



The Secret to Legal Foretelling: Generic and Inter-Generic Aspects of Vagueness in Contracts, Patents and Regulations

JAN ENGBERG *
Aarhus University
ISMAEL ARINAS PELLÓN
Universidad Politécnica de Madrid

Received: 15 February 2010 / Accepted: 8 May 2010

ABSTRACT

In this genre analysis research paper, we compare U.S. patents, contracts, and regulations on technical matters with a focus upon the relation between vagueness and communicative purposes and subpurposes of these three genres. Our main interest is the investigation of intergeneric conventions across the three genres, based on the software analysis of three corpora (one for each genre, 1 million words per corpus). The result of the investigation is that intergeneric conventions are found at the level of types of expressed linguistic vagueness, but that intergeneric conventions at the level of actual formulations are rare. The conclusion is that at this latter level the influence from the situation type underlying the individual genre is more important than the overarching legal character of the genres, when we talk about introducing explicit vagueness in the text.

KEYWORDS:

Genre analysis, corpus linguistics, generic integrity, vagueness, legal discourse, intergeneric conventions, disciplinary genres.

RESUMEN

El propósito de este artículo es comparar mediante herramientas de lingüística de corpus el uso de la vaguedad en contratos, legislación y patentes de los EE.UU. La comparación se centra en valorar hasta qué punto estos tres géneros comparten las mismas estrategias de vaguedad lingüística para lograr sus objetivos y sub-objetivos comunicativos. Los tres corpus compilados para este estudio comparten una combinación, tanto de lenguaje legal, como tecnológico y cada uno de ellos consta de algo más de un millón de tokens. El resultado de nuestra investigación apunta a que estos tres géneros comparten convenciones en el uso deliberado de tipos de vaguedad lingüística, pero la fraseología concreta de la vaguedad aplicada no es normalmente la misma. La conclusión es que, en este último nivel fraseológico, la situación subyacente en cada género es más importante que el común carácter legal de los géneros cuando se trata de introducir la vaguedad de forma explícita en el texto.

PALABRAS CLAVE:

Análisis de género, lingüística de corpus, integridad de género, vaguedad, lenguaje legal, convenciones inter-genericas, géneros legales.

**Address for correspondence:* Jan Engberg. Aarhus University, Fuglesangs Allé 4, Aarhus V, Denmark. Tel: 45 89486688; Fax: 45 86157727; E-mail: je@asb.dk Ismael Arinas. Universidad Politécnica de Madrid. c/ José Abascal, 2, 28006, Madrid. Tel: 34 913363028; ismael.arinas@upm.es

1. INTRODUCTION

Vagueness of expression and indeterminacy of content are widely recognised as a characteristic of the writing of legal texts necessary for these texts to fulfil their communicative and pragmatic functions.¹ The role of vague expressions has been investigated and documented in a number of legal genres, among them also the three genres under scrutiny in this contribution (regulations, contracts and patents from US law). However, research in this field has thus far mainly been either oriented towards a general philosophical discussion of vagueness and its role in legal systems built upon the idea of the Rule of Law, or towards types of linguistic elements applied to indicate explicit vagueness and indeterminacy within the framework of one genre. But if vagueness is a central characteristic of text formulation in the field of law as rule making, it could be interesting to know to what extent specific explicitly vague elements reoccur across the genres. Yon Maley has in a previous study claimed that behind the formulation of statutes there could be something like a legislative communicative competence in the form of “a set of communicative strategies encompassing what a specialised group of individuals in the exercise of their institutional roles know and do to produce appropriate and valid legislation and, when necessary, interpret it”. (Maley, 1987: 26). If this is true, and much work in the field of analysing individual legal genres has shown the idea to be likely to be true, an interesting second step in the investigation of legal genres could be to assume at least a partially overarching communicative competence. For drafters of all legal genres share at least to a large degree a common educational and basic philosophical background from their legal training. And the institution of law is not compartmentalised into different and separated parts containing the different genres. Instead, the same people may often be involved in drafting and interpreting more than just a few of the different genres over a lifelong career. Thus, the assumption of an overarching discourse community in the sense of Swales (1990) may be sustained.

The purpose of the research project, of which this paper reports the first results, is to establish knowledge of some of the conventionalised rhetorical strategies used in three legal language genres. More specifically, we are concentrating on strategies for the application of linguistic vagueness in drafting documents that must deal with future and difficult to predict circumstances and conflicts. We are thus looking at genres with a characteristic pre-emptive character. Such strategies may in principle be characteristic of only one of the genres (generic strategies) or may be found in all three genres (inter-generic strategies). Our assumption in this research project is that, in the case of the investigated genres, drafters of the documents resource to some extent to a common pool of available rhetoric strategies and thus apply the same strategies in all genres. A central goal of the project is to assess the size of such a common pool. Therefore, within the confines and assumptions of the overall research idea, the central goal of this paper is to assess generic and inter-generic vagueness strategies in U.S.

patents, contracts and regulations and the distribution between the two different groups of strategies on the basis of corpus linguistic and genre oriented approaches.

2. VAGUENESS IN LEGAL TEXTS

Concerning the concept of vagueness in texts as investigated in this research, Pinkal's (1981) classification of vagueness, more specifically, his definition of semantic indeterminacy has been the point of departure. He defines an utterance as semantically indeterminate if it is not possible to determine whether the sentence in which the utterance is used is true or false (Pinkal, 1981: 4). Vagueness is one type of semantic indeterminacy². It has been the focus of interest for philosophers, linguists and jurists either from the point of view of its interpretation (Dascal & Wróblewski, 1988; Kennedy, 2002; Russell, 1923; Solan, 2005), or, less frequently, its deliberate production (Barker, 2002; Endicott & Spence, 2005; Prampolini 1998). We are especially interested in the last perspective, i.e., the perspective of explicit vagueness being a common denominator of many genres in the field of law³. The three genres we have chosen to work with here all resource to vagueness because of their pre-emptive nature (cf. above). This means that people drafting these genres are describing circumstances which extend into a future which may well be somewhat different from the moment when the document became valid. Of course, there are differences in the circumstances surrounding each document: patents have a maximum life expectancy of 20 years and must be in essence different from previous ones and virtually future proof; contracts can be revised with the addition of new clauses covering previously unexpected eventualities (Roberts, 2007: viii); and legislation may be amended or revoked and re-enacted in a new corrected form to cover for unseen contingencies.

Despite these differences, Quine (1960) saw vagueness as a necessary rhetoric strategy to meet the communicative requirements of the legal language. Prampolini (1998) justifies vagueness as a drafting strategy by indicating that its purposeful use is one way to profit from the flexibility and plasticity that language has. He states that managing vagueness gives the writer the capacity of referring to concepts within a scale of approximation. Nussbaumer (2005) presents the general idea that the law lives from striking the balance between precision and vagueness, from introducing the amount of indeterminacy necessary for it to be applied to situations that could not be foreseen, but in which the law should sensibly apply. Endicott (2000: 185-192) discusses vagueness in the framework of the idea of the Rule of Law and reaches the conclusion that explicit vagueness is necessary for a legal system to be able to perform its functions. And Anesa (2007: 33) has demonstrated that contracts contain not only elements contributing to precision, but also elements contributing deliberately to introducing indeterminacy into the contractual provisions. So among people studying actual communicative practices in legal communication there is a high degree of accordance around the position that vagueness is strategically well used in order to fulfil the generic purposes of

genres in the field of law. In the empirical part of this paper (section 6) we will look more in detail upon the way vagueness is introduced in the different genres.

3. GENERIC INTEGRITY AND INTERGENERIC SIMILARITIES

3.1. Generic integrity as aspect of genre

In the framework of the purposes of this paper, the concept of generic integrity as developed among others by Vijay Bhatia acquires some centrality. Generic integrity is what is seen as keeping concrete instances of a genre together across the specificity of the situations in which individual texts are written. Bhatia (2004) defines the concept in the following way: “Generic integrity may be understood in terms of a socially constructed typical constellation of form-function correlations representing a specific professional, academic or institutional communicative construct realizing a specific communicative purpose of the genre in question” (Bhatia, 2004: 123). As is visible here, the concept of genre in this tradition is typically perceived as connected to specific communicative purpose combined with individual genres (cf. definition in Swales, 1990: 58). This may be seen as a sociologically oriented approach to genre, focusing upon fairly concrete and specific situations and situational factors. This approach has proven very valuable, not least for the purposes of acquiring text production skills in, for example, a new area of profession.

However, with its focus upon the individual genre, this approach has a tendency to overlook communicative characteristics related more to the overall situational level. It thus does as a tendency concentrate upon characteristics of a genre demonstrating the integrity of this genre, whereas it has some difficulties seeing characteristics that are common of a number of genres. However, the intergeneric level is seen as relevant by propagators of this approach to genre analysis. For example, Bhatia (2004) works with a number of different concepts for groups of genre at a gradually higher level of abstraction, going from *genre sets* via *systems of genre* to *disciplinary genres* and *genre colonies* (Bhatia, 2004: 53-56). An example of a genre colony is the overall group of promotional genres (ibid.: 59), whereas an example of disciplinary genres are genres connected to the discipline of law as performed by the domain-specific group of legal professionals (ibid.: 54). What we are interested in here, are intergeneric conventions connected to explicit vagueness between members of the disciplinary genres of US law.

3.2. Intergeneric characteristics as part of disciplinary communicative competence

As an alternative to the more sociological approach to genres presented before, genres may be seen not merely as a sociological, but as a mental fact, as a part of the disciplinary knowledge bases of a group of professionals⁴. The advantage of taking such an approach is that it is better in accordance with the idea of communicative competences mentioned above. The reason is

that focus is not upon the communicative situation, but instead upon the persons acting in the situation, and upon the different situations, in which these persons act. Consequently, focusing upon intergeneric characteristics like we do in this paper enables us to shift the focus and to model more broadly the communicative competences underlying professional text production. But it also presupposes leaving the idea of the dominating position of the global communicative function of one genre and shifting the focus to instrumental and overarching functions like in our case the application of explicit vagueness. In the following we want to show how such an analysis may be grounded in a pragmatic and communicatively oriented description of the genres involved.

4. INVESTIGATED GENRES: TEXTEXTERNAL CHARACTERISTICS

US-regulations, which is the kind of legislative texts studied in this paper, are in the words of Black's Law Dictionary 9th edition, "(a) rule or order, having legal force, usu(ally) issued by an administrative agency" and thus a sort of secondary legislation. This means that they have the pragmatic power of legislation, but they have not been adopted by the relevant parliamentary body. Instead, the parliamentary body has adopted a statute laying down the framework of the rules to be set up in order to reach a specific regulatory goal and granting the power to some administrative body to sort of fill out the details of the regulatory effort by way of a regulation. The statute adopted by the parliamentary body is called primary legislation, whereas the deducted legislation, the regulations, is called secondary legislation. However, the difference lies mainly in the different generative history of the two types of text. Pragmatically they both perform the same legislative function, i.e., to enact (and convey) a set of rules as rights and obligations (cf. Kurzon, 1986: 9; Trosborg, 1991: 74) to citizens of a country in the form of the communicative subfunctions or speech acts of permission, obligation, prohibition or constitutive rules/definitions (cf. Kurzon, 1986: 15-16; Šarčević, 2000: 138-159; Trosborg, 1991: 77-83). And this is not done towards specific individuals as such. Instead legal rules are directed towards more general classes (of persons, things, actions, and circumstances (Maley, 1987: 30-31)). Official addressee of a regulation is thus the general public (Kurzon, 1986: 27).

4.1. Regulations

These pragmatic characteristics lead to features like generality, impersonality, precision and vagueness to be typical of statutory texts like regulations (Maley, 1987: 40-41). Precision is a typical feature in statutory texts like regulations due to the quest for certainty of law. It is a general belief in statutory writing that certainty is a central prerequisite for justice. At the same time, the aimed generality of the text means that it must be applicable in many different contexts. Hence the typical use of vagueness in order to facilitate this generality (Maley, 1987: 38), or all-inclusiveness (Bhatia, 1993: 102). Vagueness is also sponsored by the fact

that the statutory text is to be seen as “a set of rules in fixed verbal form that are forever speaking” (Maley, 1987: 30), i.e., a kind of perpetual speech act to be applied in an indefinite future. In order to make it likely that a statutory text is efficient in its governing of behavior, relevant vagueness may be of value.

4.2. Contracts

Like the regulations described above, contracts are intended to lay down legal rules in the form of rights and obligations, or in the words of Black’s Law Dictionary 9th edition, “An agreement between two or more parties creating obligations that are enforceable or otherwise recognizable at law <a binding contract>”. The major difference to the statutory texts lies in the fact that the rights and obligations in general are aimed at individual and concrete parties, i.e., the parties setting up and signing the contract, as is visible in the definition. So the characteristic of generality from regulations in the form of a general group of addressees is not found here: Contracts are aimed at regulating a much more limited section of the world - not classes, but individuals.

However, Anesa (2007) shows that practical contract drafting does not consist in writing contracts afresh in every new situation. Instead, expressions that have been used before and have proven efficient are often reapplied in new contracts, either in form of contract forms to be filled out or in the format of fairly stable clauses. Thus, a certain amount of practical applicability of expressions to more contexts (like in the case of statutory texts) is normal also in contracts. Furthermore, the orientation towards the future which we saw before is also a relevant characteristic of contracts. Just like statutory texts, contracts are perpetuated speech acts in the sense that the rules laid down by the text are reinstigated every time they are needed. So by the rules the parties try to regulate their future behaviour towards each other. Consequently, also in contracts we are bound to encounter the balance between precision in order to achieve relevant certainty and vagueness in order to achieve relevant indeterminacy.

From the point of view of the pragmatic functions performed by contracts the overall function is to make an agreement between (usually two) parties concerning the creation of legal relation and the determination of (legal and illegal) actions (Voldgaard Larsen, 2008: 107). Studies have found the same subfunctions as the ones performed in statutory texts, i.e., permission, obligation, prohibition and constitutive rules/definitions, apart from the enacting function that instantiates the authoritative framework around the communicative acts performed in and by the contract (e.g., Blom & Trosborg, 1992). The individuality of the text, the fact that it is connected to a specific situation in which individual parties make agreements, means that another function, viz. the commissive act of promising, is found in contracts (Blom & Trosborg, 1992: 104). So from a functional point of view, statutes and contracts are closely related on the level of the subfunctions, but not identical.

4.3. Patents

According to Black's Law Dictionary 9th edition, a patent is "(t)he right to exclude others from making, using, marketing, selling, offering for sale, or importing an invention for a specified period (20 years from the date of filing), granted by the federal government to the inventor if the device or process is novel, useful, and nonobvious", or the document granting such a right. Thus, it is communicatively an act of granting a specific legal status to an idea (Bazerman, 1997). This status engenders specific rights for the holder of the patent concerning the commercial exploitation of the idea. The status is granted by a relevant state authority (in our case the United States Patent and Trademark Office), in the case of the US patent for the period of 20 year. In this work we will concentrate upon the patent as such and not other documents connected to the patenting process (like the Note of Allowance), as we are interested in the parts where expressions of vagueness are relevant. And this is not the case, e.g., in the Note of Allowance.

The functional role of the patent from a legal point of view is different from that of the contract in that the patent grants its owners/holders the right to exclude everyone from interfering with the patentee's intellectual property within the confines of the statutory arrangement of the relevant country. So it is a right granted by the state to an individual citizen, directed towards anyone. In contrast, contracts are agreements between two or more parties made within the confines of the statutory arrangement of the relevant country and influencing mainly the parties themselves. So where the patent is directed towards anyone, the contract has a more limited personal scope. In this respect, the patent is closer to the statute. However, it is different in that it does not set up different kinds of specific rights and obligations. It only grants a legal right to the patentee and simultaneously creates an obligation by anyone in the society to not infringe this right for the period of 20 years. In the context of this paper, the future oriented character of the patent is important: It is in the best interest of the patentee to try and cover as much ground as possible with the patent and thus extend the patentee's right to possible instantiations of the idea coming up in the period of time covered by the patent, at least to the extent this is allowed by the patent authority. And this may be achieved through the use of explicit vagueness in the text.

From a functional point of view, the pragmatic subfunctions of the patent are different from those of the other two genres treated here, apart from the subfunction of enacting the document as having legal authority, which is performed by a separate document and not treated here. The overall function of granting a specific legal status to an idea is performed by the subfunctions of Claiming (central function) and of Describing (subsidiary function) (cf. Bazerman, 1997; Arinas, 2010). The claim must be supported by different descriptions (of inventor, assignee and attorney, of field and problem and of the invention itself, to mention the main groups).

4.4. The role to be expected of vagueness in the genres

The pragmatic and text external characteristics of the three genres described above indicate that vagueness may be expected to play a role in all three genres:

- *Regulations* are directed towards laying down rights and obligations of the general public, i.e., they have an indefinite receiver group; and they are normally temporally indefinite, i.e., they have to be applicable in indefinite future situations.
- *Contracts* are often written on the basis of contract forms meant to be applied to many individual situations; and they set up rules for the future behaviour of the parties to the agreement, i.e., they intend to regulate and foresee actions and behaviour in the future, which are less indefinite than in the case of regulations, but may still not be defined fully due to lack of prevision (Shippey, 2003: 85).
- *Patents* are intended to protect the patentee's idea against possible infringements by indefinite others in the future (in our case 20 years), i.e., the patent must try to foresee instantiations coming up in the future.

Consequently, on this account we can see pragmatic similarities between the genres. Another similarity they share is that they are documents carrying legal authority, i.e., it is possible to uphold rights and obligations on the basis of all of the genres. On the other hand, from a more detailed functional point of view the genres are fairly different, as is visible through the overall communicative functions as well as through the communicative subfunctions listed above. Highest degree of similarity exists between contracts and regulations, whereas patents are fairly different from the other two genres from a functional point of view. Thus the scrutiny of the genres has confirmed our initial impression of the investigated field: A mixture of pragmatic and textexternal similarities and differences exists, giving rise to a situation in which we are likely to find overarching genre conventions concerning the expression of vagueness in the context of establishing legally binding rights and obligations.

5. METHOD

We have compiled three corpora of approximately one million tokens for this project: a U.S. regulations corpus (taken from the 2010 issue of the regulations), an international contracts corpus (covering a period between 1998 and 2007), and a U.S. patents corpus (selected from between 1999 and 2009). Their size and other descriptive data are shown in table 1 below. To analyse and compare these corpora, we have used the Antconc 3.2.1w freeware developed by Laurence Anthony (<http://www.antlab.sci.waseda.ac.jp/software.html>).

	U.S. PATENTS	CONTRACTS	REGULATIONS
Number of Docs.	135	118	10
Number of Tokens	1,009,146	1,009,192	1,017,013
Dif. Words (types)	14,498	20,788	19,590
Type / Token Ratio	69.60 (1.43%)	48.54 (2.06%)	51.91 (1.92%)
Average Words/Doc.	7475.15	8552.47	101,701.3
Keywords used	199	199	199

Table 1. Data on project corpora.

We have decided to compare our three corpora not on the basis of absolute frequencies of tokens, but on the basis of their keywords. As it has been shown in other work (see above section 4) that vagueness is present as a rhetorical instrument in the genres we are investigating here, we have not been interested in looking deeper into the use of vagueness in each of the investigated genres. Instead, we are interested in the degree, to which formulation conventions connected to the expression of vagueness are conventionalized across the three genres, i.e., we are interested in intergeneric conventionalization. Consequently, keyness is a more decisive feature than mere frequency. Our focus lies thus upon instances of vagueness characteristic enough to involve the top 199 keywords in each of the investigated genres.

Antconc 3.2.1w, among other tools, offers the possibility of extracting keywords from each corpus analysed with it. It offers two possible statistical measures of keyness: the Chi-squared test and the Log-likelihood test. Both tests require a reference corpus for extracting the keywords of the studied corpora. Taking the token frequencies in the reference corpus as a standard, the software calculates the expected frequency for each token of the studied corpus. It then compares the expected frequency against the real frequency found for each token of the studied corpus. If the frequency of a token is higher than the expected frequency, then the software assigns to this token a keyness value. According to Rayson *et al.* (2004: 3) and Rayson and Garside (2000: 2) the Chi-squared test is less reliable than the Log-likelihood test, so we have chosen this second test as the system to extract the keywords in our three corpora. When setting up the keyword lists we have not distinguished between instances of the lexems with or without a capital letter, as we are not interested in the position of the words in the sentences. We have on the other hand distinguished between instances of the lexemes in the plural or the singular, as these may be part of different collocations.

6. RESULTS AND DISCUSSION

6.1. Vague category identifiers

6.1.1. Definition

Channell (1994: 122) defines *vague category identifiers* as expressions which through their prototypicality allow the listener to identify the set to which the expression belongs. Channell's research focused on the use in spoken English of expressions such as *or something, and things like that, or whatever*, etc. to transform nouns into categories. When *translating* her approach to the written universe of the legal genres investigated here, we naturally cannot adopt these expressions from the oral register. Written legal genres, with their much more formal conventions, achieve categorization through the use of semi-technical hyperonyms. In accordance with Channell's empirical test above, candidates for vague category indicators are tested for whether they occur in the corpus in collocations like [... *or other x*] or [*such x*], in contexts where the expression indicates the character of the candidate word as referring to a (superordinate and thus vague) category rather than merely to a single referent:

If the Commission decides, or an affected resource agency requests, the Commission will conduct a meeting, telephone, or video conference, or other procedures to address issues raised by its preliminary determination of inconsistency and comments thereon. (Code of Federal Regulations Title 18; our emphasis).

Such procedures may include either: (a) The scheduling of an open public meeting under Sec. 10.65(b) at which interested persons may participate in review of and comment on the draft document; or ... (Code of Federal Regulations Title 21; our emphasis).

We have thus made a necessary distinction between categories as such and vague categories. The test is seen as categorical. Thus, we have counted all instances of the words as vague category identifiers, not only the instances where the words are used together with *such* or other similar words⁵.

6.1.2. Recurrence of vague category identifiers across the corpora

The following table shows vague category identifiers as they appear in the three investigated corpora:

Regulations	Contracts	Patents
<i>Requirements</i>	Agreement	Member
<i>Information</i>	<i>Equipment</i>	Invention
Request	<i>Services</i>	Embodiment
Statement	Schedule	Portion
Person	Obligations	<i>Device</i>
Procedures	Rights	Assembly
<i>Services</i>	<i>Service</i>	Vehicle
Facility	<i>Information</i>	Mechanism
<i>Equipment</i>	<i>Requirements</i>	Fluid
<i>Service</i>	<i>Product</i>	Actuator
Fee	Event	Housing
Action	Products	<i>System</i>
<i>Product</i>	Assets	Embodiments
<i>Device</i>	Data	Apparatus
Document	Obligation	Sensor
Entity	Expenses	Section
Certificate	Taxes	
Treatment	<i>System</i>	
Facilities	Claim	
Documents		
Container		
Persons		
Form		

Table 2. Vague category identifiers from list of keywords in the three corpora; listed in order of keyness; words appearing in more than one corpus highlighted (italics).

The words are listed here in the order of their keyness factor. Highlighted words appear in at least two corpora. As we focus upon intergeneric regularities in this paper, we will concentrate upon formulation conventions for the highlighted words in the following.

The first result of the survey of vague category identifiers is that such words are fairly numerous among the 199 most characteristic words of the investigated corpora. The lists contain all types of words, and if we calculate the vague category identifiers as a fraction of the nouns in this list, we see the following relations between vague category identifiers and other nouns:

Regulations	Contracts	Patents
23 / 97 = 24 %	19 / 97 = 20 %	16 / 110 = 15 %

Table 3. Proportions of vague category identifiers in the different corpora.

We can see that vague category indicators play a major role among the nouns in all genres, with regulations as the genre with the highest proportion of this type of indicators of vagueness.

6.1.3. Recurring patterns for the reappearing vague category indicators

As the next step, we have investigated to what extent the vague category indicators that appear in the keyword lists of at least two genres are part of recurring patterns of expressions

across the genres. In other words, we investigate to what extent collocations connected to the found words are conventionalised across genres, i.e., are the sign of intergeneric conventions.

In this connection we have applied the cluster analysis function of Antconc, searching for clusters of two to six words for the reoccurring vague category indicators. In order to be able to concentrate upon collocations with a high degree of conventionalisation we are only interested in collocations with a frequency above 25. I.e., a collocation is only interesting for us, if it occurs in at least one of the genres more than 24 times. Working with cluster lengths between two and six words mean that a number of non-relevant clusters surface (like the equipment). We have concentrated upon clusters which we have judged to be meaningful on their own in the texts (see examples below).

Requirements: This vague category indicator co-occurs in the regulations and the contracts corpus.

	Regulations	Contracts
in accordance with the requirements	36	5
comply with the requirements	49	8
compliance with the requirements	33	6

Table 4. Collocations for *requirements*.

The common cluster of these instances is the cluster with the requirements. This occurs in total 163 times in the regulations corpus and 35 times in the contracts corpus. So the found conventions make up 73% of the instances of with the requirements in the regulations corpus and 54% in the contracts corpus. Consequently, a certain degree of conventionalisation may be said to exist in both corpora, but the clusters are relatively not very frequent in the contracts corpus.

Information: This vague category indicator co-occurs in the regulations and the contracts corpus.

	Regulations	Contracts
additional information	112	15
the following information	90	31
such information	84	95
other information	65	55

Table 5. Collocations for *information*.

A small number of overlaps exist which meet the criteria set up. However, these instances are at a fairly general level (short clusters). More technical collocations from the legal register like “Information to which this section applies”, on the other hand, are found in fairly high numbers (88 instances), but only occur in the regulations corpus.

Services: This vague category indicator co-occurs in the regulations and the contracts corpus.

	Regulations	Contracts
services provided	75	96
services provided by	18	33
Human Services	49	40

Table 6. Collocations for services.

In this case, we found an example of the connections between the genres, viz. the collocation *Human Services*. It occurs with a fair frequency in both corpora, because they both refer to the same entity in the world, i.e., state Departments or Secretaries of (Health and) Human Services. This is an instance of terminological accordance between the two genres. The only other relevant collocation (*services provided*) is at a fairly general level, and although the collocation with *by* is frequent enough to meet our quantitative criteria, it may not be said to be very significantly conventionalised, as it covers only about 1/3 of the total instances.

Equipment: This vague category indicator co-occurs in the regulations and the contracts corpus. There are no relevant overlapping clusters between the two corpora.

Service: This vague category indicator co-occurs in the regulations and the contracts corpus.

	Regulations	Contracts
service providers	34	8
service agreement	6	27

Table 7. Collocations for *service*.

Some common clusters are found, but the overlaps are fairly insignificant, and the degree of conventionalisation is fairly weak.

Product: This vague category indicator co-occurs in the regulations and the contracts corpus.

	Regulations	Contracts
work product	2	34

Table 8. Collocations for *product*.

One overlapping cluster was found, but with an insignificant overlap and degree of conventionalisation. Both corpora show specific qualifications of the basic word *product* dependent on the content of texts in the corpus, like *seller product* (contract) or *biological product* (regulations).

Device: This vague category indicator co-occurs in the regulations and the patents corpus. There are no relevant overlapping collocations. It is interesting to note that in the regulations corpus *device* is mainly used in isolation (rank 1: *the device*, rank 2: *a device*; rank 5 *digital device*, rank 14: *tv interface device*), whereas in patents it is primarily qualified and only secondarily used in isolation (rank 1: *control device*, rank 2: *clutch device*; rank 9: *the*

device, rank 13: *a device*). This is an indication of differences in the generic situation behind the texts being more important for conventional formulations than the fact that both genres belong to the legal sphere.

System: This vague category indicator co-occurs in the contracts and the patents corpus.

	Patents	Contracts
system for	189	34
system to	33	41
operating system	3	62
control system	197	15

Table 9. Collocations for *system*.

Overlaps were found in the field of combining *system* with prepositions, but at a fairly general level and with major differences in the degree of conventionalisation: *system for* is relatively much more frequent in patents than in contracts. The categorical difference in use mentioned under *device* is not found here: *the system* is rank 3 in the patents corpus and rank 1 in the contracts corpus. Apart from that, what qualifications to the basic word are used is dependent on the concrete situation of the texts, not on (inter-) generic regularities.

6.1.4. Concluding remarks

By way of conclusion, it was visible from the results presented in 6.1.2 that the use of vague category indicators is common for all three genres as represented in the corpora. But it was also shown that there is little overlap in the chosen vague category indicators across the keyword lists (although more so between regulations and contracts than between any of these and patents). No vague category indicator was found in the keyword lists of all three. Finally, the results of 6.1.3 showed hardly any overlaps in the collocations of the overlapping words. This leads to the conclusion that vague category indicators play an important role in signalling relevant vagueness, but that only very few intergeneric conventions exist at the more concrete level of formulations. In this respect the genre specific aspects of the context play a more important role than the mutual characteristic of the genres of belonging to the legal field.

In his work on vagueness in statutory texts, Endicott (2000) presents indeterminate standards as one of the central ways of indicating legal indeterminacy through linguistic vagueness (Endicott, 2000: 48-49). The characteristic of linguistic means falling into this category is that they set up a standard which is used for characterising something and which is indeterminate in the way that its interpretation is dependent on the context in which it is used. Adjectives and adverbials are the prototypical conveyers of this meaning (Adams 2004: 85-89).

In the keyword lists we find the following instances of linguistic means conveying indeterminate standards:

	Regulations	Contracts	Patents
Applicable	X	X	
appropriate	X		
Preferably			X
Reasonable		X	
Reasonably		X	

Table 10. Indicators of indeterminate standards in corpora.

Due to space limitations, we will give only two examples of the introducers of indeterminate standards:

License Permits. Seller has obtained all material licenses, permits and authorizations required by *applicable* laws or regulations pertaining to the business. (Purchase Contract I; our emphasis).

As best seen in FIG. 7 and FIG. 8, proximity switches 123 are provided to sense when the lug has been displaced a predetermined distance (*preferably* short of the maximum allowable displacement) by reason of the trailer beginning to over articulate. The proximity switch is *preferably* mounted on wall 64 with a sensing wand extending upwards such that when the shoulder between shaft 90 and extension 92 is displaced to the port side, the shoulder will enter into contact with the sensing wand. (United States Patent 6,746,035; our emphasis).

From the point of view of our central question in this paper, it is important that we find the type of vague linguistic elements in the keyword lists of all corpora. It is, however, not a very dominant feature in any of the corpora, at least from the point of view of being represented in the top 199 keywords. At the same time, we hardly find any lexical overlaps across the corpora. So the conclusion is that the type of vagueness is part of the intergeneric conventions, but that there are no significant overlaps and thus no intergeneric conventions in the choice of lexems.

Joan Cutting (2007: 224) includes along her category of vague general nouns borderline cases of verbs that are “relatively empty semantically” but “heavily laden pragmatically”⁶. To this group we would count the modal auxiliary *may*⁷. It is relatively empty semantically in that it gives a broad touch of possibility to the meaning of the main verbal complex of the sentence, but does not determine the exact scope of this possibility. When used in legal texts, *may* introduces vagueness in the sense that it delegates power of interpretation of the meaning of the verbal complex to later interpreters, seen from the point of view of the time of drafting the legal text (Šarčević, 2000, 142-144). This way the text is rendered open to interpretation. Looking at the keyword lists of the three corpora gives us the following data:

	Regulations	Contracts	Patents
<i>may</i> - occurrences	9257	6241	3563
Position in keyword list	Rank 17	Rank 52	÷

Table 11. Occurrences of *may* in corpora.

Interestingly, *may* does not belong to the keywords of the patents corpus, although the modal auxiliary is fairly frequently present. Consequently, this way of introducing vagueness is not specifically frequent enough to be a main characteristic of the wordings of genre. Where *may* belongs to the keywords, it has a fairly prominent position. *Shall* is the only more frequent modal auxiliary (regulations: rank 4; contracts: rank 2).

Looking at the clusters around *may* in regulations and contracts⁸, we have made the following observations:

- In both corpora, the cluster *may be* is the most frequent one (regulations: 2314 tokens; contracts: 1213 tokens); this cluster is predominantly used for passive expressions, i.e., for enabling something to happen without emphasising who the power is given to, as in: ... *or in which a basis grade contract may be made* (Code of Federal Regulations, Title 7; our emphasis).
- *May not* is ranked 2 in regulations (305 tokens) and ranked 9 in contracts (122 tokens); we thus have a similar convention in the two corpora, although with more weight in regulations.
- In both corpora, there is a high number of instances of *actor + may*, like in *the appropriate TTB officer may authorize* (Code of Federal Regulations, Title 27). Depending on the corpus, a different set of possible actors occur. For example, rank 5 in the list of clusters for *may* in the regulations is *officer may*, and rank 2 in the list of clusters of *may* in the contracts is *party may*. There is no overlap between the two list of clusters with a frequency of at least 25 tokens in this respect. This means that we have an instance of a similar pattern, although not at the level of actual choice of words, but more at the level of fairly detailed lexical categories.

Summing up these findings, we find that we have here an example of an intergeneric convention across the genres of regulations and contracts at a more concrete level than was the case with the vague category indicators treated in 6.1. Here, we have the same word used for the same purpose (for introducing vagueness through delegating interpretive power), we have fairly similar frequencies and patterns of formulation, but still differences in choice of words in accordance with the specific situation types underlying the two genres (officers vs. parties as actors). So here we find it relevant to say that the fact of belonging to the legal field is more important than the specific situations in creating the conventional formulations.

7. CONCLUSIONS

The main underlying questions to be investigated in this paper were the following:

- To what extent is it possible to find intergeneric conventions between three legal genres pointing to an underlying legal communicative competence?

- At what level of specificity do we find these intergeneric conventions?

We have not investigated this question in a general manner in this paper. Instead, we concentrated upon one aspect with a documented pragmatic relevance in all of the genres, viz. the aspect of explicit linguistic vagueness. We have furthermore concentrated upon vagueness expressed through lexical means in order to be able to apply corpus analytic tools and thus investigate a considerable amount of text.

The general result is that within the described confines of this investigation we found the same three basic ways of expressing vagueness represented in the list of the 199 most typical keywords of all three corpora. We found some lexical overlaps across the keyword lists, although at a fairly limited degree. And we found hardly any overlaps at the level of actual conventional formulations and collocations, with the modal auxiliary *may* as a notable exception. Thus, we must conclude that intergeneric conventions in the field of expressing vagueness in the three investigated genres exist, but that they are located at a fairly high level of specificity, hardly including the level of actual conventional formulations.

This result is dependent upon a number of characteristics of our investigation:

- In our analyses, we have chosen to concentrate upon single lexemes as expressions of vagueness. It could be interesting to investigate at a later stage, whether a focus more upon keyword lists for collocations would give us a different picture.
- We have chosen to set the bar for a word to be relevant at a fairly high level, concentrating upon the first 199 typical keywords of the three corpora. In a next step it will be relevant to empirically test what happens if the list of typical keywords is extended. For example, the listed words introducing indeterminate standards are all present in all corpora, but are not frequent enough for them to get into the rather short list of keywords. Here, we need to play more with the corpora to get an overview of relative positions.

Finally, it would be interesting in future work to make a broader investigation of conventional formulations at more textual levels, instead of just concentrating upon single words expressing vagueness. For example, Engberg (2003) has demonstrated intergeneric conventions at speech act level between the German genres *Klage*, *Scheidungsantrag* and *Anklageschrift*. Equally, a number of candidates for deeper intergeneric conventions exist between the three genres investigated here.

NOTES

¹ See for example Melinkoff (1963), Tiersma (1999), Bhatia et al. (2005), Wagner and Cacciaguidi-Fahy (2006); see also section 4 below.

² For a deeper discussion of Pinkal's definitions, see Engberg and Heller (2008: 146-148).

³ For a discussion of the distinction between vagueness as an explicit, local and a more general and global characteristic of legal language, dependent on its linguistic nature, see Engberg and Heller (2008: 149-153).

⁴ See for example Berkenkotter and Huckin (1995), Paltridge (1997), and Heinemann (2000) for discussions of genres as mainly a mental and (socio-)cognitive object.

⁵ The word *section* has not been included in the list for Regulations and for Contracts although it occurs in the corpora in relevant collocations (*such section*), as at least 75 % of the instances in Regulations and 42 % of the instances in Contracts are references to concrete sections of the regulations themselves (*this section + in section*). These instances may not be seen as being vague category indicators.

⁶ Cutting (2007) distinguishes this kind of vague words from 'metonymical proper nouns' and 'superordinate nouns' (Cutting 2007: 223-225).

⁷ The modal auxiliary *can* is also used in this function, but is not part of the keyword lists of any of the three corpora. It therefore falls outside the scope of this paper.

⁸ We have followed the same procedure in the cluster analysis as described in 6.1.3.

REFERENCES

- Adams, K.A. (2004). *A Manual of Style for Contract Drafting*. Chicago: American Bar Association.
- Anesa, P. (2007). Vagueness and precision in contracts: a close relationship. *Linguistica e Filologia*, 24, 7-38.
- Arinas Pellón, I. (2010). How does a Patent Move? Genre Analysis has Something to Say about It. In M. Gotti & C. Williams (Eds.), *Legal Discourse across Languages and Cultures* (pp. 313-334). Bern: Lang.
- Barker, C. (2002). The Dynamics of Vagueness. *Linguistics and Philosophy*, 25(1), 1-36.
- Berkenkotter, C. & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition, culture, power*. Hillsdale, N. J.: L. Erlbaum.
- Bhatia, Vijay K. (1993). *Analysing Genre: Language Use in Professional Settings*. New Jersey: Pearson.
- Bhatia, Vijay K. (2004). *Worlds of Written Discourse*. London: Continuum.
- Bhatia, V. K., Engberg, J., Gotti, M. & Heller, D. (Eds.). (2005). *Vagueness in normative texts*. Bern: Lang.
- Black's Law Dictionary, 9th edition.
- Blom, B. & Trosborg, A. (1992). An Analysis of Regulative Speech Acts in English Contracts - Qualitative and Quantitative Methods. *Hermes*, 9, 83-111.
- Channell, J. (1994). *Vague Language*. Oxford: Oxford University Press.
- Cutting, J. (2007). 'Doing More Stuff - Where's It Going?' Exploring Vague Language Further. In J. Cutting (Ed.), *Vague Language Explored* (pp. 223-244). New York: Palgrave MacMillan.
- Dascal, M. & Wróblewski, J. (1988). Transparency and Doubt: Understanding and Interpretation in Pragmatics and in Law. In *Law and Philosophy* 7, 203-224.
- Endicott, T. (2000). *Vagueness in Law*. Oxford: Oxford University Press.
- Endicott, T. & Spence, M. (2005). Vagueness in the Scope of Copyright. In *121 Law Quarterly Review*, 657-680.
- Engberg, J. (2003). Textsortenkonventionen – Zum Status und zur Bedeutung für die übersetzungsbezogene Beschreibung von Rechtstexten. In H. Gerzymisch-Arbogast, E. Hajicová, P. Sgall, Z. Jettmarova, A. Rothkegel & D. Rothfuß-Bastian (Eds.), *Textologie und Translation* (pp. 61-83). Tübingen: Narr.
- Engberg, J., & Heller, D. (2008). Vagueness and Indeterminacy in Law. In V. K. Bhatia, C. N. Candlin & J. Engberg (Eds.), *Legal Discourse across Cultures and Systems* (pp. 145-168). Hong Kong: Hong Kong University Press.

- Heinemann, W. (2000). Textsorten. Zur Diskussion um Basisklassen des Kommunizierens. Rückschau und Ausblick. In K. Adamzik (Ed.), *Textsorten. Reflexionen und Analysen* (pp. 9-30). Tübingen: Stauffenburg.
- Kennedy, C. (2002). *The Landscape of Vagueness*. Ms., Northwestern University. Retrieved from <http://ocw.mit.edu/NR/rdonlyres/Linguistics-and-Philosophy/24-979Topics-in-SemanticsFall2002/669BAFE8-87D1-4938-8644-125705D071F7/0/vagueness.pdf>
- Kurzton, D. (1986). *It is hereby performed ... - Legal Speech Acts*. Amsterdam/Philadelphia: John Benjamin.
- Nussbaumer, M. (2005). Zwischen Rechtsgrundsätzen und Formularsammlung: Gesetze brauchen (gute) Vagheit zum Atmen. In V. K. Bhatia, J. Engberg, M. Gotti & D. Heller (Eds.), *Vagueness in normative texts* (pp. 49-72). Bern: Lang.
- Maley, Y. (1987). The language of legislation. *Language in Society*, 16, 25-48.
- Mellinkoff, D. (1963). *The Language of the Law*. Boston: Little Brown & Co.
- Paltridge, B. (1997). *Genre, frames and writing in research settings*. Amsterdam / Philadelphia: John Benjamins.
- Pinkal, M. (1981). Semantische Vagheit: Phänomene und Theorien, Teil I. *Linguistische Berichte* 70, 1-26.
- Prampolini, M. (1998). Il concetto di vaghezza. In F. Albano Leoni, D. Gambarara, S. Gensini, F. Lo Piparo & R. Simone (Eds.), *Ai limiti del linguaggio. Vaghezza, significato e storia, volume in onore di T. De Mauro* (pp. 107-122). Roma-Bari: Laterza.
- Quine, W. (1960). *Word and Object*. New York: John Wiley & Sons.
- Rayson, P., Berridge, D. & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies. *JADT 2004: 7es Journées Internationales d'Analyse statistique des Données Textuelles*. Retrieved from http://eprints.lancs.ac.uk/12424/1/rbf04_jadt.pdf
- Rayson, P. & Garside, R. (2000). Comparing Corpora using Frequency Profiling. *WCC '00 Proceedings of the workshop on Comparing Corpora – Volume 9*. Retrieved from http://www.comp.lancs.ac.uk/~paul/publications/rg_acl2000.pdf
- Roberts, G. (2007). *A Practical Guide to Drafting Patents*. London: Sweet & Maxwell.
- Russell, B. (1923). Vagueness. In *Australasian Journal of Philosophy*, 1, 84-92.
- Šarčević, S. (2000). *New Approach to Legal Translation*. The Hague: Kluwer Law International.
- Shippey, K. C. (2003). *A Short Course in International Contracts: Drafting the International and Sales Contract: For Attorneys and Non-attorneys*. Novato, California: World Trade Press.
- Solan, L. M. (2005). Vagueness and Ambiguity in Legal Interpretation. In V. K. Bhatia, J. Engberg, M. Gotti & D. Heller (Eds.), *Vagueness in normative texts* (pp. 73-96). Bern: Lang.
- Swales, J. (1990). *Genre Analysis*. Cambridge: Cambridge University Press.
- Tiersma, P. (1999). *Legal Language*. Chicago: The University of Chicago Press.
- Trosborg, A. (1991). An analysis of legal speech acts in English Contract Law. 'It is hereby performed'. *Hermes* 6, 65-90.
- Velluzzi, Vito (2005). Scetticismo interpretativo moderato e argomenti dell'interpretazione in *Diritto e Questione Pubblica* (5). Retrieved from http://www.dirittoequestionipubbliche.org/page/2005_n5/studi_V_Velluzzi.pdf
- Voldgaard Larsen, A. (2008). *Faglighed og forståelighed i kontrakter*. Unpublished PhD dissertation. Aalborg, Denmark.
- Wagner, A. & Cacciaguidi-Fahy, S. (Eds.). (2006). *Legal Language and the Search for Clarity*. Bern: Lang.