Homonymic and Synonymic Collisions in the
Northeastern Jiangsu Dialect*
- On the Formation of Geographically Complementary Distributions -

Ray Iwata
Kanazawa University

This paper is devoted to the classical issue of linguistic geography, “mots en collision (words in collision)”, based on the linguistic survey carried out in the dialect border area in northeast Jiangsu. Two complementary distributions are observed: one is between Map 2 (knee) and Map 3 (clod) with regard to the single form “kʰə tʰəu”, and the other between Map 4 (elbow) and Map 5 (shoulder) with regard to a set of synonyms, “kʰy” and “kuai”, both meaning “curved”. These are elucidated in terms of conflict between the two dialects, northern and southern, and are interpreted as resulting from homonymic and synonymic collisions. It is claimed that Chinese etymological study should be established on the basis of linguistic geography.

Key words: linguistic geography, homonymic collision, synonymic collision, phonetic attraction, complementary distribution, etymology

1. Words in collisions

The problem of “mots en collision (words in collision)” was a matter of the greatest concern among the founders of linguistic geography, Jules Gilliéron and his students (Gilliéron and Roques, 1912, Dauzat 1922). Sound change is a potential creator of homophonous words, and could bring about linguistic confusion among speakers, who,

* Earlier versions of this paper were presented at the CEALC regular meeting (Tokyo, March 2003) and at the lecture delivered in the Beijing Language and Culture University (Beijing, June 2004). I am indebted to the organizers, Mitsuaki Endo and Zhiyun Cao, as well as the participants, for their helpful comments. My thanks are also due to the predecessors of linguistic geography in Japan, in particular, Ookawa et al. (1991-7) and Matsubara & Yokoyama (1958). Thanks to them, we Japanese students can read the Japanese translations of the classical works by Gilliéron and Dauzat. The maps appearing in this paper are from a Window-based map drawing system, SEAL (System of Exhibition and Analysis of Linguistic Data). I thank the authors, Chitsuko Fukushima and Yusuke Fukushima, who kindly provided me with this system.
in reacting against this situation, avoid homonymic collisions by adopting various strategies. This process, which Gilliéron called “verbal pathology and therapeutics”, can be viewed as a self-protecting function of language, similar to medical treatments for the human body. To cite a well known instance from Gilliéron, the Latin (Vulgar Latin) word [gallus] meaning “cock” changed to [gat] in south France according to the regular sound change, -ll- > -ll > -t, and eventually this change resulted in a homonymic collision of this word with that for “cat” (cattus > cat > gat). The winner of this battle was “cat”, owing to its higher frequency of usage as well as to the backup of such words as “kitten” belonging to its word family. As a result, the defeated “cock” was exiled from its seat ([gat]) and was compelled to find its refuge in forms which originally meant “chicken” [pullus], “pheasant” [faisan] and “assistant priest” [vicaire], but some limited number of dialects rendered the standard French form [coq] (cf., Dauzat 1922:65-67).

In modern Chinese dialectology, the problem of homonymic collision has substantially been discussed in the studies for taboo words. In this respect it is worth referring to the works by Li Rong (李榮). In particular, Li (1994) examined the geographical distribution of the forms for “pen”, which changed to the irregular form [pei1] in Shandong (山東) by avoiding the homonymic clash with the taboo word [pi1]. However, the research done so far in China has been confined to that within the framework of Neo-grammarians/Lakglen’s comparative phonology, in which taboo words are considered as factors for preventing regular sound changes, but homonymic collisions in general have not been studied. Consequently, these studies have failed to discover the rich linguistic creativeness of the dialect speakers in this country.

An end result of homophonic collision is the elimination of either one of the two competing words, as in French “cock” or Chinese “pen”, but the dialects could also avoid the collision by forming geographically complementary distributions. For example, one particular form “P” is used in the area A for denoting the semantic category (x), while in the B area the identical form is used for the category (y) (Mase 1992:56-57).

| A | “P” (x) |
| B | “P” (y) |

**Fig. 1**
As a matter of fact, this is not a rare phenomenon and often emerges as the result of semantic changes of words belonging to related semantic categories. From a nationwide perspective, an instance of this is found in the semantic shift of “zou 走” from “to run” to “to walk” in the northern dialects and the preservation of its original meaning in the southern dialects. As a result, the two usages of the verb “zou” scarcely overlap and geographically form a complementary distribution (Miki 2006).

Homophonic collision is, so to speak, conflict between two referents of a single form. Synonymic collision, on the other hand, is conflict between two forms of a single referent, and has frequently occurred in every language, conspicuously in the case of loan words which came to coexist with native words, but also in the case of new forms which were created due to internal factors and came to compete with older forms.\(^1\)

Such conflict could end in the elimination of either one of the two competing forms, while a more frequent conclusion is a dialect retaining both forms by means of differentiating their usages. For instance, in some northern Wu (吳) dialects (e.g., Shanghai 上海 and Suzhou 蘇州) as well as in some Gan (贛) dialects (e.g., Nanchang 南昌), two kinship terms (stems), “die 爹” and “ye 爺”, competed for a single referent “father”, but compromised by differentiating their usage: “die” for vocative use and “ye” for designative use. Within the Jianghuai (江淮) Mandarin zone, the stem “ye” mostly succeeded in changing its referent from “father” to “father’s younger brothers” in a number of dialects, and eventually the original stem “shu 叔”, which had been used to refer to “father’s younger brothers” for thousands of years, disappeared completely from the system of kinship terms. However, in some dialects in central Anhui (安徽), this form “shu” is still retained as a general term for “father’s younger brothers”, while the use of “ye” is restricted to only one “father’s youngest brother”.\(^2\)

The present study will demonstrate two instances of “words in collisions” which I discovered in the Lianyungang (連雲港) area, Jiangsu (江蘇) Province. In each instance, a particular word form or word meaning shows a geographically complementary distribution between the two lexical maps. The first one can be interpreted as the result of speakers’ avoiding homophonic collisions, while the second one may be a rare but an interesting case of synonymic collision.

---

\(^1\) It is noted that the phenomenon called “wenbai yidu 文白異讀” (multiple readings of one Chinese character) can be viewed as synonymic collision, in which two competitive forms battle for getting one referent (or character). Unfortunately, this aspect seems to have been not shared so far among researchers.

\(^2\) See Iwata (2000) for the historical background and external/internal factors causing these changes.
2. Dialects in the Lianyungang area

The area under our survey is in the northeastern corner in Jiangsu Province, and linguistically is at the east edge of the long dialect borderline which demarcates the Mandarin dialects into Northern and Southern (Zavyalova 1983).³

There are two big cities in the southern zone, the traditional city L3 (Haizhou 海洲) and the newly developed city L4 (Xinpu 新浦), which are adjacent to each other.⁴ In the north and west counties, Ganyu (贛榆) and Donghai (東海), there are no big cities comparable to L3 and L4. The two county seats, G3 (Qingkou 青口) and D2 (Niushan 牛山), were smaller towns which had 20,000 or even less population when I carried out my first survey in 1980. However, the linguistic influence of the southern cities, L3 and L4, in spite of their overwhelming social dominance, is quite limited, particularly toward the north. If one walks from L4 to the north, one will encounter within 8 km the dialect with a northern basis, namely D13. The northern rural area has kept its linguistic influence, not yielding to the pressure from the southern area. As a result, the northern and southern dialectal influences have been kept in equilibrium until now. Words often move from north to south, as will be seen on our Maps 2 and 3 (“knee” and “clod”), though movement in the reverse direction also occurs, as appears on Maps 4 and 5 (“elbow” and “shoulder”).

The dialect data presented below are mostly from the joint surveys by Professor Su Xiaoqing 蘇曉青 (Xuzhou Normal University) and myself, but partly are from surveys independently performed by each since 1980 (See Iwata & Su 2000). Totally, the dialects of nearly 70 localities were explored, but the number of actual localities shown on the maps 1-5 is 51 at most, with the area limited in size to a 50-60 kilometer range both longitudinally and latitudinally.

A number of isoglosses run in this area horizontally from east to west on the south bank of the New Shu River (新沭河), while many of them turn to the southwest at about the locality D9.⁵ As a typical instance of such isoglosses, on Map 1 is drawn an

---

³ As classificatory terms, I will hereafter follow Zavyalova (1983), who divides the Mandarin (Guanhua) dialects into Northern and Southern. The latter corresponds to the two dialect groups appearing in Language Atlas of China (LAC, 1987): “Jianghuai Mandarin” and “South-western Mandarin”.

⁴ L3 had long performed the role of local center in the northeastern Jiangsu area until the 1950s, while this role was replaced by L4 after the 1960s, when the administrative government of this area moved from L3 to L4.

⁵ The New Shu River is an artificial one launched in 1952. Since then, the administrative border between the northern county, Ganyu, and the southern one, Donghai, was fixed to this river. However, the older administrative border, which is assumed to have existed for hundreds of years, ran southerly along the dialect border appearing on Map 1.
isogloss, which is relevant to the following discussion.

In this area, there exist four types of tone systems, which are indicated A, B, C, D on Map 1. Tones realized in monosyllabic citation forms are presented in Chao’s five point scale, with two or three numerals in the case of long tones and two numerals with an underline in the case of short tones. The slash indicates either minor differences among localities (for systems A and B) or free alternation (for system C).

Map 1: “Tone system”

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
</tr>
<tr>
<td>Tone 2</td>
<td>55</td>
<td>45/55</td>
<td>45</td>
<td>35</td>
</tr>
<tr>
<td>Tone 3</td>
<td>324/35</td>
<td>35</td>
<td>324</td>
<td>41</td>
</tr>
<tr>
<td>Tone 4</td>
<td>41</td>
<td>41</td>
<td>41/55</td>
<td>45</td>
</tr>
<tr>
<td>Tone 5a</td>
<td>= Tone 1</td>
<td>24</td>
<td>55</td>
<td>24</td>
</tr>
<tr>
<td>Tone 5b</td>
<td>= Tone 2</td>
<td>= Tone 2</td>
<td>55</td>
<td>24</td>
</tr>
</tbody>
</table>

The difference between types A, B and types C, D is the most significant, and is typically reflected in the behavior of Tone 5 (Rusheng 入聲).

---

6 Tones 1, 2, 3, 4, 5 are conventionally referred to as Tone Ia (Yinping), Tone Ib (Yangping), Tone II (Shangsheng), Tone III (Qusheng) and Tone IV (Rusheng), respectively, in comparative phonological terms.
Northern types A, B

In types A and B, Tone 5 underwent tone split conditioned by the type of syllable initial consonants: syllables initiated by voiceless obstruents as well as those initiated by nasals or liquids assumed Tone 5a, and those initiated by voiced obstruents assumed Tone 5b, which later merged with Tone 2. The difference between A and B lies in the behavior of Tone 5a. In type A, it merged with Tone 1; thus, Tone 5 has disappeared altogether. In type B, on the other hand, Tone 5a is distinguished from other tones. While type A is very popular in the Northern Mandarin area, B is a rare preservation of the proto-type tone system (Iwata 1996:239). Note that type B is distributed close to the eastern edge of the dialect border.7

Southern types C, D

Tone 5 has been preserved in types C and D, but unlike A and B, it did not undergo the tone split. Type C solely emerges in the locality D2. It is identical to D with respect to the tonal split of Tone 5, but Tones 3 and 4 are clearly northern in terms of their phonetic shapes. Syllables with Tone 5 end in glottal stop [ʔ] in type C, and also in type D with a lesser degree.

In all these types, tone sandhi is rich at the non-final positions of utterances, and curiously this makes the tonal realization in polysyllabic words very similar across the four tone types.

In the following sections, the term northern zone refers to the area where tone types A or B are found, and southern zone refers to the area covered by type D. Tone categories will be indicated below in numerals: 1, 2, 3, 4, 5. In addition, unstressed neutral tone is indicated by “0”. As an explanatory device, the northern morphemes that had Tone 5a are indicated by tone mark “1/5”. The same device will be applied to the case where the tone realization is unstable, e.g., “1/2” for the [liu] morpheme presented below. Syllable final glottal stop [ʔ] appearing in the southern dialects will not be notated, as it automatically appears in the syllable with Tone 5.

3. Homonymic collision: “knee” and “clod”

3.1 Complementary distribution

Maps 1 and 2 illustrate the word form distributions for “knee” and “clod.” As for

---

7 Type B is currently diminishing in the villages situated on the north bank of the New Shu River, where many speakers are unable to distinguish the monosyllabic Tone 5a from other tones. However, Tone 5a is morpho-phonologically preserved even among them, as reflected in its peculiar tone sandhi rule (Iwata 1992).
the latter, we asked speakers the general terms for “a lump of clod on the ground” (“tukuai 土塊” in standard Chinese), including the dry one and the wet one. Actually, the forms shown on Map 3 may have the optional prefixation of morphemes such as [mi2] (泥), [lan4 mi2] (爛泥) and [tʰu3] (土), meaning “mud” and “soil”, e.g., the form found in G1 was [mi2 kʰə3 lə0] (literally means “clod of soil”). Since these morphemes are empty in signification, they are omitted for simplification on Map 3.

On the two maps, one will notice a reversed distribution with regard to the single form “kʰə tʰəu”. It is found in the southern zone on Map 2 but in the northern zone on Map 3, and never overlaps with each other.
Fig. 2 is a simplified sketch of the distribution.

<table>
<thead>
<tr>
<th>knee</th>
<th>clod</th>
</tr>
</thead>
<tbody>
<tr>
<td>([k^{}\hat{a}1/5 \text{ pai4}])</td>
<td>([k^{}\hat{a}3 \text{ th\text{\text{-}au2}}])</td>
</tr>
<tr>
<td>([k^{}\hat{a}5 \text{ pai4}])</td>
<td>([k^{}\hat{o}3 \text{ liu1/2}])</td>
</tr>
<tr>
<td>([k^{}\hat{a}5 \text{ th\text{\text{-}au2}}])</td>
<td>([k^{}\hat{o}3 \text{ liu1/2}])</td>
</tr>
</tbody>
</table>

**3.2 “k^{}\hat{a} th\text{\text{-}au}” forms**

In order to show the phonetic and morphological realities of the forms, some instances of word pairs are illustrated below, with the pitch contours of the tones indicated by numerical tone letters. Forms having [liu1/2] or [la0/la0] are excluded. Among these six dialects, three (G8, D11, D13) are from the northern basis and the other three (D2, D15, D17) from the southern, while, as mentioned above, D2 has seriously been changed due to northern influence.

<table>
<thead>
<tr>
<th>Locality</th>
<th>Forms for “knee”</th>
<th>Forms for “clod”</th>
</tr>
</thead>
<tbody>
<tr>
<td>G8</td>
<td>(k\hat{a}21 \text{ par41})</td>
<td>(k^{}\hat{a}21 \text{ th\text{\text{-}au55 ts\text{\text{-}}0}\text{ ts\text{\text{-}}0}})</td>
</tr>
<tr>
<td>D11</td>
<td>(k\hat{a}21 \text{ par41})</td>
<td>(k^{}\hat{o}55 \text{ th\text{\text{-}au45}})</td>
</tr>
<tr>
<td>D13</td>
<td>(k^{}\hat{a}21 \text{ par41})</td>
<td>(k^{}\hat{o}21 \text{ th\text{\text{-}au45}})</td>
</tr>
<tr>
<td>D2</td>
<td>(k^{}\hat{o}55 \text{ pai41 ts\text{\text{-}}0})</td>
<td>(k^{}\hat{o}35 \text{ th\text{\text{-}au55}})</td>
</tr>
<tr>
<td>D15</td>
<td>(k^{}\hat{o}55 \text{ th\text{\text{-}au21 ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}})</td>
<td>(k^{}\hat{o}55 \text{ th\text{\text{-}au35}})</td>
</tr>
<tr>
<td>D17</td>
<td>(k^{}\hat{o}55 \text{ th\text{\text{-}au21 ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}}\text{ ts\text{\text{-}a0}})</td>
<td>(k^{}\hat{o}55 \text{ la\text{\text{-}a1/2}}\text{ la\text{\text{-}a1/2}}\text{ la\text{\text{-}a1/2}}\text{ la\text{\text{-}a1/2}}\text{ la\text{\text{-}a1/2}})</td>
</tr>
</tbody>
</table>

Zi-suffixation (子) tends to appear obligatorily in southern “k^{}\hat{a} th\text{\text{-}au}”s, while its appearance is optional in northern ones. R-suffixation (兒) is a feature specific to northern dialects, and it appears obligatorily in the northern forms for “knee”. The use

---

The “zi 子” suffix is pronounced as [ts\text{\text{-}0}] in the northern localities but it is [ts\text{\text{-}a0}] or [ts\text{\text{-}0}] in the southern localities. The r-affixation accompanies the deletion of the syllable final component [-i] and [-n], as in the Beijing dialect, so the second syllable in [k\hat{o} pai] changes by suffixation to [par], instead of [pair].
of these two suffixes will be omitted hereafter.

The initial syllable in southern “kʰə tʰəu” has a shorter length and is identified as having Tone 5, which undergoes tone sandhi to be realized as the high-level contour ([55]). The northern [kʰə] syllable is problematic: although it assumes a normal length and is identified as having Tone 3, its tonal realization is variable depending on localities and speakers. The expected tone, as predicted by tone sandhi, is either a high-level contour ([55]) or a rising contour ([35]); however, what is often observed is an irregular contour, namely, a low-dipping contour ([21]), as can be seen in the examples cited for G8 and D13 above.

The [kʰə] syllable includes two varieties in terms of its vowel: [a] occurs for some northern localities, e.g., G8, and [u] for some southern localities, e.g., D2 and D15. This latter vowel [u] only appears in the southern dialects, and is concurrent with the short syllable having Tone 5.

From a comparative point of view, it is important to note that northern [kʰə] corresponds to southern [kʰʊ], e.g., D13 [kʰə3 tʰəu2]:D15 [kʰʊ3 tʰəu2].9 As indicated on Map 3 by the horizontal bar, this round vowel [ʊ] is characteristic of southern dialects, being shared by such forms as [kʰʊ3 laŋ1/2], [kʰʊ3 liu1/2] and [kʰʊ3 lao0], which have in common the liquid initial [l] for the second syllable.

3.3 The hybrid form as a breakwater for protecting collision

Readers may wonder why “knee” and “clod” needed to avoid the collision. Gilliéron pointed out that serious clashes that could destroy the unexceptional regular sound changes should have only occurred to the homophonic words which were connected by association and belonged to the same or adjacent semantic domain (Gilliéron and Roques, 1912, XIII. Mots en Collision, B. Epi et Epine). The particular two objects, “knee” and “clod”, do not belong to an adjacent semantic domain, so seemingly they would tolerate homophonetic collision.

However, the fact is that dialects obey a strict rule of avoiding the co-use of “kʰə tʰəu” for the two referents. Moreover, there is an indication that this collision is intolerable for speakers, namely, the emergence of the two forms for “knee”, [kʰə5 pai4] and [kə1/5 pai4 tʰəu2]. These must have been built up by the contact between northern and southern forms.

9 This sound correspondence, i.e., [a]:[u], is valid for the syllables that have tones other than Tone 5, as confirmed for a number of morphemes, e.g., 歌, 可, 和. If it is a morpheme of Tone 5 origin, northern [a] corresponds to southern [ɔ] or [u], as is evident in the form for “knee”.
It is generally known that linguistic hybrids are frequent at dialect borders, where two competitive dialectal influences need to compromise for self-protection. In this particular case, however, it might not have been a mere accident that these forms were created exactly in the villages where northern and southern “kʰə tʰəu’s meet. As illustrated in Fig. 2, the form [kʰə pai] functions as a breakwater or buffer zone for protecting the collision between northern and southern [kʰə tʰəu].

Note that the human body part “knee” is likely to be associated with various inanimate beings in Chinese dialects.10

Although we have no means to attest this, it would not be surprising if “knee” was associated with “clod” in this area, due to resemblance in their shapes.

3.4 Scenario for changes (I)

A question is how the complementary distribution was formed historically in terms of the referents of “kʰə tʰəu”. A straightforward answer to this question is that the two “kʰə tʰəu’s came into being owing to the deletion of the second elements that had existed in the proto-forms.

Northern
- “knee”: *kə5 lə0 pai4 > kə1/5 pai4
- “clod”: *kʰə3 lə0 tʰəu2 > kʰə3 tʰəu2

Southern
- “knee”: *kʰə5 tɕia5 tʰəu2 > kʰə5 tɕia0 tʰəu2 > kʰə5 tʰəu2

The proto-forms proposed here are found in and outside the presently explored area. Northern forms for “knee” and “clod”, [kə1/5 lə0 pai4] and [kʰə3 lə0 tʰəu2], are found mainly in the west and south Shandong (Ma and Wu 2003:104-5, Qian et al. 1993:

---

10 All these instances are from Southern Chinese dialects, while Northern Mandarin dialects generally use those unmotivated forms as found in Beijing, [pə ləŋ kar]. This must be a cognate with the form [kə par] found in the northern zone, cf., Iwata (1995:206-7).
Homonymic and Synonymic Collisions in the Northeastern Jiangsu Dialect

390-1), and on Map 3, the form \[k^h3 l^0 t^h\omega u2\] appears in G2 (indicated by the black vertical bar). The Southern form for “knee”, \[k^h5 t\omega i5 t^h\omega u2\], is found in such localities as Huaiyin (淮陰), Xuyi (盱眙) and Yangzhou (揚州) within the Southern Mandarin zone (Bao 1998:432).

The scenario given above implies that the accidental encounter of the two “kʰə tʰə u”s was realized at the dialect border, owing to a trend of change shared by the two areas, northern and southern. A phonetic condition that triggered this change was a rhythmic pattern “strong (medium)-weak-strong” in the trisyllabic construction, in which the weak syllable was likely to be deleted.

However, northern and southern changes had different implications in a morpho-logic sense. In the northern area, the deletion uniquely occurred to the infix \[l^0\] (or \[la0\]). This change actually finds its parallels in such forms as “forehead”, \[i\epsilon l la0 kar4\] > \[i\epsilon l kar4\]. On the other hand, the deletion in the southern area occurred to the lexical morpheme \[t\epsilon i5\]. Exactly the same type of the deletion process occurred to the forms for “elbow” and “shoulder”, as will be mentioned below in §4.1.

From a longer historical perspective, the change that occurred to the southern dialects in this area in effect meant an irrecoverable loss of the stem \[t\epsilon i5\] (膝), which originally meant “knee” and had been an indispensable element for this word “knee” throughout the history of Chinese. But the dialects eventually achieved this change, owing to speakers’ unconscious association with the verb \[k^h5 t^h\omega u2\] (磕頭) meaning “to bow with kneeling on the ground”, a very common association which appears across languages, for instance, English “knee” and “to kneel” or Japanese “hiza” (knee) and “hizamazuku” (to kneel down). This change was accompanied by the nominalization process of this form, \[k^h5 t^h\omega u2\], by the addition of the suffix “zi 子”.

The scenario so far may be a reasonable one, but there is a crucial issue unanswered yet, that is, the history of the southern forms for “clod”.

In the southern zone, there are three forms for “clod”, \[k^h3 la\n1/2\], \[k^h3 liu1/2\] and \[k^h3 la0\]. The most popular one is \[k^h3 la\n1/2\], but according to Lianyungang Shizhi (2000:2602), it is also the form for denoting “small hole (dug) on the ground” in L4. Moreover, Liu (2001:105) and Su (1997:94) reported the same use of the form \[k^h3 la\n1/2\] existing in the northern and western localities G3 and D2. Though, to my regret, I have not been aware of this usage during my surveys, it seemingly prevails over the whole explored area. It would be reasonable to assume that this particular usage is a retention of the original meaning, and the usage as a general term for “clod” must be an extension of the semantic domain.

Why did the southern dialects proceed with this semantic change? My answer is that this is because southern “clod” once had a form \*\[k^h5 t^h\omega u2\]. It was defeated in collision with the homophonous form for “knee”, and so the southern dialects rendered
the form \([k^h\textumlaut{u}3 \textumlaut{l}\textumlaut{a}1/2]\) from the adjacent semantic domain.

**3.5 Scenario for changes (II)**

An etymology proposed for the southern “clod” is \([f\textumlaut{a}5 t^h\textumlaut{u}2]\), which later changed to \([k^h\textumlaut{u}5 t^h\textumlaut{u}2]\).

As a matter of fact, the morpheme \([f\textumlaut{a}5]\) (垡) has a long history, tracing back to at least one thousand years ago. It was registered in the rhyming dictionaries, “Guangyun 廣韻” and “Jiyun 集韻”.

垡, 耕起土也, “(It is) the soil dug out from the field”. MC: *[byat] (Jiyun)

This form as well as its usage has also been preserved in modern dialects. Inside the area shown on our maps, Su (1997:151) found it in D2, where it appears as \([f\textumlaut{a}2 t^h\textumlaut{u}2]\) but is employed in a similar semantic range with Jiyun’s description, “clod dug out from the cultivated field”. Again, I have not been aware of this lexical use until recently, so no information is available for the other localities. Outside this area, the same use of the morpheme \([f\textumlaut{a}5]\) is reported for the dialects in Yangzhou (揚州) and Nanjing (南京), cf., Wang and Huang (1996:362), Liu (1995:33).

While the usage found in Yangzhou, Nanjing and D2 is a retention of the original meaning, the morpheme \([f\textumlaut{a}5]\) is assumed to have changed to a general term for “clod” in the southern zone under our survey, undergoing the same type of semantic extension that occurred to \([k^h\textumlaut{u}3 \textumlaut{l}\textumlaut{a}1/2]\). This latter usage has been confirmed so far in two southern counties out of this area: Binhai (濱海) and Lianshui (連水) (my own survey in 1987 and Chen and Li 1996:2096). Also, according to Bao (1998:329), the form \([f\textumlaut{a}5 t^h\textumlaut{u}2]\) is currently used as a general term for “dry soil” in the two southern counties, Funing (阜寧) and Yancheng (鹽城).

Based on this etymological consideration, I will propose a revised scenario for the changes, which is composed of three stages.

**1st stage:**

It is assumed that the forms for “clod” underwent the changes at this stage, while those for “knee” remained unchanged.

1) Southern “clod”: \(f\textumlaut{a}5 t^h\textumlaut{u}2\) (垡頭) > \(k^h\textumlaut{u}5 t^h\textumlaut{u}2\)

The initial consonant in the form \([f\textumlaut{a}5 t^h\textumlaut{u}2]\) changed to \([k^h]\). The most probable factor contributing to this change may be phonetic attraction by the form for “knee”, namely, \([k^h\textumlaut{a}5 \textumlaut{t}\textumlaut{e}5 t^h\textumlaut{u}2]\). Phonetic attraction (or homophonic attraction) is, so to
Homonymic and Synonymic Collisions in the Northeastern Jiangsu Dialect

speak, a contamination of forms, which is triggered by speakers’ mispronunciations influenced by the form of another lexical item, and it could also result in the two forms becoming homophonous, cf., Dauzat (1922:72-77), Grootaers (1946:219-20). An interesting instance is the form \([k^h \text{ə}t^h \text{ə}u]\) for “clod” appearing in G5 and D14 on Map 3. This is no doubt the product of assimilation of the initial syllable to the second one in the form \([k^h \text{ə} t^h \text{ə}u]\), but plausibly this change might also have been motivated by the attraction to the verb “\(k^h \text{ə}t^h \text{ə}u \text{叩頭}\)”, meaning “to bow with hitting one’s head on the ground”, a synonym of “\(k^h \text{ə} t^h \text{ə}u \text{磕頭}\)” mentioned above. Note that this is the case of parallel change that has independently occurred in the two villages. Another instance on our maps is the forms for “knee”, \([k^h \text{ə}1/5 \text{p} \text{ə}0]\) and \([k^h \text{ə}1/5 \text{p} \text{ə}0 \text{kai4}\), appearing respectively in D12 and G3 on Map 2. These forms are the products of phonetic attraction to the form for “arm”, \([k^h1/5 \text{p} \text{ə}0]\) (胳膊), which appears on Map 4 as the constituent of the form for “elbow”.

According to the literature available, as a general term for “clod”, \(*[k^h\text{ə}5 \text{t}^h\text{ə}u2]\) has not yet been found either in or out of this area. Note that its initial syllable at this stage was \([k^h\text{ə}5]\), which had Tone 5, so that it was distinct from both northern \([k^h\text{ə}3]\) and southern \([k^h\text{ə}3]\).

(2) Northern “clod”: \(*[k^h\text{ə}3 \text{la}0 \text{t}^h\text{ə}u2] > [k^h\text{ə}3 \text{la}0] > [k^h\text{ə}3 \text{liu1/2}\]

One will see on Map 3 that the forms with \([\text{la}0/\text{la}0]\) and \([\text{liu1/2}]\) tend to cover the whole area. An assumption induced from this evidence is that these forms might have been older than \([k^h\text{ə}3 \text{t}^h\text{ə}u2]\). In other words, the deletion of the final element chronologically precedes that of the second element.

\[(i) \quad k^h\text{ə}3 \text{la}0 \text{t}^h\text{ə}u2 > k^h\text{ə}3 \text{la}0 \]
\[(ii) \quad k^h\text{ə}3 \text{la}0 \text{t}^h\text{ə}u2 > k^h\text{ə}3 \text{t}^h\text{ə}u2\]

This assumption presupposes that the birth of the new forms in the northern zone did not bring about the extinction of the older form, \([k^h\text{ə}3 \text{la}0 \text{t}^h\text{ə}u2]\). Rather, it remained an influential form in the northern zone, coexistent with the new forms. Note that it is still alive within this area in the locality G2.

It is assumed that the northern form \([k^h\text{ə}3 \text{la}0]\) extended as far as the southern zone, where, according to the inter-dialectal phonological rule, it was accepted as \([k^h\text{ə}3 \text{la}0]\).\(^{11}\) Indirect evidence to support this assumption is the distribution of the word \([\text{mi}2]\), a

\(^{11}\) This form \([k^h\text{ə} \text{la}]\) is a very popular one as the general term for “clod”, being distributed widely in the Northern Mandarin area, including Beijing. Therefore it is also possible to assume that this form has been native to this area from a very early epoch.
general term for “soil” or “mud”. This is a typical northern form which is commonly
found in the Shandong dialects, while its southern correspondent is [ni2] or [li2] (泥).
An isogloss is drawn on Map 3, indicating the southern extreme of the distribution for
[mi2]. As constituting adjacent semantic categories, the northern forms for “clod” and
“soil” must, side by side, have penetrated into the southern zone.

In the process of its dispersion, the form [kʰə3 lə0] changed to [kʰə3 liu1/2],
motivated by analogical attraction of semantically unrelated word forms, such as [tʊŋ4
liu1/2 liu0] (hail), [kʰə5 liu0] (mussel) and [tɕə3 liu0] (cicada), and was eventually
incorporated into the paradigm of these forms with the morpheme [liu].

In summary, on Fig. 3 is illustrated a reconstruction of the distribution of the
proto-forms as well as the changes that occurred at this stage.

![Fig. 3](image_url)

2**nd** stage:

At this stage, the second element in trisyllabic forms was deleted due to the
common trend for changes mentioned above. As for southern “knee” and “clod”, it is
also assumed that phonetic attraction continued from the previous stage, and the two
words, [kʰə5 tɕə5 tʰəu2] (knee) and [kʰə5 tʰəu2] (clod), interacted with each other.
Consequently, the form for “knee” tended to lose its second element, [tɕə5], and came
to be in homonymic collision with the form for “clod”. Recall that the vowel in the
syllable [kʰə5] freely alternates with the vowel [ə].

The winner of the battle in the southern area was “knee”, presumably related to a
backup of its association with “to kneel”. Thus, the form [kʰə5 tʰəu2] for “clod”
disappeared within the southern zone. At this moment, the solution of the southern
dialects with “clod” was to accept rescue parties from inside and outside.

(a) To render the form [kʰə3 laŋ1/2] from an adjacent semantic category
(b) To render the forms, \([k'ə^3 \, lə0]\) and \([k'ə^3 \, liu1/2]\), which were of northern origin but at this stage had already intruded into the southern zone.

To conclude, the particular vowel \([u]\) is evidence of an early southern form \([k'ə5 \, tʰəu2]\) for “clod”.

Fig. 4 is a summary of the changes occurred at this stage.

<table>
<thead>
<tr>
<th>knee</th>
<th>clod</th>
</tr>
</thead>
<tbody>
<tr>
<td>(*[kə5 , lə0 , pai4]) &gt; [kə1/5 , pai4]</td>
<td></td>
</tr>
<tr>
<td>(*[k'ə5 , tɕiə5 , tʰəu2]) &gt; [k'ə5 , tʰəu2]</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4**

Final stage:

Going through the changes that occurred in a preceding stage, “kʰə tʰəu” succeeded in switching its referent from “clod” to “knee” in the southern zone, but happened to face another “kʰə tʰəu” (\([k'ə^3 \, tʰəu2]\)) which emerged in the northern area as the form for “clod”. Northern “kʰə tʰəu” was in the phase of intruding into the southern zone, and as a matter of fact, some southern villages (D2, D15 etc.) around the dialect border had already become under the territory of \([k'ə^3 \, tʰəu2]\) or \([k'ə^3 \, tʰəu2]\). Naturally, the southern dialects reacted against this intrusion, and the policy they adopted to avoid the collision was to build a *breakwater* around the border, that is, to change the form for “knee” from \([k'ə^5 \, tʰəu2]\) to \([k'ə^5 \, pai4]\) by accepting the second element of the northern form \([kə1/5 \, pai4]\).

4. Avoidance of synonymic collisions: “elbow” and “shoulder”

Maps 4 and 5 illustrate the distributions of the forms for “elbow” and “shoulder”. For each map, the use of the suffix “zi 子” is eliminated, and the word forms are classified mainly with respect to the final lexical morpheme.\(^{12}\)

\(^{12}\) The r-suffix is scarcely found in the forms for “elbow” and “shoulder”.\[^{12}\]
Ray Iwata

Map 4  “elbow”

Map 5  “shoulder”
Fig. 5 is a simplified sketch of the distribution.

<table>
<thead>
<tr>
<th>elbow</th>
<th>shoulder</th>
</tr>
</thead>
<tbody>
<tr>
<td>[kə1/5 pə0 kʰy1/5] 曲</td>
<td>[kien1 pə0 tʰəu2]</td>
</tr>
<tr>
<td>[kə1/5 pə0 tʂu2]</td>
<td>拐</td>
</tr>
<tr>
<td>[kə5 pəŋ3/0 tʂəu3] [kə5 tʂəu3] [kə5 pəŋ0 kuai3]</td>
<td>[kien1 pə0 kuai3] [tɕien1 pə0 kuai3] [tɕien1 pəŋ3/0 kuai3] [tɕien1 kuai3]</td>
</tr>
</tbody>
</table>

Fig. 5

The topic that I will discuss in this section is how complementary distribution was formed between a set of synonyms, northern [kʰy1] (曲) for “elbow” and southern/intermediate [kuai3] (拐) for “shoulder”. Unlike the case of “knee” and “clod”, we will see here the dominance of a southern dialectal influence over the northern.

4.1 Forms for “elbow” and “shoulder”

Two rules of sound correspondence are introduced for readers’ better understanding of the following descriptions (refer to Iwata & Su 2004 for the details).

<table>
<thead>
<tr>
<th>Northern</th>
<th>Intermediate</th>
<th>Southern</th>
</tr>
</thead>
<tbody>
<tr>
<td>ki-, kʰi-, hi-</td>
<td>ki-, kʰi-, hi-</td>
<td>tɕi-, tɕʰi-, cɕi-</td>
</tr>
<tr>
<td>tʃi-, tʃʰi-, jʃi-</td>
<td>tʂ-, tʂʰ-, ʂ-</td>
<td>tʂ-, tʂʰ-, ʂ-</td>
</tr>
</tbody>
</table>

This rule also applies to the environment where the front-round vowel [y] follows gutturals or non-retroflex sibilants.

(1) elbow

In the southern zone, the final lexical morpheme is generally [tʂəu3], while in L4 another morpheme [kuai3] appears. Forms within the northern zone are of two types, [kʰy1/5] (or its phonetic variant [tɕʰy1/5]) and [tʂu2] (or its tonal variant [tʂu1]), the former in the north and the latter in the south of the Fan River (范河).

From a comparative point of view, southern [tʂəu3] is a regular form corresponding
to the standard Beijing form “zhou 肘”, while the other two are irregular forms created by folk-etymology. The morpheme [kʰy1/5] (曲) normally serves as an adjective, meaning “to be curved”, while [tʂu2] (軸) means “axis”.

It is improbable, however, that the morpheme [tʂʊu3] (<[tʃiu3]), which is still retained in the north of the Fan River as the character reading of “肘”, directly evolved to [kʰy1/5] or [tʂu2] (<[tʃy2]). Rather, it is more probable that particular folk-etymologies might have arisen from any other forms which more closely resembled the present forms phonetically, namely, *[tʃy5]. It is further assumed that the ultimate reason for such irregular forms was to avoid the homophonic collision with the word for “pig’s thigh (for food)”, *[tʃiʊ3 tʂ0], which is usually written with the identical Chinese character “肘” with that for “elbow” (see Iwata 2003 for the details).

The preceding component is [kə5 pa0] (胳膊) in the north, while it is [kə5 paŋ] (胳膊) in the south. These forms should have formerly meant “arm”, but this meaning is presently expressed in this area by another form [paŋ3 tʂ0/tʂə0] (膀子) in all localities but the northernmost locality, G1, where the form [kə1/5 pa0] (胳膊) is retained as well. Evidently, the form [kə1/5 pa0] has secured its wide territory in the north of this area until now. This indicates the preservation of the older form for “arm”, [kə1/5 pa0], in the compounds for “elbow” and the early intrusion of the newer form, [paŋ3 tʂ0/tʂə0], from a southern area outside the map to this area. Within the area shown on the map, this new form eventually drove away the older form as far as north of G1. On the other hand, resistance of the older form against the new one left its trace in the southern form for “elbow”, [ka5 paŋ0], which may be a hybrid made up from the contact of the forms, [ka5 pa0] and [paŋ3 tʂ0].

The form [ka5 tʂʊu3] only appears within the southern zone, and it is the product of the deletion of the second lexical morpheme, a trend shared by “knee”.

The initial syllable is realized as having the aspirated stop [kʰ] in some southern localities, namely, [kʰa5 paŋ tʂʊu3], [kʰa5 tʂʊu3] and [kʰe5 tʂʊu3]. This is due to the phonetic attraction of the syllable [kʰə5] found in such forms as “knee” and “clod”, so is a recent innovation. Also, one will notice the vowel in the initial syllable fluctuating between [ə] and [ɐ], a common phonetic tendency shared by the southern forms for “knee”.

(2) Shoulder

There are two morphemes appearing as final lexical components, northern [tʰəu2] and southern/intermediate [kuaɪ3].

The original meaning of the morpheme [kuaɪ3] (拐) is “to turn”, so it serves as a verb, as found in the standard Beijing word “guaiwan 拐彎” (to turn a corner). However, it often behaves like an adjective meaning “to be curved,” as found in the standard word,
“guaijiao 拐角” (street corner), and here in the dialects, semantic distinction is minimized to the degree that [kʰy1] is used in the sense “slow-angled” but [kuai3] in the sense “right-angled”. On the other hand, the preceding components, [tɕien1 pan3] (肩膀), [kien1 pa0], [tɕien1] etc., denote “shoulder”, so the addition of the morpheme [kuai3] is solely for strengthening the association with its referent.

Parallel with [kə5 ʦəu3], the form [tɕien1 kuai3] is due to the deletion of the second morpheme. Importantly, it spreads not only over the southern area but also extends to the north as far as the Fan River, indicating intrusion of southern forms to the northern area. Compare the distribution areas of [kə5 ʦəu3] and [tɕien1 kuai3] on Maps 4 and 5.

Forms without [tʰəu2] or [kuai3], i.e., [tɕien1 paŋ] (肩膀), are scattered throughout this area, and are considered to be the product of the deletion of the final component ([tʰəu2] or [kuai3]). Though it is identical with the standard Beijing form, it may be not a case of borrowing but an independent innovation, similar to the change occurred to “clod”, i.e., [kʰə3 la0 tʰəu2] > [kʰə3 la0].

4.2 Scenario for the changes

A hypothesis proposed is that the northern forms, [kʰy1] for “elbow” and [tʰəu2] for “shoulder”, once had a broader territory than now within the northern zone.

As for “shoulder”, [tʰəu2] has yielded their southern territories, namely, south of the Fan River, to the southern form [kuai3]. One reason for this might be that this morpheme [tʰəu2] functioned like a noun suffix, just as that in the “clod” form, [kʰə3 tʰəu2], so that it was unable to resist the intrusion of the southern [kuai3], since this latter form was well motivated in terms of association with its referent.

As for “elbow”, it was mentioned above that the two forms for northern [kʰy1/5] (曲) and intermediate [ʦəu2] (軸), might have been created from the form *tʃy5 due to folk-etymology.13

Northern: *tʃy5 > kʰy1/5 (曲)
Intermediate: *tʃy5 > tʃu1/5 > ʦu2 (軸)

13 According to my survey carried out in 1987, some villages in the north of G1 use a form [kə1/5 pa0 uan1] (胳膊彎) for “elbow”. From this, one may argue that the folk-etymology for the form [kʰy1/5] (曲) came from the influence of these northern dialects through the association with [uan1 kʰy1/5] (彎曲), a very popular word denoting “to be curved”. However, a contrary change would be more plausible, that is, the older one was [kʰy1/5] and [uan1] (彎) was an innovation due to the association with [uan1 kʰy1/5].
In the intermediate zone, the form \([tʃy5]\) should, according to the phonological rule, evolve to \([tsu1]\), as is evidently observed in three localities on Map 4. However, speakers provided this form with the folk-etymology of “axis” by changing its tone to Tone 2. Presumably, this is because the form \([tsu1]\) was homophonic with the form for “pig”.

The northern form \([kb^h y1/5]\) has its phonetic variant \([tc^b y1/5]\) in D11, remote from its northern kernel area, and this suggests that this variant might be evidence of the broader distribution of the form \([kb^h y1/5]\) at an earlier stage, and that it is older than \([tsu2]\) in this area. If true, the newer form \([tsu2]\) must have been an outcome of transmission from the area outside the map, presumably from the west.

Introduction of the form \([tsu2]\) into the intermediate zone, hence, the disappearance of \([kb^h y1/5]\) in this zone, was an advantage for the southern form \([kuai3]\) denoting “shoulder”: it could intrude into the intermediate zone without semantic interruption of its synonymic form \([kb^h y1/5]\).

Suppose that the southern form \([kuai3]\) for “shoulder” intruded into the northern area when the form \([kb^h y1/5]\) was still used for “elbow”, the following lexical system would be formed.

“elbow” \([kə1 pə0 kb^h y1]\) / “shoulder” \([kien1 pə0 kuai3]\) (< \([kien1 pə0 t^h au2]\))

Now the form for “shoulder” is well motivated in terms of its meaning, “to be curved”. However, this is actually a system in which synonymic morphemes oppose each other in final position, while maintaining the distinction of the preceding components: \([kə1 pə0]\) / \([kien1 pə0]\). Therefore, the southern form \([kuai3]\) for “shoulder” would be unlikely to intrude into the immediate zone as long as \([kb^h y1/5]\) is used there for “elbow”. Compare this system with that appearing in the southern locality L4:

“elbow” \([kə1 panʃ0 kuai3]\) / “shoulder” \([tcien1 panʃ3 kuai3]\)

This system is apparently a new one, which came into being by the rendition of the morpheme \([kuai3]\) for “shoulder” to the form for “elbow”: \([kə1 panʃ3/0 tsəu3]\) > \([kə1 panʃ3/0 kuai3]\). This is an economical system since the distinction lies just in the initial component.

The northern \([t^h au2]\) is currently in a regressive phase, actually yielding four localities to the southern form \([kuai3]\) to the north of the Fan River (see Map 5). Therefore, it can be said that synonymity does not bring about such serious effects as homonymity for words. Nevertheless, what is significant here is the fact that not a few northern localities have so far defended the original form \([t^h au2]\).
In conclusion, in this particular case, the use of synonymic forms for the two related but distinct semantic categories has been avoided.

5. Concluding remarks

In this paper, histories of the four words were reconstructed, albeit within a limited geographical area. In discussing the problems of the two collisions, I also referred to some contributing factors to the changes, such as folk-etymology and contamination by phonetic or analogical attraction.

Phonetic or analogical attraction occurs not only with semantically related lexical items but also could occur with unrelated ones. The necessary condition that triggers this process is that the given element should be semantically unmotivated, in other words, speakers become unconscious of the meaning, due to any number of factors, in particular due to the weakening process of the syllable.

For example, the northern forms for “elbow” and “shoulder”, [kə1/5 pə0 kʰy1/5], [kə1/5 pə0 tʂu2], [kien1 pə0 tʰau2], [kien1 pə0 kuai3], are characterized by an unaccented second syllable [pə0], with speakers in some localities pronouncing it as [mə0] or [m0]. This phenomenon may be explained in terms of a progressive assimilation of the initial [p-] to the nasal [-n] in the preceding syllable in the case of the forms for “shoulder”, and also in terms of an analogical infection of this phonetic behavior to semantically related lexicon “elbow”. Interestingly, the identical phenomenon also emerges in semantically unrelated lexical items. For example, the word referring to “husband’s elder brother” is generally [ta4 pə0 tʰəu2 tsə0] (大伯頭子) in this area, but in some villages around the New Shu River, it is realized as [ta4 mə0 tʰəu2 tsə0].

An eventual outcome of this process was the deletion of the second syllable in a trisyllabic rhythmic pattern “strong (medium)-weak-strong”. In the southern area, this process resulted in the deletion of lexical morphemes: [tɕia5] (膝) for “knee” and [paŋ3] (膀) for “elbow” and “shoulder”. Looking at this process from the other side of the same coin, it can also be viewed as a stemization process of prefixes, “kʰə” and “kə”. Metaphorically speaking, the form “kʰə tʰəu” is merely composed of bone (prefix) and skin (suffix), lacking meat (lexical morpheme), and since the kinds of affixes are limited in number, a statistical chance for the production of such similar forms will be much higher than others across the dialects. This is the background for the homonymic

---

14 The identical process has occurred repeatedly in historical changes of kinship terms with modifier/modified constructions, e.g., “bo fu 伯父” (elder father) > “bo 伯”; “da bo 大伯” (elder father) > “da 大”, in which original stems have been deleted and modifiers, “bo” and “da”, have been stemized (Iwata 2000:182).
In Chinese dialectology, the ideas and methodologies of linguistic geography have generally not been applied to historical studies in spite of the fact that the phenomena described in this paper are well recognized across the languages and dialects. Etymological studies so far are devoted to establishing the linearity of changes from Old Chinese to modern dialects, but the fact is that such linearity as assumed by Neo-grammarians was often destroyed due to various non-mechanical factors, which linguistic geography has discovered during the past century. Linguistic geography has grounds for contributing to a reconstruction of old language, whereas any etymological study lacking this foundation is risky.

For instance, it is tempting to think that words having the morpheme “liu”, [tʂʰliu1/2 liu0] (hail), [kə5 liu0] (mussel), [tɕɪ3 liu0] (cicada) and [kʰə3 liu1/2] (clod) etc., can be traced back to the “word family” existing in Old Chinese. In line with the scholarly tradition established by the great philologists in Qing (清) Dynasty, Cheng Yaotian (程瑶田) in particular, one would argue that there is a resemblance of the shapes among “hail”, “mussel”, “cicada” and “clod”, and would most likely insist on the existence of the word family, which might have comprised a number of polysyllabic words with k-l initials. However, I suggest that this is the product of a recent innovation motivated by analogical attraction among etymologically unrelated words, which then eventually were incorporated into the paradigm of the “liu” morpheme.

In conclusion, I would encourage Chinese etymological study to be established on the basis of the well-recognized ideas and methodologies of linguistic geography.
References


Wang, Shihua (王世華), and Jilin Huang (黃繼林). 1996. *Yangzhou Fangyan Cidian* [揚州方言詞典]. Shanghai: Shanghai Education Press (上海教育出版社).

Xu, Baohua (許寶華), and Huan Tao (陶寰). 1997. *Shanghai Fangyan Cidian* [上海方言詞典]. Shanghai: Shanghai Education Press (上海教育出版社).