Introduction

This introduction is not a direct translation from the Chinese version; rather the two versions are different and complimentary in some respects. We adopt this policy due to our consideration of readers' interests. We think that readers in the Chinese field will be more concerned with the current topics of Chinese dialectology, and that those in general linguistics will be more concerned with methodology and theoretical issues. Naturally not a few topics may be of common concern to scholars in both fields. This version focuses on two main topics: one, to demonstrate the methodology of our making and interpreting of the maps presented in this volume; two, to propose some theories on lexical change through defining some particular notions of linguistic geography. For the other topics, such as the history of Chinese dialectology, characteristics of dialect distribution and the backgrounds of its historical formation, refer to Iwata (2009), which appears in a special volume of *Dialectologia*.

Hereafter, in discussions we will refer to the maps presented in this volume for exemplification, and if necessary, we will also refer to some of the maps appearing in Cao ed. (2008). Word forms will be presented in standard Putonghua (PTH) with notating Pinyin Romanization in italics followed by Chinese characters, i.e., kanji. Word meaning will be notated inside double quotation marks, in contrast with the single quotation marks which are used for other purposes.

A. Methodology

1. Classification of word forms: morpheme based and consonant based analyses

Mapping requires us, as its prerequisite, to work with the classification of the word forms that are stocked in the Data-base for each entry. It also depends on researchers' profound insight into the variations of linguistic forms. In performing this task, we usually attempt to do classification according to synchronic features. However, the results may vary depending on different researchers' individual points of view, thus creating multiple treatments for the same set of linguistic information for each entry. It should be pointed out here that the validity of this 'classification' task is only evaluated empirically as is the case in other fields of humanities. It is empirical knowledge for us that a good classification generally succeeds in figuring out the 'well ordered' distribution on a map (Shibata 1969). In our PHD system, the classification results are spontaneously reflected in the form of the map. By utilizing this property, we usually repeat the feed-back until a rather desirable output appears on the map.

It is also known to us that the *morpheme based analysis* is effective in figuring out a 'well ordered' or 'relevant' distribution on the map (Iwata 1988, 1989). This is because modern Chinese words of higher frequency, in particular nouns in Northern Chinese, are largely polysyllabic compounds which can be decomposed into two or more morphemes. Notably this method of analysis has also been applied to the statistical study of dialect grouping (Wang and Shen 1992).

In working through our analysis, we usually decompose the word forms into two elements: the initial element is referred to either as a prefix, determiner, modifier, prepositional component or initial component, and the non-initial element is referred to either as a stem, head or final component. As an

editorial device, the terminology is not unified since the function of each element varies, depending on the morphological structure of each word. For some lexical entries in which longer forms are dominant, a bi-syllabic unit is treated as composing one element. For example, in Maps 12-1, 2 "three days from now" and Maps 14-1, 2 "three days ago", the large majority are forms of a tri-syllabic structure initiated by a monosyllabic determiner, and the non-initial part is treated as being composed of one element, which generally constitutes a bi-syllabic unit meaning "the day after tomorrow" or "the day before yesterday", e.g., *da-houtian* 大-后天, *da-qiantian* 大-前天. For some entries, the notion of 'zero morpheme' is applied to the forms of monosyllabic structure or those lacking modifier or determiner. For example in Map 3-1 "hale", the forms *bao* 雹 and *baozi* 雹子 are considered as having a 'zero modifier'. Also presence or absence of the diminutive suffixes, *zi* 子 and *er* 儿, e.g., *bao* 雹 or *baozi* 雹子, is mostly disregarded in classification because a relevant distinction in distribution is not observed in the maps.[1]

After accomplishing the synchronic classification, we often encounter the situation where etymological considerations are indispensable. As a practice in Chinese dialectology, a morpheme, if it is etymologically well identified, is glossed with one Chinese character. Actually, however, there usually appear a number of word forms which have come to be deformed to such a great extent that researchers are unable to determine their etymologies based on inter-dialectal sound correspondence. As is mentioned below in detail, this should be due to the effect of some un-mechanical factors, such as folk-etymology, and it is not realistic for us to work with etymological considerations for all these forms.[2]

In cases where etymologically unidentified forms are fully abundant, a sort of mechanical treatment, which is referred to as *consonant based analysis*, is worth attempting for classification purposes. Here 'consonant' uniquely refers to the initial one, but not the coda, in a syllable. For example, the majority of the forms for "ankle" in Northern and Central China can be classified into three major types: h-k, k-k and l-k (refer to Maps 37-1, 2). This notation indicates that those syllables initiated by 'h', 'k' and 'l' appear in the given order either successively or intermittently in a polysyllabic construction, and it implies that all forms belonging to one type must be traced back to a common etymon. Also, this analysis may assist us in investigating the mutual relationship between the types. For example, Northeastern forms for "knee", mostly tri-syllabic, can be classified into two types, p-l-k and k-l-p, regarding the initial stops appearing in each syllable, e.g., [pə luə kai] and [kə lə pai] (Map 40-3). In this case, the particular two types, which are solely distinguished from one another by the order of their stops, are apparently the product of *metathesis* and must derive from a common etymon. The same technique is applied to some other maps appearing in this collection, e.g., Map 17-1, 2 "snail", Map 21-1,2 "bedbug", Map 49 "subordinate particle". The fact that 'consonant based analysis' is effective in classification would reflect the stability of Chinese consonants even at the moment the words themselves undergo a serious attack of un-mechanical factors. In other words,

---

[1] The exceptions are Maps 2-1 and 19, in which the distribution of monosyllabic forms *yue* 月 and *shi* 虱 is relevant in contrast with the forms *yue'er* 月儿 and *shizi* 虱子, which have diminutive suffixes.

[2] Due to this reason, the kanji notation for the same morpheme is often different between researchers. Some of them may write it based on their etymological considerations, while some others may write it according to folk-etymological consciousness of the speakers, but some prudent researchers may just write a blank mark for it while noting the IPA transcription.

un-mechanical changes generally proceed as the way to preserve the initial consonants. Meanwhile vowels and tones seem to be more susceptible to un-mechanical changes, and in this volume there is only one instance, Map 42-3 "navel", which we attempt to classify word forms according to the difference of vowel quality involved in a particular morpheme.

2. Classification of word meaning: semantic map

Some maps in this collection were created according to the meaning or the referent corresponding to a particular word form. Such maps are generally referred to as *semantic maps*, but are of two types. One is the type in which all forms are classified with regard to their meanings. Maps 39-1 "armpit", 43-1 "malaria" and 44-1 "kitchen" fall into this category. For example, in Map 43-1, all forms describing 'the cases of malaria patients' are treated as belonging to one group irrespective of the difference of their forms. Meanwhile, there are some maps in which a semantically induced classification is applied only partially, e.g., Map 37-1 "ankle".

All the other semantic maps fall under the category *referent map*, which displays spatial variation of referent or semantic category (hereafter 'referent') for a single lexical form or a single morpheme. This methodology is substantially the same as the *method of semantic micro-field* which was practiced for Chinese dialects by the Russian linguist Elena Astrakhan (Astrakhan, et al. 1985, Zavjalova and Astrakhan 1998). Technically the map-creating software Wonderland enables us to extract multiple items of linguistic information from the source database. We created referent maps by utilizing this property of Wonderland.

A 'referent map' indicates the occurrence of *referential shift* or *referential extension*. This is typically observed in Maps 8 and 10, which prove that the referents of the forms for "morning" and "evening" shifted or were extended to mean "tomorrow" and "yesterday", exactly the same lexical change as what occurred in many European languages, as well as in the Japanese language (Buck 1949: 999-1000). Maps 24-4 (referents of the form *shanya*o) and 29 (referents of the forms *dadou* and *xiaodou*) are of this type. In addition, Map 47 (the verbs for eating and drinking) actually suggests the occurrence of 'referential shift', though it is not embodied in the style of a 'referent map'. Some of the referent maps indicate that a certain number of words, each of which has its own referent, come to share a single element, thus becoming paradigmatically integrated. Maps 36 (word family "BANG"), 38 (word family "GUAI") and 41 (prefix *kɑk) are such instances. The problem of 'referential shift' and 'referential extension' will be discussed in detail in Chapter B (Section 4) below.

3. Mapping

In many cases, we use two standards in mapping the result of classification. The large categories called 'types', which are notated as 'A', 'B', 'C', are distinguished from each other by the difference in colors put on symbols, and smaller categories called 'groups', which are notated as '1', '2', '3', are distinguished from each other by differences in symbol shapes.

4. Some principles for reconstructing the history of words
4.1 'ABA' distribution

Two simplified figures are presented in order to explain the principles. Here, the largest square

presents the whole explored area. Suppose that one lexical form (or linguistic feature) 'A' originally distributed consecutively over the entire area, and that later, another form (or feature) 'B' was given birth to in the central area for some particular reason. Then the expected distributional configuration should then be like Fig. 1. For example, in Map 3-1 for "hail", the distributional area for the stem *bao* 雹 is seen to be divided by the area of the stem *leng* 冷.
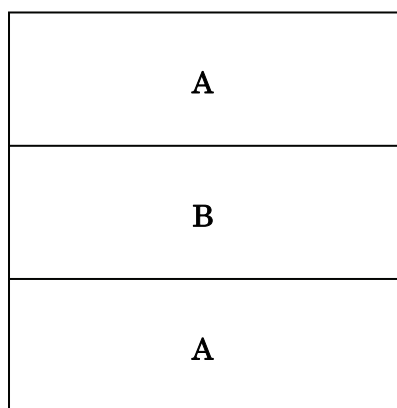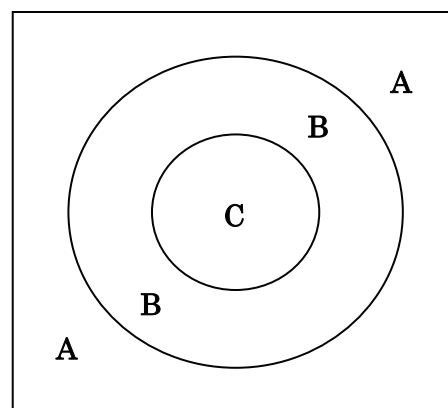

Fig. 1   Fig. 2

In traditional Japanese dialectology, the distributional type as sketched in Fig. 1 has been called '*ABA' distribution* (Shibata 1968). If again another new form 'C' was born within the area of 'B', its outcome should be the formation of 'ABCBA' distribution, then logically there could be manifold such variations, such as 'ABCDCBA' distribution.

The configuration sketched in Fig. 2 is actually a variety of 'ABCBA' distribution, where the oldest form 'A' is distributed outside, while the newest one 'C' emerges inside, with the intermediate stage form 'B' appearing in between. Japanese scholars have specifically referred to this type as *concentric distribution*. In this collection, a typical instance of the concentric type appears in Map 16-2 for "pica pica". It is seen on this map that the types E and F (comprising several kinds of folk-etymologically motivated forms, presented by symbols colored in red or orange) center around the Central Plain along the Yellow River, while it is surrounded by type B ("crow" type forms, presented by the symbols colored in green), which is distributed over vast areas. As discussed in detail in the commentary, it is legitimate here to assume the derivation of types E and F from type B.

If the area of a recent form 'B' in the 'ABA' distribution is the *kernel area*, this form is very likely to extend its territory toward its outskirts, resulting in the territory of an older form 'A' gradually diminishing and eventually disappearing altogether. Thus, logically the distribution 'ABA' should develop into 'BBB'. The term 'kernel area' refers to the politically, economically and culturally central area, usually comprising one or two big cities, from which linguistic innovation disperses consecutively to the outer areas. In the regional maps delimited in the explored area, we usually find one local city or town bearing the role of 'kernel city' (Grootaers 1957). Meanwhile, in the nationwide maps such as ALF and LAJ, the capital city Paris in France, and the older capital city Kyoto in Japan, indeed have been found to perform this particular role.

In China, however, the situation is much problematic since the present capital, Beijing, seems not to have performed this role, rather not a few maps give us an impression that the capital area is

'isolated' from its surrounds, so that it looks like a 'dialect island', as we will soon demonstrate such instances below. Instead, the role of the *Southern kernel area*, namely the *Jianghuai* area, is of high relevance as a center of lexical innovation and diffusion, as discussed in Iwata (1995, 2000, 2009). This area is situated between the Huai River and the lower reaches of the Yangtze and possesses the two kernel cities around its southern border: Nanjing and Yangzhou. This area, ever since the Six Dynasties era (222-589 AD.), has played the role of relay-station or bypass, through which linguistic features have been transmitted from the North to the South. During the pre-modern period, particularly since the beginning of the Ming Dynasty (1368-1644), linguistic radiation fanned out of this area has also been directed toward the West and the North. Some of our maps, such as 3-2 "hail", 6 "today", 24-2 "potato" and 33-2 "arm", exhibit the *Yangtze type* in terms of the distribution of a particular word form, indicating that lexical innovation born in this area is likely to have spread along the Yangtze River. Maps 37-2 "ankle" and 42-2 "navel" indicate that some word forms of Jianghuai origin crawled across the Huai River, eventually intruding into the Northern area.

In the process of the historical shift from 'ABA' to 'BBB', there could emerge a situation where the territory of the older form 'A' is seriously eroded by the newer form 'B', just being preserved in some limited zones or localities. In this collection, the term *remote distribution* is employed to refer to this situation.[3] For example, in Map 4-2 "morning", we can see that the forms containing the head *zhao* 朝 not only concentrate in the southern most area but also are witnessed in isolation in the lower Yangtze basin as well as in three localities on the Shandong Peninsula.

Since all of these terms, 'ABA', 'concentric' and 'remote', introduced above, in essence point to a single substantial, we will also express these varieties of geographical distribution through the use of the general term 'ABA distribution'.

4.2 Parallel change and migration theory

The idea of 'ABA distribution' is substantially the same as what was developed by Italian linguists in the 1920s and known in Europe as the *principle of continuity (contiguity)* or as the *principle of lateral (eccentric) areas* (Dauzat 1922:38-41, Grootaers 1982: 335-336). As a matter of fact, the counter-evidence to this 'principle' is not scarce (refer to Iwata1995: 208-210 for some Chinese evidence), so neither was the criticism for this scarce. They said that those Italian linguists "attempted to work with these principles as if they were 'laws', whereas they are really simply tendencies" (Chambers and Trudgill 1980:183-184). In our view, however, we are yet grounded to use this theory as one of the tools for interpreting the maps. This can also be exemplified by examining how the contradictory evidence is produced against the 'principle'. There are two conceivable cases for it: *parallel change* and *migration*.

(1)　Parallel change:　If the 'A' in the 'ABA' distribution were the outcome of innovation that occurred independently of one another on both sides of the area, then this principle should be considered to be invalid.

Two instances are given here. One is Map 005 cited in the Lexicon volume of Cao ed. (2008). In this map, lexical forms for "to rain" evince the simplest North vs. South opposition, *xiayu* 下雨 (literally "to come down rain") in the North and *luoyu* 落雨 (to fall rain) or *luoshui* 落水 (to fall

---

[3] This terminology is based on Yamaguchi (1992).

water) in the South, with its isogloss roughly coinciding with the course of the Yangtze River. It should be taken for granted that several localities around the isogloss are found to use both expressions, *xiayu* 下雨 and *luoyu* 落雨.[4] Curiously, however, according to our own data, such co-use of the two forms is also found in some localities within the Northern zone, far apart from the transitional area, scattered among the vast area that simply uses the single form *xiayu* 下雨. Then, one possible interpretation for this is that *luoyu* 落雨 is a retention of the old form while *xiayu* 下雨 is an innovation. However, since the verb *luo* 落 "to fall" is a basic verb which has a high frequency of usage either as a free form (e.g., *hua luole* 花落了 "the flower has fallen") or as a bound form in many compounds (e.g., *luo ye* 落叶 "fallen leaves"), it is also probable that Northern *luoyu* 落雨 is the product of recent innovation which occurred without any historical relationship with that distributed in South China.

In this collection, the probable case of parallel change appears in Map 27 "soybean", where the form *baidou* 白豆 is distributed remotely in the Northwest on the one hand and in the Southeast on the other, with the vast intermediate area being covered by some other forms, notably *huangdou* 黄豆. Judging from the fact that the form *baidou* 白豆 (white bean) tends to be adjacent in distribution with that of the more prevalent form *huangdou* 黄豆 (yellow bean) both in the North and South, we can say that speakers in two separate areas, without any contact, changed the color of this bean as it conformed to their vision or taste. Here it is also necessary for us to know about the dominant species (so the color) of the soybean cultivated in each area.

(2) Migration: If either side of 'A' in the 'ABA' distribution is the legacy of having been conveyed from one side of the area to the opposite side by migration, then the principle should be considered to be invalid.

The effect of migration often manifests itself in the formation of *remote distribution*, in which a lexical form 'A' emerges in one or some limited number of localities called 'dialect islands', far apart from its homeland where we find this form 'A' to have a broad and concentrated distributional area. Probable cases of it are the distribution of such PTH forms as *bingbao* 冰雹 and *jintian* 今天 which appear in Map 3-2 "hail" and Map 6 "today". Here these forms are observed to be scattered in the Northern area, including the capital Beijing, meanwhile they have a broad and concentrated distributional area in the Yangtze basin. It would be probable that these forms might have been conveyed via migration directly to the capital Beijing from the southern kernel area mentioned above, though the possibility of their being shipped via the Grand Canal should be considered. Logically, if these PTH forms extended their territory into the Northern area, it would have lead to the formation of 'ABA distribution'. Evidently, however, these words have not yet succeeded in their extension, thus giving us an impression that the dialect in the Capital, being isolated from its surrounds, looks like a 'dialect island'.

The influence of migration is generally disregarded in our interpreting the nationwide maps.

---

[4] According to the information provided by some speakers, the expression *xiayu* 下雨 is specifically used when 'it rains hard' while *luoyu* 落雨 is more generally employed for 'to rain', presumably indicating that *xiayu* 下雨 is a recent introduction from the Northern area. A parallel instance appears in Map 263 "earthquake" in LAJ (Linguistic Atlas of Japan); two expressions are coexistent in the Kyushu area in Japan, the original and old form *na'e* and the recent form *jishin*. In this case, the form *jishin* is used when a big earthquake occurs, while the form *na'e* denotes a relatively smaller earthquake.

However, migration theory is actually quite prevailing in Chinese dialectology as a convenient tool for accounting for dialect distribution. Not a few studies have attempted to establish this theory, based on the fragile authority available from historical documents, instead of the detailed field survey of historical and cultural legacy as demonstrated in Grootaers (1945). This is, at least in part, due to the lack of two critical views among Chinese scholars as well as Sinologists elsewhere in the world. One is the view saying 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. This is the truth that has always existed in rural farming societies everywhere in the world, where the majority of inhabitants scarcely had a chance to leave their homeland and their daily communication throughout their entire lives is mostly confined inside their villages or nearby. Another view that has been unshared by the scholars at present is that the dialect of immigrants inclines to be assimilated to the host dialect and basically fades away after three generations (Grootaers 1945, Huang 2005).

If we can exclude the possibility of these two cases, then, 'A' in the 'ABA distribution' will quite probably be older than 'B'. For example, the morpheme *zhao* 朝 we referred to above is philologically known to have been the one that constituted the Old Chinese (OC) lexicon and was used as a free form meaning "morning". But since it has lost its original usage in contemporary Northern dialects, just appearing in some compounds of a literary nature, it is improbable for the dialects in mutually inaccessible areas to produce the resembling words containing this particular morpheme without any contact. And in this case there is no evidence for the possibility of migration.

B. Mechanism of lexical change

Linguistic geography developed in the last century revealed the roles of some universal factors which could always be at work in lexical change, that could hereby destroy the 'linearity' of change (Iwata 1995a). The application of such factors to the changes is, unlike 'sound law', quite arbitrary, in other words, un-mechanical, therefore we have no 'principle' in accounting for each individual change; instead, we are obliged to treat such and such phenomena case-by-case. However, since the mechanism of changes itself is surely common across languages, it will be worthwhile for us to work with some sorts of generalizations in order to enhance further studies. This chapter will be devoted to such a purpose based on the evidence as reflected in the maps in this collection.

1. Arbitrariness of linguistic sign and motivation
1.1 Motivated and unmotivated forms

All linguistic signs in essence are arbitrary in terms of the relationship between 'signifié' and 'significant', but many words are only relatively arbitrary while some are absolutely arbitrary (De Saussure). Here are Chinese words which are employed for denoting the general categories of domestic animals/fowls:

niu 牛(cattle), *ma* 马(horse), *gou* 狗(dog), *zhu* 猪(pig), *ji* 鸡(chicken)

These monosyllabic words are absolutely arbitrary and are *unmotivated*. It is noted that the words for general categories mostly observe scanty dialectal variations, in other words, they are stable and

unchangeable. This can also be said for the words for natural phenomena as well, e.g., *feng* 风, *shui* 水, *xue* 雪, *lei* 雷. On the other hand, if there is a need to refer to a sub-category, compounds of a subordinate construction are employed. Here are Northern Chinese forms for the sub-categories of "cattle":

muniu 母牛(cow), *gongniu* 公牛(bull), *niudu* 牛犊(calf), *jianniu* 犍牛(castrated ox)

Since these compounds share a common head (e.g., *niu*) with some determiner (e.g., *mu* 母 "female", *gong* 公 "male", *du* 犊 "baby") either prefixed or suffixed, they are 'analyzable' and 'explicable' for speakers, so they are *motivated*. It is noted that such forms for sub-categorized semantic domains are generally rich in dialectal variation. In other words, these latter forms are unstable and changeable. Geographically, animal/fowl names evince a rather clear cut North vs. South opposition in terms of word formation type, namely head final and head initial structure (refer to Map 076 in the Grammar volume of Cao ed. 2008). For example, "cow" is called *muniu* 母牛 in Beijing, but it is *niuna* 牛嫲 in Hong Kong.[5] In addition to this binary feature, morphemes specifying the femininity or some other attributes are regionally very diverse.

Lexical change is, so to speak, a process of dialects' incessantly decreasing the arbitrariness of their words. This process is referred to as *motivation*, and this is exactly the main target of our analysis. The motivating process is comprised of two types; the individual type and the paradigmatic type.

1.2 Motivation: individual type (folk-etymology)

Individual type motivation occurs when one word is associated with another without any other paradigmatic relationship. This is actually one of the most frequent devices for motivating the word forms, the process of which is usually referred to as *folk-etymology*. For a simple example, in Map 3-1, the form *bingbao* 冰雹 for "hail", literally meaning "ice hail", should have been created from the unmotivated monosyllabic form *bao* 雹 via speakers' association with 'ice'.

Seemingly any unexpected and fantastic association can be permissible in the world of dialects. For example, some dialects in Zhejiang refer to the body part "knee" as *huxuntou* 猢狲头 "monkey head" (Map 40-1); some dialects in He'nan refer to the disease "malaria" as *laojian* 老犍 "castrated ox" (Map 43-4). This latter form actually dates back to the form *jianniu* 犍牛 cited in the previous section and in Maps 43-1 and 43-4 it is treated as the one describing 'the case of a malaria patient'. Presumably speakers associated the malaria patient with the behavior of castrated ox, thus creating a *personified* name *laojian* 老犍 by adding a prefix *lao* 老 "Mr.", and in this occasion they replaced the head with the determiner *jian* 犍 "castrated". Notably the end result of this change was the creation of a new stem, *jian* 犍.

If certain numbers of motivated forms, A, B, C, D etc., are distributed continuously as sketched below and the change is judged as having proceeded in this order, A > B > C > D, or vise versa, we refer to this type of distribution as *gradated distribution*.

A | B | C | D

---

[5] This peculiarity was discussed in Hashimoto (1978) in terms of 'word order', in which he argued for Altaic influence over the Southern and Thai substrata. However, the two constructions are actually coexistent in many dialects irrespective of North and South. For example, we can find in Northern dialects such words as *niudu* 牛犊, in which the head *niu* 牛 "cow/ox" precedes the determiner *du* 犊 "baby". This issue will be discussed in this volume with Map 19, though within a limited range.

For example, in Map 43-4, all symbols colored in green present the four 'personified' forms for "(to get) malaria", and these forms are distributed continuously from west to east; *niu* 牛(cow/ox), *laoniu* 老牛(old cow/ox) or *laojian* 老犍(Mr. Castrated), *laozhang* 老张(Mr. Zhang) and *laoye* 老爷 (old man), suggesting that the change might have proceeded in this order.[6] This map would also indicate that a well-motivated form, a well-recognized 'good' name, is likely to be accepted by dialect speakers. Therefore it is likely to spread over consecutive areas, while at the same time it is also likely to be deformed in the process of transmission, due to speakers' various associations with the object.

1.3 Motivation: paradigmatic type (analogy)

The paradigmatic type motivation refers to a change in which several words come to be integrated into a set paradigmatic system by sharing one particular element in common. For example, the morpheme *zhao* 朝, which originally denoted "morning", was also employed as a head of the word for "tomorrow", namely *mingzhao* 明朝, successively extended and used as well for "today", namely *jinzhao* 今朝. Eventually it was employed in some areas (refer to Maps 6, 7, 11, 13) for such related semantic categories as "the day after tomorrow" and "the day before yesterday". A similar process of extension can be seen in Map 21, where the morpheme *shi* 虱, which originally denoted "louse", came to also be employed as the head of the compounds for "flea" and "bedbug", e.g., *goushi* 狗虱 (literally "dog louse"), *bishi* 壁虱 (wall louse).

Significantly, these extensions were the products of *analogy*. It is usually induced by the decline of semantic and phonetic contents of the given element, the head of compounds in particular. It would have been impossible for the referent of the morpheme *shi* 虱 to be extended to "flea" and/or "bedbug" without eventually losing its original meaning, "louse". As for *zhao* 朝, this morpheme lost its original meaning "morning" once it was employed for "tomorrow", and subsequently must have 'declined' and changed to a fully arbitrary sign that expressed the abstract notion "day". The same can be said of a more frequent head *ri* 日, which originally denoted "sun" but had fallen in the process of decline once it came to be shared by the words relating to "day", e.g., *jinri* 今日"today", *mingri* 明日 "tomorrow".

Then, it will not be surprising if such morphemes have developed into affixes. From Maps 6 for "today", as well as in the other maps for the temporal words relating to day, we can presume that the head -*ri* 日 in Northern dialects changed to a diminutive suffix –*r* 儿 through the intermediate stage of syllabic -*er* 儿, e.g., *jinri* 今日> *jin'er*今儿> *jinr* 今儿. In the final stage of this change, the syllabic -*er* changed to a non-syllabic retroflex ending –*r* by its fusing with the preceding syllable, whereby the determiner *jin* 今, meaning "present", eventually took on the status of head. Paradoxically, the end result of this motivational process is the production of the monosyllabic unmotivated form.

Northern Chinese dialects have also produced prefixes through a motivational process. In Map 41, it can be seen that the Northern forms for such body parts as "arm", "elbow", "armpit", "knee" and "ankle" share a common prefix *ge* 圪 (*kɑk > kəʔ). For this prefix, we propose a hypothesis that it originates from some different substantial morphemes which resembled each other in their phonetic

---

6 Note here that the monosyllabic free form *niu* 牛 by itself can by no means express the meaning "malaria", but should be bounded to the verb *fa* 发"to get". Meanwhile, those compounds, *laoniu* 老牛, *laojian* 老犍, *laozhang* 老张 and l*aoye* 老爷, could denote "malaria" by themselves, but they are also likely to be bounded to the verb *fa* 发 or *da* 打.

shapes and that they were grammaticalized due to the function of analogy. Note that this prefix is also shared by the word for "flea", namely *gezao* 虼蚤 (refer to Map 19-1,2).

Meanwhile we have also observed that Northern dialects, as well as the dialects around the Yangtze basin, have created some particular morphemes which retain their original meanings but are rich in derivational ability, thus each forming a word group which is tentatively referred to as a *word family*. For example, the morpheme *guai* 拐 appearing in Map 38 can refer to at least four different objects; "elbow", "ankle", "cheek bone", and "hallux valgus", and all these body parts share a common feature 'protrusion'. Another instance of 'word' family is illustrated in Map 36.

1.4 Head formation

An end result of paradigmatic type motivation is that the determiner takes the position of head in exchange for the loss of the original head. This process is referred to as *head formation*. Two instances were introduced above; *laojian* 老犍 and *jinr* 今儿. Personification of animals and insects, as seen in the form *laojian* 老犍 for "castrated ox", is abundantly witnessed in Northern dialects.[7] For example, *laobian* 老扁 (literally "Mr. flat") appearing on Map 21-1 (bedbug) should come from the motivated form *biangezao* 扁虼蚤("flat flea") which is distributed adjacently with *laobian*.

'Head formation' is actually the process that has taken place repeatedly in the history of Chinese kinship terms. It is known that the morphemes *bo* 伯 and *shu* 叔 in the OC period served as the determiners meaning "elder" and "younger" in such compounds as *bofu* 伯父 ("father's elder brother", but literally "elder father") and *shufu* 叔父 ("father's younger brother", but literally "younger father"), but these determiners came to take on the status of a head in exchange for the loss of the stem *fu* 父 "father" by the MC period (Iwata 1995c: 210).[8] For the new stems thus created, then new determiners, *da* 大"elder" and *xiao* 小"younger", were prefixed, forming such compounds as *dabo* 大伯 and *xiaoshu* 小叔. Significantly the trend of such change has not become obsolete, even now, as we find not a few dialects which eventually saw the determiner *da* 大 become a head in exchange for the loss of the stem *bo* 伯 (Iwata 2000: 182).[9]

2. Phonetic conditions for motivation

The motivation process is often initiated via the inducement of some phonetic factor. In this section, the involvement of phonetic factors is discussed for both individual and paradigmatic types of motivation.

2.1 Phonetic conditions for the individual type of motivation

Creation of folk-etymology, at a glance, looks like an occasional happening caused by speakers'

---

[7] Grootaers (1957: 78-79) reported a fascinating instance he collected in the Hsüanhua region of North China, where one kind of grasshopper which has a knife-like tail is called *wangdaoshi* 王道士"the Taoist monk Wang". Here the homophony of the two words, *dao* 刀(knife) and *dao* 道(Tao), is utilized in creating this folk-etymology.

[8] In the case of female kinship, some new stems were created due to the phonetic fusion of the two syllables involved, instead of through the determiner changing its status to that of a head. The bi-syllabic form *shumu* 叔母(*ʃiuək mo) for "wife of father's younger brother" changed to monosyllabic *shen* 婶(ʃiəm＞ʂən); likewise *jiumu* 舅母(*kiəu mo) for "mother's elder/ younger sister" changed to *jin* 妗(*kiəm＞tɕiən). This process of phonetic fusion seems still productive in contemporary dialects, refer to Iwata (1995b:203).

[9] In contemporary dialects, this new stem, *da* 大, is again compounded, either by reduplication, namely *dada* 大大, or by prefixing the *Paihang* number to it, namely *er da* (second father), *san da* (third father) etc.

caprice. However, a careful observation of the form, as well as of the distributional characteristics appearing on the map, will reveal that some of such associations were brought about by the phonetic similarity of the two forms. For example, one of the Northern forms *geniu* 胳扭 for "elbow", appearing in Map 34-2, must have been created from the form *gezhou* 胳肘, which is geographically distributed around *geniu* 胳扭. Here it is evident that the word forms were motivated by the replacement of the unmotivated *zhou* 肘 with *niu* 扭, which means "twisted". From a phonetic point of view, this change was actually the replacement of merely one initial consonant, namely *tɕiəu>niəu, with vowels and the tones fairly unchanged. Taking another example, we find in Map 37-1 that the body part "ankle" is associated with "walnut" in some localities. It is assumed that this folk-etymology was produced based on the phonetic similarity between the form *huai* 踝[xuaʔ] for "ankle" and the form *he* 核[xəʔ] for "walnut".

2.2 Phonetic conditions for the paradigmatic type of motivation

Analogy, as a contributor to the paradigmatic type of motivation, is usually induced by the decline of semantic and phonetic contents of a particular element in the compound, actually the 'head' or 'stem' in many cases. In Northern Chinese, phonetic decline of such elements was brought about by the development of word accent, namely *stress*.

As reflected in the rhyming dictionary *Qieyun* (edited by Fayan Lu in 601 AD.), the total number of syllables existing in Middle Chinese (MC) amounted to approximately three times as many as the number that exists in the contemporary Beijing dialect.[10] This means that the phonological structure has been much simplified, and such a drastic change should lead to a drastic increase in the number of monosyllabic words that have become homophonous because of this phonological simplification. As a matter of fact, homonymic collision was usually avoided by means of producing a polysyllabic compound. For example, at the moment when the two bird names, *yan* 雁"wild goose"(MC: *ŋan) and *yan* 燕"swallow"(MC: *ien), tended to be homophonous, a remedy that the Northern dialects afforded with this collision was to prefix the modifiers *da* 大"big" and *xiao* 小"small" to each monosyllabic word.

Quite naturally, as the polysyllabic compounds increased, word accent developed, resulting in the distinction of the following stress patterns, as is observed in the present day Beijing dialect.

	Right stressed	Left stressed
Bi-syllabic structure:	Middle-strong (MS)	Strong-weak (SW)
Tri-syllabic structure:	Middle-weak-strong (MWS)	Strong-weak-weak (SWW)

Metrically each structure can form two kinds of *feet*, a right stressed foot (iamb or anapest) and a left stressed foot (trochee or dactyl). In forming a right stressed foot, the initial syllable automatically

---

[10] According to Li (1952), the number of syllables existing in the earlier version of the *Qieyun* text was 3,617, and according to Liu (1957), it has decreased to 1,256 in the present day Beijing dialect. In fact, since the *Qieyun* phonological system was somewhat artificially constructed, reflecting some archaic phonological features, it cannot be considered a synchronic system. However, we can assert that the number of syllables existing in the Tang *koine* was more than twice that of the contemporary standard language, as reflected in the commentary gloss for sutra complied by the Buddhist monk Huilin 慧琳 (737-820).

resumes its stress to a 'middle' level (Iwata 2005: 256-261).[11]

Note that accentual reduction is generally accompanied by tonal reduction in Northern dialects, and that in the left-stressed patterns ('SW' and 'SWW') the entire tonal contour of the polysyllabic structure is uniquely determined and can be predicted by the tonal feature of the 'strong' syllable, since tonal distinctions are neutralized in the 'weak' syllable, resulting in a presumed 'zero' tone (Iwata 2001:275; Iwata 2005: 256-257).[12]

Under this environment, Northern bi-syllabic words of higher frequency tended to bear a 'SW' stress, even facilitating the *suffixation* of their heads (in other words, 'head formation'), e.g., *jinri* 今日> *jin'er* 今儿> *jinr* 今儿. Taking another example, we find in Map 19-1 that the stem *zao* 蚤 in the word *gezao* 圪蚤 has changed to the diminutive suffix *zi* 子 in some Northern localities. Here the change was triggered by means of analogy with such semantically related words as *shizi* 虱子 "louse", but there are two necessary conditions for facilitating this change: the 'SW' stress and the common initial consonant [tsɿ] shared by *zao* 蚤 and *zi* 子.

As reflected in the loss of the head *fu* 父 for the kinship terms *bofu* 伯父 and *shufu* 叔父, it is probable that a sprout of the trochaic foot might have existed as early as the MC époque, or even earlier. Meanwhile, the iambic foot may have been a coherent feature which existed at some point in Chinese compounds, and would have contributed to prefixation of the initial morpheme. The prefix *ge* (*kɑk), to which we referred in Section 1.3 above, might have been a product of 'MS' stress, in which the vowel distinction existing in the initial syllables, namely *kɑk, *kiɑk and *kæk, should have tended to be neutralized by the effect of 'middle' stress (refer to the commentary for Map 41).

The right-stressed tri-syllabic pattern 'MWS' was most likely to trigger analogical changes. This is typically seen in the creation of the infix *le*. We assume that this morpheme might have come into being due to the segmental reduction of the 'weak' syllable in some tri-syllabic words. In Map 40-4, we hypothesize that the Northern form *bo luo gai* [pə luə kai] for "knee" might have originated from a form *[pɑk siət kai] existing in the pre-modern period, due to the segmental reduction of the stem *siət. Once such an infix was established in the dialects, some other semantically related words began to be incorporated in the particular word group due to analogy, so that the word *ge zhi* for "armpit" changed to *ge le zhi* (Map 39-2) and the word *ye gai* for "forehead" changed to *ye le gai*. In this particular case of analogy, however, bi-syllabic forms became tri-syllabic, and we should explain how such a change could have been possible.

3. Phonetic attraction: paronymic and analogical

Admitting the involvement of phonetic factors in analogical change, we so far have presupposed that any semantic or morphological factor is indispensable in its occurrence. As an instance which we referred to above, the word family *guai* 拐 ("elbow" "ankle", "cheek bone" and "hallux valgus") in Northern dialects is semantically characterized by its members' sharing a common feature 'protrusion'.

---

[11] Iwata (2005: 256-261) argues that the left-stressed pattern constitutes a 'phonological word' and the right-stress pattern constitutes a 'phonological compound'.

[12] According to Hayata (1999), we refer to such a tonal unit as *word tone*, which is defined by the rule 'T1+T2→T1'. The Wu dialects, such as Shanghai and Suzhou, exhibit exactly this particular tonal characteristic, but they are distinct from the Northern dialects in that tonal reduction is not linked with accentual reduction, so that even the 'zero' tone is realized in a long syllable with a clear quality of vowel (Iwata 2001: 281).

Curiously, however, we often encounter the case where analogy seems to have occurred without any semantic motivation.

In some Northern Chinese, the use of the *le* infix has extended to semantic domains other than the body parts, e.g., *ge le niu* [kə lə niu] (Map 17-2 "snail"), *e le zhu* [ə lə tʂu] ("spider"), *jie le hou* [tɕie lə xəu] ("larva of cicada").[13] Furthermore, some words share not only the infix *le* but also the prefix *ge,* e.g., *ge le zhi* ("armpit"), *ge le pai* (Map 40-3 "knee"), *ge le* ("mussel"). Let us call this particular word group 'word family *ge le*'. While it is true that these forms are the product of analogy, how did these forms employed for different semantic categories come to share a common element *ge le*?

Chinese etymological study so far, in line with that established by the distinguished Qing philologist Yaotian Cheng 程瑶田 (1728-1804), argued that all such referents share a common semantic feature 'rounded shape' and that all such forms should derive from a word family called *guoluo* 蜾蠃, which Cheng assumed to have existed 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 Cheng or by changing their clothes to the comparative method, we dare to assert that the truth does not lie in the ancient world, but rather in the lively event that is performed in the world of contemporary dialects.

Linguistic geography tells us that words can be attracted to each other unbiased by semantic factors. We refer to such a phenomenon as *phonetic attraction*, which we tentatively divide into two types: *paronymic attraction* and *analogical attraction*.

3.1 Paronymic attraction

Paronymic attraction can occur to words which resemble each other in their phonetic shapes but denote different referents. For example, it is assumed for Map 16-2 that the forms for "pica pica" underwent a series of lexical changes in the Northern area (the area where the orange colored symbols are concentrated): *yaque* 鸦鹊 or *yeque* 野鹊 > *mayeque* 麻野鹊 > *mayique* 麻衣鹊, 麻尾鹊 > *mazha* 麻喳. The lineup that affected and attracted the forms for "pica pica" were *maque* 麻雀 for "sparrow", *mayi* 蚂蚁 for "ant", and *mazha* 蚂蚱 for "grasshopper". Even though the involvement of any folk-etymological association may be assumed for some phases of change, these changes must have been triggered primarily by phonetic attraction.

One may probably insist on the commitment of semantic factors in paronymic attraction. However, such an argument is faulty (Dauzat 1922: 74). In some Northern dialects, the forms for "navel" and "Chinese water chestnut" (Eleocharis dulcis, 荸荠) after all have become homophonous. For this phenomenon, Wei (2004), without knowing about the function of phonetic attraction, hypothesized that the forms for "water chestnut" were created from the form for "navel" due to the resemblance of the configuration between the particular two objects. However, our Map 42 suggests the possibility of the contrary change; it was the form for "navel", namely *boqi* 脖脐, that underwent

---

[13] The vowel of the *le* infix actually varies depending on the particular dialect, so it could be [a] or [o], instead of [ə]. This means that the affix assuming a 'weak' stress is generally unstable in its phonetic realization, and due to this reason it is also likely to be deformed by undergoing the effect of folk-etymology and phonetic attraction. So the form for 'spider' [ə lə tʂu] has changed to [ə lan tʂu] in some dialects and the form for 'armpit' has changed from [kə lə tʂʅ] to [kə lao tʂʅ] in a number of dialects, with the latter's second syllable resuming a 'strong' stress. Meanwhile it is also true that the *le* infix is likely weakened or even disappears, so that the form *ge le pai* for "knee" changed to *ger pai* in several localities.

modification, and it changed to such forms as *buqi, puqi*, and *biqi*, due to the paronymic attraction by the form for "water chestnut", namely *biqi* 荸荠 as well as its variants. Ota (2005, 2007) demonstrated some irrefutable evidence in favor of our assumption, e.g., the form for "star", *xingsu* 星宿, was attracted by and changed to the form denoting "new bride", *xinxifu* 新媳妇; so was the form for "Chinese inkstone", *yanwa* 砚瓦, by the form for "King of Hell", *yanwang* 阎王.

For all words referred to above in this section, it is assumed that a common factor that induced the resultant paronymic attraction might have been the formation of a trochaic 'SW' pattern. The morpheme assuming a 'weak' stress is generally unstable in its phonetic realization (refer to Fn.13), and subsequently the whole word is inclined to be influenced by the other words. This implies that such a word becomes *unmotivated* and needs to resume its motivation by any means. Phonetic attraction is one of the plausible remedies that dialects can afford the word at this moment.

A natural outcome of paronymic attraction is the crisis of homonymic collision. For example, in Map 17-1 & 4 ("snail") we observe forms incorporating the morpheme *e* 蛾[ə], which also emerges as the form for "butterfly" in Map 18. Based on this evidence, we assume that the form for "snail", *wo niu* 蜗牛[uə niu], in which the second syllable presumably assumed a 'weak' stress, was attracted by the form for "butterfly", typically *e zi* 蛾子[ə tʂ]. And as a result, the two words eventually became homophonous. The end result of this collision will be mentioned in Section 4-1 below.

The crisis of homonymic collision triggered by paronymic attraction can be evaded by the given words' forming a new word group, thus being paradigmatically motivated in the end. For example, names for "gecko", "bat" and "ant" have come to share a common element, *bie fu or bie hu* in some Northern dialects, as the result of paronymic attraction (Iwata 1996), e.g.,

Datong (Shanxi Province)   *ye bie fu* "bat" / *ma bie fu* "ant"
Dezhou (Shandong Province)   *yan bie hu* "bat" / *xie bie hu* "gecko"

The form *bie fu or bie hu* undoubtedly came from the words *bi hu* 壁虎"gecko", *bian fu* 蝙蝠 "bat" *and pi fu* 蚍蜉"big ant", which all must have assumed a 'SW' stress. Consequently, they began to attract each other, but once they tended to be homophonous, the dialects afforded them a remedy for evading the crisis of homonymic collision, namely to prefix any meaningful modifier to each word, e.g., *ye* 夜"night" for "bat".

3.2 Analogical attraction

The formation of the 'word family *ge le*' may be explained in terms of the combined effect of analogy and phonetic attraction, which we tentatively refer to as *analogical attraction*. As mentioned above, the prefix *ge* and the infix *le* were the products of relatively recent innovation that occurred in Northern Chinese, and each element respectively must have diffused across semantic categories, eventually coming to be shared by many nouns. This is exactly due to the function of analogy. However, it was not the end of the change, and crucially those words having the *ge* prefix and those having the *le* infix began to be attracted to each other. For example, in Map 40-3, two types of the forms for "knee", p-l-k and k-l-p, e.g., *bo luo gai* [pə luə kai] and *ge le bai* [kə lə pai], are seen to form a concentric distribution, and based on this evidence we can judge that the p-l-k type changed to k-l-p. This change is accounted for in terms of 'metathesis', but more significant is the fact that such forms as *bo luo gai* were attracted and motivated by the other words initiated by the prefix *ge*.

Likewise, the epenthesis of the infix *le* to the form *ge zhi* for "armpit" is explained in terms of

analogical attraction by the other words having the *le* infix. Again, a phonetic factor that induced this change was the decline of the second syllable *zhi*, which originally assumed a 'strong' stress in the form *ge zhi*, but was eventually weakened to assume a 'weak' stress.

Another instance of analogical attraction is found in the temporal words which take such suffixes as possessed by pronouns, e.g., *jinrge* 今儿个, *jinmen* 今们 appearing on Map 6 "today". A more frequent one suffixed to the preceding component is the general classifier *ge*, which also appears as a suffix of demonstrative pronouns, such as *zhe ge* 这个 "this" and *na ge* 那个 "that". Another suffix, *men* 们, which is less frequent, but concentrates in distribution around west Shandong, is evidently the suffix of personal pronouns, such as *wo men* 我们 "we" and *ni men* 你们 "you (plural)".

Here the discussion is focused on the suffix *ge*. Again the change was induced by a decline in the phonetic and semantic content of the head, namely *ri* 日. As mentioned above, the temporal words having this head should have borne a 'SW' stress. At the moment of change, while one choice for the dialects was the road to the suffixation of this head *ri* 日 itself, an alternative choice was to add the suffix *ge* 个 to such forms as *jinri* 今日, thus forming a "MWS" pattern. Actually, the majority of dialects seem to have chosen both options, as seen from the coexistent two forms which are witnessed in a number of dialects, e.g., *jinrge* 今儿个 and *jinr* 今儿 in Beijing.

There should have been no reason for temporal words to be analogized with pronouns unless some other factor got involved in the process of analogy. It must have been the phonetic attraction of particular words that facilitated this change, e.g., *jin ri* 今日 > *jin ri ge* 今日个. Iwata (2007b) suggests that the most probable candidates that attracted these temporal words were pronouns of tri-syllabic structure, namely *zhe yi ge* 这一个 "this" and *na yi ge* 那一个 "that", which assumed a 'MWS' stress.[14] Here a condition that triggered this attraction might have been a phonetic similarity between the head *ri* 日 in the temporal words and the second element *yi* 一 in pronouns.[15]

A succeeding change was once again the decline of the head *ri* 日, which should have assumed a 'weak' stress at the second syllable of a 'MWS' pattern; *jin ri ge* > *jin er ge* > *jinr ge*. Here the suffix *ge* should have carried a 'heavy' stress originally, but it gradually declined and changed to carry a 'weak' stress, and the whole word changed to have a bi-syllabic structure, *jinr ge*, which assumed a 'SW' stress.

4. 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 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 so-called "dialect contact".

It is proposed that word collision is of two types: *homonymic collision* and *synonymic collision*.

---

[14] 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 came 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*.

[15] The particular two forms, *ri* and *yi*, were originally identical in rhymes (MC: *n̠iĕt / *iĕt). Moreover, since they both carried a 'weak' stress in the words, the phonetic distance between the two should have been minimized.

In the following, word forms are represented by *P* and *Q*, their semantic content 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).[16]

4.1 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 defeated words. Suppose 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. Here an arrow represents the direction of structural pressure, and a greater-than sign represents the change.

1. The victory of "x" over "y": *P*(x)→*P*(y) > *P*(x) / *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:

   ⟦*P*(x) | *P*(y)⟧ (the rectangle symbolizes an explored area and '|', indicates an isogloss)

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"(笔) 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).[17] In Map 21, we argue for the possibility that the word for 'bed bug' fell victim to this particular word *bi*, again as the result of phonological change. Seemingly, taboo words relating to sex are always stronger than others.[18] 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", *ying zi* 蝇子, came to be homophonous with that for "mosquito", i.e., *wen zi* 蚊子, presumably due to paronymic attraction (MC: 蝇*jiəŋ/蚊*myən), refer to Map 038 in the Lexicon volume of Cao ed. (2008). In this case, the distinction is generally maintained in terms of adding some sort of modifier to each, e.g., *ye wen zi* 夜蚊子(literally "night mosquito") in contrast with *fan wen zi* 饭蚊子(rice mosquito). Meanwhile, curiously some dialects seem to make no distinction between the two referents in terms of their phonetic shapes.

Case 3 often manifests itself as the result of Case 1. In Section 3-1 above, we mentioned that the form for "snail" was attracted to that for "butterfly", thus the two forms eventually came to be homophonous. In this conflict, "snail" gained victory over "butterfly", and as a result, *e zi* 蛾子, as well as those forms containing the stem *e* 蛾, scarcely appear in the area where it is employed for "snail" (compare Maps 17-4 and 18), while it still secures a broad territory in the surrounding area.

A complementary pattern of distribution could emerge as the result of *referential shift*. An instance appeared in Map 10, one of the 'referent maps' in this collection. The two forms, *yelai* 夜来

---

16 Formalization adopted here is after Mase (1992).

17 In this case, the tonal categories of these words are of relevance. However, this information is omitted for the sake of simplicity.

18 Some counterevidence was offered by Zhang (2006).

and *yeli* 夜里, denote "evening" or "night" in Central China (in the lower and mid reaches of the Yangtze), whereas in the North these forms mostly denote "yesterday". This implies that the referent of these forms shifted from "evening" or "night" to "yesterday". The occurrence of referential shift is quite frequent in Chinese kinship terms (Iwata 2000).

A complementary pattern of distribution could also be created owing to accidental changes that occurred on both sides of an isogloss. Refer to Mase (1992) and Iwata (2006) for such instances.

A rare case of homonymic collision triggered by an external factor is displayed in Map 29. 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* 大豆(big bean) and the latter *xiao dou* 小豆(small bean). This partnership, however, was destroyed due to the introduction of a new species of "broad bean", which was as large as that we eat nowadays, into the Northwestern area through the Silk Road in approximately 1200 AD. 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 his garment into colorful ones, such as *huang dou* 黄豆(yellow bean) and *bai dou* 白豆(white bean).

4.2 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 "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* : $Q\,(x) \rightarrow P\,(x) >$ disappear
2. Dividing the semantic field or usage between *P* and Q without changing referent: $P(x1)\,/\,Q(x2)$
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, in the southern reach of the Yangtze there are not a few dialects (Shanghai, Suzhou, Nanchang etc.) which co-use the two kinship stems that refer to one's father, *die* 爹 and *ye* 爷. Significantly, however, the usage of the two is differentiated: *die* for vocative use and *ye* for designative use. Based on the geographical evidence, Iwata (2000: 192) assumes that this situation was an outcome of the strife between the original form *die* and the recent form *ye*, which should have competed for a single referent "father", eventually reaching a compromise by differentiating their actual usage. Refer to Fn. 4 above for the case of the two competing forms denoting 'to rain', *xiayu* 下雨 and *luoyu* 落雨.

Here it is worth mentioning that Case 2 is by no means special in Chinese, rather this is exactly the phenomenon which researchers have discussed in terms of 'multiple readings' of Chinese characters. 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 fields 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 newly introduced products or notions, e.g., [smoke+fog]/÷2=smog. In the world of a dialect, it is usually produced by the contact of two forms which are continuously distributed but are separated by an isogloss. In Map 40-4, we will specifically discuss this issue, arguing for the existence of various types of word blending created in different areas and in different epoques. The frequent occurrence of such forms suggests that fierce dialect contact has been repeated everywhere in China (Iwata 2007a).
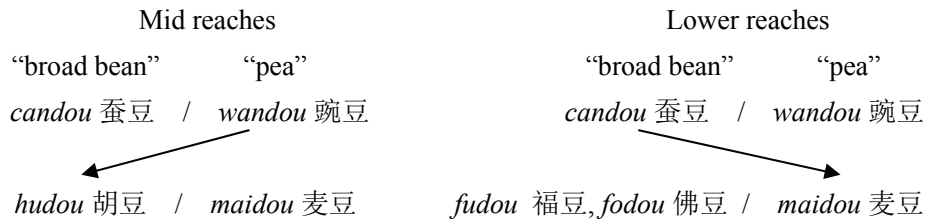
4.3 Referential shift

In the preceding section, we mentioned that the referents of the forms *yelai* 夜来 and *yeli* 夜里 shifted from "evening" or "night" to "yesterday" in North China. However, this statement just points to the result of change, and a *referential shift* might quite probably have been a *referential extension* at the beginning, substantially the same process as homonymic collision, which is formalized as $P(x) \rightarrow P(y)$. As a matter of fact, what we usually witness is that the original territory of $P(x)$ has already been occupied by a new form $Q(x)$, whereby $P(x)$ does not emerge on the map. This implies that the same process as synonymic collision, $P(x) > Q(x)$, occurred internally in the dialect, without the inducement of external factors. When the forms *yelai* 夜来 and *yeli* 夜里 shifted to denoting "yesterday", the semantic category "evening" or "night" was filled up by the new forms, such as *heiye* 黑夜 and *houshang* 后晌, refer to Map 5-1. The situation is summarized as follows;

$$Q(x) \quad \rightarrow \quad P(x) \quad \rightarrow \quad P(y)$$
heiye 黑夜(evening) → *yelai* 夜来(evening) → *yelai* 夜来(yesterday)

Two explanations are possible to account for this change. One, since the referent of *yelai* 夜来 shifted from "evening" to "yesterday", the dialect filled up the semantic *case de vide* by creating a new form *heiye* 黑夜. This can be compared to a 'pull chain' in historical phonology. Two, since a new form *heiye* 黑夜 came into being for "evening", the structural pressure forced the form *yelai* 夜来 to move to the category "yesterday". This can be compared to a 'push chain' in historical phonology. Considering the evidence that Chinese forms for "evening" and "morning" are likely to change to denote the categories "yesterday" and "tomorrow", a tendency even observable in contemporary Northern dialects, presumably the first explanation may point to the truth. At any rate, a crucial result which was brought about by this change was that the original form for the category "yesterday", presumably *zuori* 昨日, was ultimately expelled from the lexical system forever.

Referential shift could be triggered by an external factor, namely dialect contact. The most fascinating instance in this respect will be demonstrated in Map 32, where the referent of the form *candou* 蚕豆, which denotes "broad bean" in the North, has shifted to "pea" in the lower reaches of Yangtze river, whereas the form *wandou* 豌豆, which denotes "pea" in the North, has shifted to denote "broad bean" in the mid reaches of the river. Note here that the two beans in question form a partnership in terms of the cultivation season, i.e., both are autumn seeding and spring cropping in the Yangtze basin.

Though seemingly quite unusual reversals, what is of significance here is that they occurred exactly in the areas where Northern and Southern forms clashed with each other, as schematized below.

	Mid reaches			Lower reaches	
"broad bean"		"pea"		"broad bean"	"pea"

candou 蚕豆 / *wandou* 豌豆   *candou* 蚕豆 / *wandou* 豌豆

hudou 胡豆 / *maidou* 麦豆   *fudou* 福豆, *fodou* 佛豆 / *maidou* 麦豆

As a whole, the Northern force was superior to the Southern one, as seen in the evidence that the latter was never intruded into the Northern area. However, we can fairly say that Northerners' victory was only partial both in the lower and mid reaches of the river. This was because they encountered a fierce resistance of Southerners, who wished to protect their precious names from the menace of Northerners; lower reaches' *fudou* 福豆 (lucky bean) or *fodou* 佛豆 (Buddha bean) for "broad bean" and mid reaches' *maidou* 麦豆 (wheat bean) for "pea". This latter form fell as a sacrifice to Northern *candou* 蚕豆 in the lower reaches of the river, while in the mid reaches it survived at the sacrifice of an ominous word *hudou* 胡豆, which was eventually replaced by Northern *wandou* 豌豆.

From the side of Northerners, the conversion of referent was a sort of strategic compromise, by which they avoided encountering a strong defensive line, instead they attacked a relatively weaker point of the Southerners. Curiously, Northern *wandou* 豌豆 in the lower reaches and Northern *candou* 蚕豆 in the mid reaches seem to have been negative in attacking the Southern forms. In our view, this would suggest that dialectal forces merely concerned themselves with the particular words they attacked and defended, i.e., "broad bean" in the lower reaches and "pea" in the mid reaches.

All of the above represents an attempt to find regularity within a myriad of apparent irregular phenomena. Whether or not our proposed regularity represents a suitable explanation for the data still awaits confirmation from the fruits of future development in linguistic geography research.

(Ray Iwata)

References

Astrakhan, E., Zavjalova, O. and M.Sofronov.1985. *Dialekty i Nacionaljnyj Jazyk v Kitae* (Dialect and National language in China). Academy of Science, USSR.

Buck, Carl Darling. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. The University of Chicago Press.

Cao, Zhiyun 曹志耘 ed. 2008. *Linguistic Atlas of Chinese Dialects* (汉语方言地图集). Beijing, Shangwuyin Press.

Chambers, J.K. and Trudgill Peter. 1980.　*Dialectology*. Cambridge University Press.

Dauzat, Albert. 1922. *La géographie linguistique*. Paris, Librairie Ernest Flammarion. [Japanese translation by H. Matsubara and K.Yokoyama (1958) *France Gengo Chirigaku* フランス言語地理学, Tokyo, Daigaku Shorin].

Gilliéron, Jules and Mario Roques. 1912. *Etudes du Géographie Linguistique d'après I'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 ALF による言語地理学的研究, *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, Premiere Partie: La method 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. 1982. "Dialectology and Sociolinguistics: A general survey". *Lingua*, Vol. 57, 327-355.

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* 中国の方言地理学のために (Contributions to Chinese dialect geography). Translated by R. Iwata(岩田礼) and M. Hashizume(橋爪正子), Tokyo, Kobun Press.

Grootaers, Willem A. 2003. *Hanyu fangyan dilixue* 汉语方言地理学 (Contributions to Chinese dialect geography). Translated by R-J. Shi (石汝杰) and R. Iwata (岩田礼), Shanghai Jiaoyu Press.

Hashimoto, Mantaro 橋本萬太郎. 1978. *Gengo ruikei chiri ron* 言語類型地理論 (Typo-geographical linguistics). Tokyo, Kobundo.

Hayata, Teruhiro 早田輝洋. 1999.　"Accent and Tone: Towards a General Theory of Prosody". In Kaji ed., *Proceedings of the Symposium, Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*, ILCAA (Tokyo University of Foreign Studies).

Huang, Xiaodong 黄晓东. 2005. "Duo fangyan qu de yuma xuanze yu zhuanhuan 多方言区的语码选择与转换"(Code-selection and code-switching in the multi-dialects speaking district). *Contemporary Research in Modern Chinese* 现代汉语研究, No.7, Tokyo, 128-137.

Iwata, Ray 岩田礼. 1988. "Koso Anki ryo-sho ni okeru shinzoku sho'i keisiki no chiriteki bunpu to

ko-shou'i taikei no saikou 江蘇・安徽両省における親族称謂形式の地理的分布と古称謂体系の再構" (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* 漢語史の諸問題, Research Institute of Humanities, Kyoto University, 207-272.

Iwata, Ray. 1989. *Zhongguo Jiangsu, Anhui, Shanghai liang sheng yi shi jingnei qinshu chengweici de dilifenbu* 中国江苏・安徽・上海两省一市境内亲属称谓词的地理分布 (The geographical distribution of kinship terms in Jiangsu and Anhui Provinces and in Shanghai City). *Kaipian* 开篇, *Monograph Series*, No.1，Tokyo, Kobun Press.

Iwata, Ray. 1995a. "Kango hogen shi no fu-renzokusei 漢語方言史の不連続性" (On the non-linearity of the history of Chinese dialects). *Studies in Humanities* 人文論集, Faculty of Humanities, Shizuoka University 静岡大学人文学部, 45-2, 43-77.

Iwata, Ray. 1995b. "Hanyu fangyan "zufu" "waizufu" chengwei de dili fenbu – Fangyan dilixue zai lishiyuyanxue yanjiu shang de zuoyong 汉语方言"祖父""外祖父"称謂的地理分布－方言地理学在歷史语言学研究上的作用" (The geographical distribution of the kinship terms zufu 祖父 and waizufu 外祖父 in Chinese dialects). *Zhongguo Yuwen* 中国语文, 1995-3, 203-210.

Iwata, Ray. 1995c. "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. 1996. "Du Shandong fangyan zhi liu zhong 读《山东方言志》六种" (A book Review of the six dialect reports of Shandong Dialects). *Zhongguo Yuwen* 中国语文, 1996-3, 236-240.

Iwata, Ray. 2000. "The Jianghuai area as a core of linguistic innovation and diffusion: a Case of the kinship term 'ye 爺'". *In Memory of Professor Li Fang-Kuei: Essays of linguistic change and the Chinese dialects*, University of Washington /Academia Sinica, 179-196.

Iwata, Ray. 2001. "Tone and Accent in the Chinese dialects". In S. Kaji ed., *Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese, Accentology and Other Topics*, ILCAA, 267-291.

Iwata, Ray. 2005. "Comments on 'The Tone-Syntax Interface in Chinese: Some Recent Controversies' by San Duanmu". In S.Kaji ed., *Cross-Linguistic Studies of Tonal Phenomena: Historical development, Tone-Syntax Interface, and Descriptive Studies*, ILCAA, 255-267.

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 方言接触及混淆形式的产生" (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 汉语方言"明天"、"昨天"等时间词的语言地理学研究" (Geo-linguistic study of time words, such as "tomorrow" and "yesterday", in Chinese dialects). *Chuugoku Gogaku* 中国語学 (Bulletin of the Chinese Linguistic Society of Japan), No.254, 1-28.

Iwata, Ray. 2009. "Chinese Geolinguistics: history, current trend and theoretical issues". *DIALECTOLOGIA*, Special volume. (In press)

Li, Rong 李荣. 1952. *Qieyun Yinxi* 切韵音系 (The sound system of the *Qieyun*). *Yuyanxue zhuankan* 语言学专刊, Vol. 4, Chinese Academy of Sciences.

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. [Chinese translation: "Jinji ci juli 禁忌词举例", *Fangyan* 方言，1994-3，161-169].

Liu, Zexian 刘泽先. 1957. "Beijinghua li jiujing you duoshao yinjie?北京话里究竟有多少音节?" (Exactly how many syllables does Pekingese have?). *Zhongguo Yuwen* 中国语文, 1957-2,3, 1-8, 17-23.

Mase, Yoshio 馬瀬良雄. 1992. *Gengo chirigaku kenkyu* 言語地理学研究 (Studies in Linguistic Geography). Tokyo, Ohu Press.

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

Ota, Itsuku 太田斎. 2005. "'Xifu' ga kakawaru minkan gogen sandai – Chuugokugo seihoku hougen no rei wo chuusin ni "媳妇" が関わる民間語言三題－中国語西北方言の例を中心に" (Three Case Studies on Folk Etymology: "xifu" 'wife' in Northwest Mandarin). *THE KOBE GAIDAI RONSO* 神戸外大論叢 (The Kobe City University Journal), Vol. 56-7, 15-41.

Ota, Itsuku. 2009. "Beifang fangyan limian suo neng jiandao de leiyin qianyin deng yuyin xianxiang – yi Shanxi fangyan de 'yanwa' wei li   北方方言里面所能见到的类音牵引等语音现象－以山西方言的"砚瓦"为例" (Paronymic Attraction and Other Paradigmatic Phonetic Changes in Northern Chinese Dialects ——"Inkstone" in Shanxi Dialects——). *Di san jie Jin fangyan guoji xueshu yantao hui lunwenji* 第三届晋方言国际学术研讨会论文集 (Proceedings of the 3rd International symposium on the Jin dialect). (In press)

Shibata, Takeshi 柴田武. 1969. *Gengo chirigaku no hoho* 言語地理学の方法 (Method in linguistic geography). Tokyo, Chikuma Shobo.

Wang, Shiyuan 王士元 and Zhongwei Shen 沈钟伟. 1992. "Fangyan Guanxi de Jiliang Biaoshu 方言关系的计量表述" (A quantitative description of the relationship among Chinese dialects). *Zhongguo Yuwen* 中国语文, 1992-2, 81-92.

Wei, Gangqiang 魏刚强. 2004. "' biqi' kao '荸荠'考" (An etymological note on "biqi"). *Fangyan* 方言 2004-3, 260-266.

Yamaguchi, Koyo 山口幸洋. 1992. "Chuukyori enkaku bunpu no hougen 中距離遠隔分布の方言" (The dialect of the middle-distance remote distribution). *Kokugogaku* 国語学(Studies in the Japanese Language), No.168, 26-39.

Zavjalova, O. and E. Astrakhan.1998. *The Linguistic Geography of China.* Progressive Report, Vol. 1, Grant-in-Aid for Scientific Research (A), 1997-1999, Directed by Mitsuaki Endo 遠藤光暁, Linguistic Geography & Cultural-Natural Geography in China.

Zhang, Yanlai 张燕来. 2006. "Xibei Fangyan de Jihui Gaidu Xianxiang 西北方言的忌讳改读现象" (Taboo pronunciation avoidance in Northwest dialects). In X-D Xing 邢向东 ed. *Xibei fangyan yu minsu yanjiu luncong* 西北方言与民俗研究论丛(Studies in Northwestern dialects and folklore research), No. 2, 308-311.