

**CHINESE GEOLINGUISTICS:  
HISTORY, CURRENT TREND AND THEORETICAL ISSUES**

Ray IWATA

Kanazawa University, Japan

riwata@kenroku.kanazawa-u.ac.jp

**Abstract**

Although the international trend in geolinguistics has been to utilize dialect resources via computer-based GIS, research in the Chinese field is still grounded in exploring the history of words using the classical method of linguistic geography. After describing the historical background of how we Japanese researchers carry out projects on Chinese dialects, this paper will demonstrate some of our findings: 1) The distribution of modern dialects is well accounted for in terms of “Northernization” and “Southern kernel area”; 2) Linguistic geography can make it possible to reconstruct the history of words unbiased by historical documents; 3) The development of stress accent in Northern dialects has caused some word groups to acquire grammatical elements in their forms, due to the function of analogical attraction; 4) Any word can be in collision with others due to internal and external factors. Chinese cases are explained in terms of homonymic and synonymic collisions.

**Key Words**

Linguistic geography, Willem Grootaers, Bernhard Karlgren, analogical attraction, homonymic and synonymic collisions

**1. Introduction**

In this paper I intend to summarize and introduce the fruits of our project, titled the “Project on Chinese Dialects (PHD).” Launched in 1989, it aims at the revival of linguistic geography in the Chinese field, and will be formerly materialized starting with the first volume of a collection of maps, *Interpretative maps of Chinese Dialects*.<sup>1</sup> The

---

<sup>1</sup> This volume is scheduled for publication in December 2009.

term “linguistic geography” is defined here as a discipline of linguistics, and was founded by J. Gilliéron and continued with his successors in Europe.

Linguistic geography was brought into China as early as the 1940s by Father Willem Grootaers, who started his work in north Shanxi as a Catholic missionary. However, his ambitious project, which was intended to carry out a nation-wide survey from a linguistic geography perspective, was not realized when he left China in 1948 (Grootaers 1943, 1945, 1957). In 1950 Grootaers landed in Japan, where linguistic geography had already begun to gain reputation, due to the contribution of such precursors as Kunio Yanagida, as well as to the translations of its founders’ works, in particular Matsubara and Yokoyama (1958) for Dauzat (1922). Thus, the trend of linguistic geography was formed by the collaboration of Grootaers and Japanese scholars, and eventually caught on throughout Japan, coming to fruition with the *Linguistic Atlas of Japan* (1957-1966).

In China, on the other hand, linguistic geography ceased to exist after 1948. Meanwhile, mainstream linguistic study was directed toward two purposes: one, to reconstruct ancient Chinese phonology, and the other, to classify its dialects and demarcate their respective areas of distribution. While the current international trend in geolinguistics is to utilize dialect resources via computer-based GIS, we researchers in the Chinese field still have enough reasons for emphasizing the necessity of applying this classical method, linguistic geography, to the study of Chinese languages (Grootaers 1943, 1945).

## **2. Brief history and the recent trend in Chinese dialectology**

For the benefits of readers’ and to better understand the background of our undertaking, I will briefly mention the history of Chinese dialectology in this section.

### *2.1. Karlgren and Qing Philologists*

It is well known that philological studies in China made remarkable progress during the Qing Dynasty period (1644-1911). The achievements made by Qing

philologists, who explored the language of their ideal époque, the Zhou Dynasty (1020-249 BC.), by means of reconstructing the sound system into something they called “Old Sounds”, is in itself worthy of praise as a highly refined system of scientific philology. In fact, it is comparable to the achievement of comparative linguistics in the Western world that was accomplished at just about the same time. What is crucial, however, is that these Qing philologists, with few exceptions, were unaware of the practice of using modern living languages as a means to reconstruct proto languages. While they sometimes did refer to dialectal evidence, it was limited to the cases where they were concerned with etymological questions. It was completely out of their scope to study modern languages because they had in their hands a ready-made phonological framework, a rhyming dictionary *Qieyun* (edited by Fayuan Lu in 601 AD.). They called it “Modern Sounds”.

A contribution of Bernhard Karlgren (1889-1978) was to treat “Modern Sounds” (his “Ancient Chinese”) as a reference point for studying the whole history of Chinese; namely studying “Old Sounds” (his “Archaic Chinese”) in terms of the projection from Ancient Chinese. This approach was much like that of the Qing Philologists, while unlike the Chinese precursors, this explained the phonetic forms of modern dialects as reflexes of Ancient Chinese. In his masterpiece, *Études sur la phonologie chinoise* (1915-1926), he reconstructed the sound system of Ancient Chinese, based on his own survey of twenty-four dialects, thus establishing his comparative method. For this dialect survey, however, a severe criticism was offered by Grootaers (1943).

## *2.2. Academia Sinica and the tradition of dialect classification and demarcation*

Succeeding the comparative tradition established by Karlgren, young leaders gathered around the Institute of History and Philology in Academia Sinica, Yuen-ren Chao et al., and began their surveys in central China in the late 1920s (Chao 1928). These surveys concentrated again on recording Chinese character readings. Through surveys of this sort, sound correspondences among the modern dialects were given in a convenient fashion, and this facilitated researchers in finding criteria for classifying the dialects (Chao et al. 1948).

After 1949, under the PRC regime, although the task imposed on Chinese dialectology was to propagate a standard language, Putonghua, scientific and descriptive

spirits survived for around ten years, as demonstrated by the model case survey carried out by the Academy of Social Sciences in Changli County, Hebei Province (Chinese Academy of Social Sciences, 1960). Descriptive studies were revived after a long hiatus in 1979, and it was revealed shortly thereafter that the target of the scholars in the Academy was once again to classify and demarcate the dialects. Thus the atlas, *Language Atlas of China (LAC)*, was published in 1987, comprising 18 maps of the Chinese (Han) dialects and 17 maps of the minority languages.

### 2.3. Recent trends

Unlike former days, recent trends in Chinese dialectology are widely diversified. While the mainstream seems to still emphasize classification and demarcation, new methodological and theoretical trends such as “lexical diffusion” and “comparative dialectal grammar” have also taken hold. In the comparative field, Jerry Norman and his students have long endeavored to reconstruct the regional proto X dialects by comparing purely colloquial vernaculars in a bottom-up fashion. This would mean, in effect, the abandonment of Karlgren’s dogma, which set Ancient Chinese as the reference point in the historical study of Chinese (Norman 1973 & 1988: 228-244, Handel 2003, Akitani 2003).

As early as the early 1980s, Russian linguist Olyga Zavyalova reported on her discovery of the long isoglosses, which run along the Huai River in the East and the Qin-ling mountain belt in the West, dividing the whole Guanhua (Mandarin) area into northern and southern sections (Zavyalova 1983). In Japan, Grootaers’s works were translated into Japanese by the present author and M. Hashizume (Grootaers 1994). Following this, it was retranslated into Chinese by Professor Rujie Shi, thus bringing linguistic geography back to China (Grootaers 2003). Two publications which were crowned with the name of “dialect geography” appeared during the past four years: Xiang and Cao (2005) and Simmons et al. (2006). Although both studies were based on detailed surveys and provided us with abundant interesting phonetic evidence, discussions were still centered on the issues of isogloss and dialect boundary.

The most noteworthy event at the time of writing is the publication of the volume, *Linguistic Atlas of Chinese Dialects*, edited by Zhiyun Cao, Beijing Language and Culture University (Cao ed. 2008). Unlike LAC, this atlas is item based, comprising

205 phonetic maps, 203 lexical maps and 102 morphological and syntactic maps, which were compiled using a GIS based computer system. For this atlas, Cao and his colleagues surveyed 930 localities all over the Han Chinese speaking area. Most of the localities surveyed were local villages or towns, instead of big cities or county seats. This is in accordance with the policy proposed by Grootaers (1957), and the speakers selected were mostly males born during the years from 1931 to 1945. This atlas is comparable in scale to such distinguished atlases as ALF and the Wenker Atlas, and it is remarkable that the authors completed all necessary processes, including the dialect survey, data editing and cartography, within seven years.

This atlas, defined by the authors as a collection of “descriptive maps”, is aimed at providing readers in the Chinese field with basic linguistic data. However, it will not be easy for most Western readers to access its linguistic information, as the information is only notated in Chinese characters. Moreover, although an overall principle for classifying the forms or features are mentioned, no explanations are given for each map on the grounds of classification. Thus much depends on what readers themselves manage to get out of each map.

### **3. Project on Chinese Dialects**

For our own project, PHD, we take up the task of following up on the lost zone in Chinese linguistics along lines compatible with the tradition of linguistic geography. It is our belief that Chinese dialects are the crucible of a huge amount of attractive evidence, which still waits researchers’ recognition in order to obtain their real linguistic value. For instance, the name for “fly” changed to that of “mosquito” in a vast area of Southwest China, meanwhile the name for “fly” is now used for “bee” in some areas, but it is used for “ant” in another area. As another example, we witnessed that the names for “broad bean” and “pea” are reversed between one area and another, and the reversed *signifiant* and *signifié* also emerge for “house” and “room”, evoking the possibility of a North-South contrast on a nationwide scale. In spite of the fact that some of these phenomena are well known among Chinese researchers, there have been few attempts to approach the historical truth of how these referential shifts came about. Note that these puzzling phenomena find their parallels in European and Japanese languages,

and have been well explained in terms of phonetic attraction and the homonymic collision of one word with another.

Our maps, *Interpretative Maps of Chinese Dialects*, uniquely demonstrate the “interpretation” of each author for the historical change of each word. We believe that the accumulation of individual findings will lead us to discover the existence of some universal factors at work in the changes across dialects and across languages, and will shed light on the aspect of language universals and individuality as a whole. In the following sections, I will introduce briefly some of the results that have appeared so far through our research.

#### 4. Northernization and the Southern kernel area

Similar to French and German, Chinese is characterized by its dialects’ evincing a North-South opposition, with the Southwest area under a considerable degree of Northern influence. Norman (1988: 181-183) succeeded in describing this characteristic simply by setting up ten linguistic criteria. There are two main dialect boundaries which run along the two rivers: the longer one, hence historically a more significant one, is referred to as the *Huai River line*, and the shorter one is referred to as the *Yangtze River line*. Needless to say, this situation was brought about by extra-linguistic historical factors, including politics, the economy, climate, geology, and population movement. I propose two key terms in interpreting the historical formation of dialect distribution in China. One is *Northernization* and the other is the *Southern kernel area*.

##### 4.1. Northernization

*Northernization* refers to a long time process in which Northern features incessantly moved southwards, causing a varying degree of deformation to occur to the Southern dialects on an individual basis. This process was promoted in the first place by the movement of the population. Two thousands years ago, when the earliest dialect dictionary was compiled by Xiong Yang (BC. 53-AD.18), South China, and the southeast coastal area in particular, had the least Han Chinese population, while the area along the Yellow River (Central Plain) was densely populated. As evidenced by Yang’s

dictionary, the Chinese dialect at that time showed an East-West opposition within the Northern zone. Although there existed around the middle reaches of the Yangtze River a strong linguistic influence which succeeded the tradition of the large *Chu* empire and which opposed the Northwestern standard dialect of the *Qin*, it is true that the clear cut horizontal dialect boundary we witness nowadays had not yet been formed in this period (Matsue 2006). It must have taken hundreds of years for the Chinese dialects to accomplish the conversion of its distributional pattern from the East vs. West type to the North vs. South type. One of the contributors to this conversion was the repeated Chinese settlement into Southern non-Chinese areas. However, this was not the sole factor. The formation of the two horizontal dialect boundaries are also attributed to the national boundaries set up during the war-torn periods, as well as to administrative boundaries set up during relatively peaceful times, both of which were drawn in a horizontal direction along the two rivers: the Huai and the Yangtze.

The situation appears to have drastically changed during the Tang era (618-907 AD.), which was relatively stable and peaceful, and which followed the war-torn Six Dynasty (222-589 AD.) era. The population in South China increased remarkably due to large-scale cultivation (Chen 1982). This evidence seems to support Karlgren's theory, which argued for the transplantation of the *Tang koine*, the standard language spoken in the Capital Chang'an, to almost all areas inhabited by Han Chinese. Indeed this theory can account for the main body of some Southern dialects, as far as their phonetic features are concerned.

Even with the strong influence of the Tang koiné on Southern dialects, we can still confidently say that this is not the whole story. The relationship between the Tang koine and modern dialects may be comparable to that between classical Latin and the modern Latin languages, as most modern Chinese dialects are not the direct descendants of the Tang koine. We should be aware of the existence of the old linguistic layer which was transplanted into South China by early settlers who emigrated from North China before the Tang era. This is the language that constitutes the main strata of the present day dialects in the Southeast coastal area, typically Min, and this may hold true to a considerable extent for its neighboring dialects, such as Hakka and southern Wu (Norman 1989). The dialects in their surrounding areas, Yue, Gan, Xiang and northern Wu, owe their characteristics to a considerable extent either to the Tang koine or directly to the influence of neighboring Northern dialects. Meanwhile it is also true that

they more or less reveal their own characteristics, which are either of an innovative or of a conservative nature.

Northernization of the Southern dialects could not have been accomplished simply by the southern movement of population. There are two views we consider common sense but which are not shared by most Chinese dialectologists. One regards the effect of migration on the subsequent development of the host dialects. Based on historical documents, scholars in the Chinese field have interpreted dialect distribution in terms of migration. However, evidence shows that the language of immigrants has a tendency to assimilate to the host dialect and will basically fade away after three generations, if the immigration is to the village of Han Chinese inhabitants (Iwata 2007a: 125). The detailed field survey by Grootaers (1945) carried out in the rural Datong district in Shanxi Province in North China clearly shows the relevance of this evidence, denying the faulty assumption of those scholars who insist on the total replacement of the inhabitants in North China due to massive migration from one village, Hongtong in Shanxi, at the beginning of the Ming Dynasty (1368-1644), a legend prevailing even now. Our common sense must hold true for the dialects in South China as well.

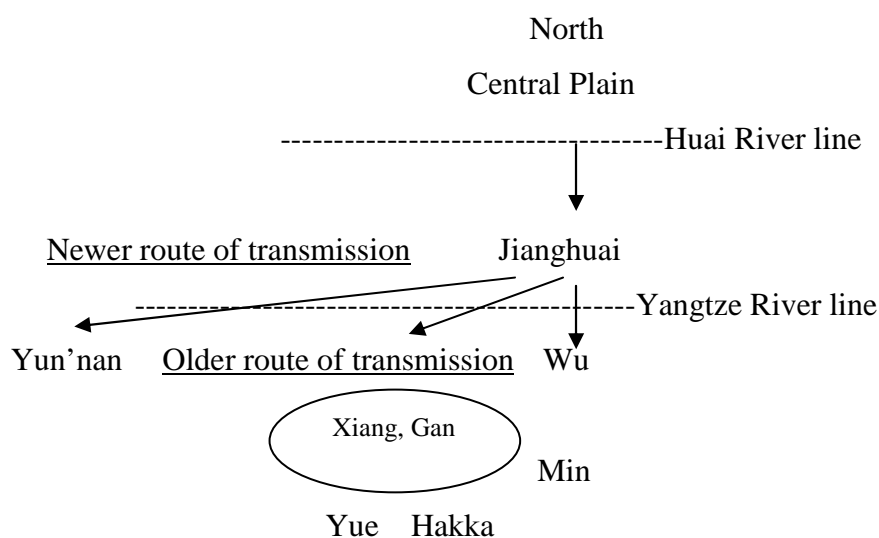
The second piece of common sense for us is the fact that *words could travel by walking*, instead of flying from one place to another by migration. This is to say, the most prevalent medium for dialect diffusion or transmission should have been daily communication of farmers living in one village with those in another, therefore it should have taken a long time for one word to move from one place to another. In Chinese dialectology, however, this view has been least recognized by researchers due to their overestimation of the factor of migration (Iwata 1995: 222-223). Meanwhile, the dispersion of the Tang koine would have been attained most likely in a top-down fashion, from the northern Capital Chang'an to the kernel city in each provincial area, then to the local city in each prefectural area, and finally to their circumferential rural areas. And all these were fulfilled through the introduction of Chinese character readings by intellectuals. From the viewpoint of host dialects, this event should be viewed as *lexical borrowing* from the upper class language (see section 7.2. below).



#### 4.2. Southern kernel area

The *Southern kernel area* specifically refers to the Jianghuai area, the area situated between the Huai River and the lower reaches of the Yangtze. It had two kernel cities around its southern border: Nanjing and Yangzhou. Nanjing was established as the Southern capital during the era of the Six Dynasty (222-589 AD.), and through to the 20<sup>th</sup> Century it performed the role of political center for the whole of the Southern area, namely the region south of the Huai River. The latter city, Yangzhou, flourished as a large economic and commercial center starting from the Tang era.

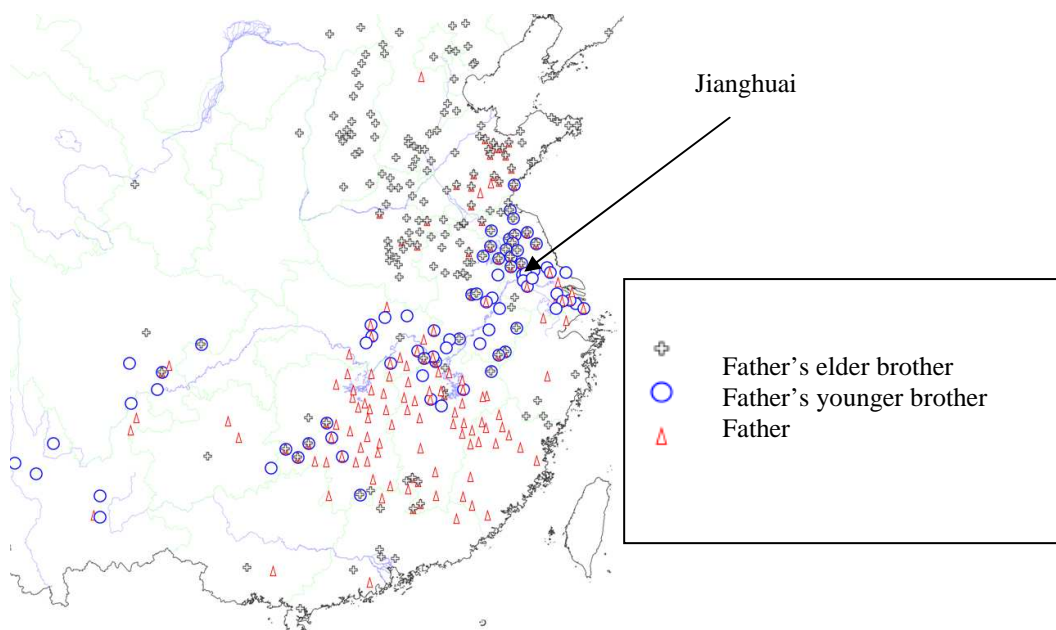
Linguistically, this area has played two roles ever since the Six Dynasty era. One has been the role of relay-station or bypass, through which linguistic features have been transmitted from the North to the South. Another has been the role of core areas, where a number of linguistic innovations were initially born and subsequently radiated out to surrounding areas (Iwata 2000: 180). It is noted here that transmission of words by no means takes place haphazardly; rather words usually take the routes that were previously determined by extra-linguistic factors (Dauzat 1922: 456). The following chart is the essence of our assumption in this respect.



By this chart we assume the existence of three main routes along which many words of Northern origin were conveyed from Jianghuai to the South. The one to the east actually represents a direct intrusion of Jianghuai words or features into the Wu

area. The other two routes contributed to the linguistic transmission toward a southwestern or western direction. The older one, which long functioned as a main route to the southern inland area, ran along the Yangtze down to its middle reaches, where it turned its direction toward the south. The newer route was brought about in the initial stage of the Ming Dynasty (1368-1644) by its military occupation in Yun'nan and the subsequent establishment of a trading route along the Yangtze. This new route eventually made it possible for Jianghuai words or features to be conveyed as far as the southwestern extreme of the river. It would have significantly accelerated the transmission of words by means of shipping from one port to another, successively to each economic hinterland that developed along the river.

Map 1 is an instance exemplifying this assumption (Iwata 2000).



Map 1. Referents of the stem *ye*

The kinship stem *ye* could be employed for three referents, “father”, “father’s elder brother” and “father’s younger brother”, and it can be seen on the map that each of the three referents respectively possesses a concentrated distribution zone of its own.<sup>2</sup>

<sup>2</sup> Since modern Chinese nouns, in particular those in northern Chinese, of higher frequency are largely polysyllabic, a morpheme-based analysis is effective in investigating the historical changes of words. Hereafter, maps 1-3 and the discussions about them will be concerned with the main component in a word, which will be referred to as “stem” or “head”. For the phonetic representation of words, it is

“father’s elder brother” : Northeastern zone

“father’s younger brother”: along the Yangtze River

“father”: Southern zone (mainly in Gan and Xiang and their surround areas)

The oldest usage of the stem *ye* is undoubtedly that for “father”. This is philologically evidenced by the famous popular poem *Mulan*, presumably the product of the Six Dynasty era or of the beginning of the Tang at the latest, but revived in America as the Walt Disney animation in the 20<sup>th</sup> Century. The leading woman, Mulan, who would have probably been born and raised in North China, called her father *ye* (she called her mother *niang*). Suppose that this original usage of the stem *ye*, crossing the Huai River line, entered the Jianghuai area. From this kernel area, it would have been conveyed along the “older route of transmission”, instead of moving southward, and eventually would have reached the Gan and Xiang areas, the inland area within the Southern zone. While the original usage of the stem *ye* has mostly been preserved in these Southern areas, there occurred innovations in its usage in the Northeastern zone and the Jianghuai area. In the Northeastern zone, the referent of *ye* shifted or extended from “father” to “father’s elder brother”, but the change was from “father” to “father’s younger brother” in Jianghuai. From the map, it is assumed that the Jianghuai type innovation was transmitted along the “newer route” to as far as the upper reaches of the Yangtze.<sup>3</sup>

The difference in the direction of referential shift or extension between the two areas would reflect the difference shown by the extra-linguistic background.<sup>4</sup> The

---

actually impossible for the present paper to give each word form in an exact phonetic style since the number of dialectal variations would amount to hundreds. Hereafter, the word forms will be substituted by the standard Putonghua form presented in Pinyin Romanization in italics, and the word meaning will be notated inside quotation marks or round brackets.

<sup>3</sup> On map 1, the distribution of the usage for “father’s younger brother” is not continuous, but is cut off in the middle reaches of the Yangtze. This is due to the lack of relevant dialect materials for this area.

<sup>4</sup> To be exact, the two notions, “referential shift” and “referential extension” are distinguished from each other in meaning. The former refers to the case where the original usage disappeared completely, while the latter refers to the case where it was retained and was coexistent with the recent innovation. However, these two types of referential change are often indistinguishable from the present status of the dialect. If a particular form denotes only one semantic category in the present status, it is not certain if it is the outcome of referential shift or if the original usage has faded away after the extension occurred. If the form denotes two semantic categories, it is probably the result of extension. For instance in Map 1, the original usage of *ye* for “father” and its innovative usage for “father’s elder brother” coexist in a number of dialects even within the Northern zone. In such cases, however, we should be aware of the possibility

example of referent shift from “father” to “father’s elder brother” would reflect the reverence for the eldest member of one generation. The original stem used to refer to “father’s elder brother” must have been *bo*, but it was eventually replaced by *ye*, because from the side ego, his father and father’s elder brother are equally the object of reverence in the clan. Likewise the stem *ye* also came to be used for the eldest member in the clan, i.e., grandfather (father’s father) in North China.

The referential shift from “father” to “father’s younger brother” that occurred in Jianghuai came from a popular belief of one’s parents (actually “grandparents” from the angle of ego) who wished to protect their younger sons from the menace of evil. The original stem employed for “father’s younger brother” would have been *shu*, but this usage was intentionally avoided by replacing it with *ye* so that it would be hard for evil to recognize this particular man (Iwata 1988: 232-241).

Referential shift or extension is not only confined to the two types mentioned above (Iwata 2000: 194). For example, in Map 1 the black cross symbol and the blue circle are duplicated in northern and mid Jianghuai, indicating that the stem *ye* is used for both “father’s elder brother” and “father’s younger brother”. It is assumed that the older usage was restricted to “father’s younger brother” but that the usage was extended to refer to “father’s elder brother”. Externally this change can be explained in terms of the influence of Northern usage which actually intruded into this area, but internally it is quite probable that the dialects in this area tended to adopt the sole stem *ye* for referring to all paternal uncles as well as one’s father, as evidenced by the fact that some dialects employ this particular stem even in referring to “father”. Readers who are unfamiliar with the Chinese kinship system may wonder how these members are distinguished with each other in terms of linguistic form. As a matter of fact, the distinction is guaranteed by prefixing the *Paihang* number to the stem, namely *da ye* (elder father), *er ye* (second father), *san ye* (third father) etc.

---

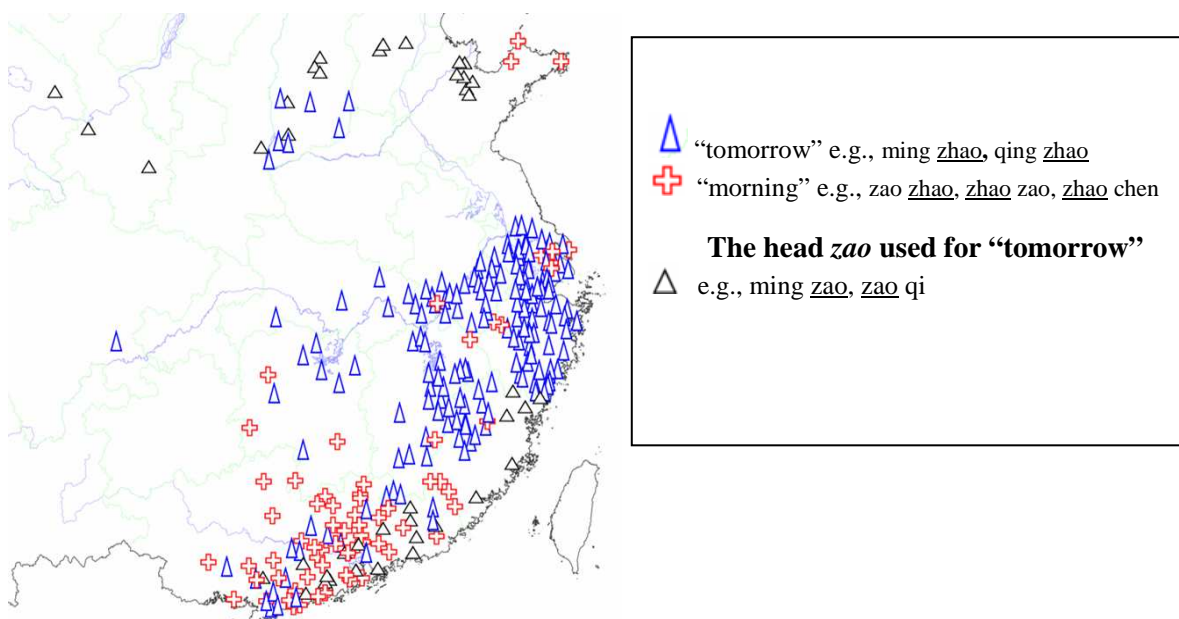
that the original usage has been revived for whatever reason after it once faded away. Refer to the example of the Jianghuai dialects, which will appear shortly in this section.

## 5. Reconstructing the history of words

Being blessed with the richest historical legacy of written texts, historical linguistics in China has depended too much on philological evidence, resulting in the purely dialectal approach that is unbiased by these texts, and not even included within the scope of linguistic inquiry until very recently. Etymological studies have so far been centered on finding a one-to-one correspondence between the form recorded in the written text and that found in the dialect.

A working hypothesis, which was referred to as the “principle of continuity” in Europe, and as the “principle of ABA distribution” or the “theory of concentricity” in Japan, may be applicable to our study, for distinguishing the older form from the newer one, if we can exclude the possibility of parallel change or that of population movement in interpreting the distribution.

Map 2 shows a typical instance for the preservation of old forms in isolated areas.



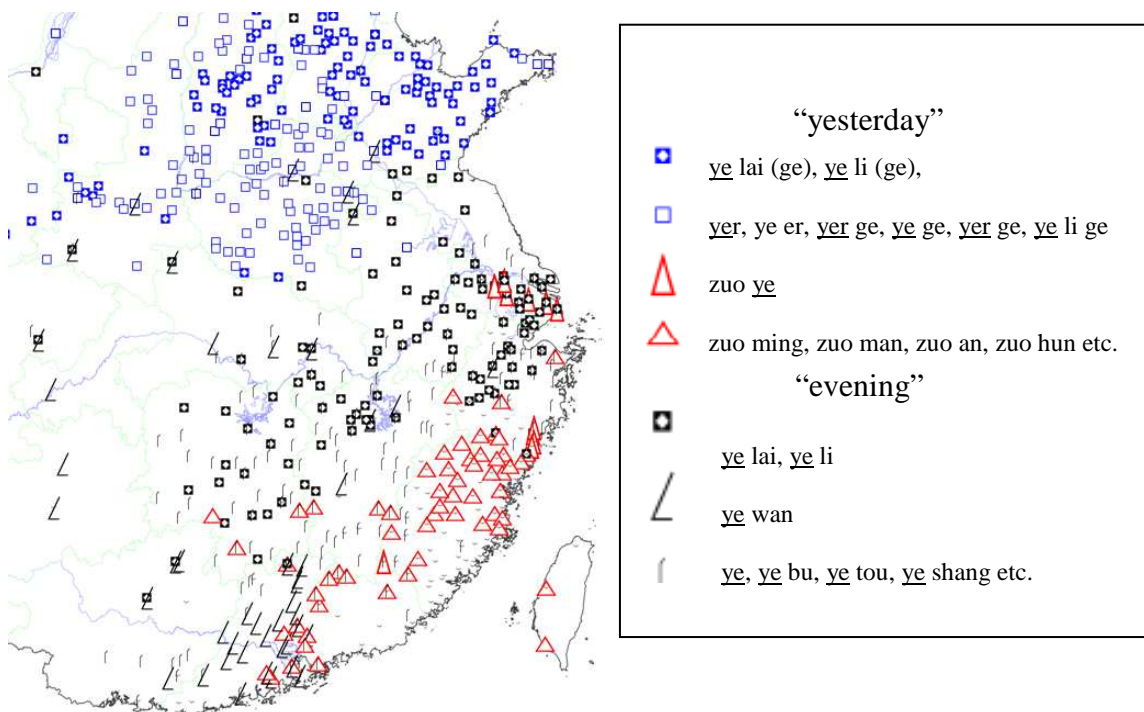
Map 2. Referents of the head *zhao*

The stem *zhao* appears on the map as the form for “morning” and “tomorrow”, not only concentrated in the South, but also emerging at some localities in the North. In this case, *zhao* was a morpheme that was present in the lexicon of Old Chinese (Kargren’s “Archaic Chinese”) and was used as a free form meaning “morning”, but

since then it has lost its original usage in modern Northern dialects, just appearing in some compounds in literary documents, thus we have no other alternative but to consider it as a preservation of an old usage.

The evidence that the identical head, *zhao*, is shared by the two semantic categories, “morning” and “tomorrow”, indicates an etymological relationship between them. Undoubtedly its usage for “tomorrow” was an outcome of later semantic extension, exactly the same extension process as what has occurred in many European languages, as well as in the Japanese language (Buck 1949: 999-1000). Map 2 also indicates that this has actually been a repeated process in Chinese dialects. In some dialects, the morpheme *zao*, which means “early” if it is used as a free form, but also appears in such compounds as *zao shang*, meaning “morning”, came to be used for denoting the semantic category “tomorrow”, apparently an outcome of recent innovation.

Its parallel was a referential extension of the head *ye* from the category “evening” to “yesterday”.<sup>5</sup> This change can be induced from Map 3.



Map 3. Referents of the head *ye*

<sup>5</sup> There is a semantic distinction between “evening” and “night” in contemporary standard Chinese, but it is doubtful if there was such a distinction in ancient dialects.

In South China, the use of the head *ye* for “yesterday” has been preserved in such forms as *zuo ye*, while the semantic category “evening” came to be denoted by such compounds as *ye wan* (literally meaning “night evening”).<sup>6</sup> In North China, on the other hand, though the use of the head *ye* for “yesterday” at one point disappeared altogether, the same type of change was revived in a relatively recent epoch, causing this head *ye* to shift to denoting the category “yesterday”, while the category “evening” came to be denoted by such compounds as *hei ye* (literally meaning “black night”).

As a matter of fact, a more popular head used for “tomorrow”, “yesterday” as well as for the other semantic categories concerning “day” is *ri*, originally meaning “sun”, and it is evident that this morpheme has been used since ancient times. However, the reconstruction unbiased by historical documents as demonstrated above leads us to assume that there had existed two lexical strata since an unknown ancient period: one was the stratum that used the head meaning “day” (<“sun”) for all time words denoting “day”, and another was that used the heads meaning “morning” and “evening” for “tomorrow” and “yesterday”.

## **6. Verbal pathology and therapeutics: analogical attraction**

The phonetic and semantic contents of words may be damaged due to internal and/or external factors, however at this time the dialect in question usually provides these words with some linguistic remedy for reconstruction. This may be one of the core ideas of Gilliéron, an idea that he metaphorically called “verbal pathology and therapeutics”. To my understanding, the situation in Chinese and French is parallel in that Northern dialects have been seriously damaged and deformed due to radical changes. Concerning Northern Chinese, although a number of researchers have assumed an Altaic influence, notably Hashimoto (1978), direct evidence for supporting this hypothesis is actually scarce, at least so far as lexical changes are concerned. Rather, changes in most cases seem to have been motivated or triggered independently of non-Chinese influences.

---

<sup>6</sup> In South China there also appear other forms like *zuoming*, *zuoman*, *zuo'an* and *zuohun*, the head of which, *ming*, *man* etc., consistently denotes “evening” or “dark”. It is noted also that the form *zuoye* is known in the standard language as the one meaning “last night”.

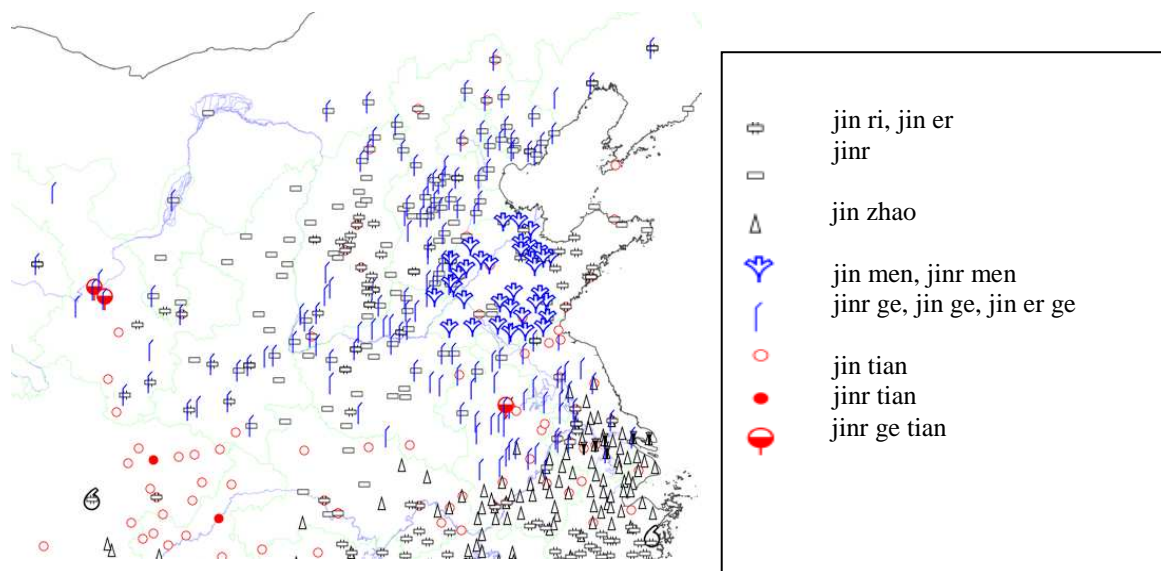
One internal factor that may have had a serious effect on lexical changes was the development of word accent, namely *stress*, which was brought about by the increase of polysyllabic words, in effect a compensation effect caused by the simplification of phonological structure. Word stress thus produced the following patterns, and is currently observed in the majority of Northern dialects, typically the dialect of Beijing.

Bi-syllabic structure: Strong-weak (trochee)

Tri-syllabic structure: Medium-weak-strong

Even though the second syllable in each structure presumably was not so weakened at the beginning, it is true that these stress patterns came to be shared by many colloquial words of higher frequency. What is of relevance here is that such words were then inclined to be attracted by other words. A manifestation of it is a phenomenon which we refer to as *analogical attraction*.

Map 4 demonstrates the distribution of the forms denoting “today”, limiting the area mainly within the Northern zone.



Map 4. Word forms for “today”

Curiously, in many dialects, word forms take suffixes that are possessed by pronouns, which were indicated on the map by blue symbols. A more frequent



morpheme suffixed to the preceding component is the general classifier *ge*, which also appears as a suffix of demonstrative pronouns, such as *zhe ge* (this) and *na ge* (that). Another one, *men*, which is less frequent, but concentrates in distribution around west Shandong, is evidently the suffix of personal pronouns, such as *wo men* (we) and *ni men* (plural “you”). Note that these changes, were not mere happenings that solely occurred to “today”, also occurred to the other time words, such as “tomorrow” and “yesterday”. In particular, the use of the general classifier *ge* has even extended to cover many time words, including those denoting “year” (e.g., this year) and “season” (e.g., spring).

For this particular change, Iwata (2007b) proposed a hypothesis that the extension was phonetically motivated by a decline in phonetic and semantic contents of the head in the time words like the *ri* in *jin ri* “today”, and that at this moment they started to be attracted to the particular pronouns of tri-syllabic structure, namely *zhe yi ge* “this” and *na yi ge* “that”, eventually having changed to such forms as *jin ri ge*.<sup>7</sup> Seemingly, this change was triggered by the function of analogy, similar to the grammaticalized process of some words. However, there should have been no reason for time words to be analogized with demonstratives, unless some other factor got involved in the process of analogy. It must have been the case that through the function of phonetic attraction, time words were analogized with particular demonstratives and became deformed as a result, resulting in the acquisition of the suffix *ge*.

In my field survey performed at the eastern edge of the Huai River line, I discovered evidence of analogical attraction. Namely, a number of bi-syllabic nouns taking a trochaic stress pattern tended to replace their head with the common noun suffix *zi*, so that the word *lao shu* (mouse), in which *lao* was a prefix, had changed to *lao zi* in many villages, having lost its stem *shu* forever. This change implies that the name for “mouse” was incorporated into the word group having the noun suffix *zi*.

The operation of paronymic attraction has been well recognized in linguistic geography. It can occur among semantically unrelated words provided that they more or less resemble each other in their phonetic shapes. Abundant examples of this phenomenon exist in Chinese dialects. For example, names for “gecko”, “bat” and “(big) ant”, *bi hu*, *bian fu* and *pi fu* in standard Chinese respectively, have been attracted

---

<sup>7</sup> These two pronouns, *zhe yi ge* and *na yi ge*, in which *yi* is a numeric for “one”, originally meant “this one” and “that one”, but became to be used as common demonstratives by changing their phonetic shapes to *zhei ge* and *nei ge*, as we find them in present day Beijing, coexisting with the authentic forms *zhe ge* and *na ge*.

to each other in Northern dialects. Note here that all these bi-syllabic words are usually uttered with a trochaic stress pattern. A natural outcome of one such attraction is the crisis of homonymic collision. In this case however, the crisis has been evaded without exception, and the particular words are distinguished from one another by prefixing any folk-etymologically motivated modifier. For example, in Datong (Shanxi) the names for “bat” and “ant” are *ye bie fu* and *ma bie fu*; in Dezhou (Shandong) the names for “bat” and “gecko” are *yan bie hu* and *xie bie hu* respectively. Here the result of change is the formation of a new word group which shares the common component *bie fu* or *bie hu*.

In this manner, words see a weakening of some of their phonetic and semantic contents, eventually putting their linguistic status at risk. Phonetic attraction, either analogical or paronymic, is so to speak a remedy that dialects afford the words at this moment, so that the words in question succeed in recovering stability by forming a new lexical system, thus decreasing the arbitrariness of linguistic signs.

Turning back to the topic of time words, we find that the tri-syllabic form thus produced by the function of analogical attraction, namely *jin ri ge*, has completely disappeared, with its phonetic variant *jin er ge* being retained in a small number of localities in North China. This is because the accent rule for tri-syllabic structure, i.e., medium-weak-strong, has been applied to the time words, giving birth to a successive change described as follows:

*jin ri ge* > *jin er ge* > *jinr ge* > *jin ge* or *ji ge*

This is considered a weakening process of the head, *ri*, which ultimately changed to the non-syllabic retroflex ending by fusing with the preceding syllable, namely *jinr ge*, the form still existing in Beijing. In Jianghuai, the change has proceeded one step further, leaving no trace of the original head in the form *jin ge*.

The following change was again motivated by the accent rule for bi-syllabic structure, i.e., strong-weak. As a result of its application to such forms as *jinr ge* and *jin ge*, the second (last) syllable was weakened to an *intolerable* degree for the speaker. For this occasion the savior for the time words denoting days was analogical substitution, instead of analogical attraction. The unstressed *ge* was substituted for by a free form (thus stressed) *tian*, meaning “sky”, due to the analogy with such compounds as *qing tian* (fine weather), *meitian* (every day) and *bai tian* (daytime), which dialects already

possessed in their respective lexicons (Iwata 2007b: 22-23). Importantly, the usage for *tian* as a verbal measure, e.g., *zou le liang tian* (walked for two days), must have assisted this substitution. As marked in red on Map 4, this standard form *jin tian* is mainly distributed along and around the Yangtze basin, including Jianghuai, and the distribution in the North is scarce except for the Capital Beijing and its surroundings. This would suggest the Jianghuai origin of this standard form, *jin tian*.

Our theory of word stress will also contribute to correcting the erroneous assumption of etymological studies. For example, Northern Chinese has in its lexicon a number of words that phonetically take the polysyllabic structure comprising k- (g- in Pinyin Romanization) and l- initiated syllables in this order, e.g., *ge le bai* (knee), *ge le niu* (snail), in which the second syllable *le* is unstressed and is morphologically considered an infix. According to an etymological study by Yaotian Cheng, a Qing philologist who as an exception did pay attention to spoken vernaculars, all these forms should derive from a word family called *guoluo*, which he assumed to exist in Old Chinese. While there are still many scholars who argue for the archaic origin of this morpheme *le* either in line with the idea advocated by Yaotian Cheng or by adopting the comparative method, we could say that such an assumption is nothing but a grand illusion, due to the fact was that this particular infix was the product of relatively recent innovation that occurred in Northern Chinese. In the first place, the second syllable in these tri-syllabic forms, many of which may have been stems, was weakened by the application of the accent rule, i.e., medium-weak-strong, and secondly they began to be incorporated into a word group which shares the common infix *le*, by attracting one another to their phonetic shapes (Iwata 2007a: 132-134).

## **7. Words in collision: homonymic and synonymic collisions**

A word may come into collision with another due to either internal or external factors (Dauzat 1922). Sound change, attraction by the other forms and folk-etymology are common factors that could internally affect the phonetic and semantic contents of words. The external factor mainly refers to the transmission of words from one locality to another, and this eventually will cause what is called “dialect contact”.

It is proposed that the word collision is of two types: *homonymic collision* and *synonymic collision*. In the following, word forms are represented by “P” and “Q”, their semantic contents or referents (hereafter “referents”) are represented by “x” and “y”, and the whole linguistic sign is represented by such devices as P(x) and Q(y).<sup>8</sup>

### 7.1. Homonymic collision

Homonymic collision is, so to speak, a conflict between different referents for a single form. It is mostly triggered by internal factors, and some sorts of remedies are usually adopted for rescuing the abandoned words. Supposing that a word form “P” for a referent “x” came to be homophonous with that for another referent “y” due to some particular reason(s), there could be at least three outcomes in this type of collision:

1. The victory of “x” and the defeat of “y” (or vice versa):  $P(x) \rightarrow P(y) > Q(y)$
2. Both “P” and “Q” partially change, thus avoiding collision :  $P1(x) / P2(y)$
3. Avoidance of conflict by forming complementary distribution in the geographical area:  $\boxed{P(x) \mid P(y)}$

A frequent outcome of homonymic collision is that the winner “x” takes the place of “P” and that the defeated “y” changes its form to “Q” (Case 1 above). In Chinese dialectology, the problem of taboo words has been discussed in terms of exceptions to “sound laws”. For instance, the word for “pen” unluckily came to be homophonous with the notorious taboo word *bi* as the result of phonological change, and as a natural consequence it was defeated and changed its shape to *bei* in western Shandong (Li 1994).<sup>9</sup> Seemingly taboo words relating to sex are always stronger than others. On the other hand, if the relative frequency of the given two referents is equally high, a sort of compromise can be attained between “x” and “y” (Case 2). For instance, in the vast area of Southwest China, the form for “fly”, i.e., *ying zi*, came to be homophonous with that for “mosquito”, i.e., *wen zi*, presumably due to paronymic attraction. In this case, distinction is generally maintained in terms of adding some sort of modifier to each, e.g., *ye wen zi* (night mosquito) in contrast with *fan wen zi* (rice mosquito), meanwhile

<sup>8</sup> Formalization adopted here is after Mase (1992).

<sup>9</sup> In this case, the tonal categories of these words are of relevance. However, this information is omitted for the sake of simplicity.

curiously some dialects seem to make no distinction between the two referents in their phonetic shapes.

Homonymic collision could be avoided by dialects forming a system of complementary distribution in their shared geographical area (Case 3). Refer to Mase (1992) and Iwata (2006). An instance appeared on Map 3 cited above. The two forms *yelai* and *yeli* denote “evening” or “night” in Jianghuai, as well as in the mid reaches of Yangtze, whereas in the North these forms mostly denote “yesterday”.

One of my colleagues in our project, Yukinobu Murakami, reported on the rare case of homonymic collision triggered by an external factor. In North China, two popular beans, “soybean” and “red bean”, had for a long time formed a counterpart in their names: the former was called *da dou* (big bean) and the latter *xiao dou* (small bean). This partnership, however, was destroyed due to the introduction of a new species of “broad bean” into the Northwestern area through the Silk Road in approximately 1200 AD, which was as large as the ones we eat nowadays. As a consequence, the two beans, “soybean” and “broad bean”, came to compete for the single name *da dou* (big bean). It was apparent which bean was victorious and formed a new partnership with “red bean”. Eventually the defeated “soybean” changed its name into more colorful ones, such as *huang dou* (yellow bean) and *bai dou* (white bean).

## 7.2. *Synonymic collision*

Synonymic collision is defined as the conflict between different forms for a single referent. It is mostly triggered by external factors. Suppose that one form “P” existing in an area encountered another form “Q” which had been transmitted from an adjacent area, and the two forms came to compete with one another for a single referent (semantic category) “x”. There could be at least three possible outcomes from this type of collision:

1. The victory of the recent form “Q” over the original form “P” (or vice versa):

$$P(x) > Q(x)$$

2. Dividing the semantic field or usage between “P” and “Q” without changing referent:  $P(x_1) / Q(x_2)$

3. Forming a blend form:  $\{(P+Q) \div 2\}(x)$

Case 1 above may be the most frequent outcome of synonymic collision, in which “Q” takes the place of “x” and the defeated “P” disappears altogether.

Case 2 is a sort of compromise attained between the recent form and the original form. For example, on Map 1 shown above, when the kinship stem *ye*, which Mulan used in referring to her “father”, reached the lower reaches of the Yangtze, it would have encountered there the original form *die*, thus the two forms should have competed for a single referent “father”. This problem, however, reached a solution by differentiating the usage of the two: *die* for vocative use and *ye* for designative use, as we nowadays find in many northern Wu dialects. In Chinese, this is by no means a special phenomenon, rather it is essentially identical with what researchers have discussed in terms of “multiple readings” of Chinese characters. Recent linguists have shared the same view that the existence of multiple readings, more specifically, colloquial and literary readings, is a reflection of the multiple linguistic strata so far formed in each area in different époques. By looking at this view from a different angle, we can see that this is actually a case of synonymic collision. Suppose that a recent pronunciation (literary reading) “Q” for one character “x” was introduced to a dialect and came to be coexistent with an original reading “P” (colloquial reading). The outcome of such coexistence was usually a differentiation of semantic field or that of actual usage, as formalized as P(x1)/Q(x2) above. Note that this is what frequently happens in the process of *word borrowing*.

Case 3 is phenomenally identical with *word blending*, which is usually produced in naming new products or notions, e.g., [smoke+fog]/÷2=smog. In the scope of a dialect, it is usually produced by the contact of two forms, which are continuously distributed but are separated by an isogloss. For example on Map 4, the two forms indicated by the filled red symbols, *jinr tian* and *jin ge tian*, would have probably been created by the contact of an original form, *jinr* or *jin ge*, with the most recent form, *jin tian*. Contaminated forms thus produced by this mechanism of word blending are also an outcome of compromise between the two dialects in contact with each other, and the frequent occurrence of such forms suggests that fierce dialect contact has been repeated everywhere in China (Iwata 2006, 2007a).

In closing, I would like to emphasize again the relevance of applying the idea and method of linguistic geography to the study of Chinese dialects. Fortunately, as is well

demonstrated in Cao's Atlas, *Linguistic Atlas of Chinese Dialects*, traditional dialects are still preserved in the vast rural area of this country, in spite of the fact that they are in the process of witnessing a diminishing of their strength due to the propagation of the standard language.

## References

- AKITANI, Hiroyuki (2003) *Wuyu Chuqu fangyan (xibei pian) guyin gouni* (A reconstruction of the proto Chu-Qu dialect (northeast sub-group) of Wu), Tokyo: Kobun Press.
- BUCK, Carl Darling (1949) *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*, The University of Chicago Press.
- CAO, Zhiyun ed. (2008) *Linguistic Atlas of Chinese Dialects*, Beijing: Shangwuyin Press.
- CHAO, Yuenren (1928) *Xiandai Wuyu de yanjiu* (Studies in the modern Wu dialects), Peking: Tsing hua College of Research Institute, Monograph No. 4.
- CHAO, Yuenren, Shengshu TING, Shihfeng YANG, Tsungchi Wu and T'unggho TUNG (1948) *Hubei fangyan diaocha baogao* (Report on a survey of the dialects of Hupeh), Shanghai: Shangwuyin Press.
- CHEN, Chengsiang (1982) *Chuugoku rekisi bunka chiri zusatsu* (a Historical and Cultural Atlas of China), Tokyo: Hara Shobo.
- CHINESE ACADEMY OF SOCIAL SCIENCES (1960) *Changli fangyan zhi* (Research report on the Changli dialect), Beijing: Kexue Press.
- CHINESE ACADEMY OF SOCIAL SCIENCES AND AUSTRALIAN ACADEMY OF HUMANITIES (1987) *Language Atlas of China*, Hong Kong: Longman.
- DAUZAT, Albert (1922) *La géographie linguistique*, Paris, Librairie Ernest Flammarion. [Japanese translation by H. Matsubara and K. Yokoyama (1958) *France Gengo Chirigaku*, Tokyo: Daigaku Shorin].
- GILLIÉRON, Jules and Mario ROQUES (1912) *Etudes du Géographie Linguistique d'après l'Atlas Linguistique de la France*, Paris, Librairie Honoré Champion. [Japanese translation by Y. OKAWA, W. GROOTAERS and H. SASAKI (1991-7) "ALF ni yoru gengo chirigaku teki kenkyu", *Proceedings of Dialectological Circle of Japan*, No. 52-57, 61-64].
- GROOTAERS, Willem A. (1943) "La géographie linguistique en Chine, Nécessité d'une nouvelle méthode pour l'étude linguistique du chinois, Première Partie: La méthode de la géographie linguistique" *Monumenta Serica*, VIII, 103-166.

- GROOTAERS, Willem A. (1945) “La géographie linguistique en Chine, Nécessité d'une nouvelle méthode pour l'étude linguistique du chinois, Seconde Partie: Une frontière dialectale dans le Nord-est du Chansi” *Monumenta Serica*, X, 389-426.
- GROOTAERS, Willem A. (1957) “Linguistic geography of the Hsüan-hua region (Chahar Province)” *Bulletin of the Institute of History and Philology*, Vol. 29-1, Studies presented to Yuen Ren Chao, Academia Sinica, 59-86.
- GROOTAERS, Willem A. (1994) *Chugoku no hogen chirigaku no tameni* (Contributions to Chinese dialect geography), translated by R. Iwata and M. Hashizume, Tokyo: Kobun Press.
- GROOTAERS, Willem A. (2003) *Hanyu fangyan dilixue* (Contributions to Chinese dialect geography), translated by R-J. Shi and R. Iwata, Shanghai Jiaoyu Press.
- HANDEL, Zev (2003) “Northern Min tone values and the reconstruction of ‘softened initials’”, *Language and Linguistics*, Vol. 4-1, Institute of Linguistics, Academia Sinica, 47-84.
- HASHIMOTO, Mantaro (1978) *Gengo ruikai chiri ron* (Typo-geographical linguistics), Tokyo: Kobundo.
- IWATA, Ray (1988) “Koso Anki ryo-sho ni okeru shinzoku sho’i keisiki no chiriteki bunpu to ko-shou’i taikai no saiko” (The geographical distribution of kinship terms in Jiangsu and Anhui Provinces – a Reconstruction of the proto kinship system), in: Y. Ozaki and S. Hirata eds. *Kango-shi no sho-mondai*, Kyoto University, 207-272.
- IWATA, Ray (1995) “Linguistic geography of Chinese dialects - Project on Han dialects (PHD)”, *Cahiers de Linguistique Asie Orientale*, Vol.24-2, Paris, EHESS-CRLAO, 195-227.
- IWATA, Ray (2000) “The Jianghuai area as a core of linguistic innovation and diffusion: a Case of the kinship term ‘ye’”, in: *Memory of Professor Li Fang-Kuei: Essays of linguistic change and the Chinese dialects*, University of Washington /Academia Sinica, 179-196.
- IWATA, Ray (2006) “Homonymic and synonymic collisions in the northeastern Jiangsu dialect - On the formation of geographically complementary distributions”, *Linguistic Studies in Chinese and its Neighboring Languages: Festschrift in honor of professor Pang-hsin Ting on his seventieth birthday*, Institute of Linguistics, Academia Sinica, 1035-1058.
- IWATA, Ray (2007a) “Fangyan jiechu ji hunxiao xingshi de chansheng – Lun Hanyu fangyan ‘xigai’ yici de lishi yanbian” (Dialect contact and the production of contaminated forms - a Reconstruction of the history of Chinese words for “knee”), *Bulletin of Chinese Linguistics*, 1.2, Hong Kong, HKUST, 182-212.
- IWATA, Ray (2007b) “Hanyu fangyan ‘mingtian’, ‘zuotian’ deng shijian-ci de yuyan dilixue yanjiu” (Geo-linguistic study of time words, such as “tomorrow” and “yesterday”, in Chinese dialects), *Chugoku Gogaku: Bulletin of the Chinese Linguistic Society of Japan*, No. 254, 1-28.



- KARLGREN, Bernhard (1915-1926) *Etudes sur la Phonologie Chinoise*, Archives d'études Orientales, Vol. 15, published by J. A. Lundell.
- LI, Rong (1994) "A tabu word in Chinese" In H. Kitamura, T. Nishida and Y. Nagano eds., *Current Issues in Sino-Tibetan Linguistics*, Osaka, 421-427.
- MASE, Yoshio (1992) *Gengo chirigaku kenkyu* (Studies in Linguistic Geography), Tokyo, The Ohu Press.
- MATSUE, Takashi (2006) "Han-dai fangyan zhong de tongyan-xianshu – ye tan genju 'Fangyan' de quhua-lun" (Isoglosses in the Han era – with discussing the dialect demarcation according to "Fangyan"), in: Xuecheng Hua ed., *Yangxiong Fangyan jiaoshi huizheng*, Beijing: Zhonghua Press, 1509-1533.
- NORMAN, Jerry (1973) "Tonal development in Min", *Journal of Chinese Linguistics*, Vol. 1-2, 222-238.
- NORMAN, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- NORMAN, Jerry (1989) "What is a Kejia dialect", in: *Proceedings of 2<sup>nd</sup> International Conference on Sinology (Section on Linguistics and Paleography)*, Academia Sinica, 323-344.
- SHIBATA, Takeshi (1969) *Gengo chirigaku no hoho* (Method in linguistic geography), Tokyo: Chikuma Shobo.
- SIMMONS, Richard, Rujie Shi and Qian Gu (2006) *Jianghuai guanhua yu Wuyu bianjie de fangyan dilixue yanjiu* (Chinese dialect geography: Distinguishing Mandarin and Wu in their boundary region), Shanghai Jiaoyu Press.
- XIANG, Mengbing and Hui CAO (2005) *Hanyu fangyan dilixue – Rumen yu Shixian* (Chinese dialect geography: Primer and Practices), Beijing: China Wenshi Press.
- ZAVYALOVA, Olyga (1983) "A Linguistic Boundary within the Guanhua Area", *Computational analyses of Asian & African languages*, No. 21, Tokyo: Institute of Asian & African languages & Culture, 149-159.