
FORENSIC DICTIONARY ANALYSIS: PRINCIPLES AND PRACTICE

Julie Coleman: *University of Leicester, UK (jmc21@leicester.ac.uk) and*
Sarah Ogilvie: *University of Oxford, UK (sarah.ogilvie@trinity.ox.ac.uk)*¹

Abstract

Lexicographers often provide an account of their working practices and policies, and reviewers and researchers generally take this on trust. Forensic dictionary analysis uses evidence-based methodologies to interrogate the dictionaries themselves about decision-making processes involved in their compilation. The version of events that this reveals is sometimes quite different from compilers' accounts. This paper builds on a variety of approaches in historical dictionary research—statistical, textual, contextual, and qualitative—to present forensic dictionary analysis as a technique that allows researchers to examine and understand the complex relationships between editorial policy and lexicographic practice.

1. Introduction

In this paper we seek to outline a methodology, which we call forensic dictionary analysis, by which the dictionary researcher can examine, understand, and reconstruct lexicographic policies and practices—policies and practices that sometimes differ from the accounts given by the lexicographers themselves. Forensic dictionary analysis provides the dictionary researcher with a methodology to combine statistical, textual, contextual, and qualitative analyses, to achieve a more complete picture of the making of the dictionary.

Much research on historical dictionaries to date has focussed on quotational evidence. Studies using the *Oxford English Dictionary (OED)*, and a few using the *Middle English Dictionary (MED)*; Kurath et al. 1952–2001), have examined the number and style of quotations, types of source, rate of citation of individual authors and works, and typographical representation (e.g. McConchie 1997, Brewer 2006, 2007a, Carter Hailey 2007). Other studies have used the quotations in OED or MED as a discrete corpus from which to comment on lexical and grammatical changes in English over time

(e.g. Dekeyser 1986, Jucker 1994, Coleman 1995, Fischer 1997, Nevalainen 1999, Bauer 2001, Mair 2001, Markus 2001, Hoffman 2004, to cite only a small selection). These studies have not always addressed the problems involved in using dictionary material as objective evidence of usage, as identified by Schäfer (1980) in his masterly study of Nashe and Shakespeare in the *OED*.

Purely qualitative dictionary studies are rare: dictionaries lend themselves to quantitative analysis. For instance, although their accounts are largely descriptive, Starnes and Noyes (1946) and Landau (1984: 35–72) both include estimates of numbers of words defined and proportions of entries derived from earlier dictionaries. Despite the long-established use of numerical description in dictionary studies, Lynch and McDermott (2005: 4) found that a dictionary as well-researched as Johnson's 'has led a kind of double life', in which selective treatment of its contents has sustained the belief that the dictionary is simultaneously prescriptive and quirky.

2. Forensic Dictionary Analysis

Like the judge in a criminal court, the dictionary researcher is faced with evidence of various types and varying reliability. Prefatory and publicity material is equivalent to witness statements: first-hand evidence is always valuable, but it is inevitably subjective and there may be motivations for distorting the truth. Archival material, where it exists and is accessible, allows the dictionary researcher to cross-examine published statements, but it will probably be both partial and incomplete. Few dictionaries are consistent in the application of lexicographic policies, but this need not be presented as a flaw: good lexicographers learn from experience, remain flexible in their practice, and adapt their policies to the needs of each entry.² Circumstantial evidence may also be useful—every dictionary is a product of its material, cultural and social circumstances, and it can only be properly understood as a result of its historical context and textual tradition—but interpretations based on circumstantial evidence are inevitably subjective and partial. The most reliable evidence is contained within the body of the dictionary itself. Statistical analysis of a dictionary's contents is equivalent to the work of the crime scene investigator: lexicographers leave traces of themselves in the dictionaries they produce, and dictionary researchers can thus reconstruct their working practice for comparison with public or private statements of principle. Forensic dictionary analysis brings together these different types of evidence to explore the complex relationships between principles and practice in dictionary production.

This paper explores various methodologies used by dictionary researchers, concentrating particularly on statistical methodologies, and identifies examples of good and bad practice. Numerical results can be unduly convincing: statistical analysis is only useful if it is rigorous, and the rigour and success of

quantitative analysis is dependent on a well thought-out strategy with regard to the parameters of the study and the sampling techniques adopted. This paper outlines the decisions dictionary researchers must make and the considerations they must take into account before they embark on a statistical analysis of a dictionary's contents. Although statistical analysis can identify inconsistencies, it cannot explain them: weighing up the various types of evidence available is what enables the forensic dictionary researcher to pronounce a reliable verdict.

2.1 Determining the parameters

The first step in analysing a dictionary statistically is to reduce its contents to countable features and to determine the parameters of the case study. Countable features with regard to the words listed include headword, pronunciation, orthography, variant forms, etymology, region, register, semantic field, word class, and age (see Table 1 for further explanation). The lexicographic treatment of these words can be analysed with reference to, for example, labelling, style, and the provision of semantically related terms, compounds, and phrases.

Parameters of analysis are dependent on the focus of the case study. To take a familiar example, an analysis of the *OED* that looked at the proportion of neologisms created by Shakespeare as opposed to those created by Milton would limit its focus solely to quotational evidence, in particular to the three parameters of author, date, and number of quotations.³ A case study focussing on a wider issue, such as the treatment of loanwords in the *OED*, would require parameters of analysis that were wider than quotational evidence alone. As summarized in Table 1, lexicographic treatment (such as labelling), as well as general presentation and treatment of features of each headword (such as pronunciation, variant forms, or etymology) are all relevant to the analysis. A comparison of these countable features would allow the researcher to comment not only on proportions of loanwords with respect to the whole lexicon, but also on more sophisticated issues of borrowing in English, such as whether there was a bias towards loans from a particular region, language, or semantic field in a particular century, or whether such patterns in the dictionary were merely a consequence of inconsistent lexicographic practice.

2.2 Selecting a sample

Once the parameters of analysis are determined, the researcher needs to decide on a case study sample. Sampling an entire dictionary is rarely an option unless electronic searching provides a reliable reflection of the dictionary's total contents. Some early or specialized dictionaries contain only a few thousand entries and are thus amenable to mechanical analysis in their entirety, but in

Table 1: Parameters of analysis for a case study focussing on the treatment of loanwords in a diachronic dictionary (Ogilvie 2008a: 34)

Features of the Word	
Pronunciation	<i>Is the pronunciation naturalized?</i>
Orthography	<i>Is the word's spelling naturalized? Does it include diacritics?</i>
Variant forms	<i>Is there more than one spelling variant?</i>
Etymology	<i>What is the word's provenance?</i>
Region	<i>From which region of the world does the word come?</i>
Semantic field	<i>To which semantic field does the word belong?</i>
Word Class	<i>Is the word a noun, verb, or adjective?</i>
Plurality	<i>Was the word borrowed with the plural marking of its donor language?</i>
Grammatical Gender	<i>Does the word demonstrate variants with the grammatical gender of its donor language?</i>
Age	<i>When did the word enter English?</i>
Quotational Evidence	
Number of quotations	<i>How many citations are in the quotation paragraph?</i>
Type of source	<i>Are all the citation sources of the same type (i.e. all travellers' tales, newspapers, encyclopedias)?</i>
Italics	<i>Does the word appear in italics in the citations?</i>
Inverted commas	<i>Does the word appear in inverted commas in the citations?</i>
Gloss in brackets	<i>Is the word followed by a gloss in brackets in the citations?</i>
Lexicographic Treatment	
Symbols	<i>Is the word marked with tramlines, asterisks, daggers, etc.?</i>
Labelling	<i>Is there a regional or subject label?</i>
Square-bracketed quotations	<i>Are any quotations square-bracketed?</i>

most cases the researcher will have to choose a sample for analysis. Any sample must ensure good coverage of the alphabet and (for multi-editor works) a balanced representation of lexicographic work by all editors over all time periods. In choosing a sample, the researcher needs to be aware of possible 'alphabet fatigue', a phenomenon by which lexicographers work with greater thoroughness at the beginning of the alphabet than the end (Starnes and Noyes 1946: 185, Osselton 2007). Conversely, some lexicographers treat individual entries with more thoroughness towards the end of the alphabet

(de Schryver 2005). In Johnson studies, dictionary researchers often sample the letter 'L', presumably on the assumption that the middle of the alphabet is more representative than either end (Miyoshi 2007: 31). Whether a study is intended to explore changing methodology or to generalize about the dictionary's contents, it is clearly necessary to sample from the whole alphabetical range. Gotti (1999: 61–67), for instance, sampled the letter A in B. E.'s *New Dictionary of the Terms Ancient and Modern of the Canting Crew* (c.1698), and used that as the basis for his description of the dictionary's contents. Using a more representative sample, Coleman (2004a: 76–126) identified several statistically significant differences between the dictionary's treatment of the first and second half of the alphabet, including the rate of provision of etymologies, proverbial sayings, usage labels, and cross-references.⁴ In this case, the dictionary's compiler was clearly learning from experience.

Allowing for changing methodology by sampling each letter of the alphabet is only the first step in identifying an appropriate sample. Depending on the focus of a study, selecting the sample range of two comparative dictionaries would be problematic unless the same sections of the alphabet were compared. Terms beginning <al->, for example, are disproportionately of Arabic origin (reflecting the definite article prefix Arabic *al-*) and if the sample from one dictionary is selected from this range, while the sample for the comparator dictionary comes from elsewhere in the letter A, an apparently significant preference for terms from Arabic will be a result of the sampling technique rather than of differences in dictionary content. To avoid introducing such biases, or in order to account for the significance of such patterns, the researcher might select the sample from one dictionary randomly and then match that sample from the comparator text.

An example of sampling bias is provided by Jespersen's ground-breaking analysis of French loanwords (1905). Using sections of the first edition of the *OED* (*OEDI*; Murray et al. 1888–1928) that had been published by that date, he took the first hundred French words under A–G and the first fifty under I and J. However, the results were skewed by the decision to exclude words listed with fewer than five quotations. This applied to many nineteenth-century loans because *OED* editors rarely gave more than four quotations per century, and thus Jespersen's results provided a misleading overview both of the coverage of French loans in the *OED* and of the continued influence of French upon English.

For multi-editor dictionaries, it may be more appropriate to sample the work of each editor rather than each letter of the alphabet. For example, in *OEDI* two portions of the letter S were edited by Henry Bradley (S–SH and ST–STY), one by William Craigie (SI–SQU), and one by Charles Onions (SU–SZM). With dictionary projects that were published gradually, it is also possible

Table 2: Various sampling techniques applied to Hotten (1859)

	entire dictionary		first 1000 entries of entire dictionary		first 10% of entire dictionary		first 50 entries for each letter		first 10% of each letter	
citations	75	3%	40	4%	10	5%	29	3%	7	3%
named authorities	134	6%	72	7%	9	4%	54	6%	18	8%
unattributed examples of use	289	13%	107	11%	25	11%	125	14%	30	14%
total	2204	100%	1000	100%	220	100%	909	100%	221	100%

to select early and later examples of each editor's work, so their development as lexicographers can be traced.

Methodological trends are not always unidirectional, and changes in a lexicographer's practice over time do not necessarily reflect changes in their individual style. For example, the final volume of Farmer and Henley's *Slang and its Analogues* (1890–1904), which was published after Henley's death, reverts in several respects to the policies of the first volume, compiled by Farmer alone. This suggests that Farmer had continued to work in the same way, but that his drafts had been edited and improved by Henley (Coleman 2007).

In order to explore the reliability of different sampling methods, various samples were extracted from Hotten's (1859) *A Dictionary of Modern Slang, Cant, and Vulgar Words*. The dictionary was also analysed in its entirety, as shown in Table 2.

The main list in this dictionary contains 2204 sense divisions for 2030 headwords in total,⁵ and each pair of columns shows the number and proportion of entries containing attributed citations, authorities named but not quoted, and unattributed examples of use. χ^2 tests show no statistically significant differences between these samples. In other words, for this dictionary all of these sampling techniques support conclusions that are in line with results for the whole dictionary.

However, there appear to be differences between the beginning and end of the dictionary. Comparison between the various samples in Table 2 suggests that Hotten included more citations and fewer unattributed examples of use at the beginning of the dictionary than the end, but the first two sampling techniques do not allow us to explore this. Table 3 shows how alphabetically distributed samples can be used to interrogate changes in lexicographic practice through time.

When results for the entire dictionary are subjected to the χ^2 test, the apparent differences between the provision of unattributed examples of use in

Table 3: Alphabetically distributed samples from Hotten (1859)

	entire dictionary				first 50 entries for each letter				first 10% of entries for each letter			
	A–M		N–Z		A–M		N–Z		A–M		N–Z	
citations	47	4%	28	3%	18	3%	11	3%	4	3%	3	3%
named authorities	83	7%	51	5%	32	6%	22	6%	12	10%	6	6%
unattributed examples of use	131	11%	158	16%	61	12%	64	17%	11	9%	19	19%
total	1209	100%	995	100%	530	100%	379	100%	121	100%	100	100%

the first and second halves of the alphabet prove to be highly significant, though the decrease in citations is not. The two sampling techniques both reflect this highly significant increase in unattributed examples, demonstrating that both approaches offer a sound basis for analyzing changes in methodology in this dictionary.

Although comparing two data points has clarity to recommend it, this is likely to be at the expense of detail. Having identified that there is a difference between the first and second half of the alphabets, the focus can be adjusted to identify when the increase in unattributed examples took place.

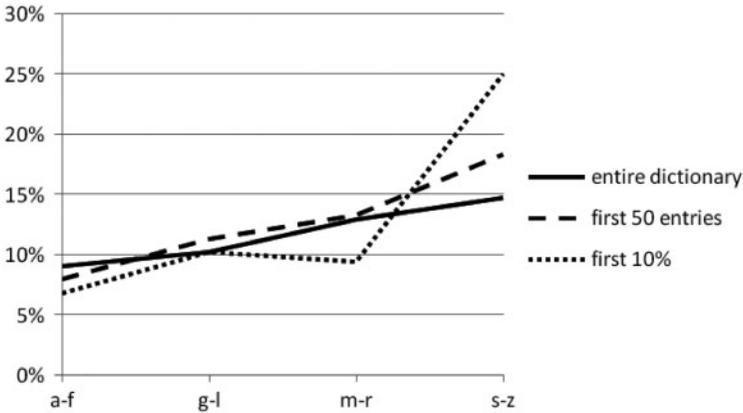


Figure 1: Percentage of entries including unattributed examples in three samples from Hotten (1859) broken into four alphabetical sequences (see Table 4).

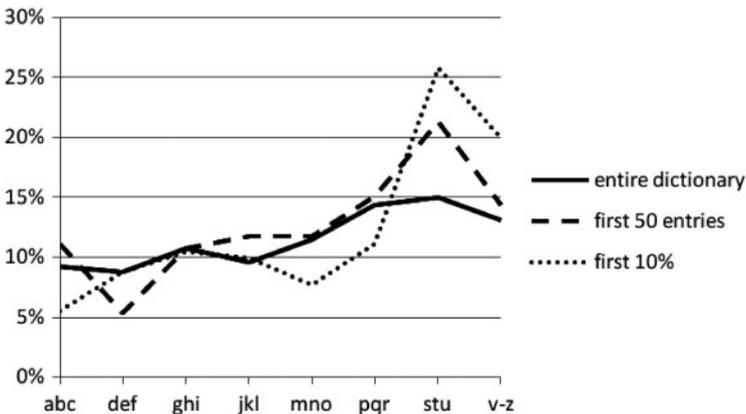


Figure 2: Percentage of entries including unattributed examples in three samples from Hotten (1859) broken into eight alphabetical sequences (see Table 5).

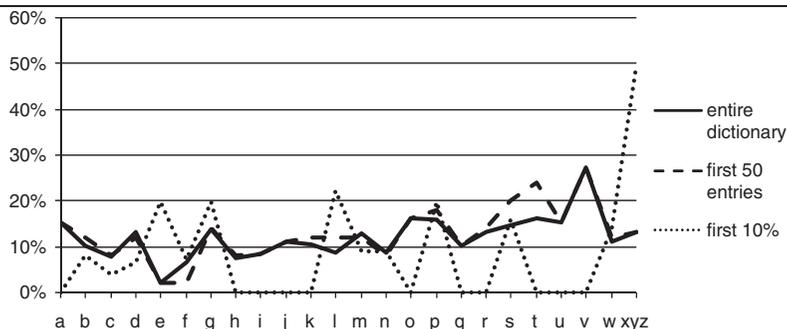


Figure 3: Percentage of entries including unattributed examples in three samples from Hotten (1859) broken into twenty-four alphabetical sequences (see Table 6).

Table 4: Entries including unattributed examples of use (grey cells) and total entries (white cells) for four alphabetical sequences of Hotten (1859)

	a–f	g–l	m–r	s–z
entire dictionary	79	40	67	108
	871	392	517	734
first 50 entries	22	29	34	36
	276	257	257	197
first 10%	6	4	5	9
	88	39	53	72

Table 5: Entries including unattributed examples of use (grey cells) and total entries (white cells) for eight alphabetical sequences of Hotten (1859)

	abc	def	ghi	jkl	mno	pqr	stu	v–z
entire dictionary	49	30	21	19	29	38	94	14
	530	341	195	197	252	265	627	107
first 50 entries	14	8	12	17	16	18	24	12
	126	150	112	145	137	120	113	84
first 10%	3	3	2	2	2	3	7	2
	54	34	19	20	26	27	62	10

Figures 1–3, based on the raw figures presented in Table 4–6, all confirm the increasing trend in the provision of unattributed examples of use, but only Figure 3 can demonstrate that this reaches a peak in the letter V and falls off towards the end of the alphabet. These results also demonstrate the importance of matching sample size to purpose: as the samples are chopped into ever

Table 6: Entries including unattributed examples of use (grey cells) and total entries (white cells) for twenty-four alphabetical sequences of Hotten (1859)

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	xyz
entire dictionary	4	25	20	20	1	9	14	6	1	5	6	8	14	9	6	23	2	13	64	28	2	3	8	3
	26	249	255	154	50	137	103	80	12	45	58	94	110	105	37	146	20	99	441	173	13	11	73	23
first 50 entries	4	6	4	6	1	1	7	4	1	5	6	6	6	4	6	9	2	7	10	12	2	3	6	3
	26	50	50	50	50	50	50	50	12	45	50	50	50	50	37	50	20	50	50	50	13	11	50	23
first 10%	0	2	1	1	1	1	2	0	0	0	0	2	1	1	0	3	0	0	7	0	0	0	1	1
	3	25	26	15	5	14	10	8	1	5	6	9	11	11	4	15	2	10	44	17	1	1	7	2

smaller pieces their reliability decreases. In a bigger dictionary 10% of entries under each letter would be a more reliable sample than the first 50 entries, but for a small dictionary a 10% sample gives unreliable results.

Presentation of results by percentages can also be misleading even when based on a complete sample. The 'entire dictionary' line in figure 3 appears to reveal considerable fluctuation in the provision of unattributed examples of use in Hotten's dictionary, but only the increase between the letters E and F and the decrease between V and W are statistically significant (both $p=0.01$). However, neither the differences between D and G nor those between U and XYZ are statistically significant. Analysing this dictionary letter by letter thus identifies inconsistencies but not trends: for a small sample, which is unavoidable when working on a small dictionary, grouping letters together can produce more meaningful results than treating them separately.

2.3 Choosing a dictionary text

Any meaningful comparative analysis of dictionaries must compare like with like, i.e. the dictionaries must be comparable in content and type (synchronic, diachronic, register, specialism, etc.) or linked by textual dependency. The choice between print and electronic versions, if available, must be made with reference to the advantages and disadvantages of each.

Searching electronically has the benefits of saving time and allowing the analysis of an entire text. However, the researcher has to be sensitive to shortcomings or inconsistencies inherent in a text which may not have been written with electronic searching in mind. The combination of inconsistent lexicographic practice and an unsophisticated search engine may result in distorted results (see footnote 3). It is therefore vital to assess all possible permutations and variations of a search term. For example, a researcher wanting to examine *OED Online* in order to determine the total number of words that have entered English from French,⁶ would first need to assess the consistency of language names, spellings, and abbreviations within the etymology field. Although the third edition of *OED* (*OED3*; Simpson 2000–) is consistent in its representation of French etymons, unedited entries from the second edition (*OED2*; Simpson and Weiner 1989) are not. *OED Online* combines material from *OED2* and *OED3*, and therefore uses the full form 'French' in etymologies (16,436 times), as well as 'Fr.' (4827), and 'F.'⁷ The results of an electronic search for French etymologies in *OED Online* might be interpreted as showing certain linguistic patterns of borrowing, or as proving that one editor favoured (or neglected) words of French provenance, whereas the differences could actually be the result of an individual preference for an abbreviation that the researcher failed to anticipate.

An added complication is that the current *OED Online* search engine cannot discriminate between words that occur as language names (eg. < French) and

those that occur in other contexts within the etymology field. For example, a search for the word 'French' in the etymology field would include the entry for *Ringelmann* among its results. This is not from French, but bears the etymology 'Of uncertain attribution; perh. the name of Maximilien Ringelmann (1861–1931), French scientist'. Similar issues exist for inconsistencies in labelling and in bibliographic representation of author names, titles, and editions. *OED Online* is working to iron out these inconsistencies, but researchers must be aware that the periodic release of updated portions produces a dictionary that is dynamic and mutable, and painstakingly compiled results will inevitably be brought into question by the next quarterly instalment.

Data collection from a print version, although time-consuming, does have certain advantages. Researchers using electronic resources will perform only those queries that occur to them and, even with careful vetting of the results, will examine only a small proportion of the data available. Print-readers have serendipity on their side, inconsistencies can be accounted for, and hunches can be explored to determine whether there is any statistical significance in casually observed trends. A print version also allows the researcher to assess markers and symbols (such as daggers, asterisks, or tramlines) that indicate style or usage and may not be searchable electronically (Osselton 2006).

2.4 Quantitative analysis

Once the contents of a dictionary have been reduced to countable features and the researcher has determined which sample and parameters of analysis are needed for the specific case study, it is possible to compare and analyze them statistically. In dictionary comparisons it is particularly important for the researcher to calibrate all calculations in order to account for differences in size and content of the respective texts. For example, if a single volume dictionary is compared with a multi-volume work, the researcher must not automatically assume that there is a direct proportional relationship between their relative sizes. For example, the second *OED Supplement* (Burchfield 1972–1986) is four volumes and the first (Craigie and Onions 1933) a single volume, but it would be wrong to assume that the former is four times the size of the latter. In fact, Burchfield's supplement has an average of 14 entries per page, compared with 28 entries per page in the 1933 supplement, amounting to a ratio not of 4:1, but 2:1. The relative sizes of each letter of the alphabet should also be factored into calculations in order to give proportional figures that account for dictionary fatigue and the fact that some letters, such as C and S, account for a disproportionate number of English words.

Databases offer convenient tools for the organization and quantitative analysis of data. Each row in the database table can represent a dictionary sense division, and each column a lexical or lexicographic feature (see 2.1).

For example, Table 7 represents a short range of entries from Partridge's *Dictionary of Slang and Unconventional English* (1937). The complete sample consists of 2086 sense divisions for 1715 headwords from the main alphabetical listing of approximately 55,800 sense divisions for 41,200 headwords:

Having constructed this table, the researcher can use the database's query language to interrogate the evidence and correlate variables. These queries are quickly constructed and instantaneously executed, making it possible to explore correlations completely beyond the scope of traditional analysis. For example, a comparison between 'name' and 'date1', sorted on 'name', exposes Partridge's dating techniques: one of his sources is Fenton's *Bush Life in Tasmania Fifty Years Ago*, published in 1891, and words supported by reference to Fenton are generally dated to the mid-nineteenth century. Partridge's practice in this respect is clearly deductive rather than evidence-based, and the format of these dates implies that they are the result of careful consideration of a wider range of documentary evidence than was actually available.

Tables constructed from individual dictionaries can be modified to allow comparison between related dictionaries. For example, the binary yes/no column 'F&H' in Table 7 indicates terms that are also in Farmer and Henley's *Slang and its Analogues*, which Partridge had been commissioned to update. The addition of this column allows the composition of the resulting dictionary to be explored with reference to its actual rather than claimed use of its sources:

it may be assumed for the period up to 1904, and where *no* author or dictionary is quoted, the debt is, in most instances, to Farmer and Henley (Partridge 1937: x)

Partridge sometimes obscures his use of *Slang and its Analogues* by re-using its citations. For instance, in the sample entries in Table 7, the references to Swift and Fielding (in the entry for *tace is Latin for a candle*) and Foote (*tack together*) are all derived from Farmer and Henley, though that dictionary is not named as a source in either case. Comparison between the two dictionaries also throws light upon Partridge's 'ineligible' words, which he lists in square brackets to justify their exclusion. Most prove to be terms that had been listed in Farmer and Henley, indicating that although Partridge was willing to disagree with his main source, he felt the need to defend these editorial decisions individually and in some detail.

Table 7 can also be modified to explore the use made of Partridge (1937) by later dictionaries. For example, the '1961' column shows which entries survived into the abridged *Smaller Slang Dictionary*. The principles underlying the abridgement are explained (Partridge 1961: vii): terms obsolete before 1900 were omitted, as were all indelicate terms. These claims are borne out by the

Table 7: Example database entries for Partridge (1937)

F&H	1961	headword	no	label	field1	field2	date1	cent1	cit	name
y	n	tace is Latin for a candle	0	coll > dial	dis	alert	1688		b	Shadwell; Grose; SwiftFielding; Apperson
y	n	tach	0	back s	clothes		1859		a	Hotten
y	n	tachs	0	Tonbridge school orig	fashion		1880		n	
y	y	tack	1	nautical > coll > SE	food		1830	20	a	Marryat; OED
y	y	tack	2	coll	food			19	b	Lyell
y	n	tack	3	Sherborne School	food	educ	1870		n	
n	n	tack	4	sol	alert			19	n	
n	n	tack or sheet, will not start	0	nautical coll	emotion			19	a	Bowen
y	n	tack together	0	jocular coll, ob	marriage		1754		a	Foote
n	n	tack wallah	0	military	alcohol			19	a	B&P
n	n	tack, on the	0	military	alcohol			19	a	F& Gibbons
n	n	tack-on	0	coll	process		1905		a	OED

Notes to Table 7:

date1 = earliest year date in the entry

cent1 = earliest century mentioned in the entry

cit = 'a' when an authority is named, but not quoted; 'y' when an example is given, but no authority cited; 'b' when a fully referenced citation is provided; 'n' when there is no example, authority, or citation

The F&H and 1961 columns are explained in section 2.4.1

evidence, but terms labelled as solecisms (such as *tack* 4) were also silently omitted regardless of their currency or offensiveness, as were ‘ineligible’ and non-lexical entries. The non-lexical entries largely comment on pronunciation and grammar, but a few are encyclopaedic in nature. Drawing attention to these omissions might have forced Partridge to address their claim to inclusion in the unabridged work (see Coleman forthcoming, ch.1).

It is unusual for lexicographers to select from earlier dictionaries entirely randomly. For example, Coleman (2004c) demonstrated that, using B. E.’s *A New Dictionary of the Terms Ancient and Modern of the Canting Crew* (c.1698) as a source, later dictionaries (Smith 1719, the anonymous *New Canting Dictionary* of 1725, and Grose 1785) all display a statistically significant preference for entries labelled as ‘cant’. *The New Canting Dictionary* and Grose also have a statistically significant preference for terms belonging to the semantic fields of CRIME AND PUNISHMENT, SEX, and POVERTY, regardless of their usage label. In addition, entries not labelled as ‘cant’ but including citations were included in Grose’s dictionary at a significantly higher rate. Similarly, Hotten (1859) demonstrates a statistically significant preference for entries including cited authorities when he makes his selection from Egan’s (1823) edition of Grose’s dictionary (Coleman 2008: 18–19). Grose and Hotten both attempted to excuse the disreputable content of their dictionaries by emphasizing their scholarly credentials, but statistical analysis demonstrates that they both preferred their research pre-packaged.

2.5 Sharing data

Although statistical analysis involves the use of more rigorous techniques than impressionistic dictionary studies, there still remains an element of subjectivity. Even where researchers describe their methodology in careful detail, there is frequently an element of judgement involved in applying that methodology: Is ‘Gypsy’ a usage label or an etymology in this entry? Does ‘Scott’ necessarily refer to the author of *Ivanhoe*? Is ‘thieves’ pidgin’ the same as ‘cant’? Is this definition based on that one in an earlier dictionary? It would, therefore, be unwise to use data compiled by another researcher as a basis for comparison without careful re-analysis to explore their methodology. For example, in her edition of a glossary of the slang used by Australian troops during the First World War, Laugesen (online) indicates terms adopted from an earlier dictionary (Downing 1919). However, upon comparison of the two lists, it is clear that Laugesen has only marked entries that were borrowed verbatim, and Downing’s influence is considerably greater than this suggests (Coleman 2008: 253). Without close re-analysis of the data, scholars run the risk of basing their conclusions on unsuspected differences in methodology and analysis.

3. The Forensic Approach: Combining statistical analysis with contextual and qualitative research

Statistical analysis can highlight trends and patterns of lexicographic practice, but it cannot always explain the policies and motivations behind them. It is therefore essential, where possible, to supplement statistical analysis with general knowledge of the textual tradition and historical period to which the dictionary belongs, and to combine this with information gleaned from a combination of other sources. These will vary from dictionary to dictionary, but may include prefaces and introductions, publicity material, reviews and responses to reviews, editorial notes, dictionary proofs, slips, unpublished lectures, diaries, marginalia, or correspondence with publishers, readers, and consultants.

The dictionary-making process can involve many layers and stages of editing before the final product emerges. This process is sometimes preserved on dictionary slips or proofs that show additions and corrections by various contributors to multi-editor dictionaries. These not only reveal who constructed the entry, or worked on different parts of it, but they can also uncover discussions between editors explaining certain editorial decisions (Morton 1994, Reddick 1996, Gilliver 2004, Ogilvie 2004, Mugglestone 2005). Quotational slips can give insights into whether or not a particular text was deliberately requested by an editor. This can indicate an editor's feeling that a subject, author, literary genre, or source language was worthy of particular attention. Dictionary proofs showing hand-written marginalia, deletions, comments, and messages between editors can show dissent amongst editors or differences in editing styles, and can provide insights into last-minute decision-making by which we can gauge editorial priorities (Mugglestone 2003, 54).

Correspondence between editors and publishers can reveal stresses between scholarship and commerce which may affect editorial practice: usually pressure to speed up progress or to save space, but sometimes to exclude certain types of words. For example, Murray's draft of the first section of the *OED*, A–ANT, was sent back to him by the OUP Delegates with instructions to omit 'Aardvark, Aardwolf, Ab², Aba, [and] Abaca' for no apparent reason other than their loan status. Gilliver (2007) used archival documents to show the pressure that OUP exerted on William Craigie to speed up progress and to save space while he edited the letter U in *OEDI*. The competitive and territorial side of lexicography, involving disputes of recognition and authorship, is also evident in correspondence among the editors of the *Dictionary of American English* (Adams 1995, 20) and the *Shorter Oxford Dictionary* (Ogilvie 2008b).

Comparison between the final text and letters from dictionary consultants can also reveal influences on policy and practice. Whether or not editors act on advice provided by a consultant can give added insight into attitudes and policies. For example, in a letter to James Murray dated 1886, the

Oxford anthropologist E. B. Tylor questioned whether or not the ‘outlandish’ word *boyuna* ‘a large black Brazilian snake’ had any place in an English dictionary. The fact that Murray ignored his consultant’s advice indicated that his definition of ‘English’ was broader than Tylor’s (Ogilvie 2008a: 28).

Delving beyond the dictionary text into archival materials can also give new insights into the reasons for certain lexicographic practice. For example, in both *OEDI* and Burchfield’s *OED Supplement*, two small parallel lines or ‘tramlines’ were placed beside headwords that the editors considered ‘alien or not yet naturalized’. The absence of tramlines in the 1933 *OED Supplement* initially suggested a printer’s error, but Ogilvie (2008b) used an internal memo found in the OUP archive to suggest that it was a conscious policy decision influenced by the Society for Pure English of which the editors were members. Similarly, the fact that Charles Onions included five times more loanwords than William Craigie in the 1933 *OED Supplement* could have been coincidence. However, letters found in the OUP archive between the Press and the American journalist H. L. Mencken show that Onions had a disposition towards focussing on loanwords, while a letter from James Wyllie provides an insight into Craigie’s more restrictive policy.⁸

While it is useful to be attentive to archival documentation in order to shed more light on policies and practices, it is also important not to read too much into such discoveries. For example, there is a three-fold increase in Chinese words in Burchfield’s *OED Supplement* (1972–1986) after the letter N. Archival documents reveal that this coincided with a trip to China made by the editor in 1979. However, it also coincided with a natural bias in the Chinese sound system for words beginning with letters in the latter half of the English alphabet, published after 1979 (*O-Scz* published in 1982 and *Se-Z* published in 1986), so further analysis would be necessary to determine how profoundly the editor’s visit to China influenced his treatment of Chinese words.⁹

Clearly there are limits to contextual and qualitative dictionary research. The researcher must be alert to the possibility that the version of events presented by a lexicographer is not always a true reflection of actual practice. In-house documents may represent good intentions or attempts to bridge differences in practice rather than applied policy. A dictionary’s preface may represent policies applied consistently throughout its production, or it may represent policies that the editor wished had always been in place, or thought that reviewers or readers might expect to have been in place. Changes in political climate may also influence the published claims made for a dictionary. For example, Burchfield wrote that:

Readers will discover by constant use of the Supplement that the written English of regions like Australia, South Africa, and India have been accorded the kind of treatment that lexicographers of a former generation might have reserved for the English of Britain alone (Burchfield 1972: xv).

These claims have generally been accepted without question, but a close examination of the dictionary by Ogilvie (2008a) revealed them to be untrue. He had, in fact, deleted 17% of loanwords and terms from World Englishes included in the *1933 Supplement*, and added only half as many words from outside Britain as his predecessor, Onions.

Lexicographers' misrepresentation of their methodology and sources is sometimes clearly mendacious. For example, Grose (1785: vii) claimed in his preface that he had omitted obscene terms wherever possible, and that where their inclusion was unavoidable he had 'endeavoured to get rid of [indecent and immodest words] in the most decent manner possible'. However, terms pertaining to SEX in Grose's sources are considerably more likely to appear in his own dictionary than non-sexual vocabulary (Coleman 2004b: 19, 31–2). Asserting the superiority of one's own dictionary is commonplace, but Hotten's (1859: 153) description of Ducange Anglicus's *Vulgar Tongue* as 'silly and childish' was presumably intended to disguise his own extensive use of it (Coleman 2008: 18–19).

4. Conclusions

Forensic dictionary analysis brings together statistical, textual and contextual approaches that allow dictionary researchers to examine, understand, and reconstruct lexicographic practices and policies. This paper has presented various techniques for forensic lexicography that allow the researcher to interrogate dictionary texts themselves about the decision-making processes involved in their compilation. It has also explored some of the pitfalls of electronic searching and of various sampling techniques. Forensic dictionary analysis builds on a variety of previous research methods, suggesting that the best way to analyze and contextualize a dictionary is by means of a synthesis of qualitative and contextual research with rigorous statistical analysis.

Notes

¹ Sarah Ogilvie would like to thank the Arts and Humanities Research Council and the Wingate Foundation for funds to pursue this research, which was carried out whilst she was a Visiting Fellow at the Research School of Humanities and the Australian National Dictionary Centre, Australian National University.

² One lexicographer's description of the evolution of his own policy and practice can be found in Botha (2005).

³ A simple search of *OED Online* (accessed 24/09/2008) for each author's name as 'first cited author' would apparently prove that Shakespeare (1881 hits) created three times more neologisms than Milton (628) (based on a search for 'Shak*' in 'first cited author' to allow for OED's spelling variants, *Shakespeare and Shakes.*). But these results would have to consider the *OED's* unusually thorough treatment of Shakespeare (Jespersen 1905: §224, Schäfer 1980), and take into account the fact that the number of

neologisms drops by 40% if hapax legomena are factored out of the equation (Goodland 2008). The results would also require manual checking to ensure that the authors corresponded to William Shakespeare (1564–1616) and John Milton (1608–1674). For example, the results included mascarpone ‘Italian cheese’, which the *OED* attests was first used in English by Arthur Milton in *Rome in Seven Days*.

⁴ Researchers might be tempted to break samples down into smaller units, such as individual letters, but in a dictionary with a few thousand headwords only broader-stroke analyses will produce statistically significant results (see Section 2.2).

⁵ Hotten’s lists of rhyming slang, back slang, and costermongers’ terms for money are omitted from these calculations.

⁶ A study of this scale was proposed by Mossé (1943: 34) long before it was practicable.

⁷ *OED Online* (accessed 24/09/2008). A search for ‘F.’ in ‘etymologies’ produced hits for both the abbreviation *F.* ‘French’ and for the abbreviation *f.* ‘from’, totalling 117,292.

⁸ A fuller comparison of the differences between Onions’ and Craigie’s treatment of loanwords is found in Ogilvie (2008a). See also Brewer (2007b: 24–33) for a general account of the differences between their working practices.

⁹ See Ogilvie (2008a: 35) for a fuller discussion of Chinese words in the *OED* Supplement (Burchfield 1972–1986). We are indebted to Mike Clark of the Chinese-English Oxford Dictionary Project for verification that the longest letters in the Chinese dictionary are S, X, Y, and Z.

References

A. Dictionaries

- Anon. 1725.** *New Canting Dictionary*. London: The Booksellers of London and Westminster.
- B. E. c. 1698.** *A New Dictionary of the Terms Ancient and Modern of the Canting Crew*. London: W. Hawes.
- Burchfield, R. W. 1972, 1976, 1982, 1986.** *A Supplement to the Oxford English Dictionary* vols. I–IV Oxford: Clarendon Press.
- Craigie, W. A. and Onions, C. T. 1933.** *A New English Dictionary on Historical Principles. Founded on the Materials Collected by the Philological Society. Edited by James A. H. Murray, Henry Bradley, William A. Craigie, C. T. Onions. Introduction, Supplement, and Bibliography*. Oxford: Clarendon Press.
- Downing, W. H. 1919.** *Digger Dialects*. Melbourne: Lothian Book Publishing.
- Ducange Anglicus. 1857.** *The Vulgar Tongue*. London: Bernard Quaritch.
- Egan, P. 1823.** *Grose’s Classical Dictionary of the Vulgar Tongue*. London: Printed for the Editor.
- Farmer, J. S. and Henley, W. E. 1890–1904.** *Slang and its Analogues Past and Present*. London/Edinburgh: Subscribers only.
- Grose, F. 1785.** *A Classical Dictionary of the Vulgar Tongue*. London: S. Hooper.
- Hotten, J. C. 1859.** *A Dictionary of Modern Slang, Cant, and Vulgar Words*. London: John Camden Hotten.
- Kurath, H., Kuhn, S. A., Reidy, J., and Lewis, R. 1952–2001.** *Middle English Dictionary*. Ann Arbor: University of Michigan Press.
- Laugesen, A.** *Glossary of Slang and Peculiar Terms in Use in the A.I.F. 1921–1924* <http://www.anu.edu.au/ANDC/res/aus_words/wwi/index.php>.
- Murray J. A. H., Bradley, H., Craigie, W. A., and Onions, C. 1888–1928.** *A New English Dictionary on Historical Principles*. 10 vols. Oxford: Clarendon Press. [OEDI].

- Partridge, E. 1937.** *A Dictionary of Slang and Unconventional English*. London: Routledge.
- Partridge, E. 1961.** *Smaller Slang Dictionary*. London: Routledge.
- Simpson, J. 2000–.** *OED Online*. 3rd edn. Available to subscribers at www.oed.com. [OED3].
- Simpson, J. and Weiner, E. 1989.** *The Oxford English Dictionary* 2nd ed. Oxford: Clarendon Press. [OED2].
- Smith, A. 1719.** *The Thieves New Canting Dictionary*. London: Sam Briscoe.

B. Other literature

- Adams, M. 1995.** ‘Sanford Brown Meech at the *Middle English Dictionary*.’ *Dictionaries: Journal of the Dictionary Society of North America* 16: 151–185.
- Adams, M. 1998.** ‘Credit Where It’s Due: Authority and Recognition at the *Dictionary of American English*.’ *Dictionaries: Journal of the Dictionary Society of North America* 19: 1–20.
- Bauer, L. 2001.** *Morphological Productivity*. Cambridge: Cambridge University Press.
- Botha, W. 2005.** ‘Concurrent Over- and Under-treatment in Dictionaries. A Response.’ *International Journal of Lexicography*, 18: 77–87.
- Brewer, C. 2006.** ‘Eighteenth-Century Quotation Searches in the *Oxford English Dictionary*’ in R. W. McConchie, O. Timofeeva, H. Tissari, and T. Saily (eds.), *Selected Proceedings of the 2005 Symposium on New Approaches in English Historical Lexis (HEL-LEX)*, Somerville, MA: Cascadilla Proceedings Project, 41–50.
- Brewer, C. 2007a.** ‘Reporting Eighteenth-Century Vocabulary in the *OED*’ in J. Considine and G. Iamartino (eds.) (2007). *Words and Dictionaries from the British Isles in Historical Perspective*. Cambridge: Cambridge Scholars Publishing: 109–135.
- Brewer, C. 2007b.** *Treasure-House of the Language: The Living OED*. New Haven: Yale University Press.
- Carter Hailey, R. 2007.** ‘To “Finde Wordes Newe”: Chaucer, Lexical Growth, and *MED* First Citations’ in J. Considine and G. Iamartino (eds.) (2007). *Words and Dictionaries from the British Isles in Historical Perspective*. Cambridge: Cambridge Scholars Publishing: 14–24.
- Coleman, J. 1995.** ‘The Chronology of French and Latin Loan-Words in English.’ *Transactions of the Philological Society* 93: 95–124.
- Coleman, J. 2004a.** *A History of Cant and Slang Dictionaries* Vol. I: 1567–1784. Oxford: Oxford University Press.
- Coleman, J. 2004b.** *A History of Cant and Slang Dictionaries* Vol. II: 1785–1858. Oxford: Oxford University Press.
- Coleman, J. 2004c.** ‘Cant and Slang Dictionaries: A Statistical Approach’ in C. Kay, C. Hough, and I. Wotherspoon (eds.), *New Perspectives on English Historical Linguistics*. Amsterdam: John Benjamins: 41–47.
- Coleman, J. 2007.** ‘Expediency and Experience: John S. Farmer and William E. Henley’s *Slang and its Analogues*’ in J. Considine and G. Iamartino (eds.) (2007). *Words and Dictionaries from the British Isles in Historical Perspective*. Cambridge: Cambridge Scholars Publishing: 136–165.
- Coleman, J. 2008.** *A History of Cant and Slang Dictionaries* Vol. III: 1859–1936. Oxford: Oxford University Press.
- Coleman, J. forthcoming.** *A History of Cant and Slang Dictionaries* Vol. IV: 1937–1984. Oxford: Oxford University Press.

- Dekeyser, X. 1986.** 'Romance Loans in Middle English: A Reassessment' in D. Kastovsky and A. Szwedek (eds.), *Linguistics Across Historical and Geographical Boundaries* vol I, Berlin: Mouton de Gruyter, 253–265.
- de Schryver, G.-M. 2005.** 'Concurrent Over- and Under-treatment in Dictionaries—*The Woordeboek van Afrikaanse Taal* as a Case in Point.' *International Journal of Lexicography*, 18: 47–75.
- Fenton, J. 1891.** *Bush Life in Tasmania Fifty Years Ago*. London: Hazell, Watson and Viney.
- Fischer, A. 1997.** 'The *Oxford English Dictionary* on CD-ROM as a Historical Corpus: *to Wed* and *to Marry* Revisited' in U. Fries, V. Müller and P. Schneider (eds.), *From Ælfric to The New York Times: Studies in English Corpus Linguistics*, Amsterdam: Rodopi, 161–72.
- Gilliver, P. 2004.** 'That Brownest of Brown Studies' *Dictionaries: Journal of the Dictionary Society of North America*, 25: 44–64.
- Gilliver, P. 2007.** 'The Great Un- Crisis: An Unknown Episode in the History of the *OED*' in J. Considine and G. Iamartino (eds.) (2007). *Words and Dictionaries from the British Isles in Historical Perspective*. Cambridge: Cambridge Scholars Publishing: 166–77.
- Goodland, G. 2008.** 'Continual Plodders: Contextualising Shakespeare's First Uses in the *OED*' unpublished paper presented at Hel-Lex2, Finland, 25–27 April, 2008.
- Gotti, M. 1999.** *The Language of Thieves and Vagabonds*. Tübingen: Max Niemeyer.
- Hoffman, S. 2004.** 'Using the *OED* Quotations Database as a Corpus—a linguistic appraisal.' *ICAME Journal*, 28: 17–30.
- Jespersen, O. 1905.** *Growth and Structure of the English Language* Leipzig: Teubner.
- Jucker, A. H. 1994.** 'New Dimensions in Vocabulary Studies: Review Article of the *Oxford English Dictionary* (2nd edition) on CD-ROM.' *Literary and Linguistic Computing* 9/2: 149–154.
- Landau, S. I. 1984.** *Dictionaries. The Art and Craft of Lexicography*. Cambridge: Cambridge University Press (1989).
- Lynch, J. and McDermott, A. (eds.) 2005.** *Anniversary Essays on Johnson's Dictionary*. Cambridge: Cambridge University Press.
- Mair, C. 2001.** 'Early or Late Origin for *Begin* + *V-ing*? Using the *OED* on CD-ROM to Settle a Dispute between Visser and Jespersen.' *Anglia* 119: 606–610.
- Markus, M. 2001.** 'Linguistic Commercialism in and around the *Paston* and *Cely Letters*. An *OED* and Corpus-Based Approach.' *Journal of English Linguistics* 29/2: 162–178.
- McConchie, R. W. 1997.** *Lexicography and Physicke* Oxford: Clarendon Press.
- Miyoshi, Kusujiro. 2007.** *Johnson's and Webster's Verbal Examples. With Special Reference to Exemplifying Usage in Dictionary Entries*. Tübingen: Niemeyer.
- Morton, H. C. 1994.** *The Story of Webster's Third*. Cambridge: Cambridge University Press.
- Mossé, F. 1943.** 'On the Chronology of French Loan-Words in English', *English Studies* 25: 33–40.
- Mugglestone, L. 2003.** 'Proof and Process: the Making of the *Oxford English Dictionary*' in M. Dossena and C. Jones (eds.), *Insights into Late Modern English* Peter Lang, 107–127.
- Mugglestone, L. 2005.** *Lost for Words*. New Haven and London: Yale University Press.
- Nevalainen, T. 1999.** 'Early Modern English Lexis and Semantics' in R. Lass (ed.) *The Cambridge History of the English Language: Early Modern English 1476–1776* vol. III. Cambridge: Cambridge University Press: 332–458.

- Ogilvie, S. 2004.** 'From "outlandish words" to "World English": The Legitimization of Global Varieties of English in the Oxford English Dictionary' in G. Williams and S. Vessier (eds.) *Proceedings of the 11th EURALEX International Conference*. Lorient, Université de Bretagne-Sud, 651–658.
- Ogilvie, S. 2008a.** 'Rethinking Burchfield and World Englishes.' *International Journal of Lexicography*, 21/1: 23–59.
- Ogilvie, S. 2008b.** 'The Mysterious Case of the Vanishing Tramlines: James Murray's Legacy and the 1933 OED Supplement' *Dictionaries: Journal of the Dictionary Society of North America*, 29: 1–22.
- Osselton, N. E. 2006.** 'Usage Guidance in Early Dictionaries of English.' *International Journal of Lexicography*, 19/1: 99–105.
- Osselton, N. E. 2007.** 'Alphabet Fatigue and Compiling Consistency in Early English Dictionaries' in J. Considine and G. Iamartino (eds.) (2007). *Words and Dictionaries from the British Isles in Historical Perspective*. Cambridge: Cambridge Scholars Publishing: 81–90.
- Reddick, A. 1996.** *The Making of Johnson's Dictionary 1746–1773*. Cambridge: Cambridge University Press, revised edition.
- Schäfer, J. 1980.** *Documentation in the O.E.D.* Oxford: Clarendon Press.
- Starnes, D. T. and G. E. Noyes. 1946.** *The English Dictionary From Cawdrey to Johnson 1604–1755*. Chapel Hill: University of North Carolina Press.