

Going out in style? *Shall* in EU legal English

Richard Foley, Researcher
University of Lapland
Rovaniemi, Finland

1. Introduction

The ambiguity of *shall* vis-à-vis *may* and *must* and its excessive and inconsistent use, or “traditional promiscuity” (Garner 1998:940), in legal English, have attracted the attention of draftspersons and reformers of legal language in the United States, Canada, the United Kingdom and Australia. The evidence brought to bear in arguments for reform in these jurisdictions typically relies on court decisions in disputes over the meaning of the word and the intuitive perceptions of usage held by the reformer. The present study describes research which expands the debate on *shall* to a new context – the legal language of the European Union – and in applying the systematic usage-based techniques of corpus linguistics (Biber 1996:172) to the issue introduces a comparatively innovative methodology as well.

The problem of *shall* can be attributed in considerable measure to tradition. This is seen as a reluctance to depart from tried and true formulae, which is described by Mellinkoff (1963:294) as a fear of change: “Lurking in the dark background is the always present, rarely voiced lawyer’s fear of what will happen if he is not “precise” in the way that the law has always been “precise.” Adherence to tokens of legalese such as *shall* not only sustains the myth of precision in legal language but also perpetuates a style and language that differentiates the genre from that of other professions (Bhatia 1993:101-2) and, by extension, general usage. The tradition of precision instils in the legal profession a prescriptivist orientation to language, exercised both consciously and subconsciously on new writers, e.g., draftspersons and practicing lawyers. While the legal profession seeks to contain complexity in response to and anticipation of litigation arising from ambiguity, the linguist in contrast, is - and must be - content to describe language practice, however complex, through a principled search for regularities in a representative corpus.

As a new legal order in which language is uniquely unfettered by tradition and, at the same time, the object of a singularly well-resourced attempt at language engineering, the EU provides an interesting crucible for language reform and descriptivist-prescriptivist perspectives. For example, a semantic analysis of *shall* and the modals in general is essential for machine translation initiatives (Svendsen 1991) and an understanding of the use of *shall* will figure prominently both stylistically and semantically in implementing the policy of transparency, i.e., making the legal instruments of the EU more readable for the average citizen. The paper undertakes to inform these concerns by addressing the following questions:

- 1) Have the problems of ambiguity identified in the literature for *shall* in common-law jurisdictions “come over” from the United Kingdom and continued in EU legal English?
- 2) Is there evidence of the use of promiscuous *shall* in EU legislation, e.g., use as a mere stylistic marker, as has also been suggested in the literature concerning other jurisdictions? In other words, is one of the ‘senses’ of *shall* null?
- 3) What is the frequency of *shall* in EU legislative language compared to that in other English-speaking jurisdictions? What is its frequency vis-à-vis general usage? To what extent does *shall* serve as a sign making legal language exclusionary of the average citizen?

The paper will proceed in the next section with a discussion of the problem of *shall* in English-speaking jurisdictions and a few of the solutions proposed. This serves to contrast the perspectives of the linguist and the lawyer and provides necessary background for the analyses. The third section describes the materials and methods used in the study. This is followed by a presentation of the analyses and a concluding section, which discusses the findings with a view to informing future research.

2. The problem of *shall*

The lawyer and linguist both strive to understand the meaning of *shall*, but all that they share in this pursuit is the token on the page. To the lawyer, *shall* is a word of authority - a word conferring rights and obligations and prohibitions - whose function is to impose an obligation (Thornton 1979:86-7) and to do so unambiguously, in keeping with the Golden Rule of drafting:

"[T]he competent draftsman makes sure that each recurring word or term has been used consistently. He carefully avoids using the same word or term in more than one sense....In brief, he always expresses the same idea in the same way and always expresses different ideas differently." Dickerson *The Fundamentals of Legal Drafting* §2.3.1, at 15-16 (2d ed. 1986) cited in Garner (1998:940).

Drafting guidelines such as the above that provide the lawyer with prescribed usage; usage deviating from this norm -- the source of the problem -- takes the form of entries in dictionaries of legal English and court cases. For example, the ambiguity of *shall* is attested in the following entry in Black's Law Dictionary, which after a paragraph substantiating obligatory *shall* continues:

But [shall] may be construed as merely permissive or directory (as equivalent to may), to carry out the legislative intention in cases where no right or benefit to any one depends on its being taken in the imperative sense. *Wisdom v. Board of Sup'rs of Polk County*, 236 Iowa 669, 19 N.W. 2d 602, 607, 608 (Black 1990:1375)

More extensive evidence can be had through an investigation of cases involving disputes over the meaning of *shall*, a source suggested in the entry above. One such study is that by Kimble (1992), who after reviewing over one hundred cases involving disputes over the meaning of *shall* concluded: "In summary, I'm afraid that *shall* has lost its modal meaning - for drafters and for courts. Drafters use it mindlessly. Courts read it any which way" (p.72).

To the linguist, *shall* is a modal auxiliary, a lexico-semantic category studied in great detail (e.g., Leech (1971), Coates (1983), Perkins (1983), Palmer (1986)) to describe and account for the variety of meanings it exhibits. With the establishment of the computer corpus linguistics paradigm (Leech 1992:107-111), the evidence invoked in such investigations will typically be derived from a systematic empirical investigation of a corpus or corpora, although researchers may still rely on their own intuition regarding language usage. The dimensions of the evidence used in the two professions are depicted below in terms of two continua: anecdotal -systematic and usage – intuition.

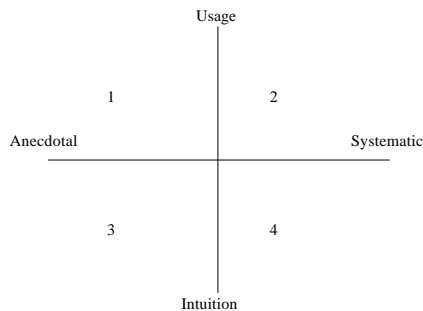


Figure 1 Dimensions of linguistic evidence

The entry in Black's Law Dictionary for *shall* could be placed in quadrant 1. It cites cases in which courts have been called on to decide whether *shall* means obligation or permission - usage –but such cases are not systematic evidence for the study of *shall*. Kimble's (1992) study can be placed somewhat more to the systematic edge of the quadrant, although a collection of cases clearly cannot represent the typical range of use. Quadrant 2 represents corpus linguistics research. Noteworthy examples relevant in the present context are the studies of Coates (1983) and Trosborg (1997). Quadrant 3 describes the approach of many studies of linguistic competence in the 1970s, in particular the introspective investigations of syntax that figured prominently in the development of transformational grammar. Quadrant 4 comprises systematic studies using native informants, which frequently complemented introspective approach.

A second crucial difference between the lawyer and linguist is that they have a different conception of meaning. To be sure, both may seek to distinguish dictionary senses of words, but linguistic studies of modality (Coates 1983) have yielded acceptable models of meaning in terms of clines, or gradients of senses running between core and peripheral meanings, which are doubtless at variance with the categorical approach seen in the drafting guideline.

In the main, solutions to the problem of *shall* have sought to restrict the word to a single sense, its original sense of obligation. Garner (1998:940) distinguishes eight senses of *shall*:

- (1a) "the court . . . *shall* enter an order for the relief prayed for. . . ."
- (1b) "Service *shall* be made on the parties"
- (1c) "The debtor *shall* be brought forthwith before the court that issued the order."
- (1d) "Such time *shall* not be further extended except for cause shown"
- (1e) "Objections to the proposed modification *shall* be filed and served on the debtor."
- (1f) "The sender *shall* have fully complied with the requirement to send notice when the sender obtains electronic confirmation."
- (1g) "The secretary *shall* be reimbursed for all expenses."
- (1h) "Any person bringing a malpractice claim *shall*, within 15 days after the date of filing the action, file a request for mediation."

Only the first of these is deemed acceptable, the criterion being that the *grammatical* subject must be person on whom the obligation has been imposed. Garner objects to *shall* in (1b) on the grounds that a duty is being imposed on an abstract thing, e.g., service, and to (1c) because a duty is being imposed on an unnamed actor; the agent is not specified. The lack of specification made possible by the agentless, short passive in English is not, however, in any way, a consequence of the author's choice of *shall*. The choice of passive voice is a syntactic, not lexico-semantic, consideration. In fact, agents are readily identifiable: both verbs, 'serve notice on' and 'bring before the court' imply action on the part of a bailiff at the request of the court. In each case a duty is being imposed and it is possible to determine the nature of that duty, the person or persons who are to perform it and the object of that performance. *Shall* tells the reader in each case that a duty is involved. The approach confounds syntactic and semantic criteria and yields three senses where there is only one.

While the foregoing will suffice to illustrate Garner's argumentation, (1d) merits comment as a case that is undoubtedly of the kind which prompted Bowers (1989:294) to conclude that *shall* is generally "used as a kind of totem, to conjure up some flavour of the law." The function of the verb in (1d) form is to indicate completed action (establishment of compliance) in future time (when the sender obtains electronic confirmation), and no interpretation of obligation is possible.

On the evidence of intuitive and empirical linguistic studies, respectively, Bowers (1989:34) and Trosborg (1997:136) propose that *shall* be restricted to indicating obligation where a human agent is specified or easily recoverable from context. This argument rests on the concept of a master speech act (Kurzon 1986:19), according to which the enactment clause in a legislative instrument establishes a global illocutionary force of obligation. Used in a non-agentive context in the enactment clause, *shall* actually undermines the purpose of the legislation in two respects: first, as the law is always speaking, if *shall* is interpreted as a future tense it creates perpetual futurity, meaning the provision will never come into force; second, "all cases where *shall* is propositional and non-agentive in fact weaken the superordinate force of the Act by suggesting that there is yet a further step to be taken before the enacted clause becomes reality" (Bowers 1989:34).

Language engineering, in particular machine translation, also can be seen to embrace a similar interest in isolating a single sense for *shall* and the modals. The following is an analysis of modality that has been presented as Euroversal by Svendsen (1991:276).

	Possible	Necessary	"others"
Epistemic	possibility	Necessity	presumption
Deontic	permission	compulsion	obligation
Subject-oriented	ability	Volition	resolution

Table 1 Eurotra matrix for English modal meanings

The senses distinguished encompass the meanings to be expressed in the official languages of the EU and presumably one form and one form only should be assigned to each for each language. *Shall* certainly would be at home in “deontic obligation.” This would amount to termification of the modals using Euroversal concepts and an agreed form – a term – in each language.

In the case of the modals, however, describing and prescribing a single sense is perhaps more daunting a task than in the case of legal terms of art proper. Significantly, the importance of context in determining meaning has been recognized by lawyers and linguists alike. For example, Asprey (1992:82) states on the question of replacing *shall* with *must*: “the reason why it is difficult to replace *shall* with a word that has all these subtle meanings is that *shall* never did it in the first place. Not on its own. It did in context.” The problems with specifying the meaning(s) of words of authority have a linguistic basis in that, as a closed system, their meanings are “reciprocally defining: it is less easy to state the meaning of any individual item that to define it in relation to the rest of the system” (Quirk et al 1972:46). Ideally, methods of investigation and evidence should accommodate both conceptions of meaning.

Another proposed solution to the problem of *shall* is that it be replaced by *must*. This would accomplish little semantically, as the entry in Black’s indicates: “but this [mandatory] meaning of the word is not the only one, and it is often used in a merely directory sense, and consequently is a synonym for the word “may” Black 1990:1019). The meaning of *must* will clearly have to be contained as well before it meets the needs of the legislator. Interestingly for the present study, Asprey (1992:77) suggests replacing *shall* with *must* in the sense of obligation on stylistic grounds, asserting that the use of *shall* “puts lawyers out of step with the language of the general community”. This proposal will be taken up in conjunction with frequency analyses below.

3. Material and methods

The present study relies on four principal sources of data comprising a variety of corpora. The first is a corpus of EU primary and secondary legislation (EULEG) compiled by the author for research on modality in EU legal English. Table 2 gives a breakdown of the current composition of the corpus.

Text type	No. of texts	Word count	Avg. word count per text
Treaty	1	47102	47102
Regulations	4	39863	9966
Directives	4	55360	13840
Decisions	2	16488	8244
Total	11	158813	14437

Table 2 Composition of EULEG

The texts were obtained in electronic form from EUR-LEX the WWW-based database of European legislation. Although the texts are not official translations, the likelihood that there are discrepancies in the use of the modal verbs vis-à-vis the authoritative version as published in the Official Journal was considered negligible. The ready availability of text in electronic form in a searchable database was seen as outweighing this criterion of authoritativeness.

Legislation in the EU can be divided into primary and secondary. Primary legislation comprises the Treaties. There are three treaties establishing the European Communities and a number of conventions and documents by which these have been amended. Complementing the Treaties are the instruments of secondary legislation, which comprise Regulations, Directives and Decisions. Regulations are directly applicable in that no national measures are required for a Regulation to become binding on the citizens of a Member State and that a Member State cannot undertake measures that would prevent the application of a Regulation. A directive is binding as to the result to be achieved; the form and methods by which these are implemented may be decided by the authorities in each Member State. Decisions are generally of an administrative nature and implement other Community rules. A decision is binding in its entirety on those to whom it is addressed

The text types chosen represent the genre of primary and secondary legislation exhaustively, whereby they should adequately represent the linguistic distribution of modal verbs. In light of Biber’s

(1993:252) criteria for representativeness, the figures in the table suggest that a larger number of samples, perhaps smaller in size, might be equally or more representative than those presently included. However, at this stage in the research, the author has opted to include the entire text, even in the case of the Treaty, for a number of reasons. First, stratified sampling without regard to the division of the texts into recitals, enactment section and annexes would have made it impossible to investigate the frequency of *shall* in the enactment section, which is crucial in light of the proposals by Bowers and Trosborg referred to above. Second, the entire text provides the most representative sample for intertextual comparison with corpora from other English-language jurisdictions. While the genre of legislation is in the main similar, no consistent structure could be observed which would have justified stratified sampling by text section across the jurisdictions.

The texts in the corpus have not been chosen entirely at random, and EULEG can be described as opportunistic (Leech 1991:10) inasmuch as several of the texts had been gathered for teaching purposes and terminological research. On the whole, however, the texts do not represent any particular field of activity or time frame; a study of the frequency of modals along such parameters is beyond the scope of the present research. The regulations span economics, social security, the directives data protection and environmental protection and the decisions trade in bananas. Given the nature of the linguistic feature - modal verbs - it is unlikely that any bias could be introduced by the subject matter. There is no evidence that one field of activity imposes more obligations and prohibitions or confers more rights than another.

The second major source of material comprises the translations of the instruments in EULEG in Swedish, Finnish, French and German. These furnish parallel corpora which are instrumental in disambiguating the occurrences of *shall*. For example, the archetypical sense of ‘obligation’ (Crystal and Davy 1969:206-207) stands out in bold relief where one finds it translated with the Agent_{gen}+ BE + present passive participle construction in Finnish. French and German, in turn, serve to reveal occurrences of *shall* in which futurity but not obligation is intended and in which the use of *shall* instead of *will* in the third person is unmotivated.

The use of the translations is predicated on the assumption that translators deverbilize the original message in the source text - whatever language this may be in - and expresses it in their native language. This is depicted as the “language-free semantic representation” in Figure 2 below:

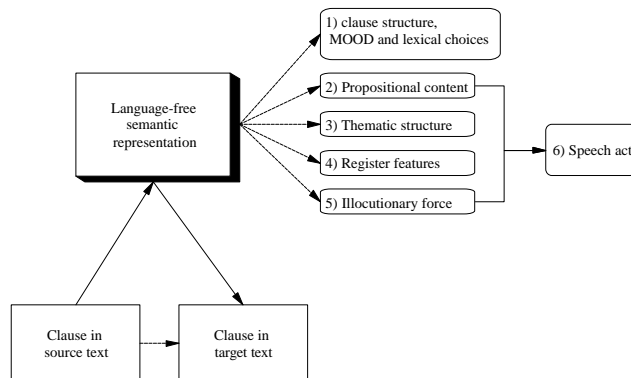


Figure 2 The Process of Translation (adapted from Bell 1991:56)

The thesis here is that, through the agency of the translator, the text of a piece of legislation in each of the official languages of the EU can be seen as encoding the same meaning. Due to lexical, syntactic and semantic differences among the languages, however, these features will be expressed in different ways. A serious qualification on the potential of deverbilization is in order, however, for translators often rely excessively on the surface structure. Evidence for translationese, represented by the dashed line between “Clause in source text” and “clause in target text” in Figure 2, has been substantiated by Schmied and Schäffler (1996:48) in translation corpora of English and Norwegian. Trosborg (1997: 159) acknowledges the phenomenon with regard to *shall* in legal language. Beeth and Fraser (1999:76) identify institutional pressures for perpetuating inadequate translations:

One of the dangers of many of the translating tools like the Translator’s Workbench (TWB) is that they provide the translator with ready-made segments of text in the target language (lifted from earlier documents), making it much easier to stay on the surface of a document. And yet in our hearts we know that what was an adequate translation for the document from which the segment originated is unlikely to be as adequate for the document we have before us now.

A further complication in using parallel corpora is that it is not possible to establish the source and target text: as all are equally authentic and ‘source’ would imply an untoward political primacy. Article 53 of the Treaty on European Union reflects this principle: “This Treaty, drawn up in a single original in the Danish...Spanish languages, the texts in each of these languages being equally authentic...” The organizational reality mirrors this political reality in that it is nowhere stated for a particular piece of legislation which language it has been first drafted in and therefore which has served as the basis for the translations found in EUR-LEX. On balance, while one must be somewhat circumspect regarding the level of abstraction at which translators operate, the parallel texts represent a significant resource to tap systematic usage.

The following examples from Regulation 95/48/EC illustrate the use of translations:

- (2a) en whereas the name given to the European currency **shall** be the 'euro'; whereas the euro as the currency of the participating Member States **shall** be divided into one hundred sub-units with the name 'cent';
- (2b) fr que le nom de la monnaie européenne **sera** «euro»; que l'euro, qui sera la monnaie des États membres participants, **sera** divisé en cent subdivisions appelées «cent»;

The corresponding Swedish text uses *skall*, a cognate of *shall*, which expresses futurity and obligation, as well as *kommer att*, a pure future.

- (2c) sv Namnet på den europeiska valutan **skall** vara "euro". Som valuta för de deltagande medlemsstaterna **kommer euron att** delas upp i ett hundra underenheter med namnet cent.

The English word ‘given’ allows an agent (‘by the participating Member States’), making obligation and thus *shall* possible. The French translation using the future tense suggests a strictly temporal use. Swedish *skall* indicates both obligation and futurity. In the second instance, the use of the future *kommer att*, coupled with the evidence from French, suggests that at least the second *shall* in (4a) is superfluous and that *is* or *will* would suffice. This is a reasonable interpretation also inasmuch as there is no recoverable agent; ‘shall be divided’ is the equivalent of ‘equals’.

The following is an example in which the use of the Finnish construction BE + Present passive participle, which indicates obligation unambiguously, argues in favor of recovering an agent and therefore accepting the use of **shall**.

- (3a) en The amount **shall** be credited to the account of the creditor in the denomination of his account, with any conversion being effected at the conversion rates.
- (3b) fi määrä **on hyvitettävä** velkojan tilille hänen tilinsä valuuttayksikön määräisenä
[the-amount is to-be-credited of-the-creditor to-the-account...]

As the third question in the introduction indicates, the present study seeks to investigate the distinctive features of the genre of legislative English and examine the characteristics of EULEG in comparison with comparable legislation from other jurisdictions. On the intuitive level, *shall* certainly seems to meet the criteria of frequency, distinctiveness and precision described by Crystal and Davy (1991:228-9). Following is a description of the several pilot corpora or representative texts that have been compiled to enable comparisons of EULEG with for the stylistic analysis.

Corpus	No. of texts	Word count	Avg word count per text
AMLEG	1	5365	5365
CANLEG	2	35213	17607
BRITLEG	1	6592	6592
FINNLEG	6	59316	9886
Tot	9	-	-

Table 3 Pilot corpora

Clearly, great caution is called for in undertaking an analysis on the basis of a single text. *Shall* is a frequent enough linear feature (Biber 1993:252), however, to sustain a tentative assumption that even a single text may have some value as a sample. At the level of delicacy sought, they do have something to offer. As the author is interested in translationese, a corpus of translations from Finnish into English has also been compiled.

The fourth principal source of data comprises the frequency of the occurrence of modals in the Brown and LOB corpora. The former comprises 1,013,737 words of American, the latter 1,013,644 words of British English, and they are used here to represent the use of modals in the standard language.

4. Analysis

In her study of *shall* in British statutes, Trosborg (1997) concluded that even though *shall* has been defined as a modal verb expressing legal obligation, “...65.1% of the observed instances of *shall* occurred with non-human subjects which could not be given orders or assigned obligations” (pp. 105-106). This semantic analysis is essential in overcoming a monolithic concept of or belief in a one-to-one correspondence of form and meaning in legal language. This is reflected in Crystal and Davy’s statement: “*Shall* is invariably used to express what is to be the obligatory consequence of a legal decision, and not simply as a marker of future tense, which is its main function in other varieties (1969:206-7) and can be inferred in studies focusing on general usage, e.g., Coates (1983), which simply cite the legal or quasi-legal meaning of *shall*.”

An examination of the frequency of *shall* in the corpus by section of text is depicted in Figure 3 below.

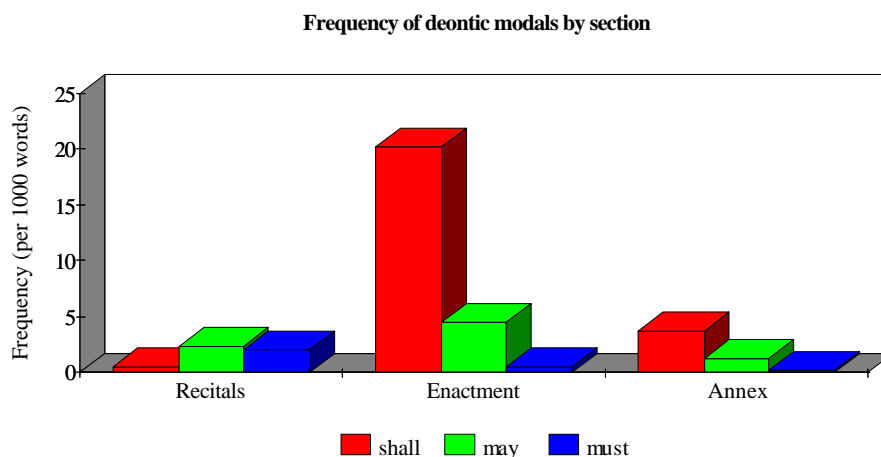


Figure 3 Frequency of deontic modals by text section

This frequency analysis of *shall* prompted examination of word frequencies overall in the corpus and revealed that the relative frequency of *shall* in EULEG places it squarely among semantically void function words. The relevant frequencies are summarized in the table below:

Word	Frequency (per 1000 words)	
	Corpus	Enactment
The	90	100
Of	55	58
To	29	30
In	26	26
And	22	23
shall	15	20
a	14	15

Table 4 Comparison of most frequent tokens in EULEG by section of text

This evidence suggests that the occurrences of *shall* in the corpus merit closer analysis along the lines of Trosborg’s study. It seems implausible that each occurrence of *shall* entails the imposing of an obligation. To investigate this hypothesis empirically, a random sample of 574 occurrences of *shall* was drawn from the enactment sections of the legislation in EULEG and analyzed interactively for the occurrence of a human subject, and where the sentence was in the passive and had a human agent, whether the agent was expressed in a *by* clause or had to be recovered from context. The percentages obtained are presented in Table 5 below:

	Sentence type		
	Active	Passive	
Subject		agentless	<i>by</i> + Agent
Human	40%	12%	3%
Non-human	39%	6 %	-

Table 5 Use of *shall* in EULEG with human subject

If one adopts relatively strict criteria for agency, e.g., that the logical and grammatical subject coincide and that the subject be human, 60% of the occurrences of *shall* analyzed are unmotivated, indicating that legislation in the EU is only slightly more fastidious in this regard than that analyzed by Trosborg. If the passives are included as agents recoverable from context, then some 45% of the occurrences of *shall* in this data are unmotivated, an improvement over British statutes but nevertheless a level of promiscuous use that is certainly in no way reconcilable with the Golden Rule of drafting cited above.

Perhaps the more significant observation to emerge in the analysis is that the criterion of human subject, which Bowers and Trosborg and, to an extent, Garner argue for, proves as difficult to specify as discrete senses of *shall*. The following are examples from the sample.

(4a) It shall be for the controller to ensure that paragraph 1 is complied with

(4a) might not fit the “human subject” criterion, especially if the text were being analyzed by a machine, but is clearly equivalent to the “controller shall ensure...”

(4b) below represents the reverse case. This main clause would presumably meet the criterion of human subject, but the wider context “if it sees fit” renders the issue of obligation irrelevant in giving the authority discretion.

(4b) If it sees fit, the authority shall seek the views of data subjects or their representative

Here, it is instructive to consult the parallel corpora. The Finnish version uses *voi* ‘can/may; Swedish uses *skall*, and German and French the present tense. The Swedish may be influenced by the English text. Obligation is not plausible.

(4c) fi Jos viranomainen katsoo aiheelliseksi, se voi hankkia [If the authority sees fit, it can/may obtain the views]
(sv skall de holt ein, fr recueille.)

The four examples that follow are also potentially problematic, and must be accounted for in any attempt to assign *shall* to a particular sense. It is unclear whether Trosborg would have included any of these as human subjects (1997:105-106). The first three were counted in the present analysis; (4g) was not. Yet the actions involved - publishing, calculating, taking into account and creating a state in which requirements are provided in written form – all require a human agent.

(4d) The report shall be made public

(4e) The turnover of an undertaking concerned within the meaning of Article 1 shall be calculated by adding together

(4f) The decision to publish shall take due account of the legitimate interest of undertakings

(4g) the requirements shall be in writing or in another equivalent form.

The analysis to follow addresses the third question in the introduction by examining the frequency of *shall* and modal in EU legislative language compared to that in other English-speaking jurisdictions and in general usage. The focal question is the extent to which modality in general and

shall in particular as the most frequently occurring modal contribute to making legal language a distinctive genre, one perhaps exclusionary of the average citizen.

Figure 4 below depicts the frequencies of modal auxiliaries in legal and general usage:

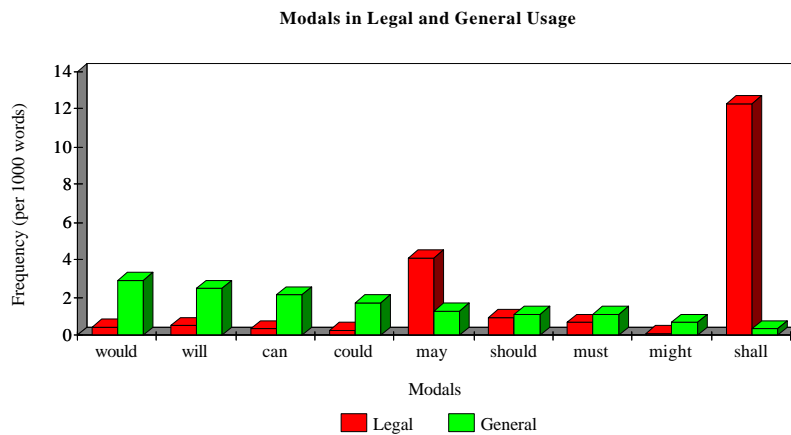


Figure 4 Frequency of modal auxiliaries in legal and general usage

General usage is represented by frequencies from the LOB and Brown corpora as described in the previous section. Legal usage is based on 200,000 words comprising EULEG, AMLEG, CANLEG and BRITLEG. Spearman’s rank correlation coefficient for the differences between the nine auxiliaries was calculated and found to be significant at the 99% level.

The frequency of *shall* here proves to be distinctive as a stylistic feature in terms of both criteria presented by Crystal and Davy (1969:21): that is, it occurs most frequently within the variety of language in question and is less shared by other varieties. Indeed, a nearly clean break can be established between the general language and legal language. In Coates’ (1983:189-9) analysis, *shall* in the sense of obligation is “virtually restricted to formal legal contexts. Its fossilisation is demonstrated by the fact that there are no examples in the informal spoken language of the Survey and only one in the more colloquial written language of the Lancaster fiction texts.”

The following analysis compares the individual corpora to one another and each to general usage.

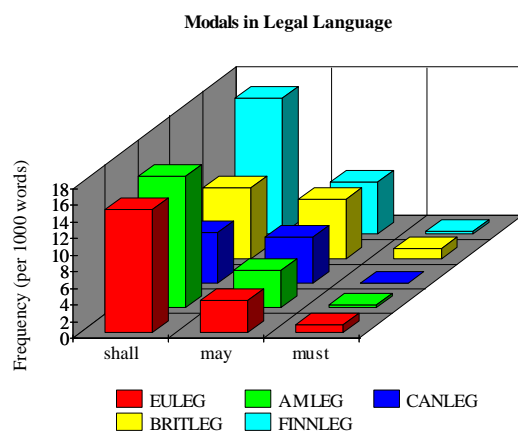


Figure 5 Deontic modals in legal corpora

Chi-squared was calculated for the frequencies of modals in each corpus vis-à-vis the Brown and LOB corpora and found to be significant. Without *shall*, however, none of the texts or corpora differ significantly from the general language. The analysis places EU legislation well within the family of English-speaking jurisdictions. The high frequency of *shall* in the translation corpus FINNLEG merits further investigation. A first observation is that the translators have learned their legal English well.

5. Discussion

The foregoing analyses have shed light on the questions presented in the introduction. From the standpoint of what might be termed legal hygiene, the research to date has revealed no ambiguities which might compromise a citizen's rights. This finding suggests the need for analysis at a greater level of delicacy or drawing more extensively on the parallel corpora for interpretations. With regard to linguistic hygiene, or the promiscuous use of *shall*, the analysis revealed superfluous use on the order of fifty percent. More significantly, the analysis revealed difficulties with the proposed solution of restricting *shall* to meaning obligation when imposed on a human subject. While there is much work yet to be done in building EULEG, this finding can be seen as an outcome of corpus linguistics abiding interest in systematic empirical investigation. The lawyer must accept what the linguist has come to realize: usage is complex and must be investigated accordingly. Yet, the descriptive analysis raises a salient prescriptive issue: Might stylistic use of *shall*, perhaps a habit perpetuating the exclusionary nature of the genre, introduce ambiguities of obligation into the text for the (monolingual) English reader where, at least on the evidence of other languages, none was intended? In light of the analyses presented here, the modals in general and *shall* in particular pose significant challenges to machine translation. The need for language engineering is patent, the prospects for success somewhat less obvious.

Descriptive analysis of the frequency of modals in EULEG and comparable texts or corpora established that legislative language in the EU has more in common with legal genres in other English jurisdictions than with the general language. Although hardly surprising as such, legislative language merits further empirical study given the European Union's express interest in transparency, which is comparable to the Plain Language Movement in the English-speaking countries. Worth noting in this regard is that it is the frequency of *shall* alone, rather than that of *may* or *must*, that binds these populations together statistically.

Obsolescent in the sense of obligation in general language, *shall* clearly signals to the reader that he or she is dealing with a distinctive genre. Replacing *shall* with *must*, as mentioned earlier, would be a problematic enterprise semantically. Even if the number of justifiable uses of a modal of obligation were cut by half, the frequency of familiar *must* would rise to the point where the word would differ distinctively from that in the general language and, in this respect, risk becoming a new *shall*.

References

- Asprey M 1992 Shall Must Go. *Scribes Journal of Legal Writing* 3(79): 79-83.
- Beeth H, Fraser B 1999 The hidden life of translators. *Translation and Terminology* 2: 76-96
- Bell R 1991 *Translation and Translating*. London, Longman
- Bhatia V 1993 *Analysing Genre: Language Use in Professional Settings*. London, Longman
- Biber D 1996 Investigating Language use through corpus-based analyses of association patterns. *International Journal of Corpus Linguistics*, 1(2): 171-197.
- Biber D 1993 Representativeness in Corpus Design. *Literary and Linguistic Computing* 8(4): 243-257.
- Bowers F 1989 *Linguistic Aspects of Legislative Expression*. Vancouver, University of British Columbia Press.
- Coates J 1983 *The Semantics of the Modal Auxiliaries*. London & Canberra, Croom Helm.
- Crystal D, Davy D 1969 *Investigating English Style*. London, Longman.
- Crystal D 1991 "Stylistic Profiling" In Aijmer K, Altenberg B (eds) *English Corpus Linguistics*. London and New York, Longman, pp. 121-238.
- Garner B 1998 *A Dictionary of Modern Legal Usage*. Oxford, OUP
- Kimble J 1992 The Many Misuses of "Shall". *Scribes Journal of Legal Writing* 3(61): 65-75
- Kurzon D 1986 *It is Hereby Performed...Explorations in Legal Speech Acts*. Amsterdam, John Benjamin Publishing.
- Leech G 1971 *Meaning and the English Verb*. London, Longman.
- Leech G 1991 The state of the art in corpus linguistics. In Aijmer K, Altenberg B (eds) *English Corpus Linguistics*. London and New York, Longman, pp. 8-29.

- Leech G 1992 Corpora and theories of linguistic performance. In Svartvik J (ed) *Directions in Corpus Linguistics*. Berlin, Mouton de Gruyter, pp. 105-122.
- Mellinkoff D 1963 *The Language of the Law*. Boston & Toronto, Little, Brown and Company.
- Palmer F 1986 *Mood and Modality*. Cambridge, CUP.
- Perkins, M 1983 *Modal Expressions in English*. London, Frances Pinter.
- Quirk R, Greenbaum S, Leech G, Svartvik J 1972 *A Grammar of Contemporary English*. London, Longman.
- Schmied J, Schäffler H 1996 Approaching translationese through parallel and translation corpora. In Percy, C, Meyer C, and Lancashire I (eds) *Synchronic corpus linguistics*. Amsterdam, Rodopi, pp. 41-55.
- Svendsen U 1991 On the translation of modality in an MT-System Part I. *Le Langage et l'Homme* 25(4): 273-280.
- Thornton G 1979 *Legislative Drafting*. London, Butterworths
- Trosborg A 1997 *Rhetorical Strategies in Legal Language*. Tübingen, Gunter Narr Verlag