New Directions in the Study of Family Names

Patrick Hanks
Professor in Lexicography, University of Wolverhampton, UK
Editor in Chief, *Dictionary of American Family Names (DAFN)*
Editor in Chief, *Oxford Dictionary of Family Names in Britain and Ireland (FaNBI)*
patrick.w.hanks@gmail.com

Abstract

This paper explores and explains recent radical developments in resources and methodology for studying the origins, cultural associations, and histories of family names (also called 'surnames'). It summarizes the current state of the art and outlines new resources and procedures that are now becoming available. It shows how such innovations can enable the correction of errors in previous work and improve the accuracy of dictionaries of family names, with a focus on the English-speaking world. Developments such as the digitization of archives are having a profound effect, not only on the interpretation and understanding of traditional, 'established' family names and their histories, but also of names in other languages and other cultures. There are literally millions of different family names in the world today, many of which have never been studied at all. What are good criteria for selection of entries in a dictionary of family names, and what can be said about them? What is the nature of the evidence? How stable (or how variable) are family names over time? What are the effects of factors such as migration? What is the relationship between family names and geographical locations, given that people can and do move around? What is the relationship between traditional philological and historical approaches to the subject and statistical analysis of newly available digitized data? The paper aims to contribute to productive discussion of such questions.

1. Goals of Genealogical Research and Surname Lexicography

It is important to recognize at the outset that the goals of research and the goals of lexicography can be – and often are – different. Academic researchers in the humanities typically choose rather a narrow topic and aim to present new findings, while saying everything that can reliably be said about the points presented. In other words, typically they aim at depth and certainty, even if what they have to say is of interest only to a very limited audience. In genealogy this tends to manifest itself as the study of one name (or even one family) at a time. Lexicographers, on the other hand, typically aim at breadth; that is, they tend to want to create works of reference that have the Benthamite goal of being of greatest usefulness to the greatest number of users. For this reason, among others, they often have to be satisfied with statements that do no more than summarize the probabilities. Nowhere is this more true than in the study of family names. A dictionary of surnames aims to say something about every name selected by systematic criteria. It comes as a shock, therefore, to discover how many names in a community are omitted from traditional dictionaries of surnames. If the author has failed to find an etymology he (for in the past such authors were always male), simply left it out.

This is in line with a general rule of scholarly behaviour. Scholars confronted with an inconvenient fact (such as a name about which they have nothing to say and about which they can find out nothing) can quietly ignore it. Conscientious lexicographers, on the other hand, will note the existence of such names, even if they can say no more about it than that it is "unexplained". Surnames lexicographers in the past, including the great P. H. Reaney, simply omitted hundreds of names for which they could find no historical evidence. Modern surnames lexicographers, by contrast, attempt to reconcile the scholarly goal of depth of treatment with the lexicographic goal of breadth of coverage. If, as sometimes happens, nothing can be said about the etymology of a modern surname, its existence and location can at least be noted, for the sake of completing the inventory and possibly prompting future research.

The lexicographic goal of presenting a complete inventory is hampered in Britain and Ireland, not only by the absence of digitized early records – a deficiency that is now gradually being remedied – but also by the legal requirement imposed on the Office of National Statistics (and its Irish counterpart) not to release census date until 100 years after a census has taken place. Fortunately, Professor Richard Webber of the University of Newcastle has been able to construct an inventory of contemporary surnames from files from a variety of public sources together containing the names of 48 million different adults in the UK.

The tension between research goals and the goals of lexicography is palpable in the study of family names. Lexicography is necessarily selective; genealogy less so. Genealogical studies typically aim to collect as much data as possible about the history of a single name or a small cluster of names and to show whether a family relationship exists among bearers of the same name, past and present. Genealogical research is a very popular pastime in the 21st century. The Federation of Family History Societies has over 160 member societies worldwide. Recognition of the fact that a surname may be borne by several different families led to the establishment of the Guild of One Name Studies. Typically, a genealogical or one-name researcher works backwards from the present day through records of past centuries. A few such researchers are lucky enough to find records of a target surname as early as the 16th century in Parish Registers, which began in England in 1538. Records from before that date are plentiful, but sporadic, offering little hope of a genealogical link to the present day. By contrast, a lexicographer compiling a dictionary of family names typically tries to find the earliest records of the target surname and to link these to an etymon (a vocabulary word, place name, or personal name from which the family name was or could have been derived).

Thus, there is a need to bridge the gap between the goals of genealogically motivated research and the goals of historical surnames lexicography. Two recent dictionaries of surnames in Britain have aimed to reconcile these goals. Redmonds' *Dictionary of Yorkshire Surnames* (2015) aims to say as much as can reliably and succinctly be said about the family names that have a historical association with that large county, including surnames that originated outside the county. If a surname has an established relationship with Yorkshire, Redmonds records its existence, even if he is obliged to acknowledge that it is unexplained. In many cases, the evidence that he has collected enables him to propose an original and convincing explanation.

An even more ambitious project is a national dictionary of family names in not one but two countries – Britain and Ireland – which have been exchanging populations for over eight hundred years. This is the *Oxford Dictionary of Family Names in Britain and Ireland*, alias *FaNBI* (Hanks, Coates, McClure and others, 2016). It is a 4-volume dictionary, also available

on line. It aims to have an entry for every surname known to have more than 100 bearers in modern Britain and Ireland (using figures compiled by Professor Webber). Inevitably, given this goal, the editors have occasionally been forced to admit that a few names are 'unexplained'. A second edition of *FaNBI,* already in preparation, will contain an entry for every 'established' family name with more than 20 bearers in the 1881 UK census, together with every recent immigrant name that has more than 100 bearers in modern Britain, using figures compiled by Richard Webber. *FaNBI* makes a distinction between 'established' family names and 'recent immigrant' names. The latter are, typically, names that have come to Britain in substantial numbers since 1945. Recent immigrant names are rare or non-existent in the 1881 census.

Each explanation in *FaNBI* is supported by statistical analysis of geographical and historical records, in particular parish registers. *FaNBI*'s goal was to summarize what I known or could be discovered about family names in Britain and Ireland, thus providing a reliable framework for future genealogical, historical, demographic, and philological research. Ths goal has been achieved.

Still, much remains to be done by way of analysing old records, for example the systematic and comparative statistical analysis of names in taxation records, insofar as these have been digitized, as has been done by Carolyn Fenwick (1997-2005) for 14th-century Poll Tax returns. Among the ludicrosities of modern scholarly publishing is the fact that Fenwick's superbly well-organized database has not yet been made publicly available. She generously shared it with the FaNBI team, but it has been published only in printed-book form, not as an electronic product.

FaNBI has developed a number of new directions for surnames research and prefigured others. It has superseded the work of P. H. Reaney (1957, 1997), which itself was a pioneering work of scholarship in its day. In recent years in Britain and Ireland, large quantities of data (hundreds of millions of historical records) from sources such as parish registers have been transcribed and made available on line, not by national record offices, which have been slow to take advantage of new technology, but by volunteers working under the auspices of organizations such as the Federation of Family History Societies and the Church of Jesus Christ of Latter-Day Saints (Mormons). Such people are to be commended for their enthusiasm, although, if they do not have training in palaeography (and even if they do, insofar as they are human and therefore fallible), they are liable to make transcription errors. Such errors are problematic for researchers only if they are systematic. In other words, a single error in among millions of accurate transcriptions does not normally matter much, but regular and repeated misreadings of one letter for another may have a serious effect on statistical analysis and hence on interpretation of the data. For example, the supposed English surname *Weaves* turns out to be a systematic misreading of records in the 1881 census for *Weaver*.

Examining Reaney's explanations in the light of newly available evidence, the *FaNBI* team coupled expertise in historical linguistics with statistical and geographical studies and recently developed computational techniques for studying large collections of data in machine-readable form. This resulted in procedures that clarified many fudges, corrected many downright errors – seemingly plausible explanations that have turned out to be untenable in the light of the newly available evidence – and filled in a surprisingly large number of gaps in earlier works. To be explicit: *FaNBI* has explained the origins of several thousand family names that had never been explained before. In a few cases, after

painstaking and often time-consuming examination of the available evidence, some family names were noted as unexplained.

FaNBI recognized that the study of family names in Britain requires competence, not only in the history of the English language, but also in Anglo-Norman French, Old Norse, Irish and Scottish Gaelic, Manx, Welsh, and Cornish. The *FaNBI* team has made contributions to scholarship in all these languages. Furthermore, influxes of economic migrants and refugees in the 19th and 20th centuries meant that additional expert consultants were needed for language and cultures as diverse as Yiddish and Hebrew, Chinese, Hindi, Urdu, Tamil, Yoruba, and Ibo, among others. In these areas, the *FaNBI* editors obtained expert advice from leading scholars.

3. The European and World Context

In most countries of Europe the binomial system of naming for individual persons (a system consisting of two elements: one or more given names plus a hereditary surname) was gradually adopted between the 11th and 15th centuries. Some places, notably Scandinavia and Welsh-speaking parts of Wales, adopted the binomial system later than others, while others, notably Ireland, were ahead of the field, having already got a system that looks remarkably like a binomial one well before the 11th century.

In some countries the study of surnames continues to be neglected even now, but in many others important new initiatives have taken place. An excellent example is the German Surname Atlas project at the University of Mainz (see Dräger and Schmuck 2009). Similar initiatives elsewhere will no doubt follow.

In today's world, human beings live in a global community. They would be well advised to accept the inevitability of this fact, with mutual respect for differences of culture and belief and (preferably) some enthusiasm for the benefits that such cultural variety brings. Among other things, a by-product of this simple fact is that the binomial system of personal naming (given name + inherited surname) in becoming a worldwide norm. In recent decades it has even begun to be adopted among people from places where different naming systems are in use, especially those who have migrated away from their place of origin. For example, some people from the Arabic-speaking world, which traditionally has a much richer system of personal naming than Europe, are now adopting something that looks remarkably like a binomial system. Immigrants to Europe from far-away places tend to select part of their traditional name as if were a surname and – lo and behold! – in the second and subsequent generations of such migrant families, the 'surname' comes to be regarded as hereditary. Something similar is happening with migrants coming to other parts of the English-speaking world from southern India, which traditionally does not have a binomial system for naming individuals.

Interest in family names and family history has never been higher in Europe and countries of large-scale immigration from Europe, such as America, Canada, Australia, and New Zealand. But sadly, this interest has, on the whole, not been well served by the available reference works. This prompts an evaluation of traditional dictionaries of surnames, along with an evaluation of future possibilities.

To deal with the first question first: as on-line publishing became popular towards the end of the 20th century and the public thirst for knowledge about personal names showed no sign of

abating, a very large number of web sites began to appear, purporting to explain the origin and meaning of personal names (forenames and surnames). Many of these sites are inaccurate (to say the least). Some contain little more than fantasy based on guesswork.

It therefore seems desirable to mention here, before moving on, the most reliable traditional printed dictionaries that explain the origins of surnames in European languages. These works have laid foundations for future research. During the 20th century (the period before corpus tools became available to surnames researchers), a few dictionaries of surnames in some European language appeared, based on philological scholarship. Such dictionaries typically aimed at breadth of coverage (large numbers of entries), rather than depth (detailed analysis of the historical evidence for each name). The main traditional dictionaries of this kind are:

- For English: Bardsley (1901); Reaney (1958; third edition 1997)
- For Irish: Woulfe (1923)*; Maclysaght (1957, 1985)*
- For Scots and Scottish Gaelic: Black (1946)
- For Welsh: Morgan and Morgan (1985)
- For French: Morlet (1991)*
- For Flemish and Belgian French: Debrabandere (1993)
- For Italian: de Felice (1978); Cafarelli and Marcato (2008)
- For Spanish: Faure, Ribes, and García (2001)*
- For Portuguese: Machado (1984)
- For German: Brechenmacher (1960); Bahlow (1967; English translation 1993); Kohlheim and Kohlheim (2005).

Full titles and publication details for these books are given in the references at the end of this paper. However, it must be said that the quality is somewhat variable, and explanations in even the best of these works are not supported by statistical analysis of the evidence of early bearers, for the simple reason that very little if any historical data had been digitized (i.e. put into machine-readable form), let alone analysed at the time they were compiled. Some of these dictionaries -- those marked with a star -- contain no citations at all for early bearers. This means that their etymologies are typically speculative -- that is, they are supported by little or no evidence in the form of records of early bearers, nor are they based on geographical or statistical analysis of data.

4. Marrying Geographical Distribution to Historical and Linguistic Scholarship

4.1 Surnames and Locations in England

The history of dictionaries of family names in Britain is discussed in Hanks (2009) and Hanks and Parkin (2016). Data for the historical development of surnames is exceptionally unstable, compared with data for the development of place-names or words in the general vocabulary. People – the bearers of surnames – not only move around and die; they also change their surnames, sometimes as a result of marriage or inheritance, but more often as a result of illiteracy (which, up to the mid 19th century, was the norm in most places) and clerical ignorance or interference. The surviving records of taxation, legal proceedings, and suchlike are voluminous -- unmanageable without the aid of a very large computer, and such resources have only recently begun to become available.

In the absence of large-scale volumes of evidence in machine-readable form – so-called 'big data' – the only option available to 20th-century scholars was to use their scholarly expertise (which in many cases was profound) to make the best guesses they could. In other words, when in doubt, they fudged. The dangers of fudging can be illustrated with an example from Reaney and Wilson (1997), where there is an entry for the following cluster of surnames (alphabetized at *Ravenshaw)*, supposedly from a place-name meaning 'the raven wood':

Ravenshaw, Ravenshear, Ramshaw, Ramshire, Ramshaw, Ramshire, Ranshaw, Renshaw, Renshall, Renshell.

The first point to note is that Webber's data shows that *Ravenshaw* itself is apparently extinct. The Mormons' International Genealogical Index shows that it was thriving in the 17th and 18th centuries, but by the time of the 1881 census it was down to 6 bearers. Evidently, it died out at the beginning of the 20th century. The phenomenon of surname obsolescence and death has been insufficiently studied. The death of a surname is not a rare event. Sturges and Haggett (1987) predicted, on the basis of a simple statistical model based on reasonable assumptions about the number of male children per family in each generation, that out of every surname borne by one individual in 1350, over half would have died out by 1950. Of course, reality is hugely more complex than a simple statistical idealization, and Sturges and Haggett's starting point took no account of a variety of factors such as the number of bearers of each surname already extant in the 14th century. Nevertheless, their prediction is a stark reminder of the caution needed by anyone trying to get to grips with surnames data. It has now become clear that many surnames recorded in previous centuries have not survived, while others have inexplicably flourished.

Let us return for a moment to Reaney and Wilson's *Ravenshaw* cluster. Webber's figures show that a rare variant, *Ravenshear*, is recorded with just 23 modern bearers, while on the other hand *Ramshaw* is thriving, with 1445 modern bearers in England.

Mention of *Ramshaw* brings us to a more serious problem with this cluster, namely the false assumption of common origin. Clusters of surnames that seem to be similar do not necessarily all have the same etymology. Recent research has shown that in this case (as in many others) lumping them all together is a mistake. Wilson's assumption that *Ramshaw* is a variant of *Ravenshaw* is phonologically plausible but wrong. *Ramshaw* is from a place of that name in County Durham, in north-east England, which was named in Old English as *hramsa halh* 'wild-garlic nook'. Nothing to do with ravens or woods. A similar problem arises with *Renshaw*. This has been shown to be from a minor place called *Renishaw* in Eckington (Derbyshire), which was named in Middle English as *Reynold's shaw* (= 'Reynold's wood or copse'). That this is the correct etymology is confirmed by a record of it as *Reynalddeschawe* in 1281 in the Assize Rolls for Derbyshire.

The geographical distribution of surnames was first studied systematically by H. P. Guppy (1891). Guppy compiled a county-by-county table of frequencies of the surnames of farmers in Kelly's directories. He noted that there is generally a strong statistical association between a surname and a locality. Of course, farmers are only a tiny subset of the population, but Guppy believed (rightly) that farming families tend to stay in one place for generations. We now know that the geographical stability of families applies to many other classes, too, while at the same time, the geographical distribution of a few names throws up some tantalizing puzzles. Guppy's methodology was impeccable. Having identified the county with which a

surnames is most associated, he went on, in many cases, to study the family history associated with the name.

Unfortunately, Guppy's insights were disregarded for almost a hundred years by leading surnames scholars, who took the views that the only sound methodology for studying surnames was to collect historical records, especially medieval records, without reference to the statistics of geographical distribution. For example, the English Surnames Series, founded by Richard McKinley, consists of books with the titles such as *The Surnames of Oxfordshire* (1977), *The Surnames of Lancashire* (1981), and *The Surnames of Sussex* (1988). These works contain important insights and record large numbers of facts drawn from medieval and later records, but they cannot be said, by any reasonable definition -- statistical, historical, linguistic, or other -- to contain an account of the surnames particularly associated with the counties mentioned in the titles. All too often, the names discussed are more associated with some county other than the one named in the title, while many of the names that we now know to have a statistically significant association with the named county are not mentioned at all. More accurate titles would have been something like "Some facts about historical records noticed while examining surnames in [the county named]". These deficiencies, reinforced by the disregard of distributional data by England's greatest surnames scholar of the period (P. H. Reaney), had far-reaching deleterious effects on the understanding of matters such as the relationship between names and dialect – and on family histories.

Literally thousands of surnames in Britain, especially locative surnames, show a strong statistical association with a particular locality. But such an association does not yield certainties of interpretation. In some cases, this statistical association may shift from one locality to another, as a cadet branch of a family thrives and multiplies while another branch (which may have been the senior line) withers and dies. But such cases are the exception rather than the rule. More often, than not a place (often a minor place such as a farm or even a single building) bearing the relevant name or something like it is found in or near the locality where the surname is most frequent today. Still more so, if the geographical data analysed is a 19th-century census. In most families, up to two dozen generations have left home and moved away since the surname was first recorded – but mostly they did not move very far. One example (one among thousands) is the surname *Copplestone*, which has been strongly associated with the county of Devon since medieval times. Bearers of this name are found scattered all over southern England, but it is still more common in the county of Devon than any other. There is a tiny place of this name in central Devon, notable for its impressive granite cross, mentioned as *Copelan Stan* in a pre-Norman charter. This tiny place is without doubt the source of the surname.

On the other hand, occasionally a surname may survive but may break away and establish itself, perhaps with a change of form, in a far-away location. An example is *Whistlecraft*, a surname that has been strongly associated with the county of Suffolk in East Anglia since at least the 16th century. What is its origin? Assiduous searching for a suitable possible source in East Anglia yielded nothing. FaNBI asserts that this name is one of several variants of the Cheshire place-name *Wolstoncroft* 'Wulfstan's enclosure'. At first sight this may seem like a wild speculation, not least because Cheshire is on the other side of the country from Suffolk. However, closer examination reveals supporting evidence for this derivation: in South Lancashire, which is next door to Cheshire, there is a surname *Whistlecroft*, which is known to be an altered form of *Wolstoncroft*. *Whistlecroft* is of course strikingly similar to *Whistlecraft*. The most plausible explanation of the Suffolk surname is that at some date

before 1524 (when the surname was first recorded in Suffolk), a bearer of the name *Whistlecroft* migrated from South Lancashire or Cheshire to Suffolk. This early migrant has not yet been identified, and may never be. Such is the case with regard to evidence for the origin of the majority of surnames: usually, no records of the very first bearer(s) of a surname exist. They may have been lost, but more often than not such records never existed.

Occasionally, evidence of such early migrations survives. For example, William de Wrotham is known to have come to Bedfordshire from Wrotham in Kent in 1206. His modern descendants bear the surname *Rootham*, which is particularly associated with Bedfordshire and Northamptonshire.

An impressive demonstration of what can be achieved by way of creating a resource showing the geographical distribution of surnames is the *Surnames Atlas of Great Britain* (Archer 2003-2015). The input data consists of all the surnames and forenames, together with their locations, recorded in the 1881 census, which had been painstakingly transcribed and digitized over several years towards the end of the 20th century by members of the LDS Church (Mormons) and the Federation of Family History Societies. Archer wrote computer programs which, among other things, map the geographical distribution statistically (in 1881) of every surname in Britain recorded in the census. The mapping is done at two levels of delicacy: by counties and by Poor Law Unions (groups of parishes), which were salient at the time of the census. The Surnames Atlas shows at a glance the geographical distribution of each surname recorded in the 1881 census. A future challenge for Archer and his ilk is to map changes in the distribution of each surname over time. Meanwhile his Surname Atlas enables researchers, among other things, to compare clusters of differently spelled surnames and to evaluate the hypothesis that they all have the same origin.

As an example of a cluster of surnames, we may take the cluster of surnames derived from the village of *Fosdyke* in South Lincolnshire, which is located near the Wash, a broad, shallow inlet of the North Sea. Two families bearing this surname are recorded in the 1381 Poll Tax, but *Fosdyke* has not survived as a surname in Lincolnshire. However, it is not extinct – far from it! Instead, it merely moved gradually away from its source, acquiring several different spellings and pronunciations along the way.

Only 42 bearers of the surname *Fosdyke* (the original place-name spelling) are recorded in the 1881 census, and almost all of them are in East Suffolk – quite a long way across East Anglia from South Lincolnshire. The most common spelling in 1881 is *Frostick*, with 202 bearers. It may not be immediately obvious that *Frostick* is a variant of *Fosdyke*, but comparison with other surnames of the region in the light of known facts about East Anglian dialects, sound changes, and spellings provides convincing evidence that this is the case. An intermediate form is *Forsdyke*, which likewise has a strong association with the county of Suffolk. In this form, the name has 147 bearers in the 1881 census. It is clearly a variant of *Fosdyke* with an intrusive orthographic *-r-*. How is this apparent oddity of spelling to be explained? The first point to note is that the dialects of south-east England (including the standard middle- and upper-class accent of Southern British English known as RP -- 'Received Pronunciation') do not pronounce pre-consonantal and word-final *-r-*. Pre-consonantal and final *-r-* was lost in these dialects of English at some period between the 14th and 18th centuries, so that the word spelled *farmer* is now pronounced /ˈfaːmə/. A second point is that in stressed syllables the East Anglian dialect has a characteristic long back vowel, One result of this was that the surname *Atherton* came to be written in East Anglia as *Arthurton* and was associated by folk etymology with the given name *Arthur*.

Another result was that the pronunciation of the first syllable of *Fosdyke* was indistinguishable from that of the word *force*. Hence the intrusive *-r-*. This change of spelling, coupled with geographical distance, meant that *Forsdyke* lost its association with the Lincolnshire place-name; the second syllable was shortened, so that the name acquired a variant spelling, *Forsdick*. This was further altered to *Frosdick* and *Frostick*. A further complication is that the form *Fosdick* (on record in 1553 in Suffolk) is now rare in Britain but has thrived and multiplied in the USA: the 2010 census records 1359 bearers of the surname *Fosdick* in America.

The hypothesis that for several centuries *Fosdyke* and its variants were surnames borne by families of coastal fishermen or seaborne carriers is supported by close study of the variant spellings, which reflect reasonable attempts by ordinary people and clerics alike to record East Anglian dialect pronunciations.
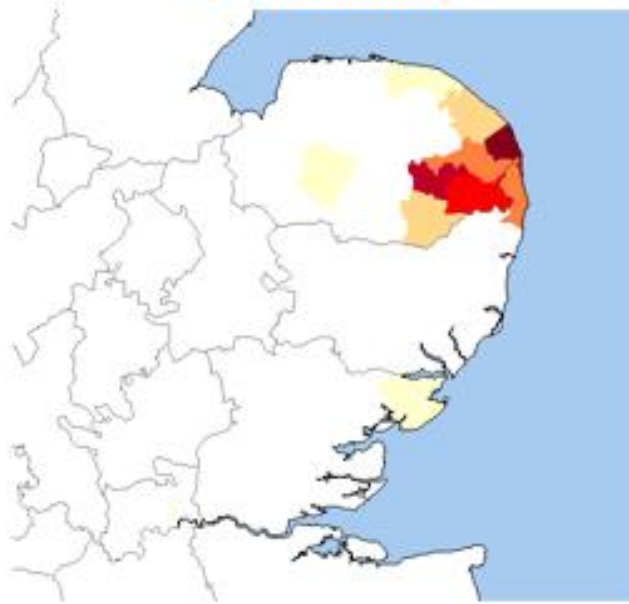

Fosdyke in 1881
(42 bearers)

(Archer: distribution by Poor Law Union)

15

9

Forsdyke in 1881
(147 bearers)
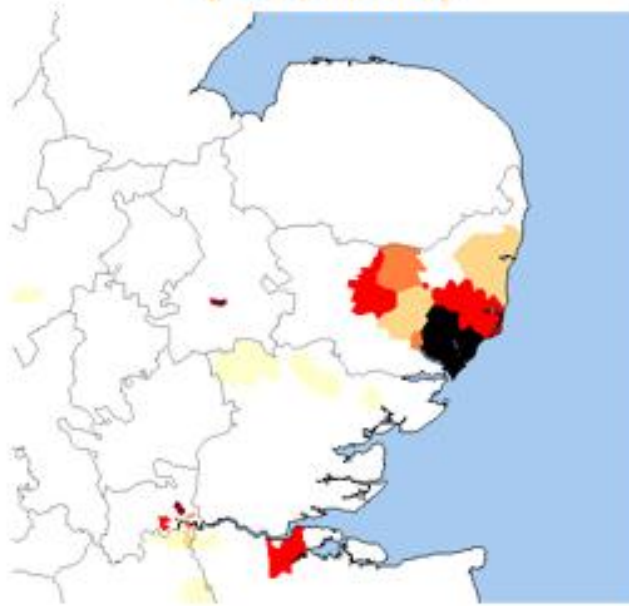
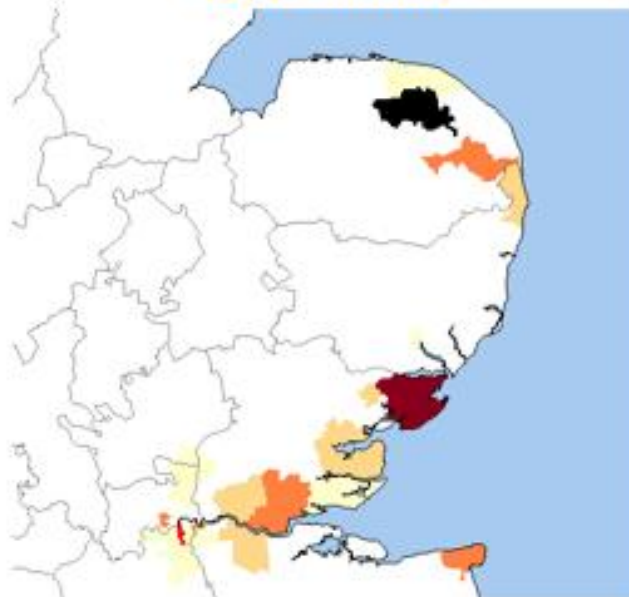(Archer: distribution by Poor Law Union)

Frosdick in 1881
(187 bearers)

(Archer: distribution by Poor Law Union)

Frostick in 1881
(202 bearers)

(Archer: distribution by Poor Law Union)

Alongside these maps, analysis of the historical records in the light of modern scholarship shows that this little cluster of East Anglian surnames do indeed all share a common origin, even in cases where the derivation is not immediately apparent. The relevant entries in *FaNBI* (taken from the Research Database, not from the published text) are as follows:

Forsdyke

Frequencies: GB 1881: 147; GB 2011: 319.
Main GB location, 1881: Suffolk.
Variants: Forsdick, Fosdyke, Fosdike, Fosdick, Frostick, Frosdick, Forsdike.
English locative name from Fosdyke (Lincs), which is recorded as *Fotesdic* in 1183, *Fotesdich* in 1196, and *Fotesdik* in 1202. The place-name derives from the Old Scandinavian personal name *Fótr* (with the Middle English genitival inflection *-es*) + Old English *dīc* or Old Scandinavian *dík* 'ditch, water-channel'.
Early Bearers: Walter *de Fotesdik*, 1202 in Assize Rolls (Lincs); Robertus *de Fossedike*, 1381 in Poll Tax (Moulton, Lincs); Emma *Fossedik*, Emma *Fosdik*, 1381 in Poll Tax (Whaplode, Lincs); John *Fosdyke*, 1524 in Subsidy Rolls (Suffolk); Margaret *Fosdyke*, 1557 in IGI (Wingfield, Suffolk); John *Frosdicke*, 1581, John *Frosdyke*, 1591, Thomas *Frostick*, 1620 in IGI (South Walsham, Norfolk); Johis *Frosdick*, 1610 in IGI (Horstead, Norfolk); Samuel *Fosdike*, 1661 in IGI (Shotley, Suffolk); Robert *Forsdick*, 1669 in IGI (London); Robert *Forsdike*, 1757 in IGI (Levington, Suffolk); Mary *Forsdyke*, 1760 in IGI (Filby, Norfolk).
References: Ekwall, DEPN, p. 185; Lincs Place-Names, pp. 45–6.

Forsdike

Frequencies: GB 1881: 171; GB 2011: 235,
Main GB location, 1881: Suffolk.
English; see **Forsdyke**.

Forsdick

Frequencies: GB 1881: 155; GB 2011: 202.
Main GB location, 1881: E Anglia, Herts, and London.
English; see **Forsdyke**.

Fosdick

Frequencies: GB 1881: 48; GB 2011: 16.
Main GB location, 1881: London and E Anglia.
English; see **Forsdyke**.

Fosdike

Frequencies: GB 1881: 44; GB 2011: 52.
Main GB location, 1881: Suffolk and Cambs.
English; see **Forsdyke**.

Fosdyke

Frequencies: GB 1881: 42; GB 2011: 0.
Main GB location, 1881: Suffolk.
English; see **Forsdyke**.

Frosdick

Frequencies: GB 1881: 187; GB 2011: 354.
Main GB location, 1881: Norfolk.
English; see **Forsdyke**.

Frostick

Frequencies: GB 1881: 204, GB 2011: 460.
Main GB location, 1881: Norfolk; Essex.
English; see **Forsdyke**.

A measure of the variability of form for surnames can be judged in this small sample by the fact that the most frequent forms (*Frostick* and *Frosdick*) are the two spellings that are furthest removed from the etymologically 'pure' form, *Fosdyke*, which appears to have died out completely in the 20th century.

4.2 Surnames Research In Ireland

In Ireland, where almost the entire population is (or was, until recently) stable and homogeneous, the effects of immigration from anywhere other than nearby parts of Britain on the stock of family names are so far negligible. Whether that will still be the case in 20 or 30 years' time remains to be seen. Ireland (unlike, say, Hungary) is clearly determined to play its part in building the emergent 'global village', not only by exporting family names to distant climes, but also by accepting incomers.

Native Irish surnames are generally more ancient than English surnames, in many cases denoting clans that already existed in Ireland early in the first millennium AD. Clan names are generally derived from traditional Irish Gaelic personal name and nicknames. Irish surnames also denote roles within a clan. For example, the surname *Ó Bairdín* means 'descendant of the bard'. Irish-language surnames are prefixed by *Ó* 'descendant of' or *Mac* 'son of'. In Tudor times, Irish Gaelic surnames regularly acquired an Anglicized form. Subsequently, in some but not all cases the patronymic prefix was dropped. Thus, *Ó Bairdín* became *Bardeen* or *Barden.*

The phonologies of Irish and English respectively are very different, and the spelling of Irish surnames is typically influenced by the etymology. As a result the English and Irish forms of a name may be very different. An extreme example is *Corish* and *Corris,* which are Anglicized forms of *Mac Fheorais* 'son of *Piers*'.

The main challenge for 20th-century Irish surname researchers was to identify the Irish-language source of each Anglicized Irish surname and to separate fact from fiction in the traditional etymologies that had been handed down in the families themselves. To this end, a magnificent job was done by two scholars in particular: Padraic (Patrick) Woulfe and Edward Maclysaght.

Woulfe's dictionary (1923) is based on deep knowledge of the history and dialects of the Irish language and the processes of Anglicization (some would say, garbling), which account

for the modern Anglicized forms of many surnames of Irish origin, which have since spread to Britain, America, Australia, and elsewhere.

Several medieval documents that shed light on early Irish surnames have survived; some have been digitized. Chief among them is the Annals of Ulster (*Annála Uladh*), which record in the Irish language events throughout Ireland from A.D. 431 to 1540.

The most valuable resource that sheds light on the Anglicization of Irish surname is known as the Tudor Fiants. From 1521 to 1603, a record was kept of the administrative decisions (issued as 'letters patent') by the government in Dublin, affecting all Ireland. The letters patent themselves have not survived, but the brief record of decisions has, in the form of Fiants (meaning roughly, 'let these things be done'). The names of over 120,000 individuals in Ireland in the Tudor period were recorded, generally along with their status and location. Fortunately, at the end of the 19th century, the Fiants were painstakingly translated into English, which what is believed to be a reliable transcription of the names of the Irish individuals, some of which seem to have baffled the Tudor clerks in Dublin, who evidently did their best to record phonetically the names that they heard. The Fiants were an invaluable resource for the *FaNBI* team, who cite them often. A decision was taken at the outset that *FaNBI* should include Irish family names as well as those found in Britain, as the two are inextricably intermingled.

Maclysaght was a historian and genealogist, who rose to be Chief Herald of Ireland. He died in 1985 at the age of 99, having just completed the fourth edition of his *Surnames of Ireland.* Maclysaght's work in particular is suitable for the general reader, being crisp, clear, and readable. However, both these works are marred by a fault that is characteristic of surnames dictionaries in some other languages too: they do not publish evidence of early bearers in support of their explanations. This defect and others will no doubt be remedied in a forthcoming dictionary of Irish surnames, being edited by Muhr and Ó hAisibeil.

Ireland and Britain have been exchanging surnames for almost as long as surnames have existed - certainly since the 'invasion' of south-east Ireland in 1170 by forces commanded by a Norman warrior lord bearing the nickname 'Strongbow'. This was Richard de Clare, Earl of Pembroke. The city of Pembroke lies at the extreme south-western tip of Wales. Its magnificent castle is of considerable strategic importance, as it controls the western approaches to the Bristol Channel and one of the two major routes between Wales and Ireland. In 1169 Diarmid, king of Wexford, who had temporarily been ousted from his kingdom, appealed to Strongbow for support, offering as an incentive the hand of his daughter in marriage. Strongbow took the bait and duly organized an invasion, taking control (with the blessing of King Henry II) of the affairs of south-east Ireland in a characteristically Norman way. (It has been said that the Normans treated warfare as a kind of joint-stock enterprise.) One effect of these events was that *Clare, Fitzgerald, Fitzmaurice, Bermingham*, and several other names borne by Norman barons and other aristocrats became established as Irish hereditary surnames.

- *Clare* is derived from a town of this name in Suffolk, eastern England, where William the Conqueror granted an estate to one the family that bore Strongbow.. *Fitzgerald* is now a very common name in Ireland; it is derived from the Norman French given name *Gerald*. (*Fitz* in Norman French means 'son'.) Gerald was the constable of Pembroke Castle in 1169; he played a major role in Strongbow's invasion, as did his son *Maurice*, who founded a separate Irish dynasty.

- The surname *Bermingham* was brought to Ireland from England in 1170 by another follower of Strongbow, Robert de Bermingham, whose son Piers became conqueror of Connacht and lord of Athenry. In this spelling, the surname is typically Irish, rather than from the English West Midlands. From Piers another branch of the family derive the surname *Corish*, mentioned above.

- The surname *Walsh*, meaning 'Welshman', has a strong statistical association with Ireland. The family name does not merely denote anyone with Welsh ancestry, but is strongly associated with Ireland. Often (but not necessarily) it denotes someone descended from one of the Welsh soldiers who followed Strongbow and his commanders on their expedition into Ireland in 1170, and who ended up settling in the country.

- *Butler*, in addition to being a English surname, is also the name of a historically important family in Ireland, where it was Gaelicized as *de Buitléir*. In origin it is a Norman status name,
- denoting an officer of highest rank in a Norman household, nominally connected with managing the supply and importation of wine. But the butler's power and influence extended far beyond the wine cellar. As head of the royal household, the chief butler of England had the ruler's ear and could be very influential. In or about 1185 Theobald FitzWalter (1165-1206) was appointed Chief Butler of Ireland by Prince John, favourite son of King Henry II, who was leading an expedition in Ireland aiming to impose royal authority on both the Norman barons and the Irish chieftains. However, there are about 47,000 bearers of the surname *Butler* in Britain and a further 7,600 in Ireland, not to mention over 200,000 in the USA and many thousands more elsewhere in the English-speaking world. No doubt only a small number of these bearers of the name, even in Ireland, are descended from Theobald FitzWalter.

To summarize the position of Irish family names: the majority of longest-established surnames in Ireland are of either native Irish or Norman origin. To these must be added surnames brought in from Wales (since the 12th century), England (especially Tudor adventurers), and Scotland – especially the names of Border reiver families and others who were beneficiaries of 'plantations' (lands confiscated in the 16th and 17th centuries from native Irish chieftains and given to incomers whose loyalty to the English crown was believed to be more reliable). A future dictionary of surnames in Ireland must not only focus on names of Irish Gaelic and Norman French origin, but must also take account of the long and often troubled relationship with the English Crown.

4.3 Surnames Research in Scotland

Like Ireland and Wales, Scotland's family names are listed in explained in *FaNBI,* and much of what has already been said about English names applies equally to Scottish family-names research. Here, therefore, I shall add only a few paragraphs commenting explicitly on the Scottish situation.

Scotland is a small country of remarkable linguistic diversity in terms of language history. The complexity and importance of its history (including the history of its languages) is out of

all proportion to its size. In 2010 it had a population of 5.3 million, which is almost exactly one tenth of that of England.

Scotland does not have a large-scale digitized resource similar to the English Parish Registers or the Irish Fiants, which would support research with copious computerized medieval and later records. Nevertheless, some early Scottish administrative documents have been digitized, and these were consulted by the compilers of *FaNBI*. Scotland also has a superb reference work, namely George F. Black's *Surname of Scotland* (1946). Black spent many years combing through Scottish historical documents, extracting citations for events, and arranging them under the names of the participants. *FaNBI* was able to add some early records and improve his philology, but without Black's pioneering efforts, the interpretation of Scottish family names would have been even more difficult than it was.

The historic languages that have affected the formation of family names in Scotland are Pictish, Cumbric, Scottish Gaelic, Old Norse, and a form of English known as Scots or Lallans (i.e. the language of the Lowlands). A few words about each of these will explain something about the linguistic context in which Scottish family-name research has to be conducted.

Pictish: There are just 45 mentions of 'Pict' or 'Pictish' in *FaNBI*. Since no records in the Pictish language have survived, there is little that can be said about it in a surnames dictionary beyond the speculation that a few surnames such as *Ogilvie* may be of Pictish origin. Perhaps the most interesting of these references is at the surname *McAlpine*. Modern bearers of this surname sometimes claim, with only orally transmitted legends as supporting evidence, to be descended from a certain *Cionaodh mac Ailpein* (known in English as *Kenneth McAlpine*), who died in 858. He was a king who, if the legends are to be believed, briefly united the Picts and the Gaels. The surname *McAlpine* was well established in Perthshire by the 13th century, but it never established itself as a coherent clan, analogous to other Scottish clans − although in recent years there have been attempts in Canada and elsewhere to revive (or create) one.

An almost equally obscure lost language was that spoken by the Strathclyde Britons, who had their capital at Dumbarton on the Firth of Clyde up to about the 9th century. This language was **Cumbric**, a Brythonic language more closely related to Old Welsh than Gaelic. Traces of it are found in place-names such as *Abernethy* (which gave to a surname) and surnames such as *Galbraith*, which is a Gaelic term meaning 'British foreigner', referring to the Strathclyde Britons.

Scottish Gaelic is spoken today by only a few thousand people in Scotland (mainly in the Highlands and Islands), though there are also speakers of Gaelic in Canada and New Zealand, loyal to their ethnic roots. In the 10th-12th centuries it was widely spoken throughout most of Scotland. Historically, the language was brought to Scotland by settlers from Ireland in about the 5th century. The two languages have since diverged considerably. Both have a rich variety of dialects, some of which are the source of distinctive surnames. For example *Carson, Cavens, Sorbie, McIlwrick, McToldridge,* and *Galloway* itself come from the region of Galloway in the extreme south-west of Scotland. Indeed, it would be a mistake to think of a sharp geographical distinction between the realms of Irish and Scottish Gaelic. There was in the Middle Ages a Gaelic-speaking dialect continuum extending from the southwest of Ireland to the north of Scotland. In the 6th-8th centuries a shadowy confederation of kingdoms called Dalriada existed, with member kingdoms on both sides of the Irish Sea.

In some cases, the Anglicized form of a surname of Gaelic origin differs in Scotland from that in Ireland. For example, the Irish Gaelic surname *Ó Dochartaigh* 'descendant of the hard-hearted one' has yielded Irish *(O')Doherty* and Scottish *Docherty*. The Irish Gaelic patronymic *Mac Amhalghaidh* 'son of *Amhalghadh*' is normally written *McAuley* in Ireland but *Macaulay* in Scotland. This surname came to be confused with *Mac Amhlaoibh* 'son of *Amhlaobh*', the Gaelic form of the Old Norse personal name *Óláfr (Olaf)*.

The latter example serves as a reminder that in a few areas of Scotland, notably the Outer Hebrides, Shetland, Orkney, and Caithness, **Old Norse** was spoken – in many places alongside Gaelic, with which it appears to have co-existed harmoniously. This bilingualism affected surname development. For example the Old Norse personal name *Sumarlíðr* 'summer warrior', denoting a viking, was adopted in Scotland and England (where its Norse origin was forgotten and it gave rise by folk etymology to the surname *Summerland* (which has nothing to do with either summer or land). The personal name was common in Scotland, where it yielded a widespread Gaelic personal name *Somharle* and the surname *Mac Somhairle*, which was Anglicized as *McSorley*.

Another Scottish name of hybrid etymology is *McLeod*, which is from Gaelic *Mac* 'son' + Old Norse *Liótr* 'the ugly one.' *McLeod* is well known as the name of a Scottish clan, with over 26,000 bearers in Britain alone. The clan system is unique to Scotland, with the result that a few surnames (originally clan names) are borne by very large numbers of people, not necessarily related to one another by bold. Not only are descendants of the original progenitor regarded as being entitled to claim clan membership, but also retainers, servants, tenants, and others. These very common Scottish clan names are mostly of Gaelic etymology (but also Old Norse, as in the case of *McLeod*, and Old Norman French. as in the case of *Sinclair)*.

Finally, in the lightning tour of the historic languages of Scotland, we come to the language that was to become dominant in Scotland for most of the period for which historical records exist in abundance. This was a northern dialect (or rather, cluster of dialects) of **English** (sometimes called **Inglis** in Scotland). This dialect is generally called **Scots** or **Lallans** (i.e. the language of the Scottish Lowlands, as distinct from Gaelic, which was seen as the language of the Highlands). Scots itself exists in a rich variety of local dialects. It achieved dominance largely for political reasons, as it was the language of the administration in Scotland from the 12th century onwards and was spoken at the royal court alongside Anglo-Norman French.

The key figure here is King David I (reigned 1124-53), who introduced Norman bureaucracy to the kingdom of Scotland. He had spent a considerable part of his youth at the court of the English king. While in England, he married the Countess of Huntingdon. When he became king of Scotland, at the age of about 40, he set about instituting a number of changes in the administration of his kingdom, sometimes referred to as the 'Davidian revolution'. One consequence of the Davidian revolution was that, during and after his reign, a number of surnames of English origin (borne, no doubt, by his retainers and those of his successors) established themselves as typically Scottish surnames from an early date: names such as *Lindsay, Ramsay, Hamilton, Coventry, Sinclair, Ainslie, Laidlaw*, and many others. In some cases, prosopographical research may be able to establish who brought each of these surnames to Scotland, when, and why. This was done in *FaNBI* for some names, but not systemically. For example, at *Ainslie*, *FaNBI* says:

The family first appears in Scotland attached to the church establishment of Glasgow during the episcopate of Walter of Saint Albans (1208–32). William of Ainslie was a steward and clerk of Bishop Walter. The first lay member of the family in Scotland was one Aymer of Ainslie, a knight, in the 1240s, and a John of Ainslie apparently held land at Crossford, Lanarks, in the 1290s; however, the family seem to have mainly been active in Roxburghs and Fife in the Middle Ages.

4.4 American Family Names

This section of this article is based on research carried out for *the Dictionary of American Family Names* (*DAFN*; Hanks, Hardcastle, and others 2003). At the time of writing (April 2018), a second, enlarged edition of this 3-volume work is in preparation. The number of entries is planned to increase from 70,000 to over 80,000. It is expected to be published in four volumes in 2021.

Some case studies of American family names are summarized in §4.4.6 below.

4.4.1 Where did each name originate?

The USA is a nation of immigrants. Almost the entire population consists of immigrants or descendants of immigrants, who arrived from somewhere else in the world. The first task, therefore, for a student of surnames, is to identify the source language, culture, and locality of origin for each surname being studied. Within that culture, an etymology can be sought.

In many cases, the origin appears to be obvious. *Smith* appears to be straightforwardly English. However in America it has also absorbed several other names, some of which are cognates (e.g. German *Schmidt*, Dutch *Smit*) while others are linguistically unrelated but have the same meaning (e.g. Italian *Ferraro*; Czech *Kovář*). Over 100 different source names meaning 'smith' are found in European languages, and several hundred more with meanings like 'silversmith' and 'goldsmith'. Of course, not all immigrants to the USA bearing the names *Schmidt*, *Smit, Ferraro*, or *Kovář* adopted the English translated surname *Smith*. Some families prefer to keep the surname in its original form. Others try to keep their original surname but have accepted that American English does not have diacritics, so *Kovář* is Americanized in some families as *Kovar*.

Where a regular pattern of Anglicization or Americanization has been observed, *DAFN* says so. For example, the English surname *Hearst* in America has absorbed at least one like-sounding Jewish name, namely *Hirsch.* However, it must be recognized that many other names have been Americanized in unexpected and unrecorded ways.

Genealogical web sites and discussion forums occasionally provide useful information of this kind. For example, the following explanation of the American surnames *Delashmit* and *Shumate* was based on a posting in a genealogical forum:

> Americanized form of French *de la Chaumette* 'from La Chaumette', name of several places in the central part of France. Pierre de la Chaumette (born 1673) came from Rochouard, France, via England to Gloucester, NJ, in about 1698; in NJ he became known as Peter Delashmet. Later he was joined by his brother Jean de la Chaumette, a

wealthy colonist from Martinique and VA. Either this Jean or his nephew (Pierre's son) became known as John Shumate.

Such information is like gold dust to the onomastic lexicographer, but extremely hard to find. Genealogical web sites need to be reorganized so that historical and linguistic information contributed by researchers is stored accessibly and separately from the vast numbers of genealogical inquiries, announcements of reunions, and general chitchat that inevitably characterizes genealogical forums. The information archived in this way would then need to be assessed for reliability – historical and linguistic reliability, not just genealogical interest.

Such an archive could beneficially be linked to a digitized archive of passenger lists, recording the names, dates, geographical places of origin, and North American destinations of every recorded passenger on the vast numbers of ships that have brought immigrants to North America since the 17th century. Such records exist in profusion (though inevitably some have been lost). Those that survive need to be transcribed, digitized, databased, and analysed – an unimaginably vast undertaking, requiring a large team to work for decades. Which may explain why so little has been done.

4.4.2 Correlation of surname with diagnostic forenames

Hanks and Tucker (2000) noticed that there is sometimes an interesting correlation between forenames that are ethnically and linguistically distinctive and particular surnames. Accordingly, we compiled a small database of 'diagnostic' forenames – forenames that are uniquely identified with just one particular culture. Accordingly, we got the computer to correlate surnames of 'unknown' origin with diagnostic forenames, and to our delight, the number of surnames of unidentified origin was reduced by half. Practices of giving names to children, even in the US, are clearly remarkably conservative, insofar as a source culture is known to and treasured by parents selecting a name for a child. Armed with this information, iin many cases we were able to obtain etymologies from relevant experts.

The correlation also prompted some surprising results. For example, we had assumed that, as a family name, *Barabas* is widely distributed in Europe, especially Spain. However, a statistically significant correlation (in a US telephone directory of 73 million entries) with the forenames *Attila, Zoltan, Akos, Bela, Gabor, Istvan, Miklos,* and *Sandor* prompted us to give greater prominence to the Hungarian origin of the American family name.

4.4.3 Selection of entries for the Dictionary of American Family Names (DAFN)

A few years after each census, when the data has been digitized and processed. the US Census Bureau publishes a list of frequently occurring surnames (defined as all surnames found with 100 or more bearers), together with their frequencies and certain other information. Information from the 2010 census, immensely valuable to genealogists and lexicographers alike, is publicly available on line at https://www.census.gov/topics/population/genealogy/data/2010_surnames.html. The census bureau is careful not to release any information that could lead to the identification of individual people.

There are 162,253 surnames with 100 or more bearers in the 2010 US census. Anyone setting out to compile a dictionary of American family names must decide on criteria for selection of entries. More than one criterion can be used. The obvious criterion is frequency, but explaining over 160,000 surnames was far more than *DAFN*'s limited resources could accommodate. Moreover, it became clear that, as the frequency of a surname decreased, the likelihood of it being unexplained – and given current resources, unexplainable – increased.

In the end it was decided to set a threshold of 300 bearers in the census. In other words, the second edition of *DAFN* (in progress) will include an entry for every name with more than 300 bearers in the census

Other criteria for selecting a family name as an entry in the dictionary include historical importance of the family or of a particular member of the family, or the philological or other interest of the name itself. Examples are given in the case studies in section 4.4.4 below.

Scribner's *Dictionary of American Biography* (1928-36) often gives information about the ancestors, including immigrant ancestors, which can shed light on sources of family names. Unfortunately, the more recent Oxford *Dictionary of American Biography* does not do this.

4.4.4 Immigration and American family names

It is the common fate of all human being that they die. Before they die, they may or may not move around, sometimes from one county, city, or village to another – often a nearby one, in other cases from continent to another (as happened with almost all present-day American family names). These simple facts are among several that contribute an element of instability to the study of surnames. Sometimes a whole family dies out, and the last member may take a surname with him or her. Such an event poses a dilemma for the lexicographer. Should some names be included in the dictionary for historical reasons, even if it no longer exists? A handful of case studies will illustrate the point.

In the 17th century Britain and the Netherlands competed for colonial settlements on the eastern seaboard of North America. The Dutch settlements of New Netherland were successful for a time, but eventually (in 1664), the Director General of New Netherland, Peter Stuyvesant, recognizing the futility of resistance without support form the other side of the Atlantic, surrendered peacefully to the British. New Amsterdam was renamed New York. At the same time, to the North of New Netherland and New England, the French were busy colonizing Quebec, while far to the south, the Spanish had already established a vast empire stretching from Mexico to Argentina. The principal colonists' languages of North America in the 17th century were therefore English, Dutch, French, and (to a lesser extent) Spanish. Each of these has left a profound impression on the surnames of the USA. Americanization was common in the 17th and 18th centuries. So, for example, *Updike* is an Americanized form of Dutch *Opdijk*, denoting someone who lived 'beside the watercourse'.

The first German settlement in North America was founded in 1683. The heyday of German-speaking settlement did not really get under way until the early 18th century. It continued in full force until the mid 20th century, when it started to slow down, According to the US Census Bureau there are approximately 44 million people of German ancestry in modern America, the largest single ethnic group.

- In the 18th century, it was common practice for German immigrants to have two surnames: one being the original German version and the other being an Americanization. This has yielded hundreds of American name that look as if they might have an English origin but are actually German. The main patterns of Americanization of German family names are described in more detail in Hanks (2003 and 2016). Three examples out of several hundred may be given:

- A famous example is *Cashdollar*, which has nothing to do with money, despite its modern form, which is due to folk etymology. It is an Americanization of the German surname *Kirchthaler* 'person from a place called Kirchthal'.

This situation regarding French in North American is more complex. The British colonial authorities discouraged immigration by Roman Catholics, with the result that French Catholic immigrants were confined mainly to Québec and adjacent regions (from where they made war with considerable enthusiasm but less success on their English-speaking neighbours for most of the 18th century) – and to Louisiana. Bearers of French family names are also found in New England – especially Maine and Vermont – from an early date A feature of Canadian French names is the existence of so-called 'dit' names or 'secondary surnames'. These originated as nicknames for soldiers. An example is *Jean Lefevre dit Leborgne,* literally 'John the smith called the cross-eyed' (or 'the one-eyed'), His descendants could bear either of two surnames: *Lefevre* or *Leborgne,*

A completely different class of French immigrants to North America were the Huguenots. Huguenots were French-speaking Protestants. The sect originated in Geneva in the 16th century, and rapidly gained massive popularity throughout the French-speaking lands. In 1598, after several massacres of Huguenots in Paris and elsewhere, King Henry IV of France signed the Edict of Nantes, which allowed freedom of conscience in matters of religion. But in 1685 his successor Louis XIV revoked the Edict in an attempt to impose Roman Catholic orthodoxy on his subjects. Huguenots were once again the victims of murderous persecution. They fled in their thousands to Protestant states in Europe, including England, where the government encouraged them to emigrate onwards to colonial North America. There, they settled mainly on the eastern seaboard and among the Dutch-speaking Protestants of the Hudson Valley.

Surnames from the Iberian Peninsula are plentiful in North America. They arrived in the Caribbean, Mexico, and South and Central America decades before the French, British, and Dutch started to North America. A disproportionately large number of them came from Galicia and the Basque Country. Many modern American surnames are of Iberian origin, often having undergone some changes of spelling. Some surnames associated with Spanish-language forenames are from a native language of Latin America. For example, *Sinchi* is from the Quechua adjective *sinchi* 'strong, brave, courageous'.

However, it must be said that the state of knowledge about Hispanic American family names at the present time is pitiful. Very little work has been done; an enormous amount is needed. The surnames of very large numbers of Hispanic people remain to be investigated and explained. Even the most basic resources (historical records, for example) exist throughout Latin America, but remain to be transcribed and digitized.

Perhaps the biggest need in American surnames research is to create databases of searchable, analyzable historical records (including, but not limited to digitization of Spanish-language records in Mexico and Latin America as well as the United States). What would be a satisfactory US equivalent of IGI's transcriptions of early bearers and Ireland's Tudor Fiants, both of which adduced as evidence bt FaNBI? The most obvious candidates are passenger lists -- records of the names and other details concerning the passengers on ships that brought them to North America, from the 17th century onwards. Unfortunately, little systematic work has been done to date on the transcription and digitization of passenger lists, although they are regular 'ports of call' for genealogists.

In a few happy cases, study of existing dictionaries of surnames in in European languages the researcher to a reliable etymology in an onomastic dictionary in the source language itself. But such is the lamentable state of surnames studies in almost every country in the world that, even if an etymology can be found, it must be treated with scepticism and re-evaluated in the light of more recent evidence. Quite often it happens that no entry for an unexplained name can be found in even the best local surname dictionary. Scholars are reluctant to admit failure, so, rather than admitting that their best efforts have resulted in failure, they prefer not to acknowledge the existence of unexplained names. They simply leave them out and say nothing abut them. This is a dereliction of duty, if only because the acknowledgment that a name is unexplained can serve as an incentive to future researchers. Even more lamentably, there are some language and cultures which have no surname dictionary at all.

These immigrants come from almost anywhere else in the world. American surnames for various parts of Europe are comparatively well understood.

Italian surnames in North America tend to preserve their original spelling, although Anglicized pronunciations are now widely heard, so that, for example the common Italian-American *Abate* is pronounced as two syllables (ə'beit), although in Italian it has three. Likewise, Polish surnames tend to preserve their original spelling in America, although spelling pronunciations according to some approximation to English text-to-sound rules. For example, the Polish American surname *Krzewicki* (Polish /krʒə'vitski/) is often pronounced as if it were spelled *Krizwicky*. Unexpected puzzles arise even with regard to English surnames. For example, the surname Greenwich is sometimes given the traditional English pronunciation /'grɛnidʒ/ but also is heard with the spelling pronunciation /'gri:nwitʃ/. However, the pronunciation of American surnames is beyond the scope of this chapter.

- Despite claims by the film director Sam Peckinpah (1925-84), *Peckinpah* (more commonly but less famously, *Peckinpaugh*) is not an American Indian name. It is an Americanization of German *Bichenbach*, the name of three places in Germany.

- The rare American family name *Knockstead* puzzled family members and professional genealogical researchers alike. It looks like an English surname, but it isn't. No such surname is known in England. The family has a strong statistical association, historically, with Pennsylvania, a state that his known to have been an area (among other things) of German immigrant settlement in the 18th century.

Almost certainly, this name is an Americanized variant of the more common German American surname *Knackstedt,* This is of North German origin, probably an altered form of *Knackster(d)t,* which is known to be a nickname from Low German *Knack* (High German *Knochen*) 'bone' + *Stert* 'tail', presumably a nickname for a bony lad.

- *Stuyvesant* is a comparatively rare Dutch surname, but of great historical importance, as Petrus (Peter) Stuyvesant (c. 1611–72) was director general of New Netherland in 1647–64. He came from Friesland and was born in 1611 or 1612 in Peperga, where the old reformed church recently has been renamed Pieter Stuyvesant Church. He entered the service of the Dutch West India Company in or before 1635 and served as governor of Curaçao in 1643–47. He was famous for his silver leg, the original having been amputated after it sustained a gunshot wound during a siege in 1644. In 1664 he was compelled to surrender New Netherland to the British, and thereafter he lived quietly until his death on his farm (Dutch *bouwerij*) on the East River, in what is now the Bowery in Lower Manhattan.

- A similar case, though the statistics are even more pronounced, is that of *Brattle* in Cambridge, Massachusetts, This was an important name in the 17th century. It comes from the village of Brattle, near Ashford in Kent, England. Captain Thomas Brattle (c.1624–83) was a seaman and ship-owner, one of the first to ply between England and the new colonies in New England. By the time of his death he was reckoned to be the wealthiest man in New England, where he and his family eventually made their home. His son, also called Thomas Brattle (1658–1713), was a learned, astute, and influential man – though not influential enough, alas, to counteract an outbreak of mass hysteria. He was appointed treasurer of Harvard College at the age of 35 (in 1693) and was noted for his rationality and humanism, which included opposition to the Salem witchcraft trials of 1692. The last known bearer of the surname in North America died in the 1990s. The dilemma facing that onomastic lexicographer is whether to include this name, for historical reasons, in a dictionary of American surnames, or whether to exclude it, on the grounds that it no longer exists in the USA.

- *Throckmorton* is an English habitational name from the tiny village of Throckmorton in Worcestershire. The place-name is derived Old English *throc* 'beam bridge' or 'drain' + *mere* 'pool' + *tūn* 'farm'. The Throckmorton family was powerful and influential in Tudor England, but the surname seems to have died out in Britain in the 20th century. Only six bearers are recorded in the 1881 census, and Richard Webber records no bearers at all in Britain in the 21st century. The variant *Throgmorton* has fared slightly better in Britain: Webber records 14 modern bearers in the UK. However, in the USA, the situation regarding this name is very different. According to figures released by the U.S. census bureau, 3238 bearers of the name Throckmorton are recorded in the 2010 census. Two 17th-century migrants from England are recorded: John Throckmorton (1601–1684) in New England and Robert Throckmorton (1609–63) in Virginia. In the 18th century the surname became established in Texas (among other places). James Webb Throckmorton (1825–94) was Governor of Texas from 1866 to 1867.

- *Griswold* has only 11 bearers in UK the 1881 census. However, there are 10,417 bearers in the US 2010 census. *Griswolds Farm* in Snitterfield (Warwicks), with its apparently possessive –*s*, sounds as if it might be one of the many farm names in England derived from the name of a family, but in fact the final –*s* in the farm's

modern name is intrusive. The place is recorded as *Groswold* in 1280 and *Grusewold* in 1321. The place-name derives from Old English *weald* 'woodland' with a first element that is disputed: probably Old English *grēosn* 'gravel'. It isvery unlikely to be Norman French *gris* 'grey', since such a mixture of French and English sources in the etymology of English place-names is not found elsewhere. DAFN's story: Edward Griswold (1607–91) and his family were Puritans who came to the American colonies from Wootton Wawen, Warwickshire, England, on the <i>Mary and John</i>, arriving on 30 May 1630. They settled first in Dorchester, MA, and in 1639 moved to Windsor, VT. Their relative, Matthew Griswold, came to New England in 1639, settling eventually in Lyme, CT.

5. Computer tools; corpus tools; database

Ambitious surnames research projects such as FaNBI and DAFN are dependent on computerization. Without the computer, they could not exist. Digitization of raw material and use of large databases of historical records, editing tools, statistical analyses, and validation are all essential components of the armoury of the modern surnames researchere. An account of the tools used by FaNBI is given separately in Rambousek, Parkin, and Horak (2018). There is no need to repeat it here.

6. Conclusion

In this paper, I have attempted to summarize severa very different kinds of research procedures iintom famliy, each of which is determined by the nature of the available data. *FaNBI* is an analysis of the copious evidence that is now available, in the form of digitized records of early bearers and their lcoations for the origins, history, and geographical distribution of family names in Britain and Ireland. As a result, it was able to provide an account of most of the the 'established' (traditional) family names in Britain and Ireland, together with supporting evidence in the form of selected records of early bearers, building on foundations laid fifty years earlier by P. H. Reaney. FaNBI has somewhat lighter coverage of the family names of recent immigrants, often stating a geographical and linguistic source (country and language of origin) and in many cases providing an etymology, but little else. It contains, for example, no supporting evidence of early bearers – neither in the source country nor of immigrants. Such tasks are for the future. The research time for the first edition of *FaNBI* was limited to four years, enough to establish the efficacy of the new research model and to make significant corrections to many previous explanations, but also revealing the size of the task if all traditional and established names were to be given equivalent research time. Two and a half further years were devoted to preparation of a second edition, extending the coverage of family names in Britain and Ireland from 45,000 entries to 60,000 entries, but with no significant changes to principles and policy.

DAFN has a much broader canvas than *FaNBI*, supported by much less satisfactory evidence. Its main task has been to identify the language and locality of origin for every name that met its criteria for entry. This itself is a new direction, which would not have been possible until much before the start of the new millennium. *DAFN*, in an American context, summarizes the best available information in traditional reference works in many languages, adding a unique contribution of its own. It can serve as a foundation for a possible future dictionary of surnames of the world, serving the 'global village' that is today's world.

Both approaches (*FaNBI* and *DAFN*) differ from traditional approaches, which were typically starved of data and tended to be reliant on guesswork as a consequence. The main difference has been the advent of computer technology, which has had two principal effects. First, it has enabled the creation of very large databases – databases containing millions (and eventually, billions) of transcribed historical records, allowing genealogists and onomasticians to consult vast numbers of historical records very rapidly and as a result to make increasingly reliable generalizations. The second main effect of computer technology has been lexicographic: compiling a major new dictionary of surnames, which previously would have taken many decades, can now be completed in a few years.

To explain the origins and history of family names in any community, competences in several different disciplines are required. Foremost among them is comparative and historical linguistics, which entails a deep knowledge of the history of the languages in which the surnames being studied originated and in particular the dialect differences affecting those languages. To this must be allied competence in other disciplines, including history, demography, geography, computer science, and statistics.

History is, of course, not just a single unified academic discipline. Many different kinds of history are needed if family names are to be adequately and accurately explained. These include not only political and social history but also local history (including a clear distinction between the origins of place names and the origins of family names). Equally important are family history and genealogy. At the present time, family history and genealogy are pre-eminently fashionable as topics of public interest in the English-speaking world, but the interaction of social and political history with family history can be regarded as equally important. Family names such as *Cecil, Cavendish,* and *Bentinck* are intimately interwoven with the political history of England, and Norman names such as *Agnew, Charters, Charteris*, and *Sinclair* with that of Scotland. Correct understanding of family names such as *Ackerman, Bond, Burgess, Butler, Reeve, Batchelor, Knight, Halfknight,* and *Squire* requires an understanding of medieval social structure.

Another aspect of history that is of the greatest importance to the student of family names is migration – both from country to country and (especially in the more distant past) within a country. Having said that, we should never allow ourselves to forget that modern boundaries and frontiers in most cases are of comparatively recent origin. For example, during the critical period of surname formation large parts of eastern France were under the rule of the English crown while during the same period, the usual language of the English aristocracy and royalty was Norman French. Territorial divisions between national groups in continental Europe were equally confused, with many localities being inhabited by diverse groups of speakers of Romance, Germanic and/or Slavic languages.

The student of family names does not have to be an expert in paleography (the interpretation of historic handwriting), but he or she must at least be aware of the problems that may have arisen in the course of transcribing and digitizing old documents.

Finally, the 21st-century anthroponomastician (the student of family names) needs either to have expertise using a computer as a tool for processing very large quantities of data, or to be well supported by an expert with these skills. This implies being able to use the sort of tools for Above all, he or she must not always expect results that offer explanations that interpret data with certainty. Part of the skill of the modern anthroponomastician is to know how to

interpret probabilities and when to call for statistics and more statistics in order to investigate a surname. Also, he or she should be able to decide when to give up on an investigation and move on. All too often I have witnessed well-intentioned scholars 'flogging a dead horse', investigating more and more data, in the hope that certainty lies 'just around the corner'. Usually, it doesn't.

Of course, it is too much to expect that all this great variety of expertise could be united in the person of a single polymath. Modern family-names research is best carried out by a team of people with different skills, integrated by mutual respect.

& & &

Surnames research has been a 'poor relation' of other kinds of research in the humanities (including place-name research) for far too long. It is not clear why this is. Perhaps scholars have been deterred by the complexity of the subject outlined in the preceding paragraphs, by the instability and variability of the data, or by an obsessional desire for certainty, to which the study of surnames does not lend itself.

Now, however, there does not seem to be any good reason why the study of surnames should not take its place alongside more hard-nosed empirical sciences that assign a clear theoretical role to the notion of probability. The origin and history of family names cnaWith the digitization of data and the development of techniques for analysing 'big data', we can hope to do better than our predecessors. But their concern for scholarly standards must not be disregarded. To take just one example, examination of early records was once regarded as the be-all and end-all of onomastic scholarship. We can now see that collection and analysis of early records is indeed a necessary **condition** for good surname research. But it is not sufficient. It must be teamed up with statistical and geographical analysis, as this paper has attempted to show.

References

Archer, Steven (2003, updated 2015): *The British 19th Century Surname Atlas*. CD-ROM. Deptford, Kent: Archer Software

Bahlow, Hans (1967; translated into English by Edda Gentry 1993): *Dictionary of German Names*. Madison: Max Kade Institute, University of Wisconsin-Madison

Bardsley, Charles Wareing (1901): *A Dictionary of English and Welsh Surnames, with Special American Instances*. London and New York: H. Frowde

Black, George F. (1946): *The Surnames of Scotland*. New York Public Library. Reissued 1993 by Birlinn Limited, Edinburgh

Brechenmacher, J. K. (1960): *Etymologisches Wörterbuch der deutschen Familiennamen*, Görlitz: C. A. Starke Verlag

Caffarelli, Enzo, and Carla Marcato (2008): *I cognomi d'Italia: dizionario storico ed etimologico*. 2 volumes, Torino: UTET

Debrabandere, Frans (1993): *Verklarend Woordenboek van de Familienamen in Belgie en Noord-Frankrijk.* Brussels: Gemeentekrediet

de Felice, Emidio (1997): *Dizionario dei cognomi italiani,* Milan: Mondadori

Dräger, Kathrin, and Mirjam Schmuck (2009): 'The German Surname Atlas Project: Computer-based surname geography'. In Wolfgang Ahrens, Sheila Embleton, and André Lapierre (eds): *Names in multi-lingual, multi-cultural and multi-ethnic contact.* Proceedings of the 23rd International Congress of Onomastic Sciences

Fenwick, Carolyn (ed., 1998-2005): *The Poll Taxes of 1377, 1379, and 1381.* 3 volumes. Oxford: Oxford University Press for the British Academy.

Faure Sabater, Roberto, María Asunción Ribes, and Antonio García (2001): *Diccionario de apellidos españoles.* Madrid: Espasa Calpe

Guppy, H. B. (1890): *The Homes of Family Names in Great Britain.* London: Harrison and Sons

Hanks, Patrick, and Flavia Hodges (1988): *A Dictionary of Surnames.* Oxford: Oxford University Press

Hanks, Patrick and D. Kenneth Tucker (2000): 'A diagnostic database of American personal names', in *Names* 48 (1)

Hanks, Patrick (2003): 'Americanization of European Family Names in the 17th and 18th Centuries'. In *Onoma, the International Journal of Onomastic Sciences* 38

Hanks, Patrick, Kate Hardcastle, and others (2003): *Dictionary of American Family Names.* New York: Oxford University Press

Hanks, Patrick (2006): 'Personal Names'. In K. Brown (ed.): *Encyclopedia of Language and Linguistics, Second Edition*, volume 9, pages 299-311. Amsterdam: Elsevier

Hanks, Patrick, Flavia Hodges, and Kate Hardcastle (2006): *The Oxford Dictionary of First Names,* fourth edition. Oxford: Oxford University Press

Hanks, Patrick (2009): 'Dictionaries of Personal Names'. In A. P. Cowie (ed.): *The Oxford History of English Lexicography.* Oxford: Oxford University Press

Hanks, Patrick (2016): 'German Family Names in North America' In Dieter Kremer (ed.), *Fremde Namen.* Leipzig: Leipziger Universitätsverlag

Hanks, Patrick, Richard Coates, and Peter McClure (2012): 'Methods for studying the origins and history of family names in Britain: philology meets statistics in a multicultural context'. In L.-G. Larsson and S. Nyström (eds.), *Facts and Findings on Personal Names: Some European Examples.* Royal Society of Arts and Sciences of Uppsala.

Hanks, Patrick, Richard Coates, Peter McClure, and others (2016): *The Oxford Dictionary of Family Names in Britain and Ireland.* Oxford: Oxford University Press

Hanks, Patrick, and Harry Parkin (2016): 'Family Names'. In C. Hough (ed.) *Oxford Handbook of Names and Naming.* Oxford: Oxford University Press

Kohlheim, Rosa, and Volker Kohlheim (2005): *Lexikon der Familiennamen.* Berlin: Duden Verlag

Machado, José Pedro (1984): *Dicionário Onomástico Etimológico da Língua Portuguesa*, 3 vols. Lisboa: Confluéncia. Lisboa: Livros Horizonte

Maclysaght, Edward (1957; 5th edition 1985): *The Surnames of Ireland.* Dublin: Irish Academic Press

McKinley, Richard (1977). *The Surnames of Oxfordshire.* London: Leopard's Head Press

McKinley, Richard (1981). *The Surnames of Lancashire.* London: Leopard's Head Press

McKinley, Richard (1988). *The Surnames of Sussex*. Oxford: Leopard's Head Press

Morgan, T. J., and Prys Morgan (1985): *Welsh Surnames*. Cardiff: University of Wales Press

Morlet, Marie-Thérèse (1991): *Dictionnaire étymologique des noms de famille*. Paris: Perrin

Rambousek, Adam, Harry Parkin, and Ales Horak (2018): Software Tools for Big Data Resources in Family Names Dictionaries' in *Names: A Journal of Onomastics*, On line at https://www.tandfonline.com/doi/abs/10.1080/00277738.2018.1453276

Reaney, P. H. (1958): *A Dictionary of British Surnames*, London: Routledge and Kegan Paul. Third edition (1997) revised and expanded by R. M. Wilson and published under the title *A Dictionary of English Surnames.*

Redmonds, George (2015): *Dictionary of Yorkshire Surnames*. Stamford: Paul Watkins Publishing

Sturges, C. M., and B. C. Haggett (1987): *Inheritance of English surnames.* London: Hawgood Computing

Woulfe, Patrick (1923): *Sloinnte Gaedheal is Gall*: *Irish Names and Surnames*. Dublin: M. H. Gill and Son