Sandro Nielsen

Legal Lexicography in Theory and Practice

1. Introduction

Dictionaries that provide help in communicative and cognitive situations are important information tools in today's society, in particular when two different languages and legal systems are involved. Researchers at Centre for Lexicography, Aarhus University, Denmark, have been involved in theoretical and practical lexicographic work related to the field of law since the late 1980s, so this work is a significant part of the Centre's output throughout the years. Researchers connected with the Centre in various ways have published several print dictionaries: an English-Danish dictionary of law, a Spanish-Danish dictionary of law, a German-Danish dictionary of law, and a Danish-German dictionary of law. In addition, one online Danish-English law dictionary has been made available and the English-Danish dictionary mentioned above is now only available as a revised and updated online dictionary. The theoretical work on legal lexicography has centred on developing principles for an English-Danish contract law dictionary (Nielsen 1994) and the lexicographic treatment of cultural differences in Danish and Spanish succession law. New insights into principles and practices of legal lexicography have been published in papers on an ongoing basis (see e.g. Nielsen 2000, Nielsen 2001, Nielsen 2010 and Nielsen 2012). The following presentation of a new project involving legal lexicography is partly based on Nielsen (2014).

2. Introducing the project on legal lexicography

The digitalization of information activities and information tools has influenced lexicographic theory and practice, in particular during the last 10 to 15 years. One outcome of this general trend is that print dictionaries have gradually been replaced by online dictionaries and many online dictionaries, sometimes completely new, have been made available. Online dictionaries compete with internet search engines as providers of data that users can process into useful information, but internet search engines tend to provide so many results from searches in unstructured data environments that the results are often irrelevant to searches for specific information. Lexicographers can remedy this situation by developing theories and principles that enable them to design online dictionaries that allow users to access structured data with targeted searches and compilers to present search results in structured and adapted ways.

The project is based on the modern theory of dictionary functions and this theoretical foundation determines all lexicographic decisions from the selection of entry words, over the selection of data types, to the way in which the data are presented. The lexicographers make their decision with due regard to the basic needs of the intended users identified through a user profiling exercise (see e.g. Bergenholtz & Nielsen 2006: 286) and attempt to match those needs with the dictionary function(s). The project aims to develop a database that can serve as the core of a set of monolingual and bilingual dictionaries designed to help Danish legal practitioners and scholars,

professional translators and law students to understand Danish and English legal texts, write legal texts in Danish and English, translate Danish legal texts into English and vice versa, and acquire knowledge about Danish and English legal matters and legal language. At present the lexicographic foundation for the database and dictionaries has been established and the dictionaries are seen as information tools that provide help to intended users in general types of communicative and cognitive situations relating to Danish and English law.

Dictionary functions are lexicographic responses to specific types of user needs in the real world, and the functions of the planned dictionaries are, in summary form (see also Section 4 below):

- to provide help to understand English and Danish legal texts;

- to provide help to write legal texts in English and Danish;

- to provide help to translate legal texts into and from English and Danish;

- to provide help to acquire general and specific knowledge about legal and linguistic matters in English and Danish.

It is important to note that usage situations occur in what is referred to as an extra-lexicographic environment. For example, lawyers who are writing texts are in an extra-lexicographic environment that exists independently of dictionaries, as they are writers of texts and, therefore, potential dictionary users. They may encounter problems related to legal text production and look for a solution in a dictionary at some point during the process and not until they consult the dictionary have they entered the lexicographic environment as actual dictionary users. Once they have found the help they need, they leave the lexicographic environment and go back to producing their texts in the extra−lexicographic environment. The writers may repeat this sequence of events several times until the task has been completed satisfactorily. In order to provide the necessary help, lexicographers must have an appropriate foundation for the online dictionaries.

3. Introducing the technical foundation

The first foundation component of the project is a database containing structured data that allows easy search in and retrieval from discrete data fields each holding a specific type of data. Examples of data types are entry words, definitions, inflectional paradigms, and phrases. These fields are connected through links that establish relationships between the different types of data such as the entry word and inflectional data. The second component is a search engine that searches the data fields in the database from where it retrieves the relevant data and the third foundation component is a user interface that presents retrieved data according to the setups for particular dictionary functions. For instance, defining dictionaries will only show entry words and definitions, while translation dictionaries will show more data types including entry words, definitions, inflectional paradigms and equivalents.

The lexicographic basis makes it possible for lexicographers to include independent lexicographic sections in the dictionaries. These sections are in effect web pages that support the function(s) and use of the dictionaries and two sections are planned. One is a subject−field section that provides introductions to Danish and English law as well as a comparative section providing data on issues where there are important differences between Danish and English law (see e.g. Nielsen 1994: 98-105, and Bergenholtz & Nielsen 2006). The second section provides help to translate linguistic structures and genre conventions by explaining the possible solutions to problems of translating relevant linguistic structures in legal communication as well as verbal and non-verbal genre conventions found in legal communication supported by appropriate examples. The plan is to include illustrative examples of documents from particular legal genres, such as contracts and statutes, as these illustrative texts may contain examples of all the verbal and non-verbal genre conventions found in contracts, statutes etc. (see Nielsen 2000: 162-165, and Nielsen 2010: 78-82). These sections will operate as integrated lexicographic components in that links between data in search results pages will be made to particular parts of the sections in functionally relevant cases. As a result, each planned online dictionary is supported by an output device (i.e. user interface) that arranges the data retrieved from the database according to type and presents them in a predetermined order that matches user needs either as articles or non-articles (e.g. individual section paragraphs and phrases). In addition to the two sections, the database will contain the following types of data:

- Entry word (including homonym index where relevant)
- Grammatical data addressed to entry word (inflection, countability, active and passive forms)
- Equivalent
- Grammatical data addressed to equivalent (inflection, countability, active and passive forms)
- Definition of entry word (including language code and polysemy index where relevant)
- Collocations (short and long combinations of words but not full sentences)
- Example sentences (full sentences)
- Antonyms and synonyms (addressed to entry word and/or equivalent)
- Source (reference and/or link)
- Grammar note (addressed to entry word or equivalent)
- Usage note (addressed to entry word or equivalent)
- Contrastive note (factual differences between entry word and equivalent)
- Cross−reference (link to relevant data)

Each jurisdiction structures its legal system in its own way and the legal language used in a jurisdiction reflects the structure of that particular system. However, as English is the (or one of the) official language(s) in several jurisdictions, the term *English* does not only refer to the legal framework of a particular country, such as UK English, but may refer to varieties of English used in communication involving different legal systems. In international contexts an "international" variety of English is used in treaties and conventions, such as the *United Nations Convention on Contracts for the International Sale of Goods*. These texts aim to describe regulatory frameworks in

a language that can be understood by as many people as possible no matter which jurisdiction they belong to. The planned online law dictionaries therefore specify the geographical and jurisdictional affiliation of terms using abbreviated language codes such as EN (England and Wales), EU (the EU), US (the USA), and INT (international English).

4. Multiple access routes to data

Users can access the data in the planned law dictionaries according to their needs. By focussing on the needs of users in various types of usage situation, lexicographers can ensure that the data satisfy a specific type of user need and that the data are presented in such a way that they can easily be turned into useful information. This goal can be achieved by making dictionaries that offer users search options related to the situations users are in when consulting the dictionaries. In a particular type of usage situation, users will need help of a specific kind and consult the dictionary they believe is most likely to help them. The monolingual law dictionaries will offer users the following kinds of assistance:

- Help to understand a legal term
- Help to write a text where the expression is known
- Help to find a legal term where the meaning is known
- Help to acquire knowledge

The bilingual law dictionaries will offer the following kinds of assistance:

- Help to translate a legal term
- Help to translate a collocation or phrase
- Help to acquire knowledge

The law dictionaries will be designed so that they are easy to use, no matter which type of help is sought. When they consult a dictionary, users go to the dictionary website where the search engine will access the database and retrieve the relevant data, which will be presented to users on the screen in a predetermined order and format. Even though the project has only reached the end of phase one, the following discussion illustrates how the dictionaries will work when fully developed and how the database will interact with the dictionaries. The examples below are prototypes and exemplify how the search engine can be manipulated by the lexicographers to retrieve the types of data that users need in various usage situations.

5. Help to translate legal texts into English

In order to fulfil the needs of intended users who translate from Danish into English, the lexicographers should compile a dictionary with a syntagmatic orientation, i.e. a dictionary providing help to translate collocations, fixed expressions, phrases and sentences in addition to terms. Taking the user group and dictionary functions into consideration, lexicographers will design a database that contains specialist (legal) terms and ordinary words that occur in legal texts, so that the lexicographic needs of users will be satisfied in the best possible way. Typical examples of words other than legal terms are those found in legal texts with their ordinary meanings as well as specific meanings (semi-technical terms), and words that are particularly difficult to translate into English, e.g. because there are several English words to choose from but only one is correct in a specific context. One example is the Danish verb *afslutte*, which may mean 1) stop doing something, 2) enter into a legally binding relationship, 3) ending a legal relationship, and 4) finalise an activity. Accordingly, the database will have to contain the following types of linguistic units:

1. Parts of speech based on grammatical categorisations, i.e. nouns, verbs (including phrasal verbs), adjectives, adverbs, prepositions, abbreviations, multi-word terms;
2. Inflexional data, i.e. comparison, derivation and conjugation;
3. Collocations, which may contain a combination of terms of art and ordinary words, e.g. adjective + noun (e.g. *omstødeligt ægteskab = voidable marriage*), verb + noun (e.g. *overdrage ejendomsretten = transfer ownership*), verb + preposition (e.g. *anvende til (et bestemt formål) = use for (a specific purpose)*), and verb + adverb (e.g. *fastsætte stiltiende = determine implicitly*)
4. Sentences showing terms in context. This type of data is often helpful, as the data may illustrate different translations of a Danish word, e.g. the Danish verb *indtræde*, in different contexts: 1) *Forsikringsselskabet indtræder i den forsikredes rettigheder = The insurance company is subrogated to the rights of the policyholder*; 2) *Konkursboet indtræder i skyldners rettigheder og pligter = The estate assumes the rights and obligations of the bankrupt*; 3) *Fra overtagelsesdagen indtræder selskabet i medarbejdernes ansættelseskontrakter = From the takeover date the company will become a party to the service contracts of the employees*.

Experience shows that non-linguists and translators with a low degree of factual legal knowledge tend to think in terms of Danish language structures when they communicate in English and the result is ungrammatical – and often also unidiomatic – English texts. The lexicographers therefore need to allow for various grammatical aspects that may help users translate texts into English.

Danish translators often need help when terms and other words in English have properties that are different from those in Danish. In order to help translators solve as many problems as possible, the Danish-English dictionary presents various types of grammatical data. Where the English inflectional paradigm differs from the corresponding Danish paradigm, the dictionary explains those differences that intended users cannot be expected to be familiar with. Compound nouns may be a case in point. As a general rule, Danish and English add a plural marker (generally: in Danish –*r*; in

English –s) to the final element of compounds but many English compounds do not follow this rule. Danish orthographic rules provide that compound nouns are generally written as single words (closed compounds), whereas English compounds are often written as strings of separate words (open compounds), and this often causes problems for Danish translators. As a result, the Danish term "*leveringsbetingelse*" forms the plural regularly by adding –*r* to the end of the word; its English equivalent, *term of delivery*, forms the plural by adding –*s* to the first constituent element. The dictionary will therefore explicitly tell users that the plural of the English word is *terms of delivery*. Similarly, Danish lawyers and novice translators cannot be expected to know that some English words change the last letter if it is a vowel when inflected for the plural. Therefore, the dictionary will explicitly inform users that the plural of the English term *remedy* is *remedies*. Finally, the spelling of the past tense of English irregular verbs will explicitly be indicated so that, for example, users are told that the verb *buy* is spelt *bought* in the past tense; again, the above types of grammatical data provide help to translate Danish texts correctly into English.

Lawyers, translators and students who translate legal texts may need to know how to translate the Danish term *varemærke* and consult the *Danish−English Law Dictionary* that has the function of providing help to translate legal terms. The output device will present various types of data addressed to the entry word (i.e. search word) and its equivalent, as shown in Figure 1.

varemærke *substantiv* <et varemærke; varemærket; varemærker; varemærkerne>

Definition
Et varemærke er et ord, symbol eller andet tegn, der af en erhvervsdrivende virksomhed benyttes som særligt kendetegn for dennes varer eller tjenesteydelser, og som nyder beskyttelse i henhold til de immaterialretlige regler.

trademark *noun* <a trademark; the trademark; trademarks; the trademarks>

Kollokationer
anmelde et varemærke til registrering
apply for registration of a trademark
erhvervede varemærker samt lignende rettigheder
acquired trademarks and similar rights
et ikke-registreret varemærke
an unregistered trademark
et registreret varemærke
a registered trademark
krænke et varemærke
infringe a trademark
retsstridig brug af et varemærke
wrongful use of a trademark
udslette et varemærke af registeret
strike a trademark off the register
vederlag for anvendelse af varemærker
charges for the use of trademarks

Eksempler

• Hvis indehaveren af et ældre varemærke rejser indsigelse, kan det varemærke, der søges registreret, være udelukket fra registrering.
 • Upon opposition by the proprietor of an earlier trade mark, the trade mark applied for may not be registerable.

Figure 1: Help to translate the Danish term *varemærke* into English.

Figure 1 shows how specific data types can provide help to Danish lawyers, translators and students who are translating legal texts into English. The definition makes it clear whether or not users have accessed the term with the relevant meaning and the equivalent with its associated inflectional data enable users to write grammatically correct texts in English. In addition to the equivalent, Danish collocations, phrases and examples with their translations into English are types of data that assist users in translating legal texts.

Users may want help to translate linguistic units larger than terms. In this case, translators can consult the *Danish−English Law Dictionary* that has the function of providing help to translate collocations and phrases, which allows users to type in a term, collocation or phrase in the search box. The search engine will search all data fields in the database containing collocations, phrases and example sentences and the output device will present the collocations and phrases retrieved from the database. Figure 2 shows the result of a search for collocations and phrases in which the Danish term *tilbagekalde* occurs.

tilbagekalde et tilbud
revoke an offer [after it has reached the offeree]

tilbagekalde et tilbud
withdraw an offer [before it has reached the offeree]

tilbagekalde en accept
revoke an acceptance

tilbagekalde et testamente
revoke a will

tilbagekalde et farligt produkt
recall a dangerous product

Et tilbud kan tilbagekaldes, såfremt tilbagekaldelsen kommer frem til tilbudsmodtageren før eller samtidig med tilbuddet.
An offer may be withdrawn if the withdrawal reaches the offeree before or at the same time as the offer.

Et tilbud kan tilbagekaldes, såfremt tilbagekaldelsen kommer frem til tilbudsmodtageren, før han har afsendt et antagende svar.
An offer may be revoked if the revocation reaches the offeree before he has dispatched an acceptance.

Figure 2: Collocations and phrases in which the Danish term *tilbagekalde* occurs

The set of collocations, phrases and example sentences in Figure 2 provides a search result that is easy to take in by users. The relevant collocations etc. are not "hidden" among other data types in full dictionary articles, something that will lead to excessive search related information costs (see Nielsen 2008: 173-174). Moreover, the dictionary offers quick access to collocations and phrases, because users do not have to guess in which of the possible articles the collocation or phrase is located.

6. Concluding remarks

Law dictionaries are important aids in communicative and cognitive usage situations. By regarding online dictionaries as information tools in a functional theoretical framework, lexicographers have a sound basis for designing online law dictionaries. Dictionaries that provide help to translate from Danish into English should contain data that match the factual and linguistic user competences, user needs and usage situations and should therefore include data about Danish, UK English, US English and international legal terms, their grammatical properties, and their potential for being combined with other words in collocations, phrases and sentences. Data items that deal with these aspects are necessary for Danish users translate subject-specific and register-specific texts into a foreign language.

The theoretical aspects discussed and supported by prototypical examples from the planned online dictionaries illustrate the importance of having a sound theoretical basis when making practical information tools. The dictionaries allow users to search the database containing carefully selected lexicographic data depending on what type of help users are looking for, and the search results provide users with different data depending on the usage situation.

7. References

BERGENHOLTZ, Henning and Sandro NIELSEN (2006): «Subject−field Components as Integrated Parts of LSP Dictionaries», *Terminology 12(2)*, pp. 281-303.

NIELSEN, Sandro (1994): *The Bilingual LSP Dictionary. Principles and Practice for Legal language*. Tübingen: Narr.

NIELSEN, Sandro (2000): «Translation Strategies for Culture-Specific Textual Conventions in Bilingual Dictionaries», *Lexicographica 16*, pp. 152-168.

NIELSEN, Sandro (2001): «Bilingual Legal Lexicography - A New Theoretical Basis», *Kalbotyra/Linguistics 50(3)*, pp. 85-92.

NIELSEN, Sandro (2008): «The Effect of Lexicographical Information Costs on Dictionary Making and Use», *Lexikos* 18, pp. 170-189.

NIELSEN, Sandro (2010): «Specialised Translation Dictionaries for Learners», in Pedro A. Fuertes-Olivera, ed., *Specialised Dictionaries for Learners*. Berlin: de Gruyter, pp. 69–82.

NIELSEN, Sandro (2014): «Database of Legal Terms for Communicative and Knowledge Information Tools», in Máirtín Mac Aodha, ed., *Legal Lexicography. A Comparative Perspective*. Farnham: Ashgate Publishing, pp. 153-175.