Surname Typology and the Problem of Inconsistent Classification

David Harry Parkin

University of the West of England, Bristol, UK

Abstract

This paper analyzes methodological inconsistency in surname classification, and the implications this has for the comparability of different works. Many studies have organized surnames by type, based on each name's "meaning", in order to identify national trends and regional differences in surnaming patterns. However, the ambiguity of "meaning" and the lack of any standard classificatory practice mean that such studies are incomparable. By reviewing P. H. Reaney's and R. A. McKinley's classifications of surnames from the same sources, and identifying discrepancies in their calculations and methods, a case for a standard method of surname classification is made. Only when there is a greater level of consistency in the classification of surnames can the findings of separate studies be reliably compared, allowing for meaningful conclusions on surnaming patterns to be drawn.

KEYWORDS Anthroponomastics, typology, methodology, surname type, classification

In order to discover national and regional trends in surname distribution, and general differences between regions, the classification of surnames is a useful approach. This method can say much about the proportions of types of by-name or surname at a particular time, and from this information a comparison of regional by-naming and surname trends can be drawn.

Using this approach, McKinley noted that ‘thirteenth- and fourteenth-century sources show that there were then marked differences between the English regions in the proportions of surnames and by-names falling into each of the main categories’ (1990: 20). Studies that have analyzed names by type, such as McKinley’s (1990) and the English Surname Series (county-based volumes, most of which dedicate individual chapters to each name type, as well as other analysis. See McKinley, 1975, 77, 81, and 88; Postles, 1998), have contributed significantly to our understanding of regional differences in by-name and surname patterns.

There are, however, a number of issues with the classification of names. Generally, the reliability of this method is likely to decrease as later records are used, with marriage and migration potentially masking or creating false patterns of surname distribution. Many records that could be used for this type of analysis do not contain all social classes, which is a problem considering that there were ‘sharp differences between one class and another in the nature of the names in use’ (McKinley, 1990: 201). There are also many records that are damaged or where some names have become illegible in some way, making it difficult to be fully confident in the reliability of name type proportions calculated from such records. These issues must all be considered when comparing the proportions of name types between regions, and their significance appreciated when interpreting any differences. However, ‘despite all these drawbacks, the method remains the best available for showing the main differences between counties or regions where surnames are concerned’ (McKinley, 1990: 21).

All of the issues mentioned above are certainly problematic for any comparison of name type proportion, but they only need to be considered once names have been accurately

categorized within a given typology. This is no simple task. Most surname scholars recognize four main classes of surname: those derived from a location, those derived from a relationship, those derived from an occupation, and those derived from a nickname, but the boundaries between these classes are not always clear. To give an example, how should the surname *Bridge* be classified? Without any sort of context it is not possible to know whether the name, in each individual and original instance, referred to someone who lived at or near a bridge, or who worked at a bridge, perhaps toll-taking. There are also names with multiple etymological origins, making it impossible to assign them a single type. The surname *Hill*, for example, may be locational, from a person who was in some way linked to that topographical feature, evidenced by forms such as 'Johannes *atte Hyll'* (1379 Wa PT)'. However, the surname may also fall into the relationship category, having its origin in a personal name as seen in the case of 'Rogerus *filius Hil* (1221 D Cur)' (Reaney, 1997: 231).

These two difficulties can cause the classification of certain surnames to be based on each scholar's own interpretation of a name, which is highly unlikely to be identical for all researchers. Some may use the etymology of the name for its classification, where others might consider the possible motivation behind its original application. To clarify this point, consider the medieval by-name, *Sheep*. Etymologically, this name refers to the animal, and nothing more can be said of it. Motivationally, it would be reasonable to suppose that the name was applied metonymically to someone who had a sheep-related occupation, perhaps a shepherd or wool-dealer, or to a person known for their timidity. With the etymological approach, the name is apparently a nickname (if using the more usual four categories mentioned above), but the motivational approach might cause the name to be categorized as having multiple possible origins in that it may have been used to refer either to occupation or

behavior. Further barriers to comparability could arise if some researchers are unaware of alternative etymologies for certain names that others know of, and some may disagree as to the most likely etymology or motivation behind a name, depending on their typological system. It is worth noting here that a decision has been made not to refer to names in terms of "meaning". "Meaning", as stated by Lyons, is a 'pre-theoretical, intuitive term', able to be split into 'a variety of theoretical terms [...] to refer to various aspects of meaning' (1977: 28). This ambiguity is sure to have caused confusion and disagreement in name classification, in that what a name might "mean", or have "meant", can be interpreted in a number of ways.

There is room for extensive speculation on how and why names might be differently classified, but one major problem that could be overcome is the lack of any standard practice for surname classification. Currently, there is no consistency in surname classification method, so no meaningful conclusions can be made by comparing two or more separate studies that organize surnames by type. If such studies were compared, it could never be clear whether apparent differences in surname type proportion were as a result of regionally specific surnaming patterns, or the classificatory choices of the researcher. Even though it has been widely recognized that 'the classification of a name is often arbitrary' (Redmonds, King and Hey, 2011: 58), no one has attempted to establish a standard practice for surname classification within the typology previously described.

The current method appears to rely on the idea that classification of names is only possible 'after their origin and meaning have been satisfactorily established' (Redmonds, 1997: 14), but there are a number of problems with this. It risks discarding a large proportion of ambiguous names from any analysis and so misrepresenting their distribution and relative

frequency. There are also issues with the establishment of the 'origin and meaning' of a name. By-name "meaning" is ambiguous and often arbitrary, and can be different depending on whether etymological or motivational origin is considered. In light of this, there is certainly a case to be made for a standard method of classification. Let us start by exemplifying and evaluating the kinds of inconsistencies that need to be rectified with a comparison of the methods of Reaney (1967) and McKinley (1990).

The most easily noticed difference between Reaney's (1967) and McKinley's (1990) method is their slightly different categories of classification. Reaney uses the more usual system with four categories as mentioned above, where McKinley (1990: 22) uses a system with six categories as follows:

1. Locative names
2. Topographical names
3. Surnames and by-names from personal names
4. Occupation names
5. Surnames and by-names from nickname
6. Names in other categories, or of uncertain origin (1990: 22)

While Reaney's and McKinley's works are different, this does not mean that their systems of classification cannot be compared. McKinley's locative and topographical names are simply subdivisions of Reaney's location names, so McKinley's values need only be added together in order to compare them with Reaney's. Excepting the sixth, McKinley's other categories are essentially the same as Reaney's, and the sixth can be disregarded as Reaney does not include those names of uncertain origin in his analysis. Those names that McKinley calls 'surnames of relationship' will therefore be omitted, but they are very few, not including any derivatives of given-names such as those ending *-son* and *-kin* (which fall into his third

category), but comprising only those names ‘such as Cousin, Brothers, Fadder, or Ayer’ (McKinley, 1990: 11). These types of name are included in Reaney’s ‘surnames of relationship’ category, and so there will be a small discrepancy in their findings here, but given the rarity of such names the effect of this difference will be negligible in a comparison of their findings.

Perhaps the most serious issue is that there seems to be no consideration of those names with multiple etymological origins or multiple motivational origins, unless these are included in McKinley’s sixth category and Reaney chose to omit them from his analysis, though neither state these explicitly. Whether or not this is the case, an assessment of their classificatory methods can be achieved by comparing only those categories which seem to be the same. If Reaney’s and McKinley’s criteria for assigning each name a particular type are the same, then the values for these categories should be almost identical. This is, however, not the case, as shown by Table 1, a table that presents the findings from Reaney’s (1967: 22) and McKinley’s (1990: 23) own analyses of the same sources.

<Table 1>

According to these figures, Reaney and McKinley only agree completely in four instances out of a possible thirty-six. Even where their percentages are identical, it is not possible to be sure that they came up with exactly the same number of names for each name type, as it is not clear whether they omitted the same number of names from their datasets due to etymological uncertainty, if any at all in McKinley’s case. The ‘total numbers of persons’ column is taken from Reaney’s analysis, and so McKinley has not necessarily analyzed the same number of names. Considering this, any comparisons made can only ever be

approximate, but in most cases the considerable differences between their percentages cannot be as a result of such minor inconsistencies, especially given the large number of names in each record.

Table 2 shows the differences, in real numbers, between their classifications. It is important to reiterate that these differences are not necessarily exact, due to the possible methodological inconsistencies in the use of the data, but they are large enough to indicate that Reaney and McKinley classified a significant number of names in different ways.

<Table 2>

Even in those cases where Reaney and McKinley differ by only 0.5% (see Nickname and Relationship in Sussex), it is clear that this still represents a large number of differently classified names. The biggest difference of 1,289 in relationship names from the 1327 Suffolk subsidy roll is quite alarming; such a considerable discrepancy is unexpected given that Reaney and McKinley have used exactly the same data.

Under each type, McKinley appears to, more often than not, have a higher number of names. This can be seen in Figure 1, being four bar charts for each name type. This appears to be due to McKinley's tendency to classify a greater number of names with more certainty, where Reaney leaves a greater number out of his analysis.

<Figure 1>

Including McKinley's sixth category, his percentages equal near enough to 100%, where Reaney's range from 68% to 85%. There even appear to be some fairly simple mistakes in

their work, most noticeably in McKinley's percentage values for the 1332 Warwickshire Subsidy Roll totaling 102%.

Whatever reasons there are for these differences in their findings, it is clear that Reaney's and McKinley's works on surname type proportion cannot be compared without a great deal of care and unproductive investigation. They appear to approach the task with irreconcilable methodologies, causing their results to be, for the most part, very different. This has been confirmed by carrying out a chi-squared test of independence, to determine whether there is any association between researcher and classification of name. The results are presented in Table 3, which shows that for each county record investigated, a null hypothesis that "there is no significant association between researcher and name classification" can be rejected, as all chi-squared values are above the critical value of 16.268, at a probability level of 0.001. In other words, the probability of these results being down to chance, rather than there being any significant association between researcher and name classification, is less than 0.1%.

<Table 3>

So, it is apparent that there is a significant relationship between name classification and researcher, or, to put it differently, name categorization is dependent on the researcher. From this, while no conclusions can be made on how their methods of name type categorization differ, it can be deduced that Reaney's and McKinley's methods are significantly different. In order to put forward a proposal for a more reliable surname classification method, it is first necessary to identify where there is any possible confusion in the current method of classification and why this confusion arises. This will, again, be discussed by comparison of

the works of Reaney (1967) and McKinley (1990) and by what they state in their works about their classificatory systems.

It is clear that Reaney and McKinley sometimes classify the same name differently. This is not just deduced from a comparison of their tables of type proportion, but from a comparison of their written explanations of surname types. To give an example mentioned above, the surname *Bridge* can be interpreted in different ways. Reaney recognizes that it is not possible to know the motivation behind each separate and original occurrence of the name, stating that '*Bridge* is local when it means "dweller by the bridge", but occupation if it refers to the keeper of the bridge and the collector of tolls there' (1967: 19). McKinley, however, treats the name as locational only, classifying it as a topographical surname, being 'surnames from terms for features of the landscape, whether natural ... or man-made' (1990: 10).

The different ways in which Reaney and McKinley explain their choice of surname type for the name *Bridge* provide a clue as to one major difference in their methods of surname classification. Reaney explains the name in terms of its application, suggesting why a person might have been known by that name. McKinley treats the name differently, as linguistic rather than onomastic item, referring to it as a being from a particular feature. To put it another way, Reaney takes the motivation behind the name into account, where McKinley takes the etymology; both have their advantages. Reaney looks into why such a name would have been given and so gets closer to its actual original use. McKinley does not speculate on the possible motivation behind the name, and so, in the case of *Bridge* at least, does not suffer from a lack of context in deciding on a surname type.

This method of McKinley's can be recognized in a number of his typological explanations. In a summary of occupational names, McKinley states that 'names from high positions have also been included, such as King, Earl, Bishop, Cannon, Archdeacon, Prior, Abbot, Sheriff, Baron, or Knight, since it is often not possible to be sure how they originated, though many seem to have begun as nicknames' (1990: 10). He later discusses these types of names further, suggesting that it is 'impossible to suppose that such names were actually the descendants of kings, bishops, etc.' and that 'there seems to be no doubt that such surnames, though apparently occupational ones, were in fact nicknames in origin' (1990: 135–6). Despite this recognition, that titles such as King and Bishop would have been used as nicknames, McKinley chooses to classify them as occupational based on their etymology.

However, this linguistic, rather than onomastic, appreciation often results in other possible etymological origins of a name being missed. McKinley has a tendency to recognize only those origins of a name that are most obvious to the modern reader, *King* being a case in point, choosing to classify the name as occupational. He fails to recognize that the name may also denote relationship, with the OE word *cyng* giving rise to a personal name, *Cyng*, as seen in Mariota filia *King* (1259 RamsCt). Reaney recognizes both possible origins of the surname, explaining how 'surnames of office such as *Abbot*, *Bishop* and *King* are often nicknames whilst the last two may also be patronymics' (1967: 20).

Further disagreement between Reaney's and McKinley's methods of name categorization is apparent in their treatment of the surname *White*, which McKinley classifies only as a nickname (see 1990: 11), where Reaney provides multiple possible origins, being from the

> OE name of *Hwīta*, a nickname from OE *hwīt* "white", or to one nicknamed "the white" from his fair hair or complexion (le white), or to one who lived by the bend or curve of a river or road (atte wyte) as at Great Whyte (Hu), or to a man from White (D), atte Wayte "a look-out post". (1967: 17)

There is clear speculation on Reaney's part, which McKinley may choose to avoid by apparently relying on the most obvious etymology of the name to the modern reader. Yet, if the possible original application or motivation of the name is the criterion for surname classification, such speculation is unavoidable.

For the sake of relative simplicity and to ensure that a large number of names are not "lost" in a "multiple possibilities" category, McKinley's method is preferable. It may seem counterintuitive to disregard the motivation behind the original bestowal of name when classifying it, given that each by-name had a particular contextual significance, yet it does allow for more certainty in classification. To restate the case of the name *Bridge*, Reaney's reliance on the possible motivation behind the name gives it either an occupational or locational origin, where McKinley's appreciation of etymology classifies the name as locational only. Reaney's method requires a greater level of interpretation and speculation, based on the unknowable context of the bestowal of each name, which is likely to cause uncertainty when classifying names.

McKinley's method does, however, need refining. While the etymological approach to classification is relatively clear for most simplex names, the reliance on etymology is not quite so simple for compounds. The name *Bridgeman*, for example, does not refer simply to a topographical feature, but neither does it refer to a particular occupation or official position. All that we can be sure of is that the name denotes a man that had some sort of connection

with a bridge. It is clear that, whether using the etymology of a name or the motivation behind its bestowal, classification is not always easy. In order to have certainty in the classification of ambiguous names, it is necessary to follow a clearly defined set of rules, the absence of which has led to the kinds of discrepancies seen in Reaney's and McKinley's classifications. In some cases it is not necessary, or always practical, to follow such rules, with surname type often being obvious. For example, locational surnames that derive from toponyms cannot be easily misinterpreted. Many surnames can also have multiple separate etymological origins that are masked by their modern forms, requiring linguistic investigation before a set of classificatory rules can be usefully followed. A new method can, however, ensure those names that are difficult to define, such as *Bridgeman*, are not also placed in this category.

A possible method of categorization is presented in flowchart form in Figure 2, however, there is still a certain amount of analysis required for each name before following this system. The etymological origin(s) of the name must be established first, with particular attention paid to individual morphemes of compound names, comparing variant medieval forms where necessary in order to ensure the philological plausibility of a possible etymology. It is unlikely that such a method as proposed will completely eliminate the possibility that two different researchers will classify a name differently, as they may disagree on etymological origin. However, in cases where etymology is agreed upon, it will ensure that such a name is categorized in the same way.

<Figure 2>

The proposed system in Figure 2 has been preliminarily tested with a sample of 100 names, taken from Reaney and Wilson's dictionary (1997). The sample was collected by using a random number generating formula in Microsoft Excel. With the function "=randbetween(1,509)", one-hundred random numbers between 1 and 509 were generated, corresponding to page numbers in the dictionary; 1 and 509 being the first and last page numbers of the dictionary respectively. The number of entries on each of these one-hundred pages was then counted, and the function "=randbetween(1,y)", where "y= the number of entries on that page", was used to generate a number corresponding to an entry on that page. This name was then used for analysis. If it was a variant, its corresponding head-form was used. In order to clarify the proposed method, a number of these one-hundred names have been selected for discussion.

The etymology of the surname *Milk* is clear, yet how to categorize it, based on previous systems, is not necessarily so. Reaney suggests that it is 'perhaps a nickname for one whose drink was milk, effeminate, spiritless', or 'for one with milk-white hair', or 'metonymic for a seller of milk' (1997: 309). None seem implausible, but these multiple interpretations risk making any categorization, based on Reaney's dictionary entry, over-complicated and confusing. The newly proposed method is more certain. Following the steps of the flowchart, the name is simplex and so we can go straight to section 2. Milk is not an occupation, official position, or rank. It is not a given name, or a word referring to relationship. It is not a toponym, topographical feature, or man-made structure. It is, however, a word in the Middle English Dictionary (MED) (Kurath, Kuhn and Lewis: 1952-2001), so this name is classified as a nickname. The criteria for classifying a name as a nickname may appear to be based on a process of elimination, yet any by-name or surname with a clear

etymological origin that does not fit into the categories of occupation, relationship or location, and that cannot then be justifiably described as a nickname, has not been found in the test of this method.

The name *Sacker* has a clear etymological origin, but is discussed here to show the morphological analysis involved in the proposed method of categorization. The final morpheme, "er*",* is bound, so the final lexeme, or entire lexeme in this case, *Sacker*, is taken for analysis in section 2 of the flowchart. The word, as defined in the MED, refers to 'a maker of sacks or sackcloth', and so the name is categorized as occupational. If the final morpheme is free then, provided it has a clear etymological origin, the name should be categorized based on that morpheme. For example, the name *Allanson* has "son" as its final morpheme. This is a word that refers to relationship and so, following the flowchart, the name *Allanson* is placed in the "relationship" category.

Finally, the name *Rowland* is an example of how the proposed method requires some names to be categorized as having multiple etymological origins. The name could originate from a given name, specifically ' OFr *Rollant, Rolant, Rolent, Roulent,* OG *Hrodland*, *Rodland',* but could also be from one of the toponyms 'Rowland (Derbys) or Rowland Wood in Slinfold (Sussex)' (Reaney, 1997). Both of these origins must be fed into the flowchart individually, and then, following their separate categorizations as "relationship" and "location" respectively, a collective name type of "multiple possibilities" can be assigned.

In the case of the name *Bridgeman*, and similar names, a loop has been incorporated into the flowchart to avoid any ambiguity associated with the classification of the word *man*. Following the flowchart from the starting point, the name *Bridgeman* is not a simplex name,

it is not a toponym, it does not contain a toponym, and it does not begin with a preposition that relates to position. It is not a given name, or a hypocoristic form of a given name, and it does not have a diminutive suffix. This leads to the box in the flowchart with dashed edges, leading back to the start of the classification process, but this time disregarding the *-man* ending, effectively feeding the name *Bridge* into the chart. This is a simplex name, and so following the processes in section 2 the name is placed in the "Location" category. In any surname, the final morpheme *-man* is etymologically ambiguous, and almost acts as a bound morpheme, often only making sense when combined with the preceding morpheme. It is for this reason that this step has been worked into the chart.

The proposed method of classification, then, might take the name away from its original application, instead utilizing a system that takes account of etymology, in some cases of individual morphemes. The sorting of names into categories is not carried out with the goal of uncovering the motivation behind the bestowal of each name, but it is a comparative tool, allowing general trends in naming to be recognized. So long as the method of classification is standardized, there can be a greater degree of confidence in the calculation of name type proportions, allowing such works to be directly compared. This may lead to some names being placed in a category that some researchers do not agree with in terms of surnaming motivation, but this is an inevitable consequence of introducing a classificatory standard.

It must be stressed that in no way is this method supposed to be a way of uncovering the origin of a name or why it was first bestowed or used, as we can rarely be certain about such things. After all, the surname type is an analytical construct which requires accuracy for the purposes of statistical comparison. As such, the uncertainty of arbitrary classification,

based on the motivation behind each name, has no place in this kind of research. This method is meant as a way of improving consistency in surname classification, so that there can be a greater level of confidence in the comparability of regional name studies, and confidence in any conclusions drawn from their comparison. To ensure that the calculation of surname type proportion is appropriate as a comparative tool, consistency in classification is essential. At the moment, the lack of any standard in surname classification renders such work invalid, shown by Reaney's and McKinley's very different results using the same data. Scientifically speaking, this would normally cause their results to be considered unreliable.

It is hoped that this discussion has established the necessity for a standard method of by-name and surname classification, allowing for future work in typology to be comparable. The proposed method is intended as a starting point for improving the accuracy of classification, with further extensive testing and revision required. It is hoped that, whether or not this method is accepted in any form, a standard classificatory system can be established and followed by all to allow for meaningful conclusions to be drawn from surname type comparison. However, any shift in the system of classification will be gradual. Only when it is sufficiently communicated, agreed upon, and widely adopted can a new system be used, otherwise there will be no value in such a change. Nevertheless, in order to ensure the reliability, validity, and comparability of research into by-name and surname types, such a new system is entirely necessary.

Acknowledgements: The author wishes to thank Richard Coates, Patrick Hanks, and Peter McClure for their comments and advice on an earlier form of this paper.

Bibliography

Kurath, H., S. M. Kuhn, and R. Lewis. 1952–2001. *The Middle English Dictionary.* Ann Arbor: University of Michigan Press.

Lyons, John. 1977. *Semantics: Volume 1.* Cambridge: Cambridge University Press

McKinley, Richard. 1975. *Norfolk and Suffolk Surnames in the Middle Ages.* London: Phillimore.

_____. 1977. *The Surnames of Oxfordshire.* London: Leopard's Head Press.

_____. 1981. *The Surnames of Lancashire.* London: Leopard's Head Press.

_____. 1988. *The Surnames of Sussex.* Oxford: Leopard's Head Press.

_____. 1990. *A History of British Surnames.* London: Longman.

Postles, David. 1998. *The Surnames of Leicestershire and Rutland.* Oxford: Leopard's Head Press.

Reaney, P. H. 1967. *The Origin of English Surnames.* London: Routledge.

Reaney, P. H. and R. M. Wilson. 1997. *A Dictionary of English Surnames.* rev. 3rd edn. Oxford: Oxford University Press.

Redmonds, George. 1997. *Surnames and Genealogy: A New Approach.* Boston, MA: New England Historic Genealogical Society.

Redmonds, George, Turi King, and David Hey. 2011. *Surnames, DNA, and Family History.* Oxford: Oxford University Press.

Abbreviations

D	- Devon
D Cur	- Devon Curia Regis Rolls
Derbys	- Derbyshire
Hu	- Huntingdonshire
Lancs	- Lancashire
MED	- Middle English Dictionary (see Bibliography)
OE	- Old English
OFr	- Old French
OG	- Old German
RamsCt	- W. O. Ault, *Court Rolls of the Abbey of Ramsey and of the Honor of Clare,* Yale 1928
Salop	- Shropshire
Wa PT	- Warwickshire Poll Tax
Warwicks	- Warwickshire
Worcs	- Worcestershire
Yorks	- Yorkshire

Notes on contributor

David Harry Parkin is a PhD student at the University of the West of England, Bristol, attached to the Family Names of the United Kingdom (FaNUK) research project. His thesis is focussed on surname development in the Cotswolds. Correspondence to: David Harry Parkin, david5.parkin@live.uwe.ac.uk

Table 1

Comparison of Reaney's (1967) and McKinley's (1990) calculations of name type proportion from the same sources

Subsidy Rolls	Date	Total number of persons	Local %		Relationship %		Occupational %		Nicknames %	
			Reaney	McKinley	Reaney	McKinley	Reaney	McKinley	Reaney	McKinley
Sussex	1332	6973	40	49	14.5	15	11	18	12.5	12
Surrey	1332	5471	42	45	14	14	15	24	10	9
Kent	1334	11016	33	42	20	22	10	19	9	11
Suffolk	1327	11720	26	26	19	30	15	23	14	5
Worcs	1327	4644	34	30	24	31	14	14	13	12
Warwicks	1332	5457	33	41	23	21	15	23	10	8
Salop	1327	4897	31	49	11	23	17	5	9	16
Yorks	1327	3848	43	32	5	13	18	24	8	8
Lancs	1332	2571	49	67	1	3	11	15	8	6

Table 2

Comparison of Reaney's (1967) and McKinley's (1990) calculations of name types from the same sources, in real numbers

Subsidy Rolls	Total number of persons	Local		Relationship		Occupational		Nicknames	
		Reaney	McKinley	Reaney	McKinley	Reaney	McKinley	Reaney	McKinley
Sussex	6,973	2789	3417	1011	1046	767	1255	872	837
		628 difference		35 difference		488 difference		35 difference	
Surrey	5,471	2298	2462	766	766	821	1313	547	492
		164 difference		no difference		492 difference		55 difference	
Kent	11,016	3635	4627	2203	2424	1102	2093	991	1212
		992 difference		221 difference		991 difference		221 difference	
Suffolk	11,720	3047	3047	2227	3516	1758	2696	1641	586
		no difference		1289 difference		938 difference		1055 difference	
Worcs	4,644	1579	1393	1115	1440	650	650	604	557
		186 difference		325 difference		no difference		47 difference	
Warwicks	5,457	1801	2237	1255	1146	819	1255	546	437
		436 difference		109 difference		436 difference		109 difference	
Salop	4,897	1518	2400	539	1126	832	245	441	784
		882 difference		587 difference		587 difference		343 difference	
Yorks	3,848	1655	1231	192	500	693	924	308	308
		424 difference		308 difference		231 difference		no difference	
Lancs	2,571	1260	1723	26	77	283	386	206	154
		463 difference		51 difference		103 difference		52 difference	

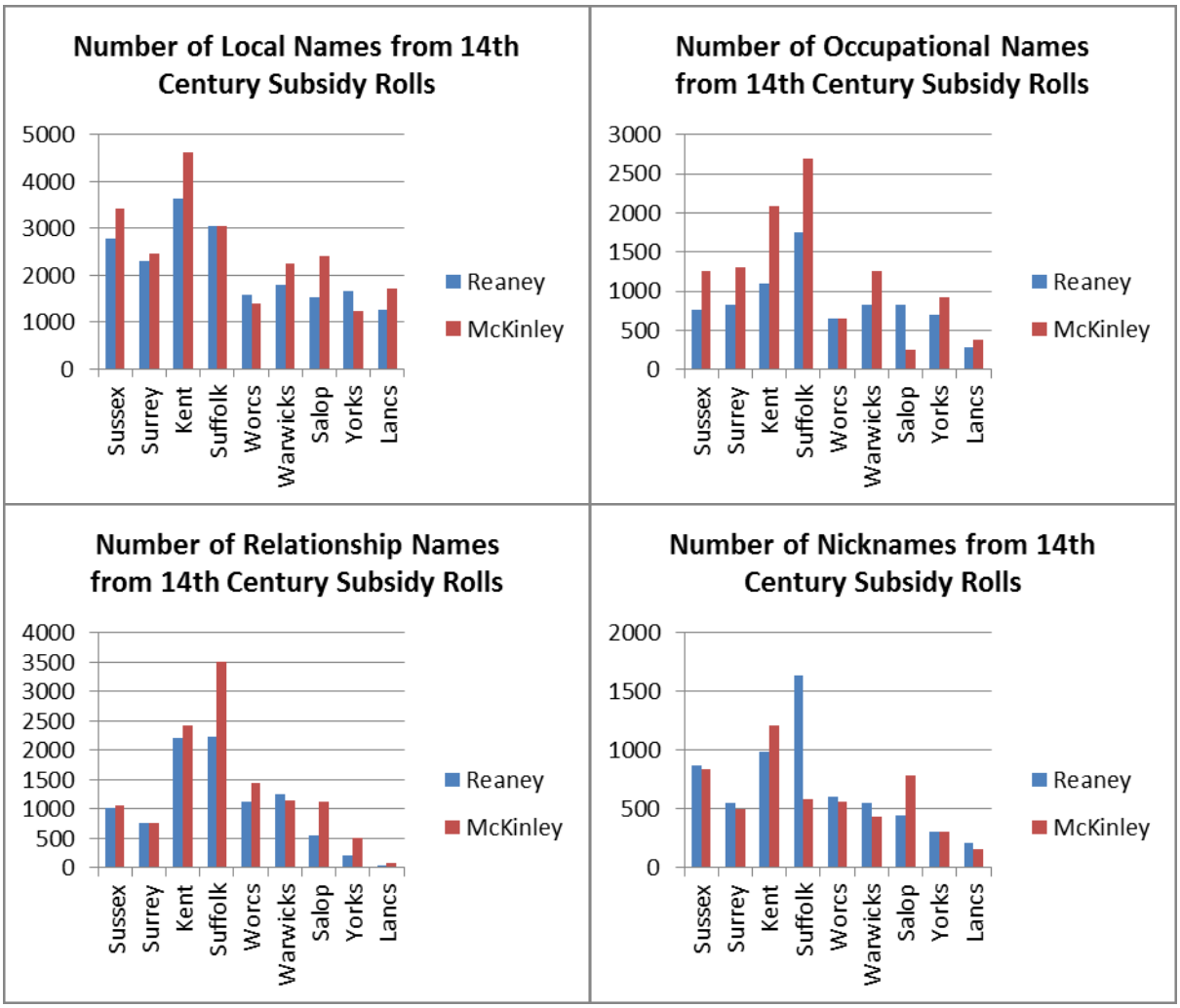

FIGURE 1 Bar charts comparing classification of surnames by type.

Table 3

Chi-squared values for test of independence, comparing surname classification and researcher

	Sussex	Surrey	Kent	Suffolk	Worcs	Warwicks	Salop	Yorks	Lancs
Chi-squared	78.6	84.2	140	916	53.4	111	647	230	43.7

* Critical value for 0.001 alpha, given 3 degrees of freedom = 16.268

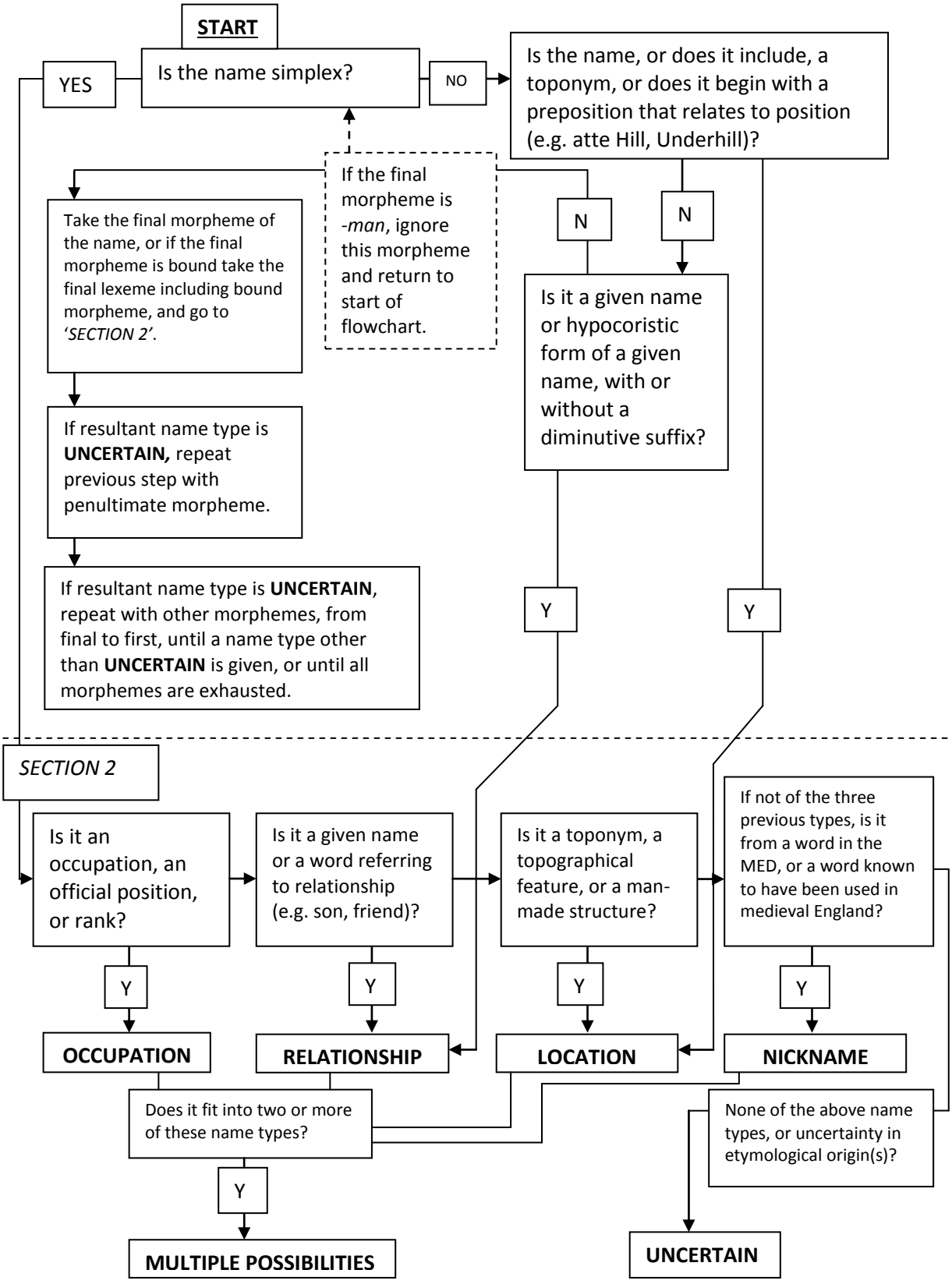

FIGURE 2 Name classification flowchart.