15 : The How and Why of a taxonomy of names

Richard Webber

Context

So far in this seminar we have examined individual names, whether in terms of their frequency, their level of localisation and the parts of the country in which they are most frequently found.

To more fully understand the relationship between names and places it is also helpful to consider each name in terms of the general class of name to which it belongs. For example names such Webber, Tucker, Dyer, Hooper and Retter all originate from distinct occupation groupings within the cloth manufacturing industry. Name historians would typically refer to these as metonyms. Within the general class of metonyms these particular names can be further classified as trade names ending in –er, distinguishing them from more recent names such as Smith or Wright. Within the group of metonyms ending in –er these particular names can be more finely classified as representing occupations associated with the textile industry. This finer level of classification would distinguish them from names such as Butcher or Baker which, though also belonging to the general class of metonyms and within that class to occupations ending in –er, would at the finest level be classified as metonyms ending in -er which are characterised by trade occupations.

There is nothing new about the idea of classifying names in this way. However we do believe that this project is the first to aim to classify so many names on a consistent basis and, in doing this, to use such a fine level of name classification.

Of those 25,730 names with more than 100 occurrences on the 1998 GB electoral roll, 19,528, or 75%, have been allocated to a specific classification. Typically it has been somewhat easier to allocate frequently occurring names (such as Smith) than infrequently occurring names (such as Tink). For this reason the proportion of electors that has been classified in this way, 89.6%, is rather higher than the proportion of names. Names which could not be effectively classified have been assigned to a specific category, ‘unclassified’

Disregarding the names in the category ‘unclassified’, the system classifies each names into one of fourteen ‘groups’. These are then broken down into a set of 67 ‘sub groups’ which themselves divide into 225 ‘fine classes’. Not every group contains more than one sub group nor is every sub group divided into fine classes. ‘Norman’ names, for example, form a group in their own right which has no further sub group or fine classes within it.

In the same manner as one can consider a name, such as Brimacombe, as having a frequency, so likewise can one consider the frequency of the name group ‘toponyms’ (names deriving from a location), of the sub group ‘names of settlements’ and of the fine

class ‘settlement names ending in –combe’. Likewise just as one can measure the level of localisation of a name such as Brimacombe one can measure the level of localisation of the set of names which are toponyms, or which are settlements or which are settlements ending in ‘-combe’. Likewise just as one can identify the location with the most significant concentration of the name Brimacombe one can identify the location with the most significant concentrations of toponyms, of settlement names or of names of settlements ending in ‘-combe’.

Objectives and benefits of classifying names

Names can and have been classified in many different ways. The particular categorisation used in this project is designed to meet a number of aims which can be summarised as follows.

One of the principal aims of the project has been to identify ‘cultural’, ‘linguistic’ and ‘economic’ divides. Whilst the distribution of individual names can give us clues as to where these fault lines may occur, the distribution of categories of name gives us clearer and more reliable indications. For example whilst maps of the occurrence of names such as Ashby, Maltby or Humby will each individually produce a reasonably clear pattern, a map of the occurrence of toponyms which are names of settlements ending in –by will give a very much clearer pattern.

Mapping of frequently occurring names, such as McDonald or McRae, will provide fairly clear evidence of the regional spread of Scottish names when shown at a coarse level of detail, for example for postal areas (such as TR for Truro). However when one wants to examine the concentration of Scottish names at a finer level of geography, such as at the level of an individual postal district (such as TR14, Camborne) or even as postcode sector level (such as TR14 9, Troon, Camborne) it is unlikely that there will be sufficient occurrences of any individual name for this to be statistically reliable. Measurements of the concentration of Celtic names which start with Mac and originate from Scotland will prove much more reliable.

The categorisation of names is also useful when one wants to measure the extent to which an area’s name structure is similar to or different from that of the country as a whole or of another area. If the name structure of the Reading postal area (RG) is more similar to that of Great Britain than the name structure of Halifax (HX) then one may be able to make interesting deductions across a wide range of applications from genetic mix to levels of community identification. But how best should we measure ‘similarity’? One perfectly valid approach would be to examine similarity at the level of the individual name – is each name represented by a similar proportion of electors in RG as it is in HX or in Great Britain? However if we measure similarity in terms of the similarity of the proportions of the population within given areas having names of a similar type, by comparing the proportions of electors with Indian names, with metonyms ending in –er which are manufacturing occupations, with Scottish names starting with ‘Mac’, with patronyms ending in ‘son’ and so on, it could be argued that a more robust measure would be created.

This is particularly the case where we wish to include overseas names as criteria for measuring similarities and differences. One postcode sector in Blackburn may contain many people with the name Chopdat whilst an adjacent one may contain many people with the name Loonat. Both names come from a similar region and caste in South Asian and their presence in the two postcode sectors shows that their populations have similar rather than differing origins.

Name categories are also useful in helping to identify names which originate from very specific locations as distinct from general cultural regions. Names such as Smith and Wright are more common in the English Midlands than in other parts of the country. However it could not be inferred from this that all people with the name of Smith or of Wright originated from a single location in the English Midlands. The name is too frequent and in any case relates to an occupation which would have been dispersed over a wide cultural region (a metonym) rather than to a specific settlement (a toponym). By contrast names such as Laity and Verran, because they are Toponyms which relate to a unique settlement, can be expected to contain members who have significant genetic similarities. These similarities are likely to be greater than those among Toponyms which are topographical, ie names which describe generic features such as Combe, Heath or Dale.

The last significant benefit of assigning each of the names to a set of classifications is that the assignment can itself provide clues to the etymological origin of the name. Take for example the class of names which are metonyms ending in –er. This class of name has a relatively low level of regionalisation, being a type which occurs extensively throughout lowland Britain. Individual names belonging to this class also have relatively low levels of regional concentration. However where we find names in the class, such as Tucker or Ridler, which have significantly higher levels of local concentration than is typical for the class as a whole, we may be identifying names, of which Tucker is a good example, which represent regional terms for an occupation called by a different name elsewhere or, in the case of Ridler, we may have an occupation which is probably associated with an industry, cider making, specific to a particular locality.

The classification system : 1 Cultural Groupings

The classification system consists both of cultural and regional groupings. Where there is a conflict the cultural grouping takes precedence.

The cultural grouping of the classification system starts with Celtic names. Within this overall grouping there are separate sub groups for Cornish, Irish, Scottish and Welsh names. These categories are based on linguistic patterns rather than the regional concentration of names per se.

There are other sub groups which distinguish names which begin with ‘Mac’ but which can not be attributed unambiguously to Scotland or Ireland, names which are valid both with an without the prefix ‘mac’ (such as Loughlin, Cormack) and names which have a

pattern whereby the generic noun precedes a descriptive adjective. Names such as these are particularly common in Cornwall, for example in names starting with Nan- (Nancarrow, Nankerris), Tre- (Trevelyan, Treloar) and Pol- (Polglase) and also occur in Cumbria and Scotland in forms such as Kirkbride, Dunleary.

Within the Irish sub-group distinctions can be made between names starting with Mc or Mac and with O'. Other variants of Irish names are distinguished by their endings, such as –han (Geoghan), -igan (Madigan) and –nan (Hannan).

Within the Scottish sub-group distinctions have been made between names starting with Mc or Mac and those which end in '-ie' (such as Fairlie). Another fine class within the Scottish subgroup are names belonging originally to clans and which are also popularly used as first names (such as Grant, Ross)

Diminutives

Diminutives are names which are usually based on first names to which one of a variety of suffixes have been applied, implying 'little'

These can be organised into three main sub-groups, those ending in –cock (such as Alcock, Peacock), those ending in –ett (such as Bartlett, Willett) and those ending in –kin (such as Watkin, Larkin). Some of these sub groups can be further divided according to whether the first name is male (Peterkin, Hodgkin) or female (such as Margett, Marriott).

Features

The general class of names described by 'features' includes indicators of a person's physical appearance, nicknames which may or may not have described a person's character as well as names which take the form of creatures. Directional descriptors, such as East, North, South and West, are also included in this category.

One important sub-group relates to the colour (Black, White, Green, Rose, Gould) of a person's appearance. Interestingly the names Red and Yellow do not appear in this list.

Many people's names are the names of animals (Bull), birds (Swift) and fishes (Salmon). It is thought that many names of this sort were acquired from the roles which people ones played in Medieval carnivals. Some of these names are of English origin, others from old Norse.

Another important sub-group are names which relate to a persons nature or character. No doubt some names derived from animals were ascribed because they fitted personal characteristics but there are other names too, such as Sly, Quick, which can be grouped on the basis of describing personal characteristics. Another interesting fine class is a set of names ending in –good (such as Toogood, Thoroughgood) which verge on nick names. Nicknames are also included as a fine class in their own right. Many of these names constitute phrases (such as Gotobed) or compound adjectives (Wellbeloved).

Foreign names

A distinction has been made between Celtic names and other non English names.

Within the group of foreign names separate sub groups have been created for Asian, Black African, European, North African and Oriental names. A further sub group is reserved for names which are known to be foreign but whose precise origin can not be identified.

Within the sub group of Asian names a number of separate fine classes differentiate Hindi, Iranian, Muslim, Pakistani/Bangladeshi, Sikh and Sri Lankan names. There are other fine classes for names which can be used by Muslims or Hindus, Muslims or Sikhs and Sikhs or Hindus.

European names are further divided on the basis of language of origin, including Dutch, French, German, Hispanic, Italian and Polish. Separate fine classes are reserved for Nordic names and for East European names. Greek Cypriot and Greek names are grouped together under Cypriot.

Within the Oriental group separate classifications exist for names of Chinese and of Vietnamese origin.

Unfortunately the ethnic group for which no classification is available is people of Caribbean origin. Though most of this population are difficult to identify on the basis of their names there is evidence of a small number of names (such as Baptiste, Gayle) which very high concentrations among Caribbean.

Genitive

The group of names described as ‘Genitive’ is easily confused with another set of name which form a sub group of patronyms, specificly those ending in –s. Both are forms of genitive but whilst in patronyms such as Matthews the name describes the first name of the parent, names in the genitive group are mostly based on the occupation of economic position of the person by whom they were employed or the status or position of their parent. The majority probably describe some attribute of the person for whom they worked.

A genitive –s following a metonym is surprising rare, though a sufficient number exist for this to constitute a separate sub group. More common is a genitive –s following a place name or topographical feature (such as Fields, Coombes), a position of rank or a place of work (Squires, Manners) or a relation (Widdows).

Jewish

Some common Jewish names are easy to identify (Gold, Goldberg, Goldstein) but many other Jewish names are commonly used in the non Jewish community. This is particularly the case of names which can also be used as first names (such as Michael, Frank) many of which were taken in preference to the use of historic family names.

A small separate sub group of Jewish names consists of Jewish patronyms (such as Wolfson, Mendelsohn).

Metonym

Metonyms are names which originally described a person's occupation. Some of these names are quite easy to identify (such as Tanner) but many others are difficult either because the name of the occupation has changed (Walker, Pothecary) or because it may be based on Old English.

One of the major sub-groups within the class of metonyms are names which end in –ard. Many of these names are derived from the suffixes –herd (Sheppard, Coward) or –ward (Steward, Appleyard) and imply a person whose job it was to look after some important resource in the Anglo Saxon or Medieval economy.

A second major sub-group, more modern in origin, are occupations ending in –er. These can be further subdivided into those involving agricultural activities (such as Thresher), industrial occupations (such as Fuller) and those involving a service trade (Carpenter, Iremonger) likely to have been undertaken in an urban setting. A separate fine class is reserved for –er endings describing tasks within the textile industry.

Another important class of metonyms are names ending in –man (such as Silverman, Coleman). Both of these examples are of names which could in certain instances have been used to describe a person's natural appearance but could equally well been used toh describe the occupation which he undertook.

Hoare, Goldsmith and Wheelwright are examples of what are described as 'non –er occupations', typically Anglo Saxon in origin, occupational names which neither end in –ard or –man as well as –er.

Name

Many people have surnames which can also be used as first names.

Quite a common sub group of such names in South West Wales are biblical names such as Aaron and Isaac which may have been adopted relatively lately.

In Cornwall by contrast it is quite common to find surnames which are names of Breton origin (such as Harvey or Gilbert)

Other names are derivative versions of common first names. For example the name Bennett as a surname is likely to be a derivation of the first name Benedict. This sub group of names is likely to have been ascribed relatively early, often before the population was literate.

Among this group these is a significant sub group of women's names but the largest group is of English male forenames excluding those that are Scottish clan names or popularly used by Jews. The names John, Walter and Richard are quite common examples of such names.

Other categories of names are forenames which originate from France (Terry), from Old English (Whipp) and from Welsh (Griffith, Owen) and forenames which have been shortened (Tedd, Thom)

Norman

A number of names are known to have Norman origins. Many of these names are corruptions of French names, for example Pomeroy and Rougemont and were derived from placenames. This group has no sub division.

Patronym

Patronyms are one of the largest groups of names. These are names which have at their root a personal name to which a prefix or suffix indicates paternity or maternity.

Particularly common in Wales are patronyms which originally were formed by prefixing the term ap- (son of). The prefix of many of these names became shorted to 'p-' (as in Probert, Pritchard). As is common in the Welsh language the 'p' often converts to 'b' before a vowel to give Bevan and Bowen in place of 'Pevan' or 'Powen'. Very occasionally one finds a double patronyms involving the prefix ap- as well as the suffix –s, as for instance in Bevans.

Another common type of name is a patronym based on a diminutive. For example there are many diminutives ending in –cock which have had a second suffix added, -s. Willcocks and Wilcox are examples of this naming practice. The same process has occurred with diminutives ending is –ett and applied to names with male (Willetts) and female (Annetts) roots. A similar but smaller group is based on the diminutive ending –pott (Philpotts).

The prefix –fitz, or bastard son of, forms a separate sub group to include names such as Fitzwalter, Fitzwilliam, most of these originally upper class names.

One of the largest sub groups are names ending in –s, such as Jones, Williams, Roberts associated with South Wales. These are relatively recent anglicised names which are typically associated with urbanisation and industrialisation and are categorised separately

from other patronyms with the suffix –s which were already common in Southern England and which are likely to have had a longer pedigree (such as Matthews, Watts).

Another very large sub group consists of names ending in –son. This sub group can be broken down into fine classes on the basis of whether the –son follows a root name (such as Robinson, Watson), a form of diminutive (such as Wilcockson, Hodgkinson, Tillotson), whether the root is male (Robson) or female (Anson) or whether the root is a first name (Johnson), that of an occupation (Clarkson) or of a relation (Widdowson).

Position

Many names are based on an ancestor's social position or status. Example of names relating to positions are Bishop, King, Lord or Pope. One presumes that not all the original holders of that name held such position. It is more likely that some of these names originate from roles undertaken in Medieval pageants.

Another sub group within the position group reflects relationship (such as Younghusband). A third group indicates economic status, for instance whether a person was bonded (Bond) or fee (Fry, Freeman).

Toponym

Toponyms form one of the largest name groups, particularly in terms of the proportion of names.

Toponyms spawn a continuum of sub groups from names denoting a country of origin (Burgoyne, Welsh), a county (Lancashire, Darbyshire), a major town (Chester, Kendall), a less important settlement (Laity, Verran) or a topographical feature, which may be a tree (Ash, Birch, Oak) or a generic feature (Edge, Hedge). Within the sub group of settlement we have identified a large number of distinct place name endings, such as –cross, - dale, -royd, -ton, most of which indicate specific topographic features. These fine classes do not include names of the topographic features themselves. Many of these settlement name types contain large numbers of low frequency names and most are themselves highly concentrated within particularly regions of the country (such as –worthy in Devon and –haugh in Yorkshire)

The classification system : 2 Regional groupings

Once all the names that could be classified into one of the groups had been identified and allocated to be appropriate classification a separate classification was created for those remaining names which had a high level of local concentration.

In this sweep we identified from the remaining pool of names those names which, bearing in mind their frequency, had an exceptionally high concentration in at least one of the 121 postal areas in Great Britain. Subject to exceeding a statistical threshold these names were allocated to a set of regional sub groups as follows: Cornwall (TR, Truro

postal area), East Midlands, Eastern England, London, North West, Northern England, Scotland, South East, South West, Wales, West Midlands and Yorkshire Humberside. A separate category is used to distinguish names of Norse origin that are particular to Scotland's Northern Isles.

These regional categories should not be used for measuring similarities between regions but they are useful for measuring the extent to which local areas have high proportions of highly regionalised and local names.

Conflicts

Given the various characteristics that have been used to create the classification there are numerous instances where classifications cross each other. For example the name Lloyd could be classified either as a colour (in Welsh it means 'grey') or as a Welsh name (in terms of linguistic origin). The name Angove is both a Celtic / Cornish name and an occupational name – meaning a blacksmith. The name Bhopal is both a Foreign / Asian /Hindu name and the location of one of the world's worst environmental disasters.

In addition many names can have more than one origin. For example Gill is both a topographical feature and a name common among the Asian Sikh community. Gold can be both a colour (though more likely written as Gould) and a Jewish name.

In general a policy of priorisation has been used whereby in conflicts of this sort priority is given to Celtic, Foreign and Jewish name categories over Metonyms, Patronyms and Toponyms.

The work involved in allocating names to categories has been undertaken with the assistance of many people including a number from minority ethnic cultures. Chances of identifying correct assignments have been improved by various intelligent sortations of the data.

For example names ending in the string –ides are mostly, though not always, of Cypriot or Greek origin. Most though not all names ending in the string –ah are African. Names more common in postal area NW than anywhere include most names of Jewish origin and most names which are clearly not English and which are more common in UB (Southall) and HA (Harrow) postal areas than anywhere else are worth careful inspection to establish whether they are of Asian origin.

Using the geodemographic classification 'Mosaic', it has been possible to identify names which are most common in areas of predominantly Asian populations and the linkage between names starting with pol-, tre-, rose- and nan- with high frequencies in the TR postal area has been helpful in identifying those with Cornish language roots.

Comparison of name frequencies in 1998 and 1881 have provided useful indicators of likely foreign names as have comparisons of the relative frequency of names in Australia and Great Britain, an analysis which has been effective in identifying a number of

Chinese, Vietnamese and Eastern European names. However there is no doubt that many of the unclassified names could in due course also be given classification codes, very often without having to created additional categories.