or Hakka dialects.

- **Home:** Linchuan uʔ<sub>7</sub> ha<sub>6</sub>, Hakka vuk<sub>8</sub> k'a<sub>1</sub>, both 'home' (apparently from *wu<sub>1</sub>xia<sub>4</sub>* 屋下). For more examples see Li and Zhang (1992:217). Standard Cantonese uk<sub>7</sub>-khei<sub>3</sub> 'home' is similar in structure, but the second morpheme (usually interpreted as *qi* 企 'to stand') is different.
- **son.** Standard Cantonese has a morpheme lai<sub>55</sub>, occurring in the compounds 孺仔 lai<sub>55</sub> tsui<sub>35</sub> 'youngest son' and 一女 lai<sub>55</sub> nœi<sub>13</sub> 'youngest daughter'. In some western Gan dialects and in many Hakka dialects, a form corresponding to this 孺 is found, also with initial ʌ. Its meaning, however, is broader: in these dialects it is the general term for 'son': Sung Him Tong (and general Hakka) lai<sub>5</sub> tsu<sub>3</sub> 'son' (departing tone); Pingxiang 平鄉 lai<sub>11</sub> 'son' (Wei Gangqiang 1990); Leiyang 耒陽, a Gan dialect of southeast Hunan (Zhong 1987: 223) has læ<sub>3</sub> ɕi<sub>0</sub> 'boy'. I assume that we are dealing with a semantic broadening from 'youngest son' to 'son', in a language ancestral to Gan and Hakka.
- **in that way.** Linchuan 臨川 en<sub>5</sub> luŋ<sub>1</sub> (Luo 1940: 211) 'in that way; so...'. Example: ...en<sub>5</sub> luŋ<sub>1</sub> yen<sub>3</sub> la "as far as that !" (Luo 1940: 231). Compare Sung Him Tong 崇謙堂 Hakka an<sub>3</sub> luŋ<sub>1</sub> (Sagart 1982: 72dubious etymology) 'in that way'. Example: an<sub>3</sub> luŋ<sub>1</sub> tɕ<sup>h</sup>iu<sub>5</sub> k<sup>h</sup>o<sub>3</sub> ji<sub>3</sub>... "that way he could...". The etymology of this form is obscure.

- **here/this** and **there/that**

A remarkable convergence between Meixian Hakka and Yongxiu 永修, a northwestern Gan dialect (my field notes of October 1985) concerns the distal and proximal deictic adjectives. Cf. Table 3:

	Meixian	Yongxiu
'here'	ke <sub>3</sub> p <sup>h</sup> ien <sub>3</sub>	ko <sub>3</sub> pien <sub>5</sub>
'there'	ke <sub>5</sub> p <sup>h</sup> ien <sub>3</sub>	ko <sub>5</sub> pien <sub>5</sub>

**Table 3: 'here' and 'there' in Meixian Hakka and Yongxiu Gan.**

The agreement, especially in tones, can hardly be a coincidence, even though this pattern only occurs in Meixian and in Yongxiu.

- **thou.** In Most Gan and Hakka dialects one finds forms for 'thou' which are relatable to the habitual Chinese pronouns *ni*<sub>3</sub> 你 and *er*<sub>3</sub> 爾. However in a few southern Gan and Hakka dialects a special form is found, cf. Table 4 :

---

<sup>11</sup> ↗ is a high rising 'changed tone' with diminutive meaning; the original tone is unknown.

	<b>2sg</b>
Dabu 大埔 Hakka (He 1993:14)	hen <sub>2</sub>
Wuping 武平 Hakka (Li and Zhang 1992: 419)	heŋ <sub>2</sub>
Qiaotou 橋頭 Hakka (w. of Taihe 泰和; my notes of June 99)	hæ̃
Dayu 大庾 Hakka (Li and Zhang 1992: 419)	heɜ <sub>3</sub>
Shaowu 邵武 Gan <sup>12</sup> (Li and Zhang 1992: 419)	xɛn <sub>6</sub>
Pingxiang 萍鄉 Gan (Wei 1995)	hẽ̃ <sub>56</sub>
W. Gan, several locations between Anfu 安福, Jishui 吉水 and Yongxin 永新 (my notes of June 99)	heĩ

**Table 4 : the 2sg pronoun in several southern Gan and Hakka dialects**

Geographically the dialects which show these forms are scattered in two broad areas along the western and eastern edges of Jiangxi: one in west Fujian and north Guangdong: Shaowu, Wuping, Dabu; an another in the hills of western Jiangxi. within these two areas, their distribution is discontinuous. We are obviously in presence of a relic form which has been submerged by newer intrusive forms except in the hills on both sides of the main communications corridor which is the Gan valley.

The /heN/-type forms for the 2nd person pronoun cited in Table 4 do not resemble any 2sg pronoun of any non-Chinese language in the area. Since they include a nasal and front vowel, they may well bear some connection to the more common Chinese 2sg pronouns. At any rate, /heN/-type forms are unknown in OC, and the special forms listed in Table 4 must be considered innovative.

Overall, then, these (so far) uniquely shared innovations argue for a particular genetic closeness between the archaic layers of Gan and Hakka<sup>13</sup>.

---

<sup>12</sup> Norman (1982) argues that Shaowu is a Min dialect. I follow Li and Zhang here.

<sup>13</sup> In the original version of this paper, it was argued that a word *ai<sub>1</sub>~oi<sub>1</sub>* for ‘mother’, which is found in a number of Hakka and Gan dialects, is a uniquely shared innovation of Hakka and Gan. However, it was pointed out by Mr. Debbin Wu at the conference that the same form (*ai<sub>1</sub>* ‘mother’) also occurs in Chaozhou, a southern Min dialect.

## did Hakka devoicing occur in northern or southern China ?

As mentioned earlier, several authors believe that aspirated devoicing in Hakka occurred at an early date, in the mid-first millennium CE, when the ancestor language of Hakka (plus Gan and/or Tong-Tai 通泰 and/or Huizhou 徽州) was allegedly still in northern China. There is actually some specific evidence that devoicing in Hakka occurred much more recently, in the south, while Hakka was in contact with She.

She is a Hmongic language spoken by about a thousand speakers in the Huizhou area in Guangdong province, in contact with Hakka. Its closest relative within Hmongic is Kiong Nai, a language spoken in Jinxiu district of Guangxi. Mao and Meng (1986:5) describe the relationship between Kiong Nai and She as being extremely close. Unlike Kiong Nai, She has been in intimate contact with Hakka for centuries. As a result, most members of the She nationality now speak a language which, despite some differences, has been described as a variety of Hakka (Luo 1980). The last speakers of She are bilingual in Hakka, and She includes a large number of Hakka loanwords. I have mentioned at the beginning of this paper the possibility that the present-day Hakka population includes a component of She language shifters, and that some words in Hakka are to be ascribed to a She substratum.

The crucial point here is that the old Hmong-Mien voiced stops and affricate initials have undergone in She the same change as in Hakka: they have become voiceless aspirated, for instance:

---

The rules of the game require me to withdraw this form, since I cannot demonstrate that its presence in Chaozhou

gloss	Proto-Hmong-Mien initial (Wang and Mao 1995)	She
pus	b	phu <sub>6</sub>
to see	b	phɣ <sub>8</sub>
to enter	bj	phɔ <sub>6</sub>
flower	bwj	phun <sub>2</sub>
hoof	d	the <sub>2</sub>
few	dz <sub>2</sub>	tshɣ <sub>6</sub>
door	dʎ	khɔŋ <sub>2</sub>
fat, grease	dʎ	khun <sub>6</sub>
little, not tall	ɣ	khje <sub>4</sub>

**Table 5: aspirated devoicing in She**

Now aspirated devoicing in She is an oddity in the Hmong-Mien context. Among the very representative sample of 23 Hmong-Mien languages and dialects treated in Wang and Mao (1995), only She has undergone aspirated devoicing. A few languages maintain some form of initial voicing, but the most common treatment is unaspirated devoicing: the old voiced stops and affricates go to voiceless unaspirated sounds. This is what happened in Kiong Nai, She's closest relative. Obviously She and Kiong Nai have devoiced independently, after their separation. The obvious explanation for the She oddity is contact with Hakka: the change was transferred from Hakka to She through bilinguals. However, in order for the Hakka mode of devoicing to be transferred to She, it was necessary for devoicing to have been in process in Hakka in the first place. Otherwise, how would She speakers have known that they were to aspirate their voiced stops? The conclusion we should draw from this is that Hakka devoiced in the south, while it was already in contact with She.

Aspirated devoicing is not a strange and unusual change. It is fairly common in China and elsewhere in East and South-East Asia. It can occur where voiced stops and affricates have acquired a breathy

---

is due to contact with Gan-Hakka.

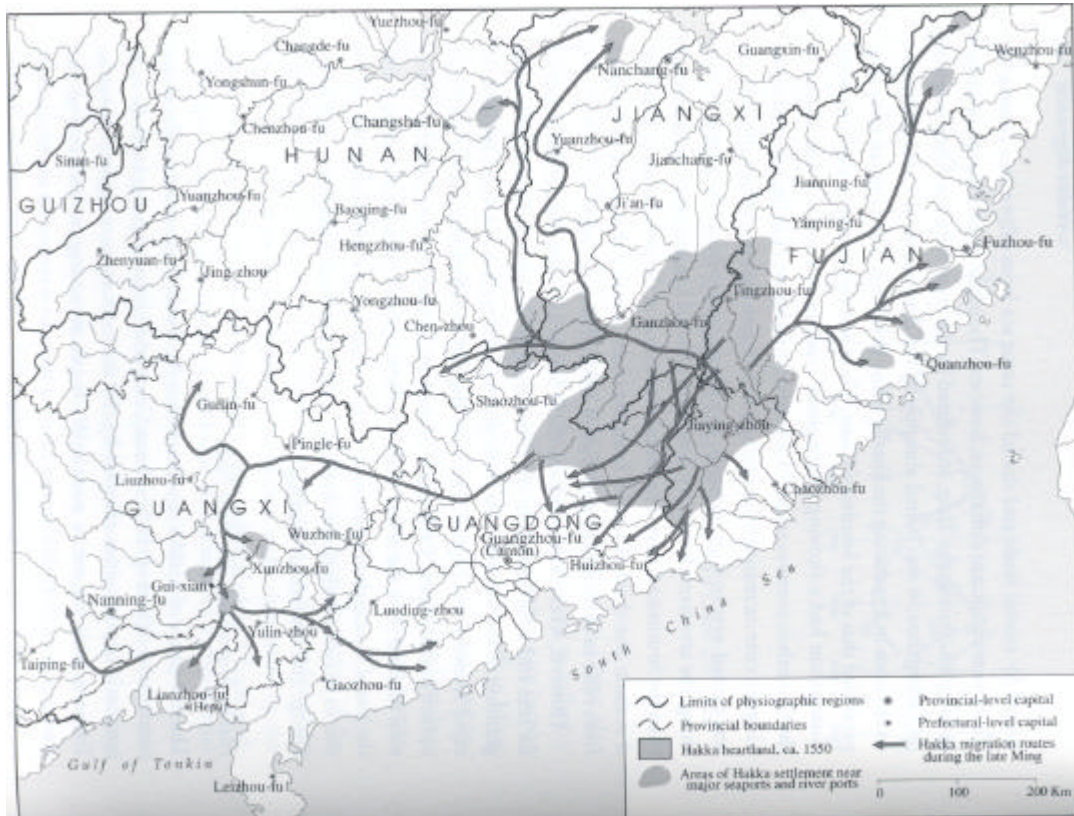
release in a low pitch context. When full devoicing occurs, breathiness is reinterpreted as voiceless aspiration. A change like this is likely to occur independently in different places when the preconditions are met. No one would argue that aspirated devoicing in Bangkok Siamese or Karen indicates a close relationship with Hakka. For the same reason there is no need to suppose that aspirated devoicing in Hakka, Gan, Huizhou, Tong-Tai, and Shanxi-Shaanxi dialects is indicative of a special relationship between these dialects. The relatively high occurrence of aspirated devoicing in Chinese dialects is probably due to the fact that Late Middle Chinese, the *koinè* of Tang times, had voiceless stops with breathy release corresponding to the voiced stops of the *Qie-Yun* system (Maspero 1920).

Although I believe that a special relationship exists between southern Gan and Hakka, and although it is a fact that they share the same type of devoicing, it is not necessarily the case that their common ancestor was a language with aspirated devoicing. It is entirely possible that Proto-Southern Gan, the common ancestor of southern Gan and Hakka, had voiceless stops with breathy release corresponding to the voiced stops of the *Qie-Yun* system; that aspirated devoicing (changing these voiceless stops with breathy release into aspirated stops) started in Hakka or in southern Gan *after* their separation, and then spread to the other language, as sound changes often will. In a future paper, I will present evidence from Nanxiong dialect that this was actually so.

## The nature of the Gan-Hakka boundary

Leong's study of Hakka migrations in the period 1550-1850 shows the Hakkas migrating out of the Hakka heartland principally along three axes: south into Guangdong province, north-east into Fujian and south Zhejiang, and north-west into the highlands separating Jiangxi and Hunan. This can be seen in map 1. These migrations occasioned expansions of the Hakka dialect as spoken in the Hakka

heartland, especially in northern Guangdong in and around Meixian: they account for the great homogeneity of modern Hakka. As a side effect, these



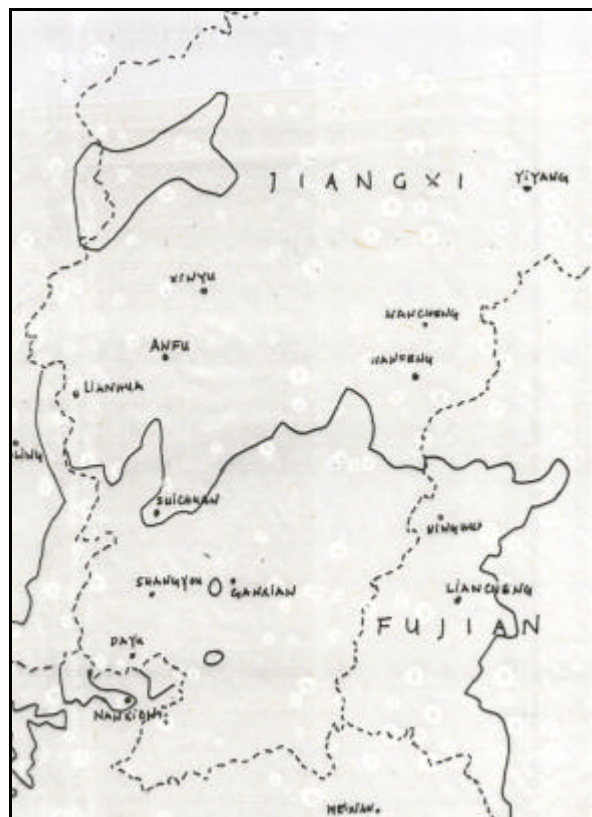
**Map 1: Hakka migrations during the late Ming (from Leong 1997:44)<sup>14</sup>**

migrations sharpened the linguistic boundaries between Hakka and neighboring dialects: the old boundaries between Hakka and its neighbours were erased in several locations. How sharp these old boundaries were, we may never know. However there is one important exception. There were apparently no significant Hakka migrations due north, into south-eastern Jiangxi. This is where the linguistic boundaries between southern Gan and Hakka have been the least disturbed by recent

<sup>14</sup> Reprinted from *Migration and Ethnicity in Chinese History*, by Sow-theng Leong, edited by Tim Wright, with the permission of the publishers, Stanford University Press. © 1997 by the Board of Trustees of the Leland Stanford Junior University.

migrations. Interestingly, the south-eastern Gan dialects (such as Nancheng 南城, Nanfeng 南豐, Guangchang 廣昌) are those which show the greatest resemblance with Hakka.

Let us now look at the boundary between Gan and Hakka as drawn on the Language Atlas of China. This is shown on Map 2. The principles according to which the boundary was drawn are not entirely explicit, but we will assume that it corresponds to some sort of reality. It is well-known that dialect boundaries sometimes correspond to old administrative or political boundaries. Such boundaries tend to reduce opportunities for speech interactions, forming an obstacle to the spread of innovations arising on either side. Even though one unified language may have been spoken at some point over the entire area, in time new isoglosses will keep bundling along the boundary, this eventually inducing dialect differentiation. This appears to be the case here. The lie of the north boundary of Hakka according to the Language Atlas of China corresponds exactly with the major administrative boundary in southern Jiangxi: the boundary between the administrative divisions of Ganzhou, Ji'an, and Fuzhou (Linchuan). This boundary doubles up as a geographical boundary: it runs along the crestline of one of the principal mountain ridges in southern Jiangxi, with summits at 1163 m. and 1454 m. Historically, as an administrative boundary, it has existed since the Sui dynasty (*Zhongguo*



## Map 2. Boundary between Gan and Hakka according to the Language Atlas of China (redrawn)

*Lishi Dituji*). The fact that some isoglosses bundle along this line is therefore entirely expected under a theory which regards southern Gan and Hakka as close relatives.

### Ganzhou city and the development of Hakka: some speculations

In his introduction to S.T. Leong's book, G.W. Skinner observes that Hakka is highly anomalous among Chinese dialects in that it lacks a lowland regional core and even a large city (Leong 1997: 3-4). The city of Ganzhou 赣州 in south Jiangxi would appear to fit the description: it is a lowland city, the largest in south Jiangxi, situated at the confluence of the Gong 貢 and Zhang 章 rivers which join there to form the Gan River. Together these two rivers and their tributaries irrigate all of south Jiangxi. Today Ganzhou city is a southwestern Mandarin isolate in Hakka-speaking territory, but this situation is the result of a language shift in Ming times, brilliantly elucidated by Han Zhenfei (1998): until early Ming Ganzhou city was Hakka-speaking. The current position of Meixian as the main Hakka cultural center is probably posterior to the defection of Ganzhou to southwestern Mandarin: one may suppose that it results from the high degree of involvement of northern Guangdong Hakkas in the Hakka migrations of the recent period. I suspect that in Song and Yuan times the dialect of Ganzhou city was the driving force behind the individualization of Hakka.

### conclusion

The theory I have just outlined accounts essentially for some of the shared innovations shared uniquely by Gan and Hakka. After the massive Tang migrations of northerners into the south, Chinese-speaking populations in south China became overwhelmingly numerous, while minority languages receded. As a result, contact between dialects became a much more important factor than

previously. Many shared innovative features of Gan and Hakka (like devoicing) were probably not features of Proto-Southern Gan but spread through the Gan-Hakka area after the separation of Gan and Hakka. Contact also accounts for many more features shared by Gan and/or Hakka and other southern dialects. More attention should be devoted to the spread of linguistic features between dialects in contact.

## References

Anfu Xian Zhi Bianzuan Weiyuanhui

1995 *Anfu Xian Zhi*. Beijing: Zhonggong Zhongyang Dangxiao Chubanshe.

Chen Manhua

1995 *Anren fangyan zhi*. Beijing: Beijing Yuyanxue Chubanshe. 256 p.

Dauzat, A.

1922 *La géographie linguistique*. Paris: Flammarion.

Deng Xiaohua

1999 Kejiahua gen Miao-yao Zhuangdongyu de Guanxi wenti. *Minzu Yuwen* 3:42-49.

Gould, Stephen J.

1989 *Wonderful life*. New York and London: Norton and Co.

Han Zhenfei

1998 Xinan Guanhua zai Gan-nan de fenbu jiqi xingcheng. In : Huang Yuzhao (ed.) *Ke Cong He Lai*, pp. 398-403. Guangzhou: Guangdong Jingji Chubanshe.

Hashimoto, M.J.

1973 *The Hakka dialect*. Cambridge: Cambridge University Press.

He Gengjiong

1993 *Kejia Fangyan Yufa Yanjiu*. Xiamen Daxue Chubanshe.

Huang Borong

1996 *Hanyu Fangyan Yufa Leibian*. Qingdao: Qingdao Chubanshe.

Huang Xuezheng

1993 *Jiangyong Fangyan yanjiu*. Beijing: Shehui Kexue Wenxian Chubanshe.

Iwata, Ray

1988 Koso anki ryosho ni okeru shinzokushoyitaikei no chiritekibunpu to koshoyitaikei no saiko [The geographical distribution of kinship terms in Jiangsu and Anhui provinces]. In Ozaki, Y. and Hirata S. (eds.) *Contributions to Chinese historical linguistics and philology*. Kyoto University: Zinbunkagaku kenkyusho, pp. 207-272.

Leong, S. T.

1997 *Migration and ethnicity in Chinese history*. Stanford: Stanford University Press.

Li Fang-kuei

1938 Languages and dialects of China. *The Chinese Yearbook*, 1938. Shanghai: Commercial Press.

Li Rulong and Xin Shibiao

1999 Jinnan, Guanzhong de “quanzhuo songqi” yu Tang-Song Xibei fangyin. *Zhongguo Yuwen* 1999, 3:197-203.

Li Rulong and Zhang Shuangqing

1992 *Ke-Gan fangyan diaocha baogao* [A report on a survey of the Kejia and Gan dialects]. Xiamen: Xiamen University Publishing House.

Liu Yongguang et. al.

1992 *Yongxin Xianzhi*. Beijing: Xinhua.

Lo Hsiang-lin

1933 *Kejia yanjiu daolun*. Xingning: Xi shan shu cang.

Luo Changpei

1940 *Linchuan Yinxi*. Changsha: Shangwu.

Lu Guoyao

1988 Taizhou fangyin shi ji Tong-Tai fangyan shi yanjiu. *Computational Analyses of Asian and African Languages* 30: 149-219.

Mao, Zongwu and Meng Zhaoji

1986 *Sheyu Jianzhi*. Beijing: Minzu Chubanshe. Maspero, H.

1920 Le dialecte de Tchang-ngan sous les T'ang. *Bulletin de l'Ecole française d'Extrême-Orient* 20: 1-124.

Norman, J.

1982 The classification of the Shaowu dialect. *Bull. Inst. Hist. Phil.* 53, 3: 543-83.

1986 *What is a Kejia dialect?* [paper circulated at the 2nd International conference on sinology. Academia Sinica, Taipei, December 1986].

1988 *Chinese*. Cambridge language surveys. Cambridge: Cambridge University Press.

Oshima, H.

1995 Nanfeng Yinxi. *Zhongshan Daxue Xuebao (Shehui kexue ban)*. 1995-3, 124-132.

Sagart, L.

1982 A list of Sung Him Tong Hakka words of dubious etymology. *Cahiers de Linguistique Asie Orientale* XI, 2: 69-86.

Sagart, L.

1988 On Gan-Hakka. *Tsinghua Journal of Chinese Studies* New Series Vol. 18 No.1 (June 1988) pp.141-160.

Sagart, L.

1993 *Les dialectes gan*. Paris: Langages Croisés. 285 p.

Sagart, L.

1999 *The Roots of Old Chinese*. Current Issues in Linguistic Theory, 184. Amsterdam: John Benjamins.

Shiba, Yoshinobu

1968 [1970] *Soodai shoogyoo-shi kenkyuu*. Tokyo: Kazama Shobo. Translated by M. Elvin as "Commerce and Society in Song China". Ann Arbor: Center for Chinese Studies, the University of Michigan, 1970.

Wang Fushi and Mao Zongwu

1995 *Miao-yao yu guyin gouni* [a reconstruction of the sound system of Proto-Miao-Yao]. Beijing: Zhongguo Shehui Kexue.

Wei Gangqiang

1990 *Pingxiang fangyan zhi*. Beijing: Yuwen.

Xu Yumo et al.

1993 *Yihuang Xianzhi*. Beijing: Xinhua.

Yuan Jiahua et al.

1962 *Hanyu fangyan gaiyao*. 2nd edition 1983, Peking: Wenzi gaige.

Zhengzhang Shangfang

1986 Wannan Fangyan de fenqu (gao). [the grouping of southern Anhui dialects]. *Fangyan* 1986, 1: 8-18.

Zhong Longlin

1987 Hunan shang Leiyang fangyan jilue. *Fangyan* 1987, 3: 215-31.