

Legal Terms in General Dictionaries of English: The Civil Procedure Mystery

Sandro Nielsen, *Centre for Lexicography, Aarhus University, Denmark* and
International Centre for Lexicography, University of Valladolid, Spain
(sn@bcom.au.dk)

Abstract: Many general language dictionaries contain specialized terms, including legal terms relating to civil lawsuits. The existing literature provides general discussions of scientific and technical terms in ordinary dictionaries but does not specifically address the inclusion of legal terms. This study examines four general dictionaries of English to see how they treat civil procedure terms used in England and Wales in the light of the change of structure of and terminology used in civil proceedings that took place in 1999. Despite being based on large, up-to-date corpora the dictionaries contain some of the old terms but fail to include the new terms that have been in use for more than 15 years. Why this is the case is a mystery. However, some clues indicate that if they pay more attention to the link between dictionary functions, corpora and the data presented in dictionaries, lexicographers may be able to work in a more focussed way that would likely ensure the inclusion of legal terms as well as terms from other subject fields in general dictionaries. This would also satisfy the needs of users.

Keywords: LEXICOGRAPHY, INFORMATION TOOLS, CORPORA, CORPUS LEXICOGRAPHY, CULTURE-DEPENDENT DOMAINS, KNOWLEDGE, LEMMATIZATION, LEGAL LANGUAGE, DICTIONARY FUNCTIONS, COMMUNICATIVE FUNCTIONS, DECODING, ENCODING

Opsomming: Regsterme in algemene Engelse woordeboeke: Die raaisel van die siviele proses. Baie algemene woordeboeke bevat gespesialiseerde terme, insluitend regsterme wat verband hou met siviele hofsake. Die bestaande literatuur verskaf algemene besprekings van wetenskaplike en tegniese terme in gewone woordeboeke, maar hanteer nie spesifiek die insluiting van regsterme nie. In hierdie artikel word vier algemene Engelse woordeboeke bestudeer om te sien hoe hulle siviele regsterme wat in Engeland en Wallis gebruik is, hanteer. Dit word gedoen teen die agtergrond van die verandering in die struktuur van siviele hofsake wat in 1999 plaasgevind het en die terminologie daarin gebruik. Alhoewel die woordeboeke op groot, bygewerkte korpusse gebaseer is, bevat hulle sommige van die ou terme, maar slaag nie daarin om die nuwe terme wat al meer as 15 jaar gebruik word, in te sluit nie. Waarom dit só is, is 'n raaisel. Sommige leidrade dui daarop dat, as leksikograwe meer aandag skenk aan die verband tussen woordeboekfunksies, korpusse en die data wat aangebied word in woordeboeke, hulle op 'n meer gefokusde manier sou kon werk, wat waarskynlik sal verseker dat regsterme sowel as terme uit ander onderwerpsvelde in algemene woordeboeke ingesluit sal word. Dit sal ook aan gebruikersbehoefes voldoen.

Sleutelwoorde: LEKSIKOGRAFIE, INLIGTINGSHULPMIDDELS, KORPUSSE, KORPUSLEKSIKOGRAFIE, KULTUUR-AFHANKLIKE DOMEINE, KENNIS, LEMMATISERING, REGSTAAL, WOORDEBOEFUNKSIES, KOMMUNIKATIEWE FUNKSIES, DEKODERING, ENKODERING

1. Introduction

Statistics show that many people come into contact with the legal system every year, in one way or another. Some are interviewed or arrested by the police, some quibble with family members over the validity of testamentary gifts, while others come into contact with the legal system in the course of their work. As members of civilized societies we are often exposed to situations in which we directly or indirectly meet legal terms, for example when we read about rules that are intended to protect consumers, and when we complain to shops because we think that the goods and services we have bought do not comply with the terms of the contract. According to the UK Ministry of Justice (2015: 8), between 1.4 million and 2.1 million cases annually were brought before the civil courts in England and Wales from 2006 to 2012 and in the fourth quarter of 2014, approximately 379,000 civil proceedings were commenced (the figures do not include family cases such as divorce). These figures suggest that we are all likely to become involved in civil proceedings at some point during our lives and it is therefore not surprising to find that general dictionaries of English contain some legal terms.

Literature discussing the treatment of terms (often called scientific and technical words) in general language dictionaries is sparse and concerns terms in general. For instance, Béjoint (1988) provides a broad discussion of the difficulties lexicographers face when they write definitions of scientific and technical terms from various domains in general language dictionaries, while Jessen (1996) gives a general analysis of the treatment of terms in selected monolingual and bilingual general language dictionaries covering English and French. Josselin-Leray and Roberts (2005) investigate the use of general dictionaries by language specialists, scientists and laypersons when they look for terms from a range of subject fields in monolingual and bilingual dictionaries of English and French, and Urbinc and Urbinc (2013) examine the inclusion and treatment of technical and scientific terms in three editions of *Oxford Advanced Learner's Dictionary* by looking at the use of subject field labels, short cuts and sense indicators. However, none of the contributions specifically examines the treatment of legal terms in general dictionaries of English so it is relevant to look closer at this aspect of lexicography.

Many people own medium-sized general dictionaries of English, either as native speakers or as learners of English as a foreign language. Taking the annual number of civil proceedings into consideration, many people will look for help in their dictionaries when they come across terms relating to civil proceedings that they do not know or if they are uncertain about the meanings of

those terms. Another relevant aspect is that the system of civil procedure in England and Wales was reformed in 1999 and part of this reform was to change the somewhat old-fashioned terminology. Even though legal terms are often regarded as relevant for expert-to-expert communication, the fact that so many ordinary people come into contact with the civil justice system indicates that a wide range of communicative instances are expert-to-layperson and layperson-to-layperson. This may be one of the reasons why general language dictionaries contain legal terms. However, as these dictionaries are intended for the general public the aim of this paper is to examine to which extent they lemmatize the specialized terminology of civil proceedings. This involves a description of the methods used (Section 2), and a study of how four internationally renowned general dictionaries of English treat the basic terms used in civil proceedings in England and Wales with particular focus on the selection of lemmas (Sections 3 and 4), and how the (non-)lemmatization of civil procedure terms affect communicative dictionary functions, and the use of corpora (Section 5).

2. Specifying legal terms and dictionaries

It may be argued that legal terminology does not belong in general language dictionaries. This is probably correct when we talk about "small" dictionaries but the idea of including legal terms in general dictionaries is not new. In his elementary reference work, Richard Mulcaster expresses the wish that "som one well learned and as laborious a man, wold gather all the words which we use in our English tung, whether naturall or incorporate, out of all professions, as well learned as not, into one dictionarie" (Mulcaster 1582: 166), and he explicitly refers to the field of law several times in his book. In a modern setting, this wish appears to have come true as Svensén (2009: 3) observes that "General-language dictionaries usually include a considerable number of technical terms, particularly those encountered by everyone in everyday life". As indicated above, some legal terms are part of the everyday life of modern citizens, so dictionaries should provide assistance with at least some of the most important terms. The question is then to find a way in which to select the legal terms to be lemmatized and here Magay (1984: 223) gives some sound advice: "The lexicographer has to differentiate between what is new and what is really important, i.e. what the user really needs". One interpretation of this statement is that lexicographers should lemmatize those terms that users really need and according to statistics, assistance with civil procedure terms are needed by users in their everyday lives.

It is beyond the scope of this paper to examine the (non-)lemmatization of all civil procedure terms in the selected dictionaries. For one thing, English civil procedure terms are found in many different geographical locations, for instance in the United Kingdom, the United States of America, Australia and Canada. Secondly, each of these countries has its own legal system that constitutes a jurisdiction of law with its own culture-dependent terminology, though

there may be the occasional overlap. Furthermore, it should be appreciated that the United Kingdom is divided into three jurisdictions with differing terminology — this study only concerns terms used in England and Wales, while terms used in the other two jurisdictions, i.e. Scotland and Northern Ireland, will be excluded for practical purposes.

The way in which civil proceedings are systematised and carried out is based on long-standing tradition but changes do occur, most significantly in 1999 when the civil procedure system was radically reformed. This reform included a change of terminology so that several core terms were replaced by new ones. The following analysis is based on an appendix to the Civil Procedure Rules that came into force in England and Wales in April 1999, which lists 33 new terms and the old ones they replaced (British and Irish Association of Law Librarians 2015). Since many of the terms represent very specific legal concepts that are predominantly used in expert-to-expert communication, general language dictionaries cannot be expected to contain all, so the point of departure will be those old civil procedure terms that are lemmatized in a benchmark dictionary: *Oxford Dictionary of English* (ODOE). The main reason for choosing this dictionary as benchmark is the following description:

The foremost single volume authority on the English language, the *Oxford Dictionary of English* is at the forefront of language research, focusing on English as it is used today. It is informed by the most up-to-date evidence from the largest language research programme in the world, including the two-billion-word Oxford English Corpus. [...] Ideal for anyone who needs a comprehensive and authoritative dictionary of current English; for professionals, students, academics and for use at work or at home. (*Oxford Dictionary of English*)

Oxford Dictionary of English contains 11 of the 33 old civil procedure terms and these lemmas are marked as belonging to the subject field *law*, i.e. 11 lemmas and definitions that relate to the legal domain (see Table 1). This seems to indicate that the 11 terms are so important for the description of English usage that they warrant inclusion in general dictionaries of English, i.e. they are part of the public domain. Consequently, these old terms and those that have replaced them should be included in that dictionary and other medium-sized general dictionaries of English.

| Old term | ODOE |
|-------------------|-------------------|
| writ (of summons) | writ of summons |
| plaintiff | plaintiff |
| discovery | discovery |
| Mareva injunction | Mareva injunction |
| in camera | camera, in camera |

| | |
|---------------------|--------------------|
| interlocutory | interlocutory |
| statement of claim | statement of claim |
| pleading(s) | pleadings |
| interrogatory /-ies | interrogatories |
| Anton Pillar order | Anton Pillar order |
| subpoena | subpoena |

Table 1: The eleven old civil procedure terms lemmatized in *Oxford Dictionary of English*

In order to be as specific as possible, it should be appreciated that the eleven terms are classified as belonging to the legal field on the basis of the concepts they refer to. This may sound obvious, but some of the lemmas have more than one meaning and only the legal meaning is relevant. For example, ODOE defines the term *plaintiff* as "a person who brings a case against another in a court of law", i.e. a legal term proper that is marked as such with the subject field label *Law*. The term *discovery* has an unmarked general sense, "the action or process of discovering or being discovered", as well as a marked legal sense: "the compulsory disclosure, by one party to an action to another, of relevant testimony or documents". ODOE treats *pleading* as a mass noun in a general sense and as a plural noun in a legal sense with the following definition: "(usu. **pleadings**) (*Law*) a formal statement of the cause of an action or defence." Finally, *in camera* is included as a cross-reference article that merely contains a cross-reference to the lemma **camera** (sense 2), which is defined as "a chamber or round building." This article contains the phrase *in camera* and the associated definition: "*chiefly (Law)* in private, in particular taking place in the private chambers of a judge, with the press and public excluded." The lemma **camera** is not a specific legal term, but the phrase *in camera* is marked as belonging to the legal domain and, therefore, qualifies as one of the eleven terms in Table 1. In their analysis of terms in three editions of OALD Vrbinč and Vrbinč (2013: 444-449) found that technical and scientific terms are not only marked directly as terms by subject field labels following lemmas, such as *law* and *physics*, but are also marked indirectly by short cuts and sense indicators in the definitions. However, the phrase *in camera* in ODOE, which has the status of a term in the legal domain, is marked directly as belonging to that domain in connection with the phrase itself and not the lemma, the sense or the definition, a practice also observed by Jessen (1996: 96).

The dictionaries selected for the analysis are all internationally well-known, medium-sized dictionaries of English. Together they occupy a considerable part of the international lexicographical landscape and are, therefore, relevant objects of study. The most recent editions of four dictionaries have

been selected for analysis and comparison with the benchmark dictionary: *Longman Dictionary of Contemporary English* (LDOCE), *Oxford Advanced Learner's Dictionary* (OALD), *Macmillan English Dictionary* (MED) and *Cambridge Advanced Learner's Dictionary* (CALD). This includes the print editions (except MED) as well as the free online editions. Furthermore, the dictionaries are all based on large electronic corpora that have been subjected to treatment by modern software and thus share important features with the benchmark dictionary. For example, MED states that its source is "a corpus, a database containing millions of examples of English as used around the world. [...] using state-of-the-art software, has allowed the dictionary writers to reveal fresh information about how and when words are used" (*Macmillan English Dictionary*) and CALD is "Based on the 1.5 bn word Cambridge English Corpus" (*Cambridge Advanced Learner's Dictionary*). Finally, some of the publishers have one edition for British English and another for American English and the dictionaries studied all cover British English, because the civil procedure terms used in England and Wales differ considerably from those used in the USA. Having specified which civil procedure terms to focus on and in which dictionaries to look, I will go on to examine which civil procedure terms are included in the four dictionaries.

3. Which old terms are in the dictionaries?

The benchmark dictionary establishes a basis for comparing the lemmatization of civil procedure terms in the four selected dictionaries. LDOCE, OALD, MED and CALD all contain lemmas that are marked with a subject label referring to the legal domain, *law* and *legal* respectively. Table 2 shows how many of the 11 old civil procedure terms lemmatized in the benchmark dictionary are also lemmatised in the four dictionaries examined. CALD scores the highest as it includes 7 of the old terms, while LDOCE has the fewest with only 4 of the eleven terms; note, however, that the terms *in camera*, *interlocutory injunction*, *statement of claim* and *pleadings* are only included in the online edition of CALD. The other two dictionaries come between these two extremes, in that OALD and MED both include 6 of the old civil procedure terms.

| Old term | LDOCE | OALD | MED | CALD |
|-------------------|---------------|---------------|-----------|-----------------------------|
| writ (of summons) | writ | writ | writ | writ |
| plaintiff | plaintiff | plaintiff | plaintiff | plaintiff |
| discovery | — | — | discovery | — |
| Mareva injunction | — | — | — | — |
| in camera | camera, in c. | camera, in c. | in camera | in camera |
| interlocutory | — | — | — | interlocutory injunction |

| | | | | |
|--------------------|----------|---------------|----------|--------------------|
| statement of claim | — | — | — | statement of claim |
| pleading(s) | — | pleading | pleading | pleadings |
| interrogatory/-ies | — | interrogatory | — | — |
| Anton Pillar order | — | — | — | — |
| subpoena | subpoena | subpoena | subpoena | subpoena |

Table 2: Old civil procedure terms lemmatized in the four medium-sized dictionaries

Two of the eleven terms are included only in the benchmark dictionary: *Anton Pillar order* and *Mareva injunction*. This means that, together, the four dictionaries cover nine of the eleven old civil procedure terms, though CALD print edition and LDOCE include less than half of the terms. The terms studied so far are those that were used before 1999, so it is relevant to examine whether the dictionaries include the new terms that have applied since 1999.

4. Which new terms are in the dictionaries?

All the dictionaries analysed are based on corpora that are intended to contain systematic collections of texts that document usage features of English, in particular British English. As indicated in Section 2, the corpus-based approach to dictionary-making is emphasised by the publishers and is a fact that should be taken into consideration. Firstly, the benchmark dictionary is examined and Table 3 shows 1) the old civil procedure terms, 2) the new terms that have replaced them, and 3) which of the new terms are lemmatized in *Oxford Dictionary of English*.

| Old term | New term | ODOE |
|--------------------|----------------------|------|
| writ (of summons) | claim form | — |
| plaintiff | claimant | (—) |
| discovery | disclosure | — |
| Mareva injunction | freezing injunction | — |
| in camera | in private | — |
| interlocutory | interim | — |
| statement of claim | particulars of claim | — |
| pleading(s) | statement of case | — |

| | | |
|---------------------|---------------------------------|---|
| interrogatory /-ies | request for further information | — |
| Anton Pillar order | search order | — |
| subpoena | witness summons | — |

Table 3: The eleven new civil procedure terms lemmatized in *Oxford Dictionary of English*

The analysis shows that ODOE lemmatizes none of the eleven new civil procedure terms. This result is surprising. Since the eleven old terms were judged relevant for inclusion, it is reasonable to expect that the new terms would also be included and given the same treatment as the old terms, in particular because the new terms have been in use for more than 15 years and, therefore, must be expected to have replaced the old terms in the public domain. The reason for the parenthesis against the term *claimant* is that ODOE treats it in a special way. The article **plaintiff** contains the following note: "USAGE: In England and Wales the term **plaintiff** was officially replaced by **claimant** in 1999". This is a proper way in which to treat this lemma but ODOE's definition of the lemma **claimant** is not optimal: "a person making a claim, especially in a lawsuit or for a state benefit." First of all, as the term *claimant* has replaced the term *plaintiff*, it is reasonable to expect that the two are defined in the same way because they both refer to the same concept in the real world. Secondly, persons claim benefits from the state whereas they usually start legal proceedings against individuals or companies. This means that the lemma **claimant** should have been treated as polysemous and the legal sense should have been marked with an appropriate subject field label, which would have been in line with the treatment given to other lemmas such as **discovery** discussed in Section 2 above. It should also be noted that the words *in private* and *interim* are lemmatized in ODOE but are described in a general sense and not in a specific legal sense, i.e. they are not treated as terms.

The next step is to examine the lemmatization of the new civil procedure terms in the four medium-sized dictionaries. Table 4 shows how many of the 11 new civil procedure terms are lemmatized in the four dictionaries and this may be compared with the findings in Table 3. Again CALD is the highest scorer as it includes 2 of the new terms, while LDOCE, OALD and MED contain none of the eleven terms. Note that the two terms lemmatized in CALD are only included in the online edition.

| New term | LDOCE | OALD | MED | CALD |
|------------|-------|------|-----|------|
| claim form | — | — | — | — |
| claimant | — | — | — | — |
| disclosure | — | — | — | — |

| | | | | |
|---------------------------------|---|---|---|----------------------|
| freezing injunction | — | — | — | — |
| in private | — | — | — | — |
| interim | — | — | — | — |
| particulars of claim | — | — | — | particulars of claim |
| statement of case | — | — | — | — |
| request for further information | — | — | — | — |
| search order | — | — | — | — |
| witness summons | — | — | — | witness summons |

Table 4: New civil procedure terms lemmatized in the four medium-sized dictionaries

As Table 4 indicates, the four dictionaries do not treat the new civil procedure terms in any significant way. LDOCE and CALD are the only dictionaries to include a lemma called **claim form** but neither defines this term in the legal sense, i.e. identical to the old term *writ*. All dictionaries include the word *claimant* but none of them defines it in a legal sense, though LDOCE provides the following definition: "someone who claims something, especially money, from the government, a court etc. because they think they have a right to it". From a legal perspective, it is unlikely that anyone will claim money from a court; they are, however, likely to claim money from a person or company in a case heard by a court. All dictionaries include the word *disclosure* but only in its ordinary meaning of "revealing a secret", and none of the dictionaries include the term *freezing injunction*. The words *in private* and *interim* are included in the dictionaries but defined in their ordinary meanings with no relation to civil proceedings. CALD online includes the term *particulars of claim*, i.e. "detailed information that you have to provide when asking a court, government department, or company to give you something such as money or property that you believe you have a legal right to", and the term *witness summons*, i.e. "a legal document ordering someone to appear in a court of law to give information about a particular person or event", both with their legal definitions and subject field labels. None of the four dictionaries includes the terms *statement of case*, *request for further information*, and *search order*.

The findings in Tables 2 and 4 show that the four general language dictionaries lemmatize some old civil procedure terms but not new terms and the difference in treatment is not clear. A discussion of lemmatization principles

and the use of electronic corpora may provide clues that can help solve the mystery.

5. Discussion of findings and implications

Users consult general language dictionaries because they think these information tools can help them solve various problems. The mainly quantitative analyses in Sections 3 and 4 reveal that the four dictionaries provide limited help with civil procedure terms but a discussion including qualitative aspects may help explain the difference in treatment. Some of the findings and their implications for dictionary making, dictionary use in communicative situations and use of corpora will be discussed in the following sub-sections.

5.1 Selecting lemmas according to dictionary function

The theory of dictionary functions is relevant when assessing the data contained in dictionaries because dictionary functions determine all active decisions from the selection of lemmas, over the selection of data types, to the way in which the data are presented. Bergenholtz and Tarp (2010: 30) define a dictionary function as "the satisfaction of the specific types of lexicographically relevant needs that may arise in a specific type of potential user in a specific type of extra-lexicographical situation". This means that lexicographers make their decisions with due regard to the basic needs of the intended users identified through user profiling (see e.g. Bergenholtz and Nielsen 2006: 286) and attempt to match those needs with the dictionary function(s).

The four dictionaries examined have the same main types of communicative function. Generally speaking, dictionaries with communicative functions are designed to provide specific types of help to specific types of user in specific types of usage situations involving ongoing or planned acts of communication, for example providing help to understand words and terms users meet when reading texts, and providing help to choose the correct word and use it correctly when writing texts. According to the informative material on the dictionaries, they have two primary communicative functions: to provide help with the meaning of words and to provide help to produce oral and written texts. Table 2 shows that the four dictionaries together include nine of the old civil procedure terms analysed, which must be regarded as a good degree of coverage compared to the eleven old terms included in the benchmark dictionary. However, a closer study reveals that the matter is not that simple.

The dictionary that provides the best help to understand the meaning of old terms is the online edition of CALD: it provides the meanings of seven of the civil procedure terms analysed. The dictionaries that provide the least help to understand terms are the print edition of CALD and LDOCE in that users will only be able to find the meanings of three and four civil procedure terms,

respectively. The implication of the study of old terms reported in Table 2 is that users will have to consult CALD online, OALD as well as MED in order to find the meanings of all nine terms included in the four dictionaries. Consequently, the four dictionaries provide varied help to understand those civil procedure terms that were in use in England and Wales before 1999.

Table 4 shows that help to understand civil procedure terms in use after 1999 is almost non-existent. CALD online is the only dictionary that lemmatizes new terms so that users will be able to find the meanings of two of the eleven terms; LDOCE, OALD, MED and CALD print edition provide no help to understand any of the eleven new civil procedure terms analysed. Nevertheless, the print editions of the four medium-sized dictionaries treat the new terms in the same way as the benchmark dictionary (CALD online scores higher). This means that LDOCE, OALD, MED and CALD print edition give users no help with the meaning of the new terms and that CALD online explains the meanings of two of the eleven new civil procedure terms that replaced the old terms in 1999 and have been in the public domain since then.

It is somewhat more complicated to assess the coverage of civil procedure terms related to the communicative function called text production. The dictionary that provides the best help to produce oral and written communication is also the online edition of CALD as it contains seven of the eleven old terms, see Table 2. Again CALD print edition and LDOCE provide the least help to produce oral and written communication because they lemmatize only three and four of the eleven old terms, respectively. That being said, the four dictionaries provide some help to produce texts which contain terms that were used pre-1999 but it is reasonable to expect that users of the dictionaries rarely need help to produce texts that contain those terms. Users are, however, more likely to look for assistance with producing texts that contain the new civil procedure terms.

Table 4 shows that the four dictionaries provide almost no help to produce oral and written communication containing civil procedure terms in current use after 1999. CALD online is the only dictionary that lemmatizes new terms so that users will be able to find help with two of the eleven terms; LDOCE, OALD, MED and CALD print edition provide no help to produce texts with any of the eleven new civil procedure terms analysed. Nonetheless, the print editions of the four medium-sized dictionaries treat the new terms in the same way as the benchmark dictionary (CALD online is the exception). This means that LDOCE, OALD, MED and CALD print edition give users no assistance in producing oral and written texts in which the new terms occur and that CALD online provides help with two of the eleven new civil procedure terms that have been part of the public domain since 1999. Since the fact that they include old civil procedure terms indicates that the dictionaries are designed to provide help to understand and use civil procedure terms (communicative functions), the lack of help with the new terms is puzzling, in particular since the dictionaries are based on large text corpora.

5.2 Corpora: what seems to be the problem?

Lexicographers have benefited greatly from the introduction of electronic corpora in dictionary making. Corpora can help lexicographers to identify such features as the meanings of words and their grammatical properties with relative ease, and finding the meanings of words is particularly relevant for this study. The four dictionaries examined are all informed by large corpora and the informative material on CALD online explains that this dictionary "includes up-to-date vocabulary" and that it is based on the Cambridge English Corpus, which is "a multi-billion word collection of written and spoken English." Similarly, OALD claims that it is "based on the authoritative Oxford English Corpus", which was also used for compiling the benchmark dictionary (see Section 2 above). So why do these dictionaries score so low?

The meanings of legal terms found in corpora can be discovered from the context in which they occur, in particular from external text-type indicators. As explained by Atkins and Rundell (2008: 299) "For lexicographers, text-type features such as domain, time and regional dialect often provide valuable evidence to support the process of identifying dictionary senses" and these features can be used to discuss the lemmatization of legal terms. As indicated in section 2, the field of law is a culture-dependent domain because each jurisdiction has its own internal structure that is reflected by the terms used. This means that texts from this domain included in corpora should be clearly marked (tagged) as belonging to the field of law. Secondly, the fact that the domain is culture-dependent means that lexicographers should carefully study the "regional dialect" used in the texts, so that texts are marked as belonging to the legal domain in the UK, the USA, Australia etc. based on the language variety in which they are written and the geographical location of the publisher. Finally, lexicographers should also pay attention to the date of the texts included in corpora because the field of law is a dynamic field in which structural as well as terminological changes occur. As far as civil procedure terms in England and Wales are concerned, the important year is 1999; for the text-reception function, lexicographers should examine corpora for both old and new civil procedure terms, while they should look for new terms for the text production function.

A look at what is going on in the world with specific focus on civil proceedings and how this is reflected in text types that are likely to be included in corpora is relevant. Texts from UK newspapers with nationwide distribution are likely to be included in lexicographical corpora for general language dictionaries so the homepages of The Times and The Guardian were visited in order to see how often the two civil procedure terms *plaintiff* and *claimant* occur in these papers. The two homepages were visited on the same day and in order to make the searches as focussed as possible, the strings "plaintiff AND court" and "claimant AND court" were typed in the search boxes; the search results are presented in Table 5.

| Legal term | The Times | The Guardian |
|------------|-----------|--------------|
| plaintiff | 923 | 5,160 |
| claimant | 2,970 | 42,400 |

Table 5: The number of texts containing legal terms in two UK newspaper databases

Table 5 shows that the new term *claimant* occurs in far more texts than the term *plaintiff*. If the term *plaintiff* is included in the dictionaries studied because of frequency of occurrence then the term *claimant* should also be lemmatized because it is much more frequent than *plaintiff*. A brief examination of the texts in which the term *plaintiff* occurs in the electronic archives of both newspapers reveals that most articles either appeared before the year 2000 (the time feature) or refer to civil proceedings outside England and Wales, particularly in the USA where the term *plaintiff* is in current use. One implication of this is that corpus builders and lexicographers should not only tag legal texts as British texts (the regional dialect feature) because they were published in the UK but they should also tag them with respect to the factual and culture-dependent contents (the domain feature), so that those texts dealing with civil proceedings in England and Wales are marked as such while texts describing civil proceedings in, for instance, the USA should be marked as such. This is one way in which to avoid the type of situation described by Magay (1984: 224): "Whether traditional human methods or computer-aided procedures are used, help must be given to the practicing lexicographer because he is often lost in a sea of words and terms, having to make haphazard choices". Another way in which to avoid haphazard choices is to have practicing lexicographers with factual and domain-specific knowledge about the subject fields and topics that are represented in dictionaries, because they possess the knowledge required to resolve issues regarding domain, time and regional dialect features.

In connection with legal terms in general dictionaries, it should be appreciated that there is a direct link between dictionary functions and underlying corpora. As mentioned in sub-section 5.1 above, dictionary functions affect all decisions for selecting lemmas and this includes decisions for compiling the corpus that is to form the basis of the selection. An assessment of the lemmatization of civil procedure terms reported in Tables 2 and 4 in the light of the extensive use of corpora, begs the following questions: 1) Do the corpora reasonably represent what is going on in the real world? 2) Are the corpora used in a reasonable way? The above discussion only answers these questions tentatively but at the same time highlights some important facts to be taken into consideration when including civil procedure terms — and other legal terms — in general language dictionaries. First of all, the corpora should be compiled so that they include texts that can help lexicographers find data that enable them to make dictionaries with, in this case, communicative functions, i.e. contain

data that help users understand civil procedure terms and use those terms in contemporary settings. This implies that lexicographers should possess the relevant factual and domain-specific knowledge in order to identify and lemmatize those civil procedure terms users really need in everyday usage situations so users can understand and talk about what is going on in the real world. It seems reasonable to suggest that the use of corpora for general language dictionaries treating legal terms, and perhaps terms from other culture-dependent subject fields, should be examined more extensively.

Secondly, the lemmatization of primarily old terms assists users who need help to understand old and new texts dealing with pre-1999 subject matter, while the lemmatization of primarily old terms does not help users talk and write about contemporary situations involving civil proceedings in England and Wales. Finally, one consequence of the practice followed by lexicographers of general language dictionaries when they do not lemmatize terms, when they remove subject field labels, and when they attempt to write very generalized definitions that (attempt to) cover both general and domain-specific meanings of words and terms is that the dictionaries become less authoritative. There are significant lemma lacunae and definitions become vague because the dictionaries contain definitions written for everyone and no-one with users ending up being uncertain or confused (see Nielsen 2012 for a similar discussion of the treatment of legal terms in *Collins COBUILD Advanced Dictionary*).

6. Concluding remarks

One of the problems facing lexicographers is that dictionaries can never be completely up to date due to various time lags involved in the dictionary-making process; though many time lags in online dictionaries are considerably shorter than those for print dictionaries. This means that when they consult general language dictionaries, users look at the present (the dictionary in front of them) but see the past. This is often no problem if the past still applies to the present but in connection with civil procedure terms some of the past does not apply to the present, as discussed above. The fact that the general language dictionaries examined do not lemmatize the new civil procedure terms that have applied since 1999 seems to indicate that a time lag of 15 years is too long. Furthermore, the fact that the four dictionaries include several of the old terms but not the new terms is a mystery, in particular because the dictionaries claim to contain up-to-date vocabulary.

The above findings and discussion indicate that lexicographers should pay more attention to the link between dictionary functions and underlying corpora. One way in which to ensure the lemmatization of new legal terms, in this case civil procedure terms, is to compile electronic corpora that contain texts that include relevant data, i.e. data that can help lexicographers make dictionaries that have the two communicative functions: provide help to understand terms and provide help to produce oral and written communication using

those terms. In addition, lexicographers should have the necessary factual and culture-dependent knowledge in order to identify and select the terms to be lemmatized. The discussion shows that the inclusion of old civil procedure terms is appropriate since the dictionaries provide help to understand those terms and to write about the concepts referred to by the terms. However, the failure to include the new civil procedure terms does not satisfy the needs of users because it prevents users from understanding and communicating about things that go on in the real civil procedure world today even though many people are likely to be involved in civil lawsuits at one time or another. A more stringent adherence to the principle of lexicographical functions would be one way in which to make dictionaries that contain up-to-date legal terms that provide help in communicative usage situations.

References

Dictionaries

- Deuter, M., J. Bradbery and J. Turnbull (Managing Eds.).** 2015. *Oxford Advanced Learner's Dictionary*. Ninth edition. Oxford: Oxford University Press.
- Deuter, M., J. Bradbery and J. Turnbull (Managing Eds.).** 2015. *Oxford Advanced Learner's Dictionary*. Ninth edition. [Online]. <http://www.oxfordlearnersdictionaries.com/> [24 April 2015].
- Fox, C. and R. Combley (Managing Eds.).** 2014. *Longman Dictionary of Contemporary English*. Sixth edition. Harlow: Pearson.
- Fox, C. and R. Combley (Managing Eds.).** 2015. *Longman Dictionary of Contemporary English*. Updated edition. [Online]. <http://www.ldoceonline.com/> [25 April 2015].
- Hands, P. (Senior Ed.).** 2009. *Collins COBUILD Advanced Dictionary of English*. Glasgow/ Boston: Harper Collins Publishers/Heinle Cengage Learning.
- MacIntosh, C. (Editor-in-chief).** 2013. *Cambridge Advanced Learner's Dictionary*. Fourth edition. Cambridge: Cambridge University Press.
- MacIntosh, C. (Editor-in-chief).** 2013. *Cambridge Advanced Learner's Dictionary*. [Online]. <http://dictionary.cambridge.org/dictionary/british> [24 April 2015].
- Rundell, M. (Editor-in-chief).** 2015. *Macmillan English Dictionary*. [Online]. <http://www.macmillandictionary.com/> [25 April 2015].
- Stevenson, A. (Ed.).** 2015. *Oxford Dictionary of English. Third edition. 2015 Online Version*. [Oxford Reference Online]. DOI: 10.1093/acref/9780199571123.001.0001 [25 April 2015].

Other literature

- Atkins, B.T.S. and M. Rundell.** 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.
- Béjoint, H.** 1988. Scientific and Technical Words in General Dictionaries. *International Journal of Lexicography* 1(4): 354-368.
- Bergenholtz, H. and S. Nielsen.** 2006. Subject-field Components as Integrated Parts of LSP Dictionaries. *Terminology* 12(2): 281-303.

- Bergenholtz, H. and S. Tarp.** 2010. LSP Lexicography or Terminography? The Lexicographer's Point of View. Pedro A. Fuertes-Olivera (Ed.). 2010. *Specialised Dictionaries for Learners*: 27-38. Lexicographica Series Maior 136. Berlin/New York: Walter de Gruyter.
- British and Irish Association of Law Librarians.** 2015. *Appendix Two. Civil Procedure Rules Terminology*. [Online]. Available at: <http://www.biall.org.uk/data/files/Publications/Moys%20Classification,%20App%20,%20extractpages.pdf> [25 April 2015].
- Jessen, A.** 1996. *The Presence and Treatment of Terms in General Dictionaries*. M.A. thesis, University of Ottawa, School of Translation and Interpretation, Ottawa. [Online]. Available at: <https://www.ruor.uottawa.ca/handle/10393/4427> [29 April 2015].
- Josselin-Leray, A. and R.P. Roberts.** 2005. In Search of Terms: An Empirical Approach to Lexicography. *Meta* 50(4): 256-265. doi: 10.7202/019920ar.
- Magay, T.** 1984. Technical or General: Problems of Vocabulary Selection in a Medium-size Bilingual Dictionary. Hartmann, R.R.K. (Ed). 1984. *LEXeter '83 Proceedings: Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*: 221-225. Lexicographica. Series Maior. Tübingen: Max Niemeyer.
- Ministry of Justice.** 2015. *Civil Justice Statistics Quarterly October to December 2014*. Available at: <https://www.gov.uk/government/statistics/civil-justice-statistics-quarterly-october-to-december-2014> [20 April 2015].
- Mulcaster, R.** (1582). *The First Part of the Elementarie vvchich Entreateth Chefelie of the Right Writing of our English Tung, Set Furth by Richard Mulcaster*. London: Thomas Vautroullier dwelling in the blak-friers by Lud-gate. Available at: <http://quod.lib.umich.edu/cgi/t/text/text-idx?c=eebo;idno=A07881.0001.001> [26 April 2015].
- Nielsen, S.** 2012. Domain-specific Knowledge in Lexicography: How it Helps Lexicographers and Users of Accounting Dictionaries Intended for Communicative Usage Situations. *Hermes - Journal of Language and Communication in Business* 50: 51-60.
- Svensén, B.** 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary-making*. Cambridge: Cambridge University Press.
- The Guardian* [online]. <http://www.theguardian.com> [29 April 2015].
- The Times* [online]. <http://www.thetimes.co.uk> [29 April 2015].
- Vrbinc, M. and A. Vrbinc.** 2013. Differences in the Inclusion and Treatment of Terminology in OALD3, OALD4 and OALD8. *Lexikos* 23: 440-455.