

The Family Names of the United Kingdom (FaNUK) Project: Retrospect and Prospect

Richard Coates
United Kingdom

Abstract

The Family Names of the United Kingdom (FaNUK) project database was delivered to Oxford University Press in June 2014, and is scheduled to be published both online and in print in November 2016 as *The Oxford Dictionary of Family Names in Britain and Ireland* (FaNBI). Here, some reflections are offered on the process of creating a resource of this kind, including an assessment of methodological and factual advances that have been achieved and an outline of some possible directions for future research. Many questions have arisen about the reliability and utility of sources of data, and programmatic answers are offered for some whilst acknowledging the emergence of new ones and the persistence of others. Progress has been made in understanding the origin of many surnames, and some choice specimens will be fully discussed illustrating either philological or methodological novelties. Some names have continued to defy explanation, and some of the broader questions that these raise are explored.

This lecture was given on behalf of the whole FaNUK team: Patrick Hanks (lead researcher and chief editor), Paul Cullen, Simon Draper, Duncan Probert, Kate Hardcastle, Harry Parkin and Deborah Cole, and also a range of specialist onomastic and linguistic consultants, too many to name here individually – but especially Peter McClure. The lecturer takes responsibility for the contents, however, and any personal reflections (of which there are indeed some) leading to controversial ideas should not be assumed to be endorsed by the other members of the team.¹

* * *

What Can a Surname Project Do?

Anthroponomastics is a multi-faceted discipline, and accordingly there are many different types of project about surnames that can be envisaged. Surnames are first and foremost linguistic objects, and they can therefore be investigated linguistically, especially from a historical viewpoint, but also pragmatically: they are used to perform tasks in language use. They are distributed in time and space: their changing form and usage can be studied and the varying forms that result can be mapped dialectologically. They are, evidently, associated with real people, and they can (at least in principle) be investigated demographically according to the great sociological variables such as class, occupation and gender. They play a role in social practices which determine or influence their usage. Accordingly, they have an anthropological dimension which can be investigated. They are essential in most European countries as a tool in genealogical (family-historical) research. Whether they have a role to play in association with genetic studies is a moot point. Clearly there has, for some seven centuries, been some relation between the transmission of surnames and that of Y-chromosomes, but it is hazardous to think that their transmission can act as a proxy for

¹ It should be noted that some names regarded as unexplained when the lecture was delivered have now been elucidated, and that some explanations proffered here may differ in the final published outcome of the project.

genetic transmission without considerable reservations, given the prevalence and disguisability of extramarital births, paternity uncertainties, and the possibility of arbitrary surname change in some jurisdictions (notably England). Surnames also have a psychological or psychosocial dimension; it has been claimed that the alphabetical position or other aspects of one's surname can have life-affecting consequences (e.g. Einav and Yariv 2006, Kirchler 2007: 174).²

The FaNUK project is essentially a philological one, whose main goal is to establish the language of origin and the detailed etymology of surnames in the countries occupying the islands of Britain and Ireland. The *Dictionary* (FaNBI) will also include data on the frequency of surnames in Britain and Ireland in 1881 and in recent decades, and on their principal modern geographical distribution(s). We believe that it will be a repository of facts of great interest for their own sake and of potential interest to historians of the relevant languages and general onomasticians, as well as being, of course, a useful tool for family historians, though we are aware that the clustering of bearers of a surname in some place is not proof of the genealogical (and certainly not the genetic) relatedness of its bearers.

Relation to Earlier Work

The stimulus for the FaNUK project came originally from the perception that the current standard reference work, Reaney (1958; third edition by Wilson, 1991), despite certain strengths in its philological detail, was no longer fit for purpose overall, or at any rate not fit for the purpose which seems to be implied by the title *A Dictionary of British (English) Surnames*. Moreover, it is not what it apparently claims to be, namely a dictionary of British (English) surnames; or rather, that is what it is only in a restricted sense. It is a dictionary which explains the origin of some British (English) surnames, but (i) makes no pretensions to completeness, and (ii) includes many surnames which are no longer current but which were evidenced in the Middle Ages. It might more accurately have been titled *A Dictionary of Some Medieval English Surnames, Some of which Have Survived to the Present Day*. That is not to make fun of it, but to spell it out like this highlights effectively what the market was lacking, namely a comprehensive resource dealing with surnames that are actually still borne and therefore of interest to living bearers. It was therefore decided to instigate a project, originally conceived and motivated by Patrick Hanks, to fill this gap, intending that it should be an online resource (though a print edition will appear at the same time as the online one, as noted above). It goes without saying that knowledge moves on, and that some of the conclusions reached by Reaney some 60 years ago, or by Wilson about 40, are no longer tenable. FaNUK needed, therefore, to re-research those names for which Reaney's explanations no longer passed muster. Reaney's conclusions were based on a hugely impressive range of medieval and early-modern documents, but many more such documents have been published in the last half-century, and many of those are now available

² Einav and Yariv: 'Faculty with earlier surname initials are significantly more likely to receive tenure at top ten economics departments, are significantly more likely to become fellows of the Econometric Society, and, to a lesser extent, are more likely to receive the Clark Medal and the Nobel Prize.' Kirchler: 'Tax evaders in the United States were shown to be significantly more likely to be persons with surnames starting with the letters B and W ...'

in electronic form, making the searching, selection and transcribing or uploading of information into a research database much easier, and of course much faster. Amounts of accessible data which would have been unthinkable to Reaney were available to us, much of it in spreadsheet or database form and already processed by palaeographically sophisticated scholars. These included names in medieval feet of fines, 14th-century poll taxes, and early-modern probate records. Other material available included machine-searchable texts of already-published material such as the medieval royal patent rolls. Crucially, we also had access to truly vast amounts of material in spreadsheet format transcribed from parish registers, that is, material from the modern era, starting in 1538. Much of this material needed to be, and was, assessed critically, and some was discarded. However, much was also assessed as being reliable and was therefore able to provide the foundations of a robust bridge between the medieval material and the modern distributions of names derived from official statistics.

A dimension which was missing from Reaney's work was that of the geographical distribution of the surnames analysed. Little was offered, except implicitly (for example, regarding those derived from place names), about where a surname might have arisen, about where it predominates in more recent times, and about what the relation is, if any, between the earlier and the later distributional facts. FaNUK takes geographical distribution seriously, and attempts to provide a bridge between medieval distributions deducible from where names are recorded in documents which are geographically indexed, such as the 14th-century poll taxes, and distributions in the modern era, exemplified in the evidence of parish records. FaNUK has attempted, wherever possible, to tell a coherent story connecting medieval and modern documentary evidence, in the belief that that story may often serve as a research tool for those pursuing genealogical studies, a geographically oriented guide. The conclusions reached are suggestive and probabilistic in many cases, and they are not a substitute for genealogy. Perhaps, though, they are the best that can be achieved until solid genealogical evidence turns up, if it ever appears at all. A full discussion of the relation of FaNUK to earlier work, and of other matters dealt with in a summary way below, is set out in the lengthy introduction to FaNBI. It should be noted that much of this lecture deals with general issues using English names as evidence, but the issues raised apply in principle to surnames of any origin.

The Scope of the Project

FaNUK, as readers will deduce, is hugely ambitious in many respects. The scale of its ambition can be seen in its initial goal of explaining etymologically (with due regard for geography and history) the just under 46,000 names which had more than 100 bearers in Great Britain in one or both of the two reference years of 1881 or 1997. (For comparison, the third edition of Reaney's dictionary explains about 16,000.) Names with fewer bearers were placed in a reserve database with a view to extending the project into a second phase to explain at least some of them. The initial 46,000 were found, on preliminary analysis (which was refined as the project progressed) to cluster into some 19,000 groups, meaning that some

27,000 were treated as etymological variants of those 19,000-odd names which were chosen as main entries, i.e. as head-forms or lexicographical addresses for their cluster of variants.

A second ambitious decision was taken at the outset: to include recent immigrant names, defined for these purposes as those arriving in the country after the reference-year 1881. This was completely new in a work of this kind dealing with Britain and Ireland, but had been foreshadowed in Patrick Hanks' earlier *Dictionary of American Family Names* (2003; DAFN), in which the question of immigration could not only hardly be avoided, but provided the main *raison d'être* of the work. England is no longer populated only by the 'English', understood to include Norman blood (or rather genes), but has accepted an inflow of Welsh, Scottish and Irish people; Britain is no longer populated only by the traditional 'British', defined broadly, but has accepted since early-modern times an ever-expanding inflow of Dutch and Irish people, Huguenots, Jews, Italians, Indians (whether Hindu, Muslim or Sikh), Chinese, Arabs, Cypriots, West Africans and many others, whose names are included in FaNBI not so much to be of interest to members of their own communities (where such exist), but increasingly to people of 'traditional British' stock who deal ever more intimately with people of these most diverse ethnic and religious backgrounds, and for whom the ethnic and cultural barriers of earlier times are of diminishing relevance and importance. Our decision meant that we had to assemble expertise in the form of a battery of expert consultants for the languages and cultures represented, whether for relatively familiar western European names, indigenous names from lands settled, colonized or otherwise 'protected' by Britain (South Africa, Hong Kong, India, Cyprus...), or names representing many post-colonial and non-colonial diasporas (India, Nigeria, Ghana, Muslim countries of the Middle East, Vietnam, Poland...). Some such migrations have returned names of British origin to these islands, with the result that some may now be more characteristically West Indian or Irish (for example) than British (**Walcott**, **Brathwaite**, **Pennant**; **Buggle**, **Stapleton**), though our researches have not produced definite figures bearing on this point, and were not designed to.

The scale of immigrant contribution to the surname stock is enormous. It has also served to highlight the fact that the large number of rare traditional names is counterweighted by some extremely frequent immigrant names; indeed **Patel**, **Khan** and **Singh** are now among the 80 most frequent surnames in the UK.

The consultants we engaged have provided a massive amount of valuable information on surnames formed in many languages, but we acknowledge that we still have some way to go with certain languages or cultures, for example non-Yoruba Nigerian, Ghanaian and some categories of Hindu names.

Historical Data Sources

FaNUK has drawn on a very large range of historical data sources. Some of those are conventionally published. We were fortunate enough to get the publisher's permission, in advance of the project, to process electronically the evidence from England used by Reaney (and Wilson), without necessarily accepting their conclusions about the origin of the names in question. For the other countries of the islands, data has been manually drawn from

handbooks by Black (1946; Scotland), McLysaght (1997) and Woulfe (1923; Ireland), Kneen (1937; Isle of Man) and Morgan and Morgan (1985; Wales). Work already done for DAFN was invaluable in backing up these sources, and in preparing the ground for the processing of a large range of immigrant names. In addition to this material, FaNUK had the advantage of access to important large-scale electronic and online resources. Online resources included the medieval patent rolls, made web-accessible in pdf format and searchable by researchers at Illinois University, and the abstracts of feet of fines, a searchable database online at medievalgenealogy.org.uk. The names in the medieval poll taxes were made available to us in searchable spreadsheet form through the courtesy of their editor Carolyn Fenwick, the Canterbury probate records in the PROB-11 series by The National Archives, and the Irish Fiants of the Tudor period through the good offices of Paul Ell of Queen's University, Belfast. By far the largest single source of data is that generously made available to the project by the Church of Jesus Christ of Latter Day Saints (the Mormons), in a form referred to in FaNBI as the International Genealogical Index (IGI). This consists of a dataset of transcribed parish records of more than sufficient volume to enable the establishment of statistically verifiable connections between surnames and geographical areas, from parish to county to region level. This vast resource required some quality control before use, since it contained material which could not be used exactly as it stood, for a range of reasons. But when its most obvious defects had been purged, it remained a resource in which we felt a further range of typical scribal or transcriptional errors could be identified with confidence, and which therefore could be used with equal confidence as a data source. Such large resources, published and unpublished, printed and online, provided the backbone of the dataset used by the project. What they had to offer was supplemented *ad hoc* by occasional use of other material, for example general web-searches revealing usable information about individual surnames from a large range of cultures, and a range of specialist records, including for example synagogue records for Jewish names.

The Analysis Proper

As previously stated, the surname evidence was grouped into sets of name-forms which are putatively related. This process was, to say the least, not entirely straightforward. There are many names which evidently have more than one etymology, which may be difficult to tease apart even in medieval sources. FaNUK's entry for **Butter**, which may be assigned to five (or six, depending on how one counts them) different origins, reads as follows:

1 Language/Culture: English

- (i) Locative name, occupational name: for someone who was in charge of provisions, the keeper of a buttery or wine store (Old French **boter*). Compare Stephanus del Butere, 1377 in Poll Tax (Hucknall Torkard, Notts).
- (ii) Occupational name: from Middle English *but(t)er*, a reduced form of Middle English *buterer* 'maker or seller of butter'.

Some early bearers may belong with (2).

Early bearers: Turchetillus Butere, 1130 in Pipe Rolls (Dorset); William Butere, 1198 in Feet of Fines (Northants); William le Buter, 1243 in Assize Rolls (Somerset); John le Buttare, 1275 in Subsidy Rolls (Worcs); William le Buttere, 14th cent. in Ancient Deeds v (Warwicks); William le Buterar', 1327 in Subsidy Rolls (Woolbeding, Sussex); Henry le Butter', Richard le Buttar', 1332 in Subsidy Rolls (Selsey, Sussex); Geoffrey Butter, 1327 in Subsidy Rolls (Worcs); John Buttere, 1327 in Subsidy Rolls (Essex); Elinore Butter, 1541 in IGI (Monk Frystone, WR Yorks); Rafe Butter, 1543 in IGI (Colyton, Devon); Jacobus Butter, 1543 in IGI (Audley, Staffs); Jonne Butter, 1551 in IGI (Fletching, Sussex); John Butter, 1561 in IGI (Baumber, Lincs); Thomas Butter, 1597 in IGI (Canterbury, Kent); John Butter, 1717 in IGI (Bolton, Lancs).

2 Language/Culture: English

Nickname: from Middle English *buter(e)* 'butter'.

3 Language/Culture: Norman, English

Nickname: possibly a variant of *Bultitude* through a late Middle English reduction of Anglo-Norman French *Botetorte* 'twisted boot' to *Butter*.

Early bearers: John Butter, 1560 in IGI (Wissett, Suffolk); Henry Butter, 1589 in IGI (Norwich, Norfolk); Frances Butter, 1706 in IGI (Shouldham, Norfolk).

4 Language/Culture: English

Nickname: occasionally perhaps from Middle English *botor*, Old French *butor* 'bittern', noted for its boom in the breeding season and sometimes called 'bull of the bog'. The 1332 form may alternatively belong with **Boater**.

Early bearers: Henry Butor, 1169 in Pipe Rolls (Yorks); Henry le Butor, 13th century in Curia Regis Rolls (Devon); John Botour, 1332 in Subsidy Rolls (Towsington, Devon).

5 Language/Culture: Scottish

possibly identical with (1, i) above, which may be the first element of the village name *Buttergask*, in the parish of Ardoch (Perths); see **Buttercase**.

Early bearers: Adam Butir, 1331 in Exchequer Rolls of Scotland; William Butyr and Patrick Butirr, 1360 in Exchequer Rolls of Scotland (Gowrie, Perths); James Buttir, 1511 in Dunkeld Rentale; Alexander Buttar, 1692 in IGI (Kirkmichael, Perths); Archbald Butter, 1704 in IGI (Bencochy, Perths); John Buttar, 1715 in IGI (Lundie, Angus); Thomas Buttar, 1723 in IGI (Kettins, Angus).

We have done our best to allocate bearers of the surname to particular original senses, but in many cases this can only be done probabilistically or not at all, hence such remarks as that under (1): 'Some early bearers may belong with (2).' For reasons of space, our reasoning

about such allocations may be left implicit, depending for example on inferences from distributions of other variants of a name (as in the case of explanation (3) of **Butter**).

The Problem of Choosing Headforms

A decision had to be taken about which form from a putatively related set should be chosen as headform. This was by no means straightforward, as conflicting criteria can easily be imagined:

- (i) *Choosing the form nearest to the etymological source, or which most clearly reveals the etymological source, where this is reasonably transparent*

Doing this would have been multiply problematic. The most transparent form may be much rarer than its relatives, and may even sometimes be extinct, even if recorded at some quite recent point in the documentary record. **Coltherd**, for example, had 98 bearers, whilst **Coulthard** had 2,247. **Bakehouse** had 54, whilst **Backhouse** had 2,874 and even the fanciful reinterpreted form **Bacchus**, with 624, was nearly 12 times more frequent than the apparently ‘etymological’ form. In any case, which is the appropriate etymological form, and what is the virtue of transparency? **Backhouse**, with a short vowel, represents the Middle English form well in this respect, whilst **Bakehouse** shows the effects of a post-medieval analogical reformation on the basis of the verb *bake* with its long vowel. The otherwise unexplained **Lipton** may be associated with the place name *Lepton* (WR Yorks), in which case compare *Johannes de Lepton*, 1377 in Poll Tax (Wortley, WR Yorks); but *Lepton* itself, whilst clearly existing quite recently (*John Lepton*, 1862 in IGI (Crewe, Cheshire)), now seems to be extinct as a surname.

- (ii) *Choosing the form nearest the current standard English form, where there is one*

Sometimes a decision would need to be made which took into account dialect differentiation. Should **Muir** or **Moor** be taken as the headform for these names which share an etymology? In this case, the difficulty was resolved by treating Scots and English as different languages and giving the names separate entries. The problem remained for e.g. **Fairhall**, a relatively uncommon form by comparison with **Verrall**, and one which seems at least in some cases to have been restored to its etymological (or standard) form by someone with appropriate historical toponomastic knowledge of a place in Sussex. It remained more acutely for e.g. mainly Kentish **Fagg** versus West Country **Vagg**, which are probably derived from a word now obsolete, for which therefore no commonly known standard form exists.

- (iii) *Choosing the form nearest to the orthography of the modern lexical form, where one was available*

This tactic, closely related to (ii) but not identical with it, would have had the effect of downplaying some of the more frequent variants that are most familiar as names, as in the case of **Sergeant**, which had 1393 bearers in 1997 whilst **Sargent** had 5954 and **Sargeant** 4280.

- (iv) *Choosing the form most frequent at the most recent date for which we had evidence*

This tactic would have resulted in a presentational order of variants which, in some cases, would have made for opacity in the narrative of historical developments.

This matter was a Gordian knot, and we had to cut it for practical lexicographical reasons. Generally speaking we have chosen the most frequent modern variant as the headform, but have made *ad hoc* decisions in other cases. The choice of (semi-)transparent **Fairhall** as the headform was justified on the basis that many of its large range of variants (including the twice as frequent **Verrall**) are conspicuously dialectal and appear later. But exceptions were made for e.g. morphological reasons: if a form with final *-s* is more frequent than the corresponding form without *-s*, the simpler form was nevertheless always selected as the headform. Many surnames of the modern period appear to consist of a more basic surname + *-s*. We have taken such forms with *-s*, which generally first appear in the post-medieval period, as variants of the unsuffixed form rather than derivatives (as in the cases of **Butters**, **Johns**, **Smithers** and **Woods**). This takes account of the fact that the same name, of whatever typological category, often appears in suffixed and unsuffixed form in the same area, and may even name the same individual. We still do not know the reason for this early-modern trend in naming, but to treat the suffixed forms as derivatives, i.e. as distinct surnames, would do violence to their nature at the point of their origin. It seems unlikely that they can all be put down to the same origin (some appear to have been pluralized as the apparent etymology demands, such as **Twelvetrees** – on which see further below – for earlier forms without *-s*). We describe all such cases simply as having post-medieval genitival or excrescent *-s*.

The Problem of Obscurer Connections

What is to be done when related forms have drifted apart to the extent that they can only be linked philologically and not in a way which makes sense genealogically for current bearers? It will not be difficult for lay or expert FaNBI users to accept that **Brown** and **Browne** or **Shepherd** and **Shephard** are variants. The following sets are a different kettle of fish:

Dust, Doust
Phalp, Philip
Baffin, Boughen

Pharaoh, Farrar
Whittlestone, Whittingstall
Verrill, Verrall, Fairhall
Twelftree, Queldrick, Wheldrake, Weldrick

All of these pairs or sets can be shown, with varying degrees of confidence, to have a common origin; that is, the differences between or among them can be explained in terms of known linguistic variation and change and/or known orthographic practices. FaNUK/FaNBI is a historical enterprise. Wherever it can be achieved credibly, etymology is used to unify the name-forms of an entry. But in any individual case where a connection might reasonably be suspected but uncertainty persists, the possible variants are treated as separate headforms, as in the cases of **Banting** and **Bunting**, **Eburne** and **Hepburn**, **Thurling** and **Thorning**.

The Problem of Data Availability

The evidential landscape might in itself be viewed as a serious problem for the enterprise because of the inconsistency of data availability. Notoriously, returns for the 14th-century poll taxes are incomplete for certain counties and totally absent for the north of England; there is no comparable single data-source at all for the same period in Wales or Scotland. Data for Cornwall is deeply problematic. It presents many uncertainties even where there are no palaeographical difficulties because of the county's bilingualism in the medieval period. Clerks trained to write in Latin and English have managed to mangle many Cornish names beyond interpretability, which adds to the difficulty caused by the fact that surnames of place name origin – which in Cornwall constitute the vast majority – may be ambiguous because of the duplication of certain place names. At the other end of the spectrum is the data for Yorkshire, where an enormous amount of medieval and other source material has been collected and rigorously analysed by George Redmonds (2015). Nevertheless we are grateful for any material surviving from remote periods. It is of course indispensable for the etymological concerns of the project, even though its patchiness presents problems for any subsequent statistical analysis, especially analysis comparing the onomastic landscape in medieval and more recent times.

The Problem of the Status of the Data

Any project using collected data from a range of sources has to take a decision about the trustworthiness of the evidence. A decision might be taken anywhere on a range from taking everything provided on trust, to the elimination of obviously faulty forms, right through to the rejection of any source that contained a significant number or proportion of forms judged to be erroneous. Faults may enter the chain of transmission at a number of points: a speaker-informant might misspeak him- or herself, a clerk might mishear, accidentally miswrite or rationalize, and any copyist, from medieval legal clerks to modern genealogists, might miscopy, for a range of reasons from struggling with difficult handwriting to inattention to wishful thinking. Publicly-sourced data in general might be thought to offer

difficulties of quality control in view of the fact that many genealogical practitioners – the most frequent source of transcribed data from modern sources – learn palaeography on the job. Take the record of variants we have amassed for **Shepherd**. Many variants are credible in terms of orthographic practices at various times, and will have become fossilized in the chain of transmission, or credible in terms of known dialect variation. These include **Shephard**, **Shepherd**, **Sheppard** and **Shipperd**, ranging from frequent to rare. We can easily discard the occasional **Shephand**; it cannot reasonably be taken as a variant of any other name we have identified, the instances on record are not found in the same geographical area, and if it is taken to be for **Shepherd**, it can illustrate easily understood letter confusions. Other spellings are more problematic. Four instances of **Shepherd** are found in 1881 (all in Lancashire) and six in 1997. This might suggest a single family with an unusual deliberately chosen or retained spelling, or a single clerk in Blackburn with a quirk, but the persistence of the name in 1997 speaks in its favour as a genuine variant. On the other hand, seven **Sheeparads** turn up in Medway Poor Law Union in 1881, and are not found in later sources; suspicion turns either onto a local clerk or a later transcriber, and the form is discounted. The occasional record of **Sheephead** or **Shephead** in 1881 would not be discounted on the grounds of being etymologically implausible as it stands. We have **Sheepshanks** as a parallel, after all; it is not phonologically implausible as a variant of **Shepherd** since preconsonantal [r] is lost in many accents of England; and many other names (whether of nickname or locative origin) end in *-head*. It might on the other hand be discounted because instances of it are scattered (one **Shephead** each in Kendal and Bradford and three in the London area, although the three close together in London give one pause for thought, along with the single **Sheephead** in West Ham). Fortunately, in numerous cases of this sort, the rarity of the forms in question takes them below our threshold for inclusion, and we have not, in the first phase of the project, needed to arbitrate on whether they are likely to be erroneous.

Some Problems of Interpretation

A number of important problems remain to be addressed by name scholarship. The following notes offer some reflections on the problems as they affected FaNUK, and on how, in some cases, FaNUK has been able to arrive at, or at least suggest, solutions, guiding assumptions or ways forward.

(i) *Monogenesis*

A standing issue of great importance in modern anthroponymy is the issue of monogenesis. It is very clear, of course, that many surnames cannot have a single origin, if only because the current number of their bearers would be implausibly large if that were the case (**Johnson**, **Smith**, **Little**, **Young**, **Gray**, **Wood**), even taking into account the possibilities of differential male fertility and the statistical chances of a strong imbalance in the gender of offspring in favour of males. On the other hand, many names appear likely candidates for a unique origin. The benchmark case is that of **Sykes** ‘ditches’ (Sykes and Irven 2000); the work done on this name established that certain surnames that could be expected to emerge in more than one

place where conditions (e.g. and especially those of topography) were suitable were nevertheless probably borne originally by a single bearer who is in principle identifiable if the documentary record is rich enough. This is not an issue which we set out to solve or on which to establish a definitive position, but we took it as a guiding principle that surnames of all types may have a single origin, and this was one of the important reasons to highlight any geographical continuity between the medieval and modern evidence for a name.

(ii) *Priority of toponymic over topographical terms*

There are many surnames deriving from place names, too many in fact to need illustration; there are also surnames which we describe as topographic rather than toponymic, such as **Green**, **Field** and **Town**, which describe or categorize a place without naming it. An intermediate category is offered by names transparently featuring topographic terms but with a qualifier, such as **Longhurst**, **Northfield** or **Highmore**. It now seems likely that such names will rarely be topographic, and that the default position to take is that they are toponymic; that is, they derive from a place having the expression as its name. **Northfield**, for example, is rarely if ever to be taken as a name for someone who lived ‘by the north field’, and may in many cases be from the major and long-established Northfield in Worcestershire; however, its current prevalence in Cambridgeshire suggests that a more local Northfield, e.g. one in Bedfordshire, Hertfordshire or Cambridgeshire, can also be responsible. At all events, a place name is likely to be responsible for such names.

(iii) *First early bearers out of place when compared with later data*

In some cases we have a reasonable, sometimes unassailable, philological interpretation of a surname, but surprisingly recalcitrant early bearer data. **Potton**, for example, often derives from the place of that name in Bedfordshire. But the earliest attestations we have include Hugh de Potton, 1227 in the Melrose Chronicle (Glasgow), and William de Poton, 1289 in the Patent Rolls (Rolvenden, Kent), as well as a decent and comforting spread of 14th-century bearers in Bedfordshire and Essex. Readers will inevitably note the tension between the methodological endeavour to associate a toponymic surname with a place by following a geographical trail of bearers backwards through time, and an apparently perverse early distribution. But laying on one side the question of possible alternative origins, examples like this open wider questions of mobility and migration, and their relation to lordship and trade, in medieval society, for which the datasets of FaNUK can only begin to provide some raw material. They also suggest directions for future research into the relation between linguistic variation and personal and societal mobility.

(iv) *Latin humanistic translations*

It is a familiar matter to historians of continental European surnames that they may sometimes be formed in Latin (and sometimes Greek). Among those with a humanist, classical education, it had become fashionable in the 16th and 17th centuries to adopt such forms to translate native ones; **Agricola** and **Mercator**, for example, often translate Dutch or

German surnames with the etymological meaning ‘farmer’ and ‘trader’ respectively. So when such names appear in Britain, can we be sure that humanistic translation was involved here also? The answer is: apparently not. **Faber** never appears to translate **Smith** in the way that it does **Schmidt**, for example; it is found in Britain only from 1624 and is probably always of European immigrant origin. That said, there is occasional evidence that the situation is rather more complex than this implies. **Pictor** – Latin for ‘painter’ (or indeed **Painter**) – was apparently used by clerks even in medieval times to render Middle English *peintour*. There seems to be no other way of understanding John *Pictor* in a Leicestershire document of 1245 in The National Archives, or a man of the same name, a bailiff of Bedford in 1299. But where it survives as a modern surname, the lack of evidence for continuity suggests that the name did not become established as a hereditary family name in the medieval period but may have been adopted in the 17th or 18th century, under the influence of migrants from northern Europe, as exemplified by this Wiltshire strand: Thomas Pictor, 1722 in Urchfont, Betty Picktor, 1780 in Ogbourne Saint Andrew, and John Pickter, 1803, and Emily Pictor, 1828 in Hilperton.

(v) *Notable regionalities*

Some names illustrate distributions which might not be expected on purely linguistic grounds. The case of **Thomas** is a striking example, and it could be paralleled by others deriving from given names. By the end of the 13th century *Thomas* had become one of the most frequently used given names in all classes of society, and it gave rise to many different surnames. The basic surname form **Thomas** or **Tomas** is widespread in England but it occurs most frequently in the southern counties, from Kent to Cornwall, and most notably of all in the counties bordering Wales. In Wales itself *Thomas* (often as an anglicization of Welsh *Tomos*) became exceptionally popular as a given name and therefore surname in the post-medieval period, although the surname was not generally hereditary until well into the 19th century. It was especially common in south Wales, notably Carmarthenshire, and is now one of the most frequent Welsh surnames, far outnumbering its equivalent in England.

(vi) *The emergence of late variants*

It has become axiomatic that surnames were no longer being created by 1700 or so (in most of England, at least; the latecomers to surname-creation, mainly in the far north-west, are well known). But there are some late-emerging variants of known names that are clearly not scribal errors, in the sense they are not one-offs which disappear when a clerk more literate in the emerging standard English orthography next records a family member; rather, they have the effect of standing at the fountainhead of a new name which then persists. A good example is **Hucklesby**, which is by all appearances a variant of **Huckerby**, a name originating in the North Riding of Yorkshire and now typical of Nottinghamshire. **Hucklesby** turns up borne by William Hucklesby in 1782 in Great Gaddesden, Hertfordshire; James Hucklesby, 1808 in Wheathampstead in the same county testifies to the local variability of the name in its new habitat; and its relation to the Yorkshire original is secured by Mary Huckerby, 1812 also in Wheathampstead. **Hucklesby** has gone forth and multiplied to the extent of having 168

bearers in 1997. There may be an unappreciated number of such local innovations. Family historians are invited to step in to help identify them. Sometimes the new name is sufficiently different from the presumed etymon that, if the variant had been medieval, we might well have treated it provisionally as a separate surname, as in the case of **Twelftree** mentioned above. There is much arbitrary-looking creation of new names in modern times, never systematically investigated; the nature of clerical practices in times of an expanding literate population deserves closer investigation. A favourite example of a clerical aberration is offered by William *Tortoiseshell*, 1790 in Derby, whose name is a desperate attempt (mediated by earlier efforts such as *Tortershell*) to render **Tattershall**, which originates in a Lincolnshire place name – or is the clerk indulging in a little sport at the bearer's expense? The new name, possibly viewed as more elegant and refined than its ancestor, continued to have currency during the 19th century at least, though it now appears to be extinct.

There is a deep problem in the analysis of variation. What is its relation to geographical distribution and population movement? **Ledgeway** appears as late as the early 19th century in Knaresborough (West Riding of Yorkshire). We cannot state *a priori* how a surname arose at such a late date; neither local alteration of some other name nor migration can be ruled out as a matter of principle, though the probabilities of one or the other might be attempted depending on time and place. **Ledgeway** invites comparison with Joane *Ledghey*, 1700 in Redruth, Cornwall, on the one hand, and with Willm. *Ledgear*, 1575 in Mirfield, Petrus *Ledyard*, 1639 in Almondbury, both in the West Riding, on the other. From the linguistic point of view, comparison with the Cornwall name is easier; comparison with the Yorkshire names raises greater linguistic difficulties (notably because all variants of **Ledgard** that have been established with confidence have an <r> in the second syllable, or a final consonant <d> or <t>, or both), but offers greater hope for family historians by sharing a county with **Ledgeway**, even though the earlier names appear some 25-30 miles from Knaresborough. Consideration of names such as this leads to consideration of the causes of variation in general.

A Fundamental Problem in Accounting for Variation and Its Causes

The case of **Ledgeway** highlights a major issue concerning what assumptions should be made when deciding whether to treat some surname *X* as a variant of a phonetically similar name *Y*. Where there is an overlap in their geographical distribution, the impulse to connect *X* and *Y* is driven by the assumption that variants arise *in situ*, rather like variant pronunciations in classic Neogrammarian views of phonetic change, supplemented by variationist sociolinguistics. On this assumption, variation begins 'under the radar'. This question is balanced by that of the extent to which one can or should treat *X* as a variant of a phonetically similar *Y* if they are separated in distribution. Here, the impulse to connect *X* and *Y* rests on the presumption that variation arises where a name is unfamiliar in a new locality, a process rather like mangling foreign words in the process of borrowing. In this scenario, variation begins 'on the radar' and involves clerks in a battle with an unfamiliar phenomenon, creating variants, which may or may not be accepted without demur by the intended bearer. Presumably both types of innovation occur, and that is easy enough to say; what is harder to

decide is whether the available evidence is strong enough in a particular case to support one or other of these explanatory tactics.

Variation arising *in situ*, or in any event not very far from the point of origin, is not at all hard to illustrate. **Habergham** (from a Lancashire place name) and its eleven variants are all still concentrated in Yorkshire, Derbyshire and adjacent counties, and we can confidently say that closeness to the point of origin is no barrier to the proliferation of orthographic, and also phonological, variants.

Variation probably arising as a result of migration can also be illustrated with relative ease, but can only be securely appealed to, in the longer run, when backed by evidence of relevant variability in the homeland, or better still by genealogical evidence.

Brokenshire (found from 1711 in Cornwall) appears beyond reasonable doubt to be a variant of the northern (and highly variable) **Birkinshaw** despite the disparity of distribution, supported by the Lancashire variant *Brockenshaw*.

Blackshire (found from 1764 in Essex and south-east England, also the USA) might be analysed as a variant of **Blackshaw**, a clearly northern English name. Given the late appearance and the difficulty of connecting it with any other name, does the difference of distribution in itself allow the inference that it is a variant of **Blackshaw**? Assuming a pronunciation with final schwa, there are no pressing dialectal phonetic objections. The conclusion would therefore be in favour of its variant status. This is reinforced by the appearance of a *Blackshaw* in Essex in 1726, but of course no genealogical connection is proven.

Flamson is a post-medieval name largely confined to Leicestershire in its early days (from 1566) and now. There is an earlier, medieval, name **Flamstead** (now extinct) which must be from the place of that name in Hertfordshire, but one early bearer is found in Leicestershire. It must be presumed that **Flamson** is an irregular development of **Flamstead** that arose *in situ* in Leicestershire.

Each of these examples suggests the emergence of the new form in the new habitat, either fuelled by clerical unfamiliarity with the name on its arrival, or precipitated by not fully understood changes, sometimes phonetic, sometimes analogical, but to some degree local. Post-medieval local variants of **Birkinshaw** in distant Sussex include **Beconsawe**, **Buttenshaw**, **Buttinger** and **Burtinshall**, which amply illustrate the hazards faced by local clerks in dealing with names from elsewhere.

Honeysett represents an intermediate, and perhaps undecidable, case. It appears in 1669 in Sussex with no obvious antecedents locally. However, a plausible etymon is the surname of Willelmus *Honyswet* in the 1381 poll tax in Lincolnshire, now apparently extinct. **Honeysett** may represent a local attempt to rationalize the name to some extent, or it may illustrate the very frequent loss of <w> after a consonant in many English dialects of both north and south, suggesting the (unattested) possibility of its loss in the Lincolnshire homeland before the surname died out.

‘Unexplaineds’

At the time of the delivery of this lecture, 1,324 entries in FaNUK contained the word *unexplained*. Things were not, however, as bad as a simplistic reading might imply! Sometimes it is truly the case that we have no explanation through lack of any relevant knowledge (**Allotey**, **Beschizza**, **Dhillon**, **Hua** (sense 5), **Kinchin**, **Ledgeway** (see above), **Manktelow**, **Piff**). On the other hand, many names labelled ‘unexplained’ have at least one full explanation (**Doust**, **Totten**), even if that explanation does not appear to account appropriately for all the bearers of which we are aware. For some we have a pretty good explanation but admit some (minor) aspect of it is unexplained (**Candlin**, **Hollingsworth**). Moreover, we have proceeded with a certain conservatism; many of these ‘unexplaineds’ come equipped with a full discussion and at least one defensible suggestion, even if a degree of diffidence is expressed (**Dungey**, **Gingell**, **Kibblewhite**, **Nimmo**, **Strugnell**, **Sussams**), and some of these have been cleared up in subsequent work during the second phase of the FaNUK project (**Pressdee** < **Prestidge** < **Prestwich**, **Ticktum** < **Titcombe**, **Tincknell** < **Tintinhull**). A curiosity is the small set of apparently toponymic surnames for which no place name source has been found, including **Bosomworth**, **Fingleton** and **Yallop**. For such items, the strategic direction includes: (1) keep looking, especially for deserted medieval villages – which depends on the availability of good toponymic surveys; (2) continue to assess the possibility of ‘corruption’ in transmission (especially in relation to unstressed second elements, which are notoriously volatile).

Conclusion: Advances Made During the Project

A summary of positive developments made during the project, as regards both methodology and outcomes, might include:

- Advance 1: FaNBI is based on masses of evidence derived from digitized or electronically searchable versions of many resources, resulting in strongly evidence-based conclusions;
- Advance 2: FaNBI is modern and up-to-date: it includes a very large number of surnames, many of relevance to present-day “consumers”, not just medieval ones, and many never explained before;
- Advance 3: FaNBI’s scope is international and inclusive;
- Advance 4: FaNBI has established, improved, or pointed to the probability of, connections between medieval names and modern names (for example **Annakin**, **Balch**, **Brabazon**, **Gaukroger**, **Goodlass**, **I’Anson**, **Nutkins**, **Stringfellow**, **Waddilove**, **Whamond**);
- Advance 5: FaNBI has achieved, or pointed to the possibility of, more precise localization of the origin of certain surnames, serving to assist in elucidating the possibility of their monogenesis;
- Advance 6: FaNBI has drawn (unsystematically for the present) on community expertise, notably that of members of the Guild of One-Name Studies.

Richard Coates
University of the West of England
United Kingdom
richard.coates@uwe.ac.uk

References

- Black, G.F. (1946) *The Surnames of Scotland*. New York: New York Public Library.
[Republished (1996) Edinburgh: Birlinn.]
- DAFN = Hanks, P. (ed.) (2003; online edition 2006) *Dictionary of American Family Names*.
Oxford: Oxford University Press.
- Einav, L. and Yariv, L. (2006) 'What's in a Surname? The Effects of Surname Initials on
Academic Success. *Journal of Economic Perspectives* 20.1. 175-188.
- Kirchler, E. (2007) *The Economic Psychology of Tax Behaviour*. Cambridge: Cambridge
University Press.
- Kneen, J.J. (1937) *Personal Names of the Isle of Man*. Oxford: Oxford University Press.
- MacLysaght, E. (1997) *The Surnames of Ireland*. 6th edn. Dublin: Irish Academic Press.
- Morgan, T.J. and Morgan, P. (1985) *Dictionary of Welsh Surnames*. Cardiff: University of
Wales Press.
- Reaney, P.H. (1958) *A Dictionary of British Surnames*. Oxford: Oxford University Press.
[3rd edn. (1991) by R.M. Wilson, retitled *A Dictionary of English Surnames*.]
- Redmonds, G. (2015) *A Dictionary of Yorkshire Surnames*. Donington: Shaun Tyas.
- Sykes, B. and Irven, C. (2000) 'Surnames and the Y Chromosome'. *American Journal of
Human Genetics* 66.4. 1417-1419.
- Woulfe, P. (1923) *Irish Names and Surnames*. Dublin: M.H. Gill.