How Vague Are International Agreements?
Introducing a Method for Systematic Comparison

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors der Philosophie (Dr. phil)

am Institut für Politikwissenschaft

der

Technischen Universität Darmstadt

vorgelegt von

Hanne Weismann

2017

Erstgutachter: Prof. Dr. Jens Steffek

Zweitgutachter Prof. Dr. Bernhard Zangl

Weismann, Hanne: How Vague Are International Agreements? Introducing a Method for Systematic Comparison

Darmstadt, Technische Universität Darmstadt

Jahr der Veröffentlichung der Dissertation auf TUprints: 2018

Tag der mündlichen Prüfung: 28.04.2017

Acknowledgements

First and foremost, I would like to thank Prof. Dr. Jens Steffek for his support and advice in writing this thesis.

I would also like to thank the Members of the defence committee, Prof. Dr. Nina Janich, Prof. Dr. Markus Lederer and Prof. Dr. Bernhard Zangl.

Thanks are also owed to my colleagues, whose constructive criticism has helped me greatly: Sabrina Engelmann, Nadia El Ouerghemmi, Leonie Holthaus, Johannes Klassen, Julia Liebermann, Miranda Loli and Ronja Schütz.

Finally, a special thank you to my colleague and friend Johannes Karl Schmees, without whose unwavering encouragement this thesis would not exist today.

Abbreviations

CoP – Conference of the Parties

IL – International Law

IR – International Relations

NPT – Nuclear Non-Proliferation Treaty

RevCon – Review Conference

UNFCCC – United Nations Framework Convention on Climate Change

Table of Contents

1 Introduction

More than at any point in history, international agreements[1] govern our lives. Deciding which country may possess nuclear weapons (NPT 1986), the future of the world's climate (UNFCCC 1992), which species of plants and animals are protected from international trade (CITES 1973), the amount of fish to be taken out of international waters (UNCLOS 1995), or which goods may be subjected to which tariff levels (e.g. GATT 1994), international agreements shape not only the landscape of states' relations with one another, but also have an immediate impact on the daily lives of their citizens.

Their importance is reflected in numerous academic studies on the subject. In addition to research on specific thematic fields, such as Human Rights (see e.g. Camp Keith 1999 or Lutz/Sikkink 2000), weapons (see e.g. Müller 2005), trade (see e.g. Tussie 2005), or the environment (see e.g. Bernauer 1995, or Tamiotti/Finger 2001), research is being conducted on the phenomenon of international agreements in general. A large amount of literature is concerned with their legitimacy, studying both how international agreements can fit into democratic structures, as well as alternative sources of legitimacy (see e.g. Franck 1988, or Kumm 2004). Other researchers are concerned with the effectiveness of international agreements (see e.g. Kingsbury 1998), asking how these agreements are enforced (see e.g. Koh 1998), under what circumstances states are more likely comply with the agreements (see e.g. Simmons 2000), and what the agreement itself actually changes from the status quo (see e.g. Hisschemöller/Gupta 1999). The negotiation of international agreements is a field of study in its own right, and is often looked at separately from the effects of agreements. Strategic concerns for state representatives, such as how to negotiate effectively for their own interest (see e.g. Mastenbroek 2002) and more general questions on how negotiations are conducted (see e.g. Jönsson 2002, or Zartman 1994b) are subjects of investigation. So is the way in which international agreements interact with domestic politics (see e.g. Putnam 1988). International organisations are considered as to their impact on the negotiation process and implementation of international agreements (see e.g. Alvarez 2006). The study of international norm dynamics also features a significant overlap with the study of international agreements (see e.g. Finnemore/Sikkink 1998). The research on regimes incorporates aspects of all the fields of study listed above, additionally studying the interplay

[1] In the Vienna Convention on the Law of Treaties, a treaty is defined as "an international agreement, concluded between States [sic] in written form and governed by international law, whether embodied in a single instrument or in two or more related instruments and whatever its particular designation" (Vienna Convention 1969: Art. 2). I follow this understanding, but use the term 'international agreement' to avoid confusion with treaties among individuals or corporations.

of different agreements (see e.g. Charnovitz 2003) as well as their development over time (see e.g. Steffek 2005).

The language in which international agreements are formulated underlies all these areas of study. Some studies, such as Bhatia et al. (2003) and Olsen et al. (2009) address issues of translation. Beaulac (2004) dedicates an entire book to the implications of a single word (sovereignty), its origin and its meaning in international agreements. Handbooks for negotiators offer detailed guidelines on how to draft international agreements (see e.g. Aust 2000). Wessel (2010) has even proposed to understand International Law (IL) as a language.

Vagueness[2] is a phenomenon at the intersection of language and international agreements which is often referred to in passing but rarely expanded upon. Starting with Morgenthau (1954), International Relations (IR) scholars have made claims about the provenance and effects of vagueness in international agreements. They assert a correlation of the precision of an international treaty with its legitimacy (Franck 1988: 741, Stokke 2001: 16), use the precision or specificity of a treaty as an indicator for its effectiveness, (see e.g. Goldstein et al. 2001, Wettestad 1999), or postulate that vague provisions may be a chance for a treaty to come into existence in the first place (cf. Abbot/Snidal 2000).

To be able to know whether these claims are true, it is necessary to systematically compare the vagueness of international agreements. At present, such comparisons are commonly ad-hoc and unsystematic, which weakens their conclusions and can lead to confirmation bias in their results. Very few authors elaborate on indicators that might be used to assess the relative vagueness of a treaty. Most merely describe some international treaties as vague and others as more precise, and from this draw conclusions about their effectiveness or legitimacy. While readers generally have intuitions about whether a treaty is vaguely formulated or not, the lack of a systematic method of comparison is unacceptable. The research question driving this dissertation is therefore: *How can the vagueness of different international agreements be systematically compared?*

In this thesis, I will introduce a method to reliably compare instances of vague language in international agreements. By systematically applying the same set of criteria in the same way to multiple agreements, the results will be consistent, reproducible, and directly linked to linguistic theory.

Identifying these criteria is a major part of this study. It requires insights from IL and linguistics, as well as the perspective of IR. While I incorporate indicators drawn from

[2] Vagueness is a complex phenomenon which eludes easy definition. Section 2.1 is dedicated to a discussion on this topic.

literature on vagueness in everyday language, the research focusses on collecting and analysing empirical data, because the assessment of experts IL experts is part of what determines its meaning. Based on this data, I conduct several statistical tests to find out which possible indicators of vagueness are relevant in the field of IL. Through the tests, I establish four criteria which indicate that a sentence in a legal agreement will probably be perceived as vague.

I then apply these indicators to a total of seventeen international agreements - six under the Nuclear Non-Proliferation Treaty (NPT) and eleven under the Framework Convention on Climate Change (UNFCCC). They are used as examples to present comparisons of vague language within and across regimes. In a final step, the developed method is tested by comparing the results of its application to the independent assessments of international legal experts.

There are limits to the scope of this study. A systematic comparison of vagueness in international agreements is a necessary first step to be able to test claims on its effects or causes, but it is only the first step: I do not actually investigate the motives that negotiators had nor do I consider the consequences of vagueness. Additionally, the indicators used in the tool are strictly quantitative and should be supplemented with qualitative analyses in order to gain a more nuanced picture. Finally, the application of the tool is currently limited to those agreements which were used in its development.

Once the method is established, it can give much more detailed and targeted practical advice for the drafters of the agreements who wish a provision to be read as precise or vague. Researchers can assess the theories on causes and effects of vagueness in international agreements, and the results may provide relevant insights for the study of vagueness in general. The method also allows to check for the development of vagueness in regimes over time and to see if agreements of different status are also written in more precise or vague language than others. While the perspective of this dissertation remains firmly anchored in the field of IR, it provides insights relevant for several other disciplines, most notably linguistics and IL. Ultimately, a deeper understanding of the type and degree of vagueness present in international agreements will lead to a better comprehension of the effects of vague agreements on international cooperation.

This dissertation proceeds as follows: The second chapter differentiates some key concepts this thesis is working with and situates the thesis in the philosophy of law. Discussing the differences and similarities of vagueness, vague language and indeterminacy

enables both a more in-depth understanding of each of the concepts, and also gives some insights into the variety of disciplinary perspectives. The second part of the chapter focuses on one of these disciplines, namely legal theory, by giving an overview of the indeterminacy debate. In doing so, the chapter also clarifies some of the assumptions underlying the premise of this thesis: that there is vagueness in international agreements, and that it can be there to differing degrees.

In the third chapter, I provide the theoretical framing of this thesis. It anchors the question of vagueness in international agreements in IR theory and gives the theoretical background on the setting in which international agreements are negotiated and (not) complied with. To this end, it briefly discusses regime theory, the negotiation and implementation of IL, its mechanisms and functions as well as some general characteristics of rules. The second part of the chapter provides a detailed look at the causes and effects that vagueness in IL is thought to have according to scholars of IR. On the basis of this literature, I argue that a means to compare the vagueness of international agreements is needed: Without one, how could it be possible to ever actually study its causes and effects?

The fourth chapter emphasizes the linguistic aspects. It outlines current methods of identifying vague language in different contexts. Since this literature is usually concerned with finding vagueness in everyday speech, chapter 4.1 briefly departs from the legal context established in the previous chapters. The vast majority of these methods defines some indicators of vagueness and then looks for it in the text. Section 4.2 then seeks to re-incorporate the international legal context by discussing legal language as a specialized language. Finally, section 4.3 briefly discusses the very few instances in which indicators for vagueness have been already sought out specifically in the field of IL. By its very scarcity, the section demonstrates the need for further work on the subject.

The fifth chapter is the methodological heart of the study. It introduces and defends the original methods used to find out how to compare vagueness in international agreements. It also presents the results of this inquiry. Part 5.1 explains the methods used to conduct a survey among professionals of IL. It gives some detail on the criteria that were important to the study, and discusses compromises between conflicting goals of the survey. Part 5.2 presents the results and discusses comments of the respondents. The third part of the chapter is dedicated to the analysis of the data gained by the survey. Eight characteristics are tested as to whether or not their presence leads legal language to be perceived as particularly vague or precise. In the end, the indicators which proved to make a difference in

the perceived vagueness are assembled to form a tool for the detection of vagueness in international agreements.

In the sixth chapter, the insights gained in the previous part of the thesis are used to study two international regimes: The United Nations Framework Convention of Climate Change (UNFCCC) and the Nuclear Non-Proliferation Treaty (NPT). Applying the developed method to seventeen different agreements yields insights into strength and weaknesses of the method and points to possible further refinements. Seeing how the results play out in the variety of agreements presented – including differences in legal status, development over time, and of course regime – supports these insights. The chapter incidentally also provides new insights into the agreements and regimes themselves. In a last step, a comparison with previous assessments of IR scholars shows that the results of the new method largely correspond to their judgements.

The conclusion highlights the results generated by the tool development process: The presence of hedge words, a low number of determiners, the presence of weak explicitly performative verbs and the absence of strong explicitly performative verbs make it more likely that a sentence will be perceived as vague. When applied to the sentences of a UNFCCC or NPT agreement, these four indicators are fairly reliable in matching expert opinion on whether that agreement is vague or not. According to my analysis, the NPT and the Kyoto Protocol are the most precise of the documents, while the NPT RevCon decisions of 1995 and the Delhi Ministerial Declaration (2002) are the vaguest. In general, legally binding documents are more precise than decisions or declarations.

2 The legal perspective: concepts and background

This chapter provides the conceptual background and introduces the terminology to the empirical findings which will be discussed further on. In the first part of the chapter, I disambiguate three key concepts of this thesis: vagueness, vague language, and indeterminacy. The three terms represent their respective disciplines of provenance: vagueness is a philosophical concept, vague language is a phenomenon discussed in the field of linguistics and indeterminacy is a term used specifically in legal theory and philosophy. By discussing the three concepts individually and jointly, I hope to show the different aspects the disciplines bring to the foreground, and in turn provide a multifaceted view of the concepts themselves.

The second part of the chapter will provide an overview over the so-called 'indeterminacy debate'. In doing so, I will be led by two main questions: Is there indeterminacy in the law? And if there is, what are the possible sources and consequences of this indeterminacy? In its entirety, the chapter is aimed at equipping the reader with the conceptual foundations and terminology for the theoretical and methodological concerns of the following chapters.

2.1 Key concepts

The first part of this chapter is aimed at disambiguating three key terms, which are at the core of the following thesis: vagueness, vague language, and indeterminacy, as well as showing important connections and differences between them. This not a trivial undertaking, because each of these terms in themselves encompass many facets and are not easily defined. Adding in the substantial overlap between the concepts means that the disambiguation, while all the more important, becomes quite challenging.

Authors are generally agreed that vagueness is not an easily defined term (see for example Cheng/Warren 2003: 382, Eklund 2005: 27). However, the reasons given for this difficulty are varied. Burns (1991:3) attributes responsibility to many different conceptions of vagueness, as no two authors generally have the exact same definition in mind when talking about vagueness

> *"Vagueness is not easy to characterize or define. One reason for this difficulty is that there appears to be a number of different conceptions of vagueness, and it is not clear just what they have in common".*

Ullmann (1962: 118), on the other hand, refers to characteristics inherent in the concept of vagueness as a reason for difficulties with defining the term.

> *"If one looks more closely at this vagueness one soon discovers that the term is itself rather vague and ambiguous: the condition it refers to is not a uniform*

feature but has many aspects and may result from a variety of causes. Some of these are inherent in the very nature of language, whereas others come into play only in special circumstances".

Most likely a combination of these factors is at play. In order to nevertheless disambiguate the terms of this thesis, I will describe different facets of each of the three key terms, while also devoting some time to exploring what does not fall inside the boundaries of the concepts.

2.1.1 Vagueness

Vagueness is a linguistic and logical concept, and it is a property of language and/or thought. It is problematic in epistemological, ontological and methodological dimensions, all of which will become relevant in the course of this thesis.

I will proceed to clear up the concept by first providing an overview on what vagueness is by going into more detail of four of its characteristics: Sorites-susceptibility, immeasurability, incommensurability, and higher-order vagueness, and then distinguish the phenomenon from several related, but distinct concepts, namely context-dependency, ambiguity and incompleteness. By doing this, I intend to clarify the concept so that further uses in the article can be more easily understood, as well as disambiguate the phenomenon from others that it can easily be confused with. This is especially necessary as the actual use of the concepts is by no means standardized, and each author discussed in this chapter may have his or her own conception of the terms used.

The Sorites Paradox

The Sorites Paradox (which can also be translated as the paradox of the heap, from the Greek word Soros, the heap) is based on a puzzle often attributed to Eubulides of Melitus (Hyde 2014). Nowadays, it is used to explain the logical conundrum of vague words and concepts.

In a nutshell, it explains the paradoxical relationship between a difference of one grain of sand and a heap. If you have one grain of sand, what you have is clearly not a heap. If you add one grain of sand to it, you still do not have a heap: one grain does not make a difference. However, if you repeat the exercise often enough, at some point you will end up with something that clearly is a heap of sand. It is a paradox because, while there are two clear states at either end, there is no point on the continuum that could be found to clearly mark the distinction between the two states. No one grain of sand that clearly makes the

difference between 'this is a heap' and 'this is not a heap'. Heap, therefore, is a vague concept.[3]

Sorites susceptibility is related to the presence of borderline cases and the lack of clear boundaries (Keefe 2000: 6): There will be collections of sand which will not be classifiable as being a heap or not being a heap, and reasonable competent language speakers with all the information about the case can disagree on the categorization. These cases, where no clear classification seems possible, are called borderline cases. The possibility of borderline cases, in turn, is what makes concepts have unclear or fuzzy boundaries (Keefe 2000: 7). The existence or even possibility of vagueness poses a problem for formal logic, because the premises depend on binary truth values: either a proposition is true, or it is not true, and if neither of these states apply, the proposition becomes merely senseless. Neither of the three is the case when it comes to vague concepts.

While it may appear at first glance that the number of words and concepts susceptible to the Sorites Paradox is relatively limited, this is not in fact the case. In fact, most concepts used in everyday language have unclear boundaries and borderline cases (Goguen 1975: 325). Nor is the phenomenon limited to nouns (as the example of the heap might suggest): Sorites susceptibility can equally occur in verbs (how fast does someone have to move to be considered hurrying?), adjectives (a favourite example is bald, see i.e. Greenawalt 2001: 434), and correspondingly, adverbs. It also (or especially) applies to normative concepts like good (see Endicott 2001: 43).

There are several different attempts to solve the Sorites Paradox (see Manor 2006: 172). An epistemic solution has been proposed by Williamson (1994). In a nutshell, the approach suggests that "vagueness consists in our ignorance of the sharp boundaries of our concepts" (Williamson 1994: xi). Basically, the argument is that vagueness is not an ontological but an epistemic phenomenon, i.e. all concepts have sharp boundaries, the problem is that we do not know what they are. A supervaluation solution has been proposed by Fine (1975: 265): "Very roughly, it says that a vague sentence is true if and only if it is true for all ways of making it completely precise." Another approach has been put forward in the context of fuzzy logic, and is basically concerned with a particular modification of traditional logic to suit inexact concepts (which occur in ordinary live but not in settings of traditional logic)

[3] The Paradox has been described and explained in different ways by many authors. Particularly enlightening examples include a legal iteration in Endicott 2001: 57-8, Hyde 2014, Keefe 2000:6, and (in logical terminology, and also for a succinct differentiation between a true paradox and a fallacy) Manor 2006: 171-2.

(Goguen 1975: 326) with the aim of eventually extending it to encompass all natural language (Goguen 1975: 372). Manor herself proposes a pragmatic resolution to the paradox, which "assumes that the meaning of vague terms is determined by their use" (Manor 2006: 172).

The number of proposed solutions itself suggests that the Sorites Paradox is far from uncontroversial. While it sets the basic of understanding the phenomenon and captures the core of the problematic concept of vagueness, two other important characteristics will be discussed in the next two sections.

Immensurability

When faced with the Sorites Paradox, one simple answer may be to arbitrarily fix numerical values to specific concepts. To stay with the example of the heap, a person, or group of persons, could decide that, for a specific context, that 'heap of sand' is something that consists of 1000 or more individual grains of sand (all touching each other and forming a conical shape). While the distinction would be arbitrary, in that there is nothing substantially different about a 999 grains of sand arranged in the same way, it seems at least a possibility to introduce such a distinction, and thus escape the Sorites Paradox – at least for technical areas, if not for everyday language. However, this solution ignores the very real possibility that some concepts cannot be measured at all.

> *"We can invent the term 'immensurate' to describe criteria of application that do not correspond to a scale. 'Immensurability' is the property that something has if and only if it can be assessed in some respect in which it cannot be measured" (Endicott 2001: 46).*

He gives the examples of humour and imagination (Endicott 2001: 46) but the same is true of many key concepts of human lives such as morality (it would seem absurd to claim that, for something to be considered moral, it must contain a specific amount of units of morality), kindness, responsibility, redness, tastiness and so on.

Incommensurability

A related phenomenon to the one termed 'Immensurability' (Endicott 2001: 46) above is the concept of incommensurability. In the Stanford Encyclopedia of Philosophy, Oberheim and Hoyningen-Huene (2013) define: "The term 'incommensurable' means 'no common measure', having its origins in Ancient Greek mathematics, where it meant no common measure between magnitudes." In the sense used here, it describes objects and concepts

with different characteristics, which cannot be compared to each other. Endicott gives the example of weight and strength in a rope:

> *"But if some rope is slightly stronger and slightly heavier than another, and if weight and strength are incommensurate, then the way in which the formula ranks the two ropes will be stipulative: there is no answer to the question that a good formula will reveal" (Endicott 2001: 42).*

It may seem like the concept of Immensurability could be subsumed under the phenomenon of Incommensurability. For example, I mentioned before that the term 'redness' could be seen as immensurable. One could argue that 'redness' simply consists of at least two characteristics – like lightness and shade, for example – that are by themselves measurable, but incommensurable when taken together. However, while this explanation might be applicable to some terms, others (like the above-mentioned morality or kindness) have no units and cannot be quantified at all, showing that this explanation does not cover all concepts (see Endicott 2001: 46). The distinction between incommensurability and immensurability thus seems relevant, and both concepts are important to fully appreciate the complexity of the problems that the phenomenon of vagueness entails.

Higher order vagueness

One of the reasons vagueness is such a tricky concept is that it includes higher order vagueness. Even when considering the implications of immensurability and incommensurability, it might still be relatively easy to deal with the problem of 'simple' vagueness by introducing a middle category – there are clearly positive cases, clearly negative cases, and here is a clearly delineated space of borderline cases, which are neither fall clearly inside nor clearly outside the boundaries of a certain concept (see Sainsbury 1991: 167-8). However, as Sainsbury (1991: 168) continues to explain, this characterization only describes the lowest level of vagueness. He goes on:

> *"However, with most or even all vague predicates it soon appears that the idea that there is a sharp division between the positive cases and the borderline ones, and between the borderline cases and the negative ones, can no more be sustained than can the idea that there is a sharp division between positive and negative cases".*

This exercise can be continued ad infinitum, such that the introduction of ever more categories can never keep up with the actual vagueness of the concept. [4]

[4] Because Sainsbury's point in his 1991 article is on the inadequacy of applying classical logic to real-world problems, he suggests to do away with the distinction between simple and higher-order

Taken together, the characteristics described in the previous sections make vagueness such a theoretically interesting and problematic phenomenon. The previous sections were aimed at describing four core properties or characteristics which constitute the phenomenon vagueness. To fully grasp the concept, and to disambiguate it from other terms that will be mentioned over the course of this thesis, I believe it useful to spend a few paragraphs describing what vagueness is *not*. In the following, I will briefly differentiate vagueness from context-dependence, ambiguity and incompleteness in turn.

Epistemic and ontological vagueness

One fundamental distinction between different conceptions of vagueness is whether one believes vagueness to be ontological or epistemic. If vagueness is ontological, the paradox of the heap means that it is actually unclear when a collection of grains of sand constitutes a heap. If it is an epistemic phenomenon, this is not so: there is be a precise point when adding a grain of sand will turn the collection from not-a-heap into a heap, we are simply not capable of knowing when exactly this point occurs. This has been discussed as a possible solution to the Sorites Paradox above: it is not really a Paradox, we just cannot know the exact solution. Endicott points out some of the problems with this second approach:

> *The epistemic theory has the attraction of simplicity. It also has the attractions of vindicating classical logic, of maintaining a simple relationship between classical logic and the meaning of words in a natural language, and of doing nothing obviously innovative with the notion of truth. Its problem is that it makes a bizarre claim about the correct application of vague words: it claims that their meanings determine sharp boundaries, which we cannot locate. The epistemic theory has the consequence that you could lay a hair on the ground so that, all along its length, half of it is close to New York and half of it is not. Or you could, except that you have no way of knowing where the boundary is. The theory solves the paradox at the price of an apparently outlandish account of the meaning of words" (Endicott 2001: 100).*

I tend to agree with Endicott's arguments here: it does not seem likely that there would be a predetermined rule in the universe of how many grains of sand constitute a heap, which humans simply cannot grasp and therefore resort to vague language. However, it is actually quite possible to remain agnostic about the nature of vagueness and still research its

vagueness altogether, and proposes boundarylessness as an alternative way of conceptualizing classifications. I do not disagree, but I do believe that the distinction still serves a purpose in explaining the full scope of the problem that the phenomenon of vagueness poses. He also states "On my view, the phenomena which those of classical inclinations classify as 'higher-order vagueness' are real enough; it is just that there is no real hierarchy here. Nothing gives rise to substantive issues about the level of vagueness appropriate to our familiar examples of vagueness, and there is no multiplication of sets" (1991: 179), indicating that his issues are with the terminology and use of the concept rather than its existence itself.

practical implications. As long as it is truly unknowable where a boundary is, the effects of vagueness in many cases will still be the same as if there truly were no clear boundary.

Vagueness vs. ambiguity

While the concepts of ambiguity and vagueness may seem similar at first glance, they are actually quite easy to differentiate. Solum (2010: 97-8) defines both terms:

> *"In the technical sense, ambiguity refers to the multiplicity of sense: a term is ambiguous if it has more than one sense. [...] The technical sense of vagueness refers to the existence of borderline cases: a term is vague if there are cases where the term might or might not apply".*

As it is, a word can be both ambiguous and vague, as exemplified by the word cool: it is ambiguous because it can refer to style, mannerism or temperature, and it is vague because there is no clear boundary between a room with a cool temperature and a room with a non-cool temperature. Ambiguity is almost always solvable by context: Saying 'that is a cool poster!' is unlikely to be taken as a comment on its temperature.

Vagueness vs. context-dependence

Bosch (1983) has argued that vagueness can be reduced to context-dependence and properly defined terms do not pose a problem for formal reasoning. However, while many terms are both vague and context-dependent (it may well be argued that no language can have any meaning when taken entirely out of context), the two phenomena are not the same. Keefe (2000: 10) explains:

> *"Vagueness must not be straightforwardly identified with context-dependence [...], even though many terms have both features (e.g. 'tall'). Fix on a context which can be made as definite as you like (in particular, choose a specific comparison class, e.g. current professional American baseball players): 'tall' will remain vague, with borderline cases and fuzzy boundaries, and the sorites paradox will retain its force. This indicates that we are unlikely to understand vagueness or solve the paradox by concentrating on context-dependence".*

Moreover, Bosch's attempt does not adequately address the problems posed by immensurability and incommensurability. While some of the problems of vagueness and context-dependence can be related in practice (namely when the exact context is unknown or itself under debate, which, as we will see, can happen in legal contexts), and while – also in practice – lots of the conundrums posed by vagueness are in fact solved by context (because for most ordinary purposes, it tends not to be of immediate importance if a certain assembly of grains of sand is *clearly* a heap or not, as long as all participants in the conversation get a close enough picture of what is being talked about), some problematic

areas persists, and one of the areas where problems still arise, and may be of huge practical significance to individuals is the sphere of jurisprudence.

Vagueness vs. open texture

As Frederick Schauer (2011: 3) points out, open texture means

> *"the possibility of vagueness – the potential vagueness – of even those terms that appear to have no uncertainties with respect to known or imagined applications".*

Open texture is used to cover the fact that we can never have complete certainty over what happens in the future, so that our language can never be precise with regards to it even if we have somehow perfectly defined everything in the present.

Vagueness vs. incompleteness of information

The problem of incomplete information is closely related to that of context dependence. The reason why I include it here is that in everyday usage, vagueness is often equated with incompleteness of information.

> *"Ordinarily, people use 'vague' to describe not expressions or concepts, but uses of language, such as promises, allegations, descriptions, threats, and insinuations. 'Vague' in this sense means something like uninformative or incomplete: wanting in useful details" (Endicott 2001: 32).*

I will further address the problem of philosophical and ordinary use of the term vagueness in section 5.2.3, and I consider it very important for the empirical purposes of this thesis. However, it is not generally the sense which theorists and philosophers[5] mean when they speak of vagueness. Rather, they refer to the concept in the technical terms described above.

Vagueness vs. generality

While many communalities exist, generality is also distinct from vagueness. As Bennett explains

> *"Vagueness is also sometimes confused with generality. If I say 'I shall see you again later this month', this is an example of generality, since the claim can be fulfilled in many alternative ways. However, it is not vague, since 'later this month' refers to a precise period of time (I assume all relevant events take place within a single time-zone). But if I say 'I shall see you in a few weeks', this is vague (as well as general) since there is no hard and fast definition of what periods of*

[5] By this I mean theorists and philosophers of vagueness. In legal theory, this is not always the case, see for example Singer (1984: 17/18)

time can be described as 'a few weeks'. Generality per se poses no real problems for the logician, since it is handled perfectly well by classical logic" (Bennett 2001: 190).

This distinction perhaps illustrates most clearly the difference between the word vagueness as used in everyday discourse, and the philosophical concept. Bennet's first example is not vague in the philosophical sense, but may well be called vague in casual conversation. This distinction will be important in chapter 5, when considering how experts in IL understand vagueness.

2.1.2 Vague language

Vague language has been an important topic in the study of linguistics (for a brief overview, see Xi 2013: 1596). While semantic studies on tend to be concerned with the more philosophical concept of vagueness, in this section I will focus on the field of pragmatics, which is concerned with what I call here vague language. There are some characteristics that most of the authors on this topic have in common. For one, the vast majority of studies is concerned with spoken language[6]. This is a marked difference to both the philosophy of vagueness, which is mostly concerned with made-up, out-of-context example sentences and legal conceptions, which are primarily concerned with written legal documents. Another characteristic of the study of vague language is that it is mainly concerned with the practical applications of the findings. Emphasis has been laid on the question how vague language affects language (see Drave 2001, Lin 2013), as well as why people employ vague language in everyday conversations. Common areas of interest here are politeness and face saving[7] measures (cf. Ruzaitė 2007).

Since it is a more practice-oriented discipline, less emphasis is placed on definitions and disambiguation and more on application and testing. This is very helpful when it comes to actually devising measurements of vagueness and deciding which words, phrases and characteristic of language are considered more or less vague. I will elaborate on these features in section chapter 3. On the other hand, it is sometimes less helpful when trying to describe the theoretical distinctions between different but related phenomena. This is why, after a very brief initial definition, I will contrast vague language to related terms and concepts, noting along the way which authors see them as different parts of the same phenomenon, and which draw more cutting distinctions.

[6] See for example Channell 1994, Cutting 2007b and Zhang 2015. A notable exception is Cutting 2012.
[7] "Your face, in pragmatics, is your public self-image. […] Whenever you say something that lessens the possible threat to another's face, it can be described as a face-saving act." (Yule 2014: 135)

One of the standard texts on linguistic approaches to vagueness appears to be Channell (1994). The classifications of what makes language vague are taken on (sometimes with slight modifications) by Cheng/Warren (2001) and Drave (2001), among others. She uses the term vague language to differentiate from vagueness as a philosophical concept (8). Her definition of vague language reads as follows:

> *"An expression or word is vague if: a. it can be contrasted with another word or expression which appears to render the same proposition; b. it is 'purposely and unabashedly vague'; c. its meaning arises from the intrinsic uncertainty' referred to by Pierce".*

However, for all its popularity, it is not at first glance clear what vague language is according to Channell. Nevertheless, her definition establishes several benchmarks upon closer examination. For one, it clearly distinguishes vague language from the existential proposition that all concepts and all language in the world is always vague. While that may or may not be the case, studies of vague language focus on instances where the vagueness is at least to some degree intentional, i.e. another expression exists that would make the utterance less vague. Part of the core interest of the study of vague language is then to see for what reasons the vaguer version was uttered. The definition itself also underscores the desire to identify vague words rather than explore the phenomenon for its own sake.

Channell's importance in the field of vague language notwithstanding, her work is a starting point rather than the final conclusion. Building on her theories, Cheng and Warren propose another definition:

> *"Vague language functions to signal to the listener that the utterance, or part of it, is not to be interpreted precisely. Vague language consists of a closed set of identifiable lexical items that can be interpreted without recourse to judgements based on the particular context in which they occur" (Cheng/Warren 2001: 82).*

Again, the definition is based on the assumption that vague language is intentional and emphasizes its functions. Interesting with regards to this definition is the inter-contextual dimension of vagueness posited: Cheng and Warren claim that vague language can be identified across linguistic contexts.

Generality

As I have shown in section 2.1.1, generality is seen as distinct from vagueness. The question how generality and vague language interact is more debated. Some authors clearly distinguish the two. Zhang, for example, defines generality as:

> *"The meaning of an expression is general in the sense that it does not specify certain details; i.e. generality is a matter of unspecification" (Zhang 1998: 16).*

Generality is a property of a word or sentence to refer to a broad category rather than a specific, clearly defined concept (Zhang 1998: 16). For instance, "bird" is more general than

"robin", while "robin" could be specified by saying "male robin", "robin sitting on a fence", etc. Very nearly all of the language used in everyday life is general, in the sense that it could be made even more specific by adding another detail. As discussed above with regards to the effectiveness of language, different levels of detail are appropriate for different situations. The distinction to vague language becomes clearer in the following quote:

> *"There are two types of general expressions: one is context-resolvable and the other is not. For example, things is a general word, usually referring to a vague inferential meaning and not context-resolvable. The pronoun she is also general, but once it is put in 'she is John's wife, and her name is Jane`, the unspecified meaning is contextually resolved. Things is both general and vague; she is general but not necessarily vague" (Zhang 2014: 24-5).*

On the other hand, the indicators used for vague language often include categories for generality. For instance, Cutting (2012: 283-4) discusses

> *"vague language (VL), defined as forms that are intentionally fuzzy, general, and imprecise, have a low semantic content and are heavily dependent on shared contextual knowledge for their meaning".*

More specifically, three of her main categories when analysing vagueness are general nouns, general verbs and general extenders (Cutting 2012: 285). The term general extenders, in turn, has been coined by Overstreet and Yule and refers to expressions which "have nonspecific or 'general' & reference, and they 'extend' otherwise grammatically complete utterances" (1997: 250-1). So while the concepts of vague language and generality may be seen as distinct, the practical orientation of studies of vague language tend to blur these differences when it comes to identifying them.

Elasticity

Zhang (2015: 5) makes an argument distinguishing vague language from elastic language (abbreviated VL and EL, respectively, in the quote):

> *"The definition of EL is developed from that of VL., which has been in use for decades [....] While EL and VL are similar linguistic phenomena, the differences are twofold. First, they carry different connotations, with the term 'EL' seeming more positive and VL more negative. Second, they have different focuses: VL gives prominence to the uncertainty and under-specification of language, EL to its fluid and elastic nature".*

Contrary to Zhang herself, I do not consider the differences important enough to classify vague and elastic language as different phenomena. The first reason she gives seems more of a rebranding – which may be understandable given that several authors mention the bad reputation of vagueness, but it does not in effect alter the content of the concept of this study. Zhangs second point seems more persuasive at first glance, but since she really does base her definition heavily on vague language, and as it merely accentuates the differences

to philosophical vagueness and indeterminacy that is already inherent in the concept of vague language, the difference is negligible for the purposes of this study.

Hedge words

The relationship between hedge words and vagueness is also quite complex. According to Lakoff (1973: 471), who coined the term,

> *"some of the most interesting questions are raised by the study of words whose meaning implicitly involves fuzziness words whose job is to make things fuzzier or less fuzzy. I will refer to such words as 'hedges'".*

For him, hedges do seem to be words that convey vagueness. However, Zhang argues that hedging is not quite the same thing as vagueness. According to her,

> *"VL and hedging may serve similar functions and purposes, but hedging can use language forms in addition to VL, while VL is not just about hedging but has many other functions" (Zhang 2015: 21).*

In this view, hedges and vague language are more like different circles that overlap to a certain extent. As with some of the other concepts discussed here, hedge words are nonetheless often used as indicators for vagueness. The question on which indicators are best used to measure vagueness is discussed in chapter 3.

Indirectness and implicitness

The Connection of vague language to these to phenomena stems from its focus on every-day, spoken language. In the words of Searle indirectness is described as follows:

> *"But notoriously, not all cases of meaning are this simple: In hints, insinuations, irony, and metaphor – to mention a few examples – the speaker's utterance meaning and the sentence meaning come apart in various ways. One important class of such cases is that in which the speaker utters a sentence, means what he says, but also means something more. For example, a speaker may utter the sentence I want you to do it by way of requesting the hearer to do something. The utterance is incidentally meant as a statement, but it is also meant primarily as a request, a request made by way of making a statement. In such cases a sentence that contains the illocutionary force indicators for one kind of illocutionary act can be uttered to perform, IN ADDITION, another type of illocutionary act. There are also cases in which the speaker may utter a sentence and mean what he says and also mean another illocution with a different propositional content. For example, a speaker may utter the sentence Can you reach the salt? and mean it not merely as a question but a request to pass the salt" (Searle 1975: 59-60).*

The example above shows that indirect or implicit sentences can often be completely explained by their context, which is not the case with vague language. Cutting (2007a: 4) makes the difference explicit in the following quote:

> *"First, I must differentiate between the terms 'VL' [vague language] and 'implicitness'. Studies of VL look at language that is inherently and intentionally imprecise, describing lexical and grammatical surface features themselves that*

> *may refer either to specific entities or to nothing in particular. Studies of implicitness mention whole bodies of underlying meaning, and language dependent on context, based on unspoken assumptions and unstated meaning".*

While she focusses on the intentionality of vague language, she also mentions context as a key difference.

Uncertainty

A last concept to discuss in relation to vague language is uncertainty. While it is not a linguistic concept, it can be a reason to use vague language.

> "Vagueness should be sharply differentiated from uncertainty, which is a distinct (though interacting) phenomenon. Uncertainty arises from lack of exact knowledge about an object or situation, and is thus an epistemic state rather than a feature of language" *(Bennett 2001: 190).*

In other words, while speakers may use vague language to express uncertainty, the uncertainty is a reason for employing vague language, not vague language itself (see also Cheng/Warren 2001: 89).

2.1.3 Indeterminacy

The academic debate on Indeterminacy and its counterpart determinacy is controversial and often heated. It is therefore difficult to provide an account of their meaning without already adopting a theoretical direction. I will nevertheless attempt to at least clarify it enough to establish a common concept of what will be talked about here.

As opposed to vagueness, indeterminacy – in the sense it is used in this thesis – is a legal concept, which is only applicable to legal situations. It has been the subject of academic debate because it touches on some of the core foundations on what law is and how it works, as well as the relationship between law and democracy. In the following, I will briefly outline what it means if a rule is indeterminate and what it means if a legal system is indeterminate, before moving on to disambiguate the concept from a few others it might be confused with.

Indeterminacy of a rule

A (legal) rule is determinate, if, when someone knows a rule, and also knows the circumstances of a case, they will then be able to determine with certainty whether the rule applies to the case or not. Any uncertainties come from insufficient information about either the rule or the case at hand, and can be eliminated by finding out that information. Indeterminacy means the opposite: if all relevant information is available and yet it is not

certain whether or not the rule applies to a certain case, the rule is indeterminate with regards to that case. Endicott (2001: 9) defines indeterminacy thus:

> *"Legal theorists say that the law is indeterminate when a question of law, or of how the law applies to facts, has no single right answer".*

While this may pose immediate and in some cases crucial problems – if one is on one side of a lawsuit and it is not clear what the law says about it, for example – it is still an issue which is localized to one or some specific rules, and could thus be dealt with on a case-by-case basis.

Indeterminacy of the law

When it comes to the indeterminacy of an entire legal system, the problem gains a new dimension. Analogous to the premises of formal logic, law can easily be perceived as being bivalent (see Madry 1999: 581). Either an accused is guilty or she is not. Either a case falls under a certain legal provision or it does not. If indeterminacy is present, this bivalence is not possible anymore.

In legal systems, the issue is even more complex because determinacy is often associated with normative conceptions on how law should be: predictability, coherence and objectivity are all values that legitimize the existence of law, and all of them are called into question if indeterminacy comes into play. The issue of indeterminacy of the law is further discussed in section 2.2.

Indeterminacy vs. discretion

Some legal texts include provisions that enable other actors to make decisions on some parts of the law. This can be called delegation or (judicial) discretion. While the outcome may seem to be the same, the theoretical implications of the two cases are very different. For one thing, this sort of intentional granting of discretion is easily preventable when unwanted: The law simply needs not to grant it. The difference is also sometimes perceived as the difference between legitimate (since expressly granted by a democratic institution which has the right to do so) and illegitimate decision-making power.

Indeterminacy vs. contestability

In modern legal systems, most rules do have an inherent contestability: appellate courts exists specifically for the purpose to solve problems arising when a legal rule – or the application of a legal rule to a case is contested. The claim that legal rules can be contested

is much less controversial than the claim that they can be indeterminate. In a way, every case that comes before the law is contested, at least by the two opposing lawyers. That in itself, however, would not undermine the responsibility and/or capability of the judge to find the one answer the law has determined. Besson makes the distinction that contestability refers to "normative disagreement, i.e. disagreement about the moral values underlying the law" (Besson 2005: 67).

Indeterminacy vs. contradictions

The difference between these two concepts are best thought of as the difference between source and outcome. Contradictions in law can be one source of indeterminacy: If two rules contradict each other, it is then indeterminate which one will be used, thus making the law indeterminate even if each rule by itself would be clear in the case at hand. For a discussion on contradictions in international agreements, see Ranganathan (2014).

Indeterminacy vs. underdeterminacy

Solum (2010: 480-1) distinguishes the three phenomena of determinacy, underdeterminacy and indeterminacy.

> *"The law is determinate with respect to a given case if and only if the set of legally acceptable outcomes contains on and only one member. The law is underdeterminate with respect to a given case if and only if the set of legally acceptable outcomes is a nonidentical subset of the set of all possible results. The law is indeterminate with respect to a given case if the set of legally acceptable outcomes is identical with the set of all possible results".*

This nomenclature is a little different from how the terms seem to be commonly used in the jurisprudential discourse. In fact, most authors call indeterminacy what Solum describes here as 'underdeterminacy'. To keep with the dominant terminology, I will use the term 'indeterminacy' to mean the second of Solum's possibilities, and speak of 'radical indeterminacy' if I want to refer to the concept that Solum here calls indeterminacy.

2.1.4 Comparing vagueness, vague language and indeterminacy

After having given a brief overview on the three concepts central to this chapter, this section is meant to very shortly describe the relationship between vagueness, vague language, and indeterminacy.

The relationship between vagueness and vague language has already been hinted at above and is relatively straightforward. I will understand vague language as the way in which

vagueness is expressed through language. Waldron (1994: 512) takes indeterminacy to be an umbrella term comprising ambiguity, vagueness and contestability. I agree with this definition only in part: As I have shown above, indeterminacy is a legal concept while vagueness is a linguistic or logic phenomenon. I take their relationship to be one of source and effect: when vague language is present in legal texts it can lead to indeterminacy of the legal outcome (for a similar argument, see Solum 2010).

It is possible to imagine the three concepts as forming a simple variation on a Venn diagram[8]:

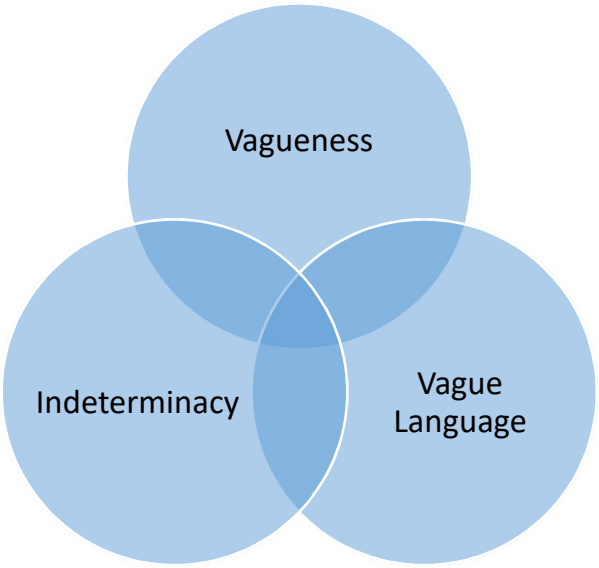


Figure 1: Relationship of vagueness, indeterminacy, and vague language

The intersection of vagueness and indeterminacy occurs when vague concepts cause a legal rule to be indeterminate. Vagueness can be expressed in vague language, and when vague language occurs in a legal setting this can in turn lead to indeterminacy.

In this thesis, I am interested in the intersection of the three concepts: vagueness interests me in so far as it is a source of legal indeterminacy, and I focus specifically on those legal provisions whose (possible) indeterminacy is caused by vague formulations of language. Within the intersection of vagueness, indeterminacy, and vague language, I understand vagueness as an ontological[9] phenomenon, indeterminacy as that vagueness directly related

[8] See Venn 1880:6, here called Eulerian circle. It is a variation because I want the mental image to signify not that in the cases of interest here, vagueness *is* indeterminacy, but simply that there occurrence overlaps.

[9] As explained in section 1.1.1, while there is a large philosophical difference between understanding vagueness as an ontological or epistemological phenomenon, the distinction actually does not matter in the context of this thesis. Whether the world itself is vague or vagueness stems from our incomplete understanding of it, the consequence when vagueness enters our consciousness is still the same.

to the propositional content of legal norms and provisions, and vague language as linguistic means and language features through which this vagueness is expressed. This intersection includes some of the most interesting features of all three phenomena.

The relationship is also invoked by Endicott (2001: 1) when he explains why he focusses on vagueness in his book:

> *"The reason for that focus is that vagueness is a paradigmatic source of indeterminacy in law, and a very important source. Along with express grants of discretion and Conventions giving judges power to develop the law, it is one of the most important sources of judicial discretion. And unlike other sources of indeterminacy such as ambiguity, it is a necessary feature of law".*

In the remainder of this chapter, it is not possible to exclusively focus on the indeterminacy in law that is caused by vagueness, or expressed in vague language, simply because some theorists important for the discussion and exploration of indeterminacy base their accounts mainly or in part on other sources. However, I will take note in particular of those who combine the ideas.

2.2 Indeterminacy debate

In this section of the paper, I will attempt to present a debate on indeterminacy in legal philosophy. In order to make the complex debate more easily accessible, I will use this section to present two rather extreme – but nonetheless important – positions at the ends of the indeterminacy debate. Each of them will start with a simplistic caricature to outline the general direction of the argument before adding in the complexities which make them actual arguments proposed by academics. Each section is followed by a discussion of the problems and inconsistencies the respective positons face. This should outline the general scope of the debate, as well as show the main arguments under contention.

During this and the next section, it is useful to keep in mind two key questions which are asked and answered in some way or another by all the authors discussed here: Is there indeterminacy in the law? And, if indeterminacy is present, what kind of consequences does that have for the concept of law? The range of possible answers to question one spans from 'no, none' over 'some' and 'sometimes' to 'yes, always'. For the second question, answers range from 'indeterminacy, if it exists, has no relevant consequences on the law as such' to 'if law is indeterminate, it is illegitimate and self-defeating. Indeterminacy is a fatal flaw for all legal systems'. Authors can potentially be arranged on a coordinate plane according to where on the spectrum their answers to these questions fall.

To understand the controversial and heated character of the debate, it is very important to keep in mind that the question of indeterminacy in law is almost always accompanied by normative consideration. As Kress (1989: 285) so succinctly puts it: "Indeterminacy matters because legitimacy matters."

2.2.1 There is no vagueness in the law: formalism and the 'Right Answer Thesis'

On one extreme of what has been called the 'Indeterminacy Debate'[10], Dworkin maintains that law is completely determinate, and that each legal case has only one right answer. One of the cornerstones of this argument is made in his 1963 essay 'Judicial Discretion'. The caricature of his argument is provided by himself:

> *"To the layman a lawsuit or a trial is an event in which a judge determines a controversy by application of established principles, rather than new principles invented to dispose of the case. He knows that individual judges may fail this ideal of justice; but he believes such failures to be aberrations, their occurrence marking injustice rather than its opposite. To him judges should and in general do, in the words of the admittedly metaphorical maxim, find the law and not make it. The layman's respect for law is founded in large part on his view that this is a fair method of deciding controversies" (Dworkin 1963: 624).*

The view that the law consists of a system of rules that only need to be mechanically applied by a judge – who functions essentially like a judging machine – has been called 'vulgar formalism' and "it is not a view to which anyone today cares to subscribe" (Leiter 2010: 111). But while Dworkin rejects the argument such as described above, according to him "the layman's opinion, while deceptive insofar as it may suggest that adjudication is simple, is, so far as it goes, closer to the truth" (Dworkin 1963: 625).

Before delving straight into his argument, I would like to note three important areas he excepts from this early version of his argument: decisions of the supreme court, what happens when judges are explicitly granted discretion by the law, and most interestingly and problematically, the question of interpretation of the facts of the case – i.e. whether or not judges may or do interpret and how much information the can have about the case they are applying their rules and standards to (Dworkin 1963: FN 6).

That said, Dworkins argument proceeds by first distinguishing rules from standards. Rules are like game rules, and like the legal rules imagined by the layman invoked above: rigid, hopefully clear, only certain events are touched by rules, and if two rules conflict, there is no

[10] For example by D'Amato 2010 (189), or Solum 2010 (481).

real resolution to the dilemma because rules aren't weighted (Dworkin 1963: 629). Standards, according to Dworkin, are another matter. Anything may become of relevance under a standard, and most importantly, they do have relative weight – although that weight is often a matter of current public opinion. (Dworkin 1963: 632) He states:

> *"My purpose thus far has been to challenge the unarticulated premise of the argument from hard cases: the assumption that the only true standards are rules and that, if the explanations given by judges in hard cases are to be regarded as standards at all, they must be regarded as rules manques, deteriorated versions which cannot be applied without a measure of discretion. I suggest, although I cannot here demonstrate, that hard eases are decided in law, as in Policies, by the application of standards other than rules" (Dworkin 1963: 634).*

He does acknowledge the complexity of the decision-making process, but nonetheless contends that while this makes the job of the people charged with applying the law harder, it neither absolves them from the responsibility to find the right answer, nor does it deny the possibility of doing so (1975: 1082). An important point shining through in Dworkins work is that although laws as expressed in language will be vague – since language can sometimes be vague – but there are other sources of law that are actually independent from language, which can provide the formerly missing clarity. Therefore, vagueness in the language need not lead to indeterminacy in the law.

The legal reasoning required from the judge allows him or her to draw from three types of public standards (as opposed to private prejudice, which would turn their legitimate decision-making process into illegitimate choosing). Old sources are standards which "have already been used in a variety of cases" (635). New sources mean standards introduced by competent institutions of the community. Dworkin goes on:

> *"But [the judge] must sometimes recognize a third source: judgments of the community at large or some identifiable segment thereof. The court refers to such judgments when it rejects a particular result or rule as unjust, as well as when it more explicitly invokes the ideals of the society. To acknowledge this source is not to say that all legal standards must be popular, for it is only one source; nor to say that widespread community opinion must become law, for other standards may conflict which must prevail. There are two crucial questions that any society must face as to this source of law: How great a consensus is required before partisan position becomes community judgment? Shall any segment of the community be competent, on the basis of special interest, special ability, or special privilege, to form standards for the whole?" (Dworkin 1963: 635-636).*

The core of Dworkin's thesis is not only descriptive, but also normative. He claims that legal doctrine must be determinate because the people bound by the law have the right to a right answer. In an example invoking the imagined judge Hercules, he illustrates his ideal version of adjudication:

> *"He knows that the question he must decide is the question of the parties' institutional rights. He knows that if he decides wrongly, as he would do if he followed the ordinary man's lead, he cheats the parties of what they are entitled to have" (Dworkin 1975: 1108).*

It is plain that this 'rights thesis' is a normative statement, although it is not always completely distinct from the descriptive claims Dworkin makes. The normative nature of his project becomes even more apparent when he concludes his 1963 paper by admonishing the reader:

> *"The layman's conception of law depends upon the principle that the judge is never entitled to offer private prejudice rather than public standards as justification. If the profession and the public should be led on the one hand not to attempt, and on the other not to expect, adherence to this principle in difficult cases, that conception would be defeated. The misdescription in question is dangerous because it has the power to correct itself, by changing the thing described" (Dworkin 1963: 638).*

To critically assess – and ultimately reject – Dworkins thesis, I will develop a fourfold argument: I contend that while Dworkin talks about indeterminacy, he ultimately ignores vagueness. I also argue that his account creates an epistemological problem which he does not sufficiently address. On more of a side note, a critical critique of his argument could read that he treats rationality as a completely uncontested concept. A fourth and final problematic aspect is the reliance his argument places on the collective identity of the community under the legal system in question.

Even though his arguments are about indeterminacy, Dworkin ignores the paradox of the heap that vagueness creates. The problem is not that the law does not attempt to be determinate, the problem is, that since it is based in language, it can never be completely determinate because all vague concepts are susceptible to the Sorites Paradox and all language contains at least some vague concepts (or, some might argue, all concepts expressed in language are vague)[11]. When invoking the proposition that standards, unlike rules, can be weighed against each other, he fails to acknowledge the twin problems of immensurability and incommensurability as described above.

When Dworkin talks about 'very hard cases', he contends

> *"Judges often acknowledge decisions to be hard, and sometimes, though usually extracurricularly, admit that some decisions are based on 'guesses' or 'hunches'. I understand such a judge to mean that at times he believes one decision to be the right decision, but has difficulty in stating his reasons for so believing and has no great confidence that such reasons would convince others. This does not mean that he has abandoned the attempt to reach the right decision and is now simply reporting his personal preference" (Dworkin 1963: 637).*

[11] See discussion on ontological versus epistemic vagueness in section 2.1.1.

If no one, not even the judge himself, can know whether a decision made is 'the' right decision, how does this differ in practice from judicial discretion (and legal indeterminacy)? In a way, Dworkin here lets what he believes normatively right take over the description of what happens in reality.

One interesting point to keep in mind, that I will come back to in the next chapter, is that for a lot of his analysis, Dworkin seems to imply a strong collective identity, in the sense that he posits

> *"[Even when his decision is difficult or debatable, the iudex] intends his reasoning to be based on public, not private, standards of good argument, and the community receives and criticizes his decision on that basis" (Dworkin 1963: 632, see also page 635).*

This is one of the features which plays an even greater role in the international sphere than the national one, because the existence, and even the possibility of an international or global collective identity remains highly contested.[12]

2.2.2 All law is vague: radical indeterminacy and the Critical Legal Studies

On the other side of the debate, academics belonging to the Critical Legal Studies (CLS) are often said to defend an argument of radical indeterminacy. I hope to show in the following that while their conclusions may be radical, the indeterminacy thesis they purport actually is not as fundamental as it may seem. However, even though I found no one actually making the claim, the radical indeterminacy thesis is an interesting thought experiment in its own right.

Endicott (2001: 7) writes that the law is radically indeterminate if "no question of the application of a linguistic expression has a single right answer". He rejects the claim of radical indeterminacy, but more interestingly, he also proposes that no one actually makes that claim. Solum goes even further: as we have seen above, he claims that "[t]he law is indeterminate with respect to a given case if the set of legally acceptable outcomes is identical with the set of all possible results" (2010: 481). This last formulation, I think, is the one that goes too far even for the proponents of the critical indeterminacy claim.

This 'Indeterminacy claim' tends to be part of a broader deconstructivist/critical agenda: most articles are not exclusively or predominantly concerned with legal indeterminacy, but

[12] See e.g. Greenhill 2008: 251, critically Abizadeh 2005.

use it as only one of several examples for logical inconsistencies and legitimacy issues in the law.

The closest formulation to the radical indeterminacy argument – the one that is usually cited to illustrate it – comes from David Kairys (1983: 244). He writes:

> *"The starting point of critical theory is that legal reasoning does not provide concrete, real answers to particular legal or social problems. Legal reasoning is not a method or process that leads reasonable, competent and fair-minded people to particular results in particular cases".*

It is perhaps not surprising that this strong statement has incited debate, even though it still falls short of what Endicott (2001:7) defined as radical indeterminacy above. The following summary of the critical agenda by Singer (1984: 5) is slightly more cautious:

> *"Those of us associated with Critical Legal Studies believe that the law is not apolitical and objective: Lawyers, judges, and scholars make highly controversial political choices, but use the ideology of legal reasoning to make our institutions appear natural and our rules appear neutral. This view of the legal system raises the possibility that there are no rational, objective criteria that can govern how we describe that system, or how we chose governmental institutions, or how we make legal decisions".*

However, when looking specifically at his indeterminacy argument, he is certainly on the same page with Kairys. It reads like this:

> *"[W]e have shown that legal reasoning is indeterminate and contradictory. By its own criteria, legal reasoning cannot resolve legal questions in an "objective" manner; nor can it explain how the legal system works or how judges decide cases. [...] If legal reasoning is internally contradictory and therefore indeterminate, there are no objective limits on what judges or other government officials can do" (Singer 1984: 6-7).*

These sweeping statements need a bit of unpacking to fully understand and appreciate them, as well as to show how they fit into the problems posed by vagueness and indeterminacy.

According to Singer,

> *"A legal theory or a legal rule is determinate if it tells us what to do. A completely determinate theory or rule will leave us no choice; a relatively determinate theory or rule will constrain our choices, more or less narrowly, within boundaries. The claim that a legal doctrine is indeterminate means that the doctrine allows choice rather than constraining or compelling it" (Singer 1984: 11).*

He goes on to identify four criteria legal doctrine[13] must fulfil to be completely determinate: Comprehensiveness, consistency, directivity and self-revision (1984: 14). He continues to go through the criteria and explains why legal doctrine habitually meets none

[13] Legal doctrine is defined by Singer (1984: 11) as including both legal rules and arguments, including standards and principles. I will use the term in the same way in the course of this thesis.

of them. It is important to note that much of the indeterminacy Singer identifies does not have its root in vagueness. When he invokes the concept (without clearly distinguishing between vagueness and ambiguity), he also seems to be referring more to the common uses of the terms (like generality and incompleteness) instead of the technical sense laid out above.

Crucially he also clearly distinguishes 'indeterminacy' from 'arbitrariness', which is how the argument is often understood by critics of the Critical Legal Studies.

> *"Indeterminacy does not mean that the choices that are made by individuals are arbitrary or capricious. It also does not mean that all outcomes are equally likely to be considered or chosen by the decisionmaker [sic]. The indeterminacy of arguments is logically distinct from the arbitrariness of choices [....] Saying that decision-making is both indeterminate and nonarbitrary simply means that we can explain judicial decisions only by reference to criteria outside the scope of the judge's formal justifications" (Singer 1984: 20).*

This comment highlights the crucial distinction between legally inherent context and extra-legal context. Dworkin's argument above was not that legal determinacy depends exclusively on the written text of laws, but that the determinacy of them comes from circumstances which are inherent in the legal system and through and with it legitimized. The difference in position here is that Singer explicitly includes in the decision-making process of the judge what Dworkin calls private prejudice.

This then has effects on the predictability of legal outcomes, which also no longer depends on the notion of legal reasoning, but simply on personal convictions and patterns of behaviour. "Custom, rather than reason, narrows the choices and suggests the result" (Singer 1984: 25), or, in other words:

> *"Legal doctrines are always potentially indeterminate. Judges can move the line between rules and exceptions. [...] Ultimately, judges always have the power to revise rules. That judges may do these things, however, does not mean they will do them. Because judges participate in a legal culture that suggests how they are to act as judges, we can often predict how they will act" (Singer 1984: 22).*

In contrast to the above straw version of the argument, however, Singer does not actually claim that all legal doctrine is radically indeterminate, but rather that what he calls traditional legal theorists need too much determinacy in the law for their theories to be consistent.

> *"Legal doctrine is far more indeterminate than traditional theorists realize it is. If traditional legal theorists are correct about the importance of determinacy to the rule of law, then – by their own criteria – the rule of law has never existed anywhere. This is the real bite of the critique" (Singer 1984: 14).*

Here it can be seen that the indeterminacy thesis actually put forward by CLS scholars is not actually as radical as it may seem at first glance with regards to the actual indeterminacy

of law. For the critique to work, law does not need to be so indeterminate as to be senseless, nor does all law need to be indeterminate in all cases. In fact, Kennedy (1997: 169) argues that the proposition that law is determinate or indeterminate only makes sense with regards to a specific case, in a specific context. However, context dependence alone is not the solution to the critical indeterminacy problem, as Peller points out here:

> *"One response to the inadequacy of author's intent as a source of meaning is to conclude that meaning is contextual and can be determined by specifying the context of a text. This approach supposes that the context "exists" around the text, to be discovered by the interpreter as the source for the meaning of the text. But the problems of indeterminacy are not avoided by this approach. Any attempt to fix the meaning of a text by the specification of context runs up against the problem that any given context is open to further description. Context does not exist somewhere. Context is constructed by the interpreter according to her calculus of relevance and irrelevance. A particular description of the context involves screening the text through representational terms used by the interpreter" (1985: 1173).*

Peller here points out quite a large flaw in the attempt to solve the indeterminacy debate by contextualization. Since the relevant context is as liable to differing interpretations as the text itself, the question of which bits of context to take into account will become all the more difficult the more indeterminate the law itself is.

It is easy to show in which way the most radical version of the indeterminacy claim falls short. Solum does it by finding simply one case in which one outcome of a specific case is definitely not legal (2010: 482). Endicott (2001: 10) goes another route when he says that radical indeterminacy claims about language are nonsensical, since they ultimately apply to themselves as well, meaning that the claim 'no language has any kind of determinate meaning' is nonsensical in and of itself, because as soon as someone can understand what it says, it has been proven false.

Arguments put forward by most proponents of CLS are much harder to criticize substantially within the scope of this thesis, because their argument relies on much more than indeterminacy caused by linguistic vagueness. Luckily, however, there is also no need: if the critical argument is not that language (and with it, law), is radically indeterminate, and thus completely meaningless, then the problem on which legal provisions are clear and which are indeterminate – and to what degree – is as relevant as ever. As it turns out, the radicalism of the critical argument seems to lie not in absolute indeterminacy but in their critique of the existing legal doctrines.

2.2.3 Looking for the middle ground

After I have shown the determinacy claim by Dworkin and the supposed 'radical' indeterminacy claim often attributed to CLS, I now turn to more moderate versions of the argument. Specifically, I will discuss the Core/Penumbra distinction introduced by Hart (1961: 123), and then show some of the debate on the differences between law and language, which is evidently of particular interest in the context of this study.

The core/penumbra distinction

Postema (2011: 226) identifies several axes along which indeterminacy claims may be more or less accurate: they can be global (meaning touching all areas of law) or local in scope, they can be limited or radical in their indeterminacy (touching on the difference between underdeterminacy and indeterminacy pointed out by Solum in section 1.2.2.5), and they can be counter-factual or narrow in relation to their imminence: counterfactual indeterminacy claims hold even "when [law] is determinate under existing conditions, if it would become indeterminate, if certain (relevant) conditions, supposed to be external to the law itself, were to change" (Postema 2011: 226). In this section, I will mainly be concerned with the limited versions of indeterminacy, because this is the one where the phenomenon of vagueness are most relevant and it is also the main focus of the existing debate.

Most famously, Hart (1961: 123) suggested the idea of a 'core of certainty' and a 'penumbra of doubt'. He starts by describing the premise that law based on precedents is indeterminate, while law based on general rules is determinate to a much higher degree. He then sets out to dismantle the distinction:

> *"Much of the jurisprudence of this century has consisted of the progressive realization (and sometimes exaggeration) of the important fact that the distinction between the uncertainties of communication by authoritative example (precedent), and the certainties of communication by authoritative general language (legislation) is far less firm than this naive contrast suggests. Even when verbally formulated general rules are used, uncertainties as to the form of behaviour required by them may break out in particular concrete cases" (Hart 1961: 123).*

Already, he makes clear how indeterminacy in law is inseparably interwoven with the language it is expressed in. His position on the issue gets even clearer as he continues:

> *"In all fields of experience, not only that of rules, there is a limit, inherent in the nature of language, to the guidance which general language can provide. There will be plain cases constantly recurring in similar contexts to which general expressions were clearly applicable … but there will also be cases where it is not clear whether they apply or not" (Hart 1961: 123).*

He then goes on to make the well-known assertion that rules have a 'core of certainty' (123) and a 'penumbra of doubt' (123). Explaining the former, he makes an interesting observation on the dependence of rules on the social context they are conceived in:

> *"the plain case, where the general terms seem to need no interpretation and where the recognition of instances seems unproblematic or 'automatic', are only the familiar ones, constantly recurring in similar contexts, where there is general agreement in judgements as to the applicability of the classifying terms" (Hart 1961: 123).*

However, Harts deliberations on vagueness in the law do not stop at the positive level. He ascribes a normative value to vagueness, possibly in response to widespread understandings that all uncertainty in law undermines it legitimacy: ("Put shortly, the reason is that the necessity for such choice is thrust upon us because we are men, not gods" (Hart 1961: 125)). Complete determinacy is impossible, but also normatively undesirable, because of 'ignorance of fact' and 'indeterminacy of aim'. Trying to formulate exact laws as a 'vice' (126), arguing for balanced approach (127), according to varying degree of possibility to predict future in different areas.

Hart even touches on the existence of higher-order vagueness:

> *"Canons of 'interpretation' cannot eliminate, though they can diminish, these uncertainties; for these canons are themselves general rules for the use of language, and make use of general terms which themselves require interpretation" (Hart 1961: 123).*

This is the level of detail he goes into on this subject though, so the more intricate complexities of the phenomenon remain somewhat unclear in his work.

Greenawalt (1992: 34) formulates his own version of the indeterminacy claim in the following way:

> *"many legal questions have determinate answers that (1) would be arrived at by virtually all those with an understanding of the legal system, and (2) are unopposed by powerful arguments, consonant with the premises of the system, for contrary results".*

While this version is certainly not contradictory to a moderate indeterminacy thesis, the emphasis clearly lies on the argument that actually, most legal rules are 'determinate enough' for all practical accounts.

Endicott equally endorses a 'middle position' in the question of whether law is indeterminate or not. But he also offers a more nuanced perspective on the desirability of indeterminacy in the law.

> *"We cannot say in general that even a very vague legal rule represents a deficit in the rule of law. But vagueness is a deficit when it enables authorities to exempt their actions from the reason of the law, or when it makes it impossible to*

> *conceive of the law as having any reason distinguishable from the will of the officials" (Endicott 2001: 5).*

According to an argument he makes in a 2005 article, it is the not the vagueness, but the arbitrariness of a norm that has to be minimized. There have to be clear reasons why a norm should be obeyed, and with implemented and accepted norms, these reasons are inherent in the norm itself. The interesting part is that according to him, arbitrariness can in some cases be reduced by vagueness, while other cases are better served with precise norms. This perspective is interesting because it elevates vagueness above the status of a necessary evil to a feature of law with a value of its own. According to his argument, precision has guidance value (i.e. the advantage of people knowing exactly what they are allowed to or have to do) and process value, which means that executives of the law also are bound by precise standards they cannot bend according to their own preferences. Vagueness, on the other hand, has what Endicott calls private ordering value, meaning that people can judge on their own whether or not it is reasonable for them to apply a certain law, power allocation value, which enables law enforcers and judges to execute a law in a way corresponding to its purpose, and fidelity value, which designates the capacity to regulate issues that are impossible to order through precise provisions (Endicott 2005: 45).

The differences between law and language

One more recent area of investigation within the context of the indeterminacy debate concerns the relationship between law and the language the law is written in.

Schauer (2011) asks whether there is an extra quality of indeterminacy – different from the vagueness or open texture of all language – in the law. He comes to the conclusion that

> *"The open texture of language thus produces open texture in law, but this conclusion is neither startling nor particularly interesting. Insofar as language is open-textured, it follows that any use of language is pro tanto open-textured. Law is thus open-textured, but so too are novels, theatrical productions, everyday conversations, restaurant menus, and the instructions I give to the automobile mechanic" (Schauer 2011: 7-8).*

He does not convincingly, or indeed at all, make the case why, to be of interest, legal indeterminacy must be different from the linguistic one. On the one hand, I disagree that there is no indeterminacy of law which, in Schauer's words, "exist[s] within the area of linguistic determinacy" (2011: 14). The fact that far from all of the indeterminacy invoked by the arguments of Dworkin and the CLS described above is linguistic should be a first hint in that direction. Endicott goes even further when he proposes (though later comes to reject

the position) that legal indeterminacy may be completely detached from linguistic vagueness:

> *"But once Hart admits the distinction between the application of the words of a rule and the application of the rule, he seems to have conceded what Dworkin claims. Given the distinction, it seems that no claim about indeterminacy in the law follows from the vagueness of, for example, 'serious distress'. Linguistic indeterminacy does not entail legal indeterminacy" (Endicott 2010: 163).*

But even more interestingly, even if we were to accept Schauer's proposition, that would not make indeterminacy uninteresting for legal theory. Because law has certain features that distinguish it from general language, and those make the vagueness inherent in all language more problematic when it is found in legal texts. The consequences can be very different. Even though all language is performative, legal language is so in a very immediate way (see section 4.2.2 for further discussion of this). That is to say, even if the indeterminacy of law and language were the same, the *consequences* of this indeterminacy would still very likely be extremely different. As Liebwald (2012: 392) observes: "Unlike every day conversation, the impact of vagueness in law is a different one, due to the legally binding force and the associated legal consequences."

Solum (2010) strikes at the heart of this issue when he talks about the interpretation/construction distinction. He distinguishes semantic content from legal effect (2011: 111) and points out the differences between the two concepts. According to him, language can be interpreted, while legal effect must always be constructed. There are certainly cases when semantic and legal content overlap: these cases are then seen as determinate, or 'easy' cases. But "[w]hen a legal provision is vague, then semantic content underdetermines legal content." (Solum 2010: 117).

2.3 Conclusion: the importance of vagueness for legal theory

As I have shown in his chapter, legal theories on indeterminacy are concerned with very global, absolute problems. Theorists and philosophers are asking deep questions about the nature and legitimacy of all law. But when it comes to a head, and I have successfully established, that yes, vagueness exists in (any) legal system, and no, the legal system is not utterly undermined and delegitimized by this fact, the question now becomes: so what? If vagueness is an inevitable fact of (legal) language, but does not fatally hurt the system, why think about it at all?

Here we come to a set of smaller questions concerned with the role of vagueness *within* more or less functioning legal systems. I argue that while no legal system can ever be completely vague (then it would not exist at all) or completely determinate, legal systems as

well as legal rules or provisions can vary in the degree to which they are vague. And these differences, with which the empirical part of this thesis will concern itself, can have important consequences, even if they are below the level of 'complete loss of legitimacy' that was so often invoked in the arguments this chapter is concerned with. But while the exposition in this chapter is first and foremost meant to provide a conceptual background and some context for the empirical questions addressed below, I do think that they, too, could profit from a confrontation with empirical studies even though these are by necessity much more limited in scope. I will provide at least some insights into which theoretical assumptions are supported or made more plausible by my empirical findings, and perhaps suggest a small alteration in some cases.

3 The IR perspective: why is there vagueness in international agreements?

In this chapter, I will discuss the causal assumptions that have been made by scholars in the field of IR about the effects and causes of vagueness in international agreements. In order to do so, I will first briefly sketch some aspects of the setting in which such agreements are made. I will not cover any of these topics in all their scope and complexity, as this would by far exceed the possibilities of this thesis. Instead, I will focus on those aspects that affect the causal assumptions made about vagueness which I will describe and assess in detail in the second part of this chapter. I will show that they can only be assessed in any meaningful way by putting them in their respective theoretical contexts. While I borrow from several IR theories in this chapter, the central focus lies on Regime Theory. It is well suited to provide the main theoretical framework of this thesis because it emphasises actor's interests and the dynamics of cooperation. I see regimes as constituted by and providing a background for the international agreements whose language I analyse in the context of the dissertation.

This chapter is aimed at two things: First, give an overview over the literature that has touched on the topic of vagueness in the field of IR and IL, and place it in a wider theoretical context. Second, I intend to show that (1) there are numerous important assumptions made about the vagueness in international agreements which (2) cannot be embraced or refused if it is unclear when a legal document is actually vague or not. While the postulated causes and effects of vagueness are diverse, they all have in common that in order to sustain or refute them, it must be possible to compare the vagueness of treaty texts.

This chapter proceeds in two steps. First, I provide contextual background in four different areas: characteristics of regimes, the two parts of the international legal process – negotiation or the making of IL and its implementation, mechanisms and functions of IL as well as the different characteristics of rules. After a brief note on causality and some remarks about the normativity of vagueness apparent in much of the literature, I then proceed to discuss causes and effects of vagueness as proposed in the relevant literature. I conclude by putting the causal claims in the broader context of this dissertation.

3.1 Providing context for international agreements

This section is aimed at providing theoretical background in four different areas essential to the understandings of causal claims about vagueness in IR. Regimes provide the conceptual framework in which international agreements are written. The twin processes of negotiating and implementing – and therefore interpreting – international legal agreements produce and use the agreements. Understanding them can provide critical insights into which, and why, language features in them. A broader picture on how IL works is necessary to understand the effects that vagueness may have on it. Finally, rules themselves have a particular relationship to language – and therefore to vagueness – which will be touched on briefly.

3.1.1 Regimes

Regime theory tries to explain why and how cooperation in the international system emerges. The question of how to define regimes has not been uncontroversial. Krasner (1982: 186) understands regimes as "implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations". Hasenclever et al (1996: 178) have argued that the phrase *around which actors' expectations converge* does not describe how we can identify a regime in a given issue-area and identify three different approaches to do so. According to Keohane (1993: 26-29), regimes should be treated as existent if an explicit rule is agreed upon by actors and is embodied in a formal treaty or document. Regardless of the exact definition, all of these authors argue that regimes establish and shape the context in which international agreements are negotiated, adopted, and implemented. Therefore, they play an important role in understanding these agreements.

In the following, I will illustrate some characteristics of regimes. While it is impossible here to summarize the enormous amount of literature on regimes inside and outside of

regime theory, I will draw attention to some of the characteristics central to international regimes which provide pertinent background to the causal assumptions made about vagueness in international agreements. In particular, I will briefly outline issues related to the stability and development of regimes over time, their inherent complexity, their issue specificity, the role regimes play in providing forums for communication and the sharing of information as well as the specific problems they face with regard to principal-agent issues.

Regime Development

Once regimes have been established, one characteristic which distinguishing them from other forms of state cooperation is their continuity over time. Hasenclever et al. (1997: 185) claim that one of the basic functions of international regimes is that they "stabilize mutual expectations". According to liberal theorists, that is one of their main assets: By rendering any form of cooperation into a repeated pattern, the incentive to cooperate grows as trust is built and the potential of defections becomes more easily detect- and punishable. According to game theoretic models, (indefinitely) repeated interaction solves many central problems like the prisoner's dilemma (Goldsmith/Posner 1999: 1125). By institutionalizing interactions, transaction costs are reduced as the amount of information that has to be collected and reviewed before each decision to cooperate decreases (Keohane 1982: 339).

The continuity of regimes also has other, sometimes less intentional effects. For one, once issues become institutionalized, they develop dynamics of their own. Levy et al. (1995: 287) get to the heart of this effect when they write

> *"We can show that once established, regimes display a persistence and robustness that cannot be explained fully in terms of the conditions under which they formed in the first place".*

Additionally, repeated interaction may become normative in character over time:

> *"Patterned behavior, originally generated purely by considerations of interest or power, has a strong tendency to lead to shared expectations. Patterned behavior accompanied by shared expectations is likely to become infused with normative significance: actions based purely on instrumental calculations can come to be regarded as rule-like or principled behavior" (Krasner 1982: 202).*

Since every institution has an interest in persisting, there is a lot of incentive to stabilize and expand the existing communication and actions taken within the system (Levy et al. 1995: 290).

A related, but distinct consequence of the longstanding nature of regimes is that they develop to be seen as institutions in their own right by the public. The accountability

becomes at least slightly removed from the individual participating states towards the regime itself, including all news coverage and PR questions. In that way, regimes can both spiral out of the control of their principals (see Zürn et al. 2006: 11) or become a smokescreen for inaction on the part of their member states. At the same time, this dynamic can be beneficial for the solving of issues the regime is supposed to address.

A third feature concerns the timing of progress – or measures taken in regimes – over time. For Setear, regimes have the advantage of adding stability to state interactions by making repeated, institutionalized exchanges possible. He discusses the specific case of the 'Convention-Protocol approach', stating

> *"an initial "convention" identifies the subject matter of the relevant discussions, creates an administrative and procedural machinery, and sets forth vague substantive principles to guide future negotiations. In later treaties (the "protocols"), the parties, under the general framework erected by the convention, undertake the specific obligations that constitute significant limits on their behavior" (1996:84).*

He also gives reasons for the effectiveness of this approach:

> *"In the framework conventions, the parties therefore Seek [sic] a cooperative solution to a problem with respect to which they have not cooperated much previously. The evolution of cooperation in international relations does not occur overnight, and the unfamiliarity of the issue area may mean that the substantive outlines of a solution are unclear–especially when the scientific or engineering information underlying rational policymaking is absent or rapidly changing" (Setear 1996: FN 316).*

These examples show that regimes develop and change over time, affecting the interactions and behaviours of their members in the process.

Complexity

Another feature of international regimes is their complexity (Young 1982: 278). It results not only from the often large number of member states, but also from the range of issues discussed. Krasner (1982: 196) illustrates that complexity can be both a cause and a consequence for regimes, which are "designed to manage complexity and complexity increases with interconnectedness".

While regimes are organizationally separated from other thematic issues, its member states are not – and it is not unusual that conflict between states originating elsewhere will also have a bearing on the regime (for a discussion on the complexities of intersecting international regimes, see Alter/Meunier 2007). The Complexity of regimes often surpasses the ability of individual negotiators to overview all at once. From this fact can arise

interesting dynamics none of the participants could have foreseen. As Young (1982: 287) puts it,

> *"In very large systems, it is hard for the participants to play a meaningful role in the negotiation of regimes, and eventually even the idea of explicit consent will begin to lose significance".*

Therefore, the issue of complexity also relates to the question of unintentional outcomes of international negotiations. Even assuming a high degree of rationality among negotiators, which of course is the case from several, but by no means all theoretical perspectives, complexity can lead to consequences – and treaty formulations – unintended by any of the negotiators. Additionally, it creates problems for states with less resources – both in personnel and in finances – because to try and keep track of all the issues and under–issues experts may be needed in many fields at once.

Issue Specificity

One of the core characteristics of the current systems of international regimes is that they are issue specific. According to Levy et al. (1995: 291): "International regimes commonly emerge in response to particular problems".

This creates an interesting governmental structure, because when different regimes issue conflicting legislation, there is no clear legal way of deciding which rule has precedence. The growing importance of issue-specific regimes has led to the international legal structure being described as fragmented (see for example Dupuy 1999: 792 and Fischer-Lescano/Teubner 2004). This issue may be amplified by the fact that membership in regimes can vary – even though there is substantial overlap, not all states are represented in each regime and thus outcomes can vary even more widely. In practice, certain regimes are usually given precedence and others will then amend their provisions accordingly. Still, conflicting legislations can potentially exist for a long time until the discrepancies are noticed, especially when relatively minor issues are concerned (Ranganathan 2014: 5-6).

While the broad theme of a regime stays stable over time, goals can certainly evolve. For example, the non-proliferation regime may be seen as having extended its original intent by becoming the forum in which the abolition of all nuclear weapons is discussed. Other times, under-issues may gain importance and break away into a regime of their own. Most commonly, regimes gain in scope as they grow older (see Levy et al. 1995: 290)

Communication, Learning, and Information

One of the main functions of international Regimes is that they can be forums of communication for their members.

> *"Another means of reducing problems of uncertainty is to increase the quantity and quality of communication, thus alleviating the information problems that create risk and uncertainty in the first place" (Keohane 1982: 346).*

Since the interaction already takes place in an institutionalized setting, potential conflicts may be discussed directly and without much additional cost. Over the course of negotiations within a regime, a sort of internal language develops – with shorthand and acronyms and slang that may make it harder for outsiders to understand what exactly is going on, but it also facilitates communication between the initiated[14]. Describing problems in a shared language can be an important factor in finding a solution, and some of it is definitely happening in most regimes. As Haas (1989: 377) puts it,

> *"Regimes are not simply static summaries of rules and norms; they may also serve as important vehicles for international learning that produce convergent state policies".*

This feature of regimes can be understood in conjunction with the acculturation mechanism of IL discussed below.

On a more technical level, regimes vastly facilitate the exchange of information between parties (Keohane 1982: 344). Most regimes actively collect the information relevant to their causes from member states, but the states also get the opportunity to exchange information through topical, institutionalized channels, and, most importantly, a relatively impartial organization overseeing the information, so they can be relatively sure it is correct.

Two-level games

Putnam (1988: 427) argues that domestic and international politics are entangled. He suggests two-level games as a metaphor for the central dynamic of international regimes:

> *„At the national level, domestic groups pursue their interests by pressuring the government to adopt favorable policies, and politicians seek power by constructing coalitions among those groups. At the international level, national governments seek to maximize their own ability to satisfy domestic pressures, while minimizing the adverse consequences of foreign developments. Neither of the two games can be ignored by central decision-makers, so long as their countries remain interdependent, yet sovereign"(Putnam 1988: 434).*

[14] As illustrated by some aspects of the nukespeak debate, see for example Hilgartner et al. 1982.

This dynamic accounts for some of the complexity discussed above, as one actor may have different sets of interests within the same regime.

According to Zangl (1994: 281-2), two-level games underlie not only the functioning, but also the emergence of regimes. Interest groups within a state influence the creation of a regime just as much as shared interests of the states do. Müller has emphasized that cultural differences and the character of international politics as a two-level game hamper the emergence of a shared life world (Müller 2004: 419–425), which could in turn foster cooperation. Because of their mediated decision structures, international regimes also give rise to the principle-agent problematic. The government of a state is technically acting on behalf of its people (the principle) but may well pursue diverging interests (for a detailed discussion of international organizations in the principal-agent Framework see Elsig 2011: 498).

In summary, regimes provide stable environments for state interaction and negotiation over time. They create complex structures while at the same time helping to manage complexity. While usually issue-specific – and sometimes quite focused – regimes interact and overlap on account of their often similar membership. In addition to their formal outcomes, they provide a valuable space for developing shared understandings and languages, not lastly by means of the exchange of information between members. Due to their mediated decision-making structures, regimes are also prone to principal-agent problematics. All these characteristics shape the environment in which international agreements are negotiated and concluded, and they impact the causes and effects of the language which is used to write them. I will draw on the background sketched above when I discuss the causal assumptions made about vagueness in international agreements in the second part of this chapter.

3.1.2 Negotiation and implementation of IL

IR as a discipline is particularly interested in two aspects of IL: Negotiation, when international agreements are conceived of and IL is made, and implementation, which covers the time after the conclusion of the treaty and the main concerns are compliance and the incorporation of international into national law. Since assumptions on the causes and effects of vagueness in international agreements vary significantly according to the part of the process they relate to, I will in the following briefly outline a few of the most pertinent characteristics of each of them.

Negotiation

The negotiation phase of international agreements involves the actual drafting and creation of international treaties, which usually happens at international conferences. According to Jönsson (2002: 217), "In the international arena, dominated by sovereign states, negotiation is the primary and predominant mode of reaching joint decisions". When looking to understand how treaties are drafted, and which sort of language is used in a treaty and why, it is therefore crucial to study international negotiations. Even though international negotiations can potentially include many types of actors, like NGOs, corporations, or intergovernmental actors, in the context of this thesis I focus on regimes in which state representatives are the only official participants of negotiations.

While international negotiations have been understood in terms of fairly simplified models where rational actors try to maximise their clearly defined interests (see e.g. Fearon 1998: 269-70), the reality tends to be much more complex. For example, the negotiators' interests are not always clearly defined or coherent (see section 3.1.1 for a discussion of two-level games). The possibility of norm development (Finnemore/Sikkink 1998), or the effects of different institutional settings (Underdal 1983: 193) are other examples of areas where simpler explanations struggle, because they fail to include imperfect rationality and are based on too few variables.

For Zartman, complexity even is the defining quality of multilateral negotiations (see section 3.1.1 for a discussion on complexity in regimes):

> *"If the basic analytical question for any negotiation analysis is 'How to explain outcomes?', the question is answered for multilateral negotiations in the form of another question: How did/do the parties manage the characteristic complexity of their encounter in order to produce outcomes?" (1994a: 222).*

This complexity is rooted in their nature as multi-party, multi-issue, multi-role negotiations with values, parties and roles capable of varying throughout the negotiations (Zartman 1994b: 4-6).

In an attempt to reduce complexity and increase their own negotiating power, parties often form coalitions. Pěchouček et al. (2000: 161)

> *"define a coalition as a set of agents, which agreed to fulfill a single, well-specified goal. Coalition members committed themselves to collaborate on the within-coalition-shared goal".*

In the international sphere, some coalitions, like the G-77[15], or AOSIS[16], are well established and stable, working together in several different regimes. Other coalitions fit the above definition more closely and are built more spontaneously around specific issues. Dupont (1996: 49) points out that while coalitions may well be useful, they also bear the risk of increasing polarization. For Zartman, not only state parties can be said to form coalitions:

> *"But coalition is a mechanism not only available to the many parties but also applicable to the many issues. Because the goal of negotiations is to arrive at decisions on issues, it is necessary to reduce their complexity and make them, as well as the number of parties, manageable. Packaging, linkages, and trade-offs – the basic devices of the negotiation process – are all ways of making coalitions among issues, interests, and positions" (Zartman 1994b: 6).*

In essence, multiple issues are often discussed in relationship to one another, in order to enable compromises. This can, as Zartman states here, reduce complexity, but it can also heighten the bar for new negotiators who do not necessarily know all the issues at stake.

Framing and Agenda-setting are methods often employed by NGOs and other non-state actors in order to influence negotiations in which the only formal participants are state representatives. "The concept of framing captures the processes by which actors influence the interpretations of reality among various audiences" (Fiss/Hirsch 2005: 30). Initially, the concept of frames was developed by Erving Goffman who described them as "*schemata of interpretation*" used by individuals to "*locate, perceive, identify, and label*" events in people's life and in the world (Goffman 1974: 21).

The relationship between diplomats and the governments who employ them can lead to an additional layer of complexity. As Pfister et al. (unpublished paper) have shown, negotiators, especially in long-standing regimes like the climate regime, develop a deeper understanding of the issue at hand that can often differ significantly from the position of the country they are representing. This can lead to several problems: a discontent among negotiators, making them less likely to come up with creative solutions, or the need to repeat all the preparatory work they do when the heads of states arrive at the last stages of a conference.

A lot of the negotiation takes place over very specific instances of language. Disputed provisions or formulations of a negotiation text are usually put into brackets, and negotiators then work together to alter the formulations such that everyone can agree on it (FIELD 2006: 46-7). Negotiation chairs often play a central role in these drafting procedures (Tallberg

[15] Group of 77
[16] Alliance of Small Island States

2010: 242), attaching influence to roles that might be considered ceremonial by some. When the bracketed texts are available to the public, they can provide valuable insights into the current state negotiation process (see e.g. Depledge 2000).

Implementation and compliance

After negotiations have ended and an agreement has been concluded, the second part of the international legal process begins. The question then is about the implementation of or compliance with the agreement: Are the concluded agreements actually changing the parties' behaviour? This question presents major conceptual and methodological challenges, which I will briefly address in this section. First, I will attempt a disambiguation of several terms which intersect with each other in the implementation phase of an international agreement. These are compliance, implementation, effectiveness and enforcement. I will then briefly touch on the methodological challenges researchers face when trying to tackle the implementation of IL empirically.

On the subject of compliance, Young (1979: 104) states that

> *"Compliance can be said to occur when the actual behavior of a given subject conforms to prescribed behavior, and non-compliance or violation occurs when actual behavior departs significantly from prescribed behavior".*

This definition seems relatively straightforward. However, Kingsbury (1998: 346) makes the argument that

> *"the concept of 'compliance' with law does not have, and cannot have, any meaning except as a function of prior theories of the nature and operation of the law to which it pertains. "Compliance" is thus not a free-standing concept, but derives meaning and utility from theories, so that different theories lead to significantly different notions of what is meant by "compliance".*

The question of whether states comply with IL has been a crucial one, even leading scholars to dispute that IL is actually law. This claim is made especially often from the perspective that to be considered law, rules must possess the capacity to be enforced (see D'Amato 1984: 1293). Regardless of whether or not one believes IL to be law, the question of enforcement seems inextricably linked with questions about the nature of law in general. Oppenheim (2005: 4), for example, states that "the characteristic of rules of law is that they shall, if necessary, be enforced by external power".

When scholars describe the enforcement possibilities of IL, it is not unusual for them to start with an account of the deficits:

> *"International law has no international police force to oversee obedience to the international legal standards to which States agree or that develop as*

> *international standards of behaviour. Similarly, there is no compulsory enforcement mechanism for the settlement of disputes" (Stratton 2009: 2).*

According to Chayes and Handler Chayes (1995: 2), enforcement in the realm of IL can consist of either economic or military sanctions. They take a critical view of assessing IL primarily in terms of its enforcement capabilities. As we shall see in section 3.1.3, the significance of enforcement depends on the mechanism by which one believes IL to function.

Implementation means the incorporation of IL into domestic law. The process involves many challenges, because international agreements are written in a specific environment that is different from domestic legal systems. It might be necessary to translate the text into another language, with all the challenges that entails. Moreover, the larger context of the specific legal system will be different from the international legal system. Implementation may or may not be a prerequisite for compliance – it is possible for a country to comply with IL without incorporating the specific provisions into domestic law, although usually at least some measure of implementation takes place.

The effectiveness of an international regime may well be related to its rate of compliance, but the two terms nevertheless should not be confused. Mitchell (2001: 221-2) illustrates the scope of the term effectiveness by saying that some authors understand it to mean compliance, others efficiency, and yet others as reaching the objective of the agreement. He goes on to argue that

> *"The question 'Is this regime effective?' is often simply a shorthand for 'Did this regime accomplish certain goals?' Answering the question, therefore, requires the analyst to define, implicitly or explicitly, the goals against which she will evaluate the regime's performance".*

Depending on which goal is taken as the baseline, a regime could have very high rates of compliance and yet not be especially effective, simply by not demanding much. On the other hand, even when compliance is relatively low, the effectiveness could theoretically still be on the high side if the terms of the regime are extraordinarily demanding (though there must still be some compliance).

Even after disambiguating and narrowing down the definitions of all these terms, the question of compliance with IL remains a challenging one, not least because of the extreme methodological difficulties of empirical research in the field.

> *"Theoretically it is possible to demonstrate the agreement or coincidence between a certain state behaviour and a certain rule of international law. […] it would be difficult to demonstrate that the state adopts a particular policy because it intends to comply with the rule of international law" (Onuma 2003: 129).*

The effectiveness of international agreements can only be approximated by contra-factual reasoning as well, since once an agreement is made, it becomes impossible to know what the world would have looked like without that agreement. Of course, information can be gained by looking at similar cases and parallel developments, but in the complex world of IR that is no easy or certain undertaking.

To summarize, the two parts of the international legal process – the negotiation phase and the compliance or implementation phase – differ significantly both in their subject matter as well as in the ways they are usually approached, theorized and researched. In light of these substantial differences, and often segregation in research, it is important to keep in mind that they represent two sides of the same coin and are extremely interrelated. Negotiators draft agreements with a view to their implementation, and just because an agreement or international rule exists does not mean the negotiation about it ends – far from it. It is equally fallacious to posit a differentiation in terms of a political and a legal part of IL. The making of law is a deeply legal issue and questions of compliance and implementation are often highly political. Nevertheless, the different perspectives in theoretical approaches highlight important characteristics of both parts of the legal process, which I will use in sections 3.2.3 and 3.2.4 to put the causal claims about vagueness into context.

3.1.3 How IL works: mechanisms and functions

To determine what kind of impact vagueness may have on an international agreement – and what it may be caused by – it is essential to be clear on one's perspective about how IL works. Depending on the point of view, language in general and vagueness in particular may play very different roles. In this section, I will first take a look at four different functions IL may fulfil, and then pay attention to four mechanisms by means of which it can accomplish them. Of course, functional approaches or those based on mechanisms are by no means the only perspective one can use to examine IL. However, I contend that they do, one, present an adequately broad spectrum of ideas on how IL may work and, two, are significant for the question of vagueness in the law.

Functions of IL

In a first step, I will look at the functions IL is supposed to fulfil. Onuma (2003: 108) identifies four of them: "the binding, communicative, value-declaratory, and justifying and legitimating functions."

The binding function is perhaps the most obvious one: that the rules made by the law will be binding to, and thus complied with by the signatories. Onuma himself declares it law's primary function, although he does note a point of contention, namely that bindingness is a property, not a function of law (Onuf 2010: 317). The difference between the two is that in the first case, non-binding rules may also be legal, while in the second case they would not be. In any event, the issue of bindingness is clearly central to the concept of IL.

The communicative function of IL mirrors the function of international regimes as forums for communication (see section 3.1.1). The communicative function,

> *"is to embody and express shared ideas and understandings of the constitutional structure and legitimate aspirations of international society to its diverse members" (Onuma 2003: 134).*

For Sunstein (1996: 2024), the expressive – what Onuma calls value-declaratory – function of law is the "function of law in 'making statements' as opposed to controlling behavior directly". Van der Burg (2001: 41) elaborates on this point:

> *"Because legislation expresses certain values and communicates them effectively, these values can be taken as a common point of reference in interpretation and communication processes".*

Going from a mere communicative to a sort of socializing function, Wessel (2010: 136) argues:

> *"International law is but one important way in which the process of socialization, and the formation of intersubjective agreement, occurs amongst the speech community of diplomats. In the same way that communication is one of the leading factors of socialization of the individual in his state, the communicative medium of international law socializes the state vis-a`-vis other states".*

Lastly, the justifying function is the most openly political function of law, which is in this case serving as an instrument to legitimize state practice (Onuma 2003: 136). In Koskenniemi's terms, this may be seen as the apologetic character of IL (Koskenniemi 1990: 11).

Mechanisms of IL

As we have seen, IL can be said to have different functions. It may fulfil these functions predominantly by means of four mechanisms. Ginsburg and Shaffer (2010: 11) identify them

as reciprocity, coercion, persuasion, and acculturation. Goodman and Jinks (2004: 630) make similar distinctions, but only recognize three mechanisms: coercion, persuasion and acculturation. Reciprocity follows the ultimate logic of bargaining: Parties to an agreement can trade legal provisions between themselves, or comprise on issues until they find a solution advantageous to all. Coercion is the means by which law is most often presumed to fulfil its purpose. It relies on sanctions for any behaviour that is not conforming to the law. Persuasion, according to Goodman and Jinks (2004: 635), is "the active, often strategic, inculcation of norms." Finnemore and Sikkink (1998: 914-5) explain when persuasion is likely to happen:

> *"The persuasiveness of a normative claim in law is explicitly tied to the "fit" of that claim within existing normative frameworks; legal arguments are persuasive when they are grounded in precedent, and there are complex rules about the creation of precedent–such as which judgments trump which and how the accretion of judgments is to be aggregated over time"*

Acculturation is defined by Goodman and Jinks (2004: 638) as

> *"the general process of adopting the beliefs and behavioral patterns of the surrounding culture. This mechanism induces behavioral changes through pressures to assimilate– some imposed by other actors and some imposed by the self".*

The mechanism through which one presumes IL to function has consequences for one's view of compliance.

> *"in fact, powerful states sometimes find they can make more progress through measures to build capacity and otherwise assist weaker states in their efforts to implement the terms of international regimes than they can through threats or sanctions intended to force weaker states to comply with the terms of international regimes" (Levy et al 1995: 283).*

Summing up, IL may well have functions beyond binding states to its statutes, and it likely functions through mechanisms other than coercion. According to the functions and mechanisms one focusses on, causal assumptions about the sources and effects of vagueness in international agreements change (see sections 3.2.3 and 3.2.4).

3.1.4 Rules

A last topic I want to sketch in this overview are the different conceptual types of rules that make up IL. They are important because different types of rules may serve different functions or rely on different mechanisms of IL. Additionally, different types of rules may call for different languages. Finally, it is entirely possible that different types of rules become more or less important during different stages of a regime life-cycle:

> *"International Rules are prescriptive statements that forbid, require or permit some action or outcome. One of the three deontic operators – forbid, require, permit – must be contained in a statement for it to be considered as a rule" (Ostrom 1990: 139).*

The quote clearly shows the relationship between performative language (as discussed in section 4.2.2 and rules. Rules that forbid are usually what we think of first when we hear the word 'rule'. It may be persuasive to imagine that especially in the context of IR, everything is allowed that is not expressly forbidden, in which case it may seem like the only rules which make sense are prohibitive rules. In the context of international regimes, however, rules that require parties to an agreement to do a certain thing might be even more common. And that, also makes sense from a certain point of view: The most common enforcement mechanism of laws in the international arena is recognition ('blaming and shaming'), and it is often easier to find out that parties have not done something they were required to than that they have done something forbidden. This is because in the former case, the burden of proof lies on the party: If it has obeyed the request, it then needs to show that it did so. In the latter case, the burden of proof lies with the other parties: they need to show that the party in question did something they were not supposed to.

Permissive rules may seem redundant in the international context, but they are actually crucial for the functioning of international cooperation, since they provide a framework of action for the parties. They may contain new ideas on how to interact, and provide novel means of mutually beneficial compromises. They also do not require any kind of enforcement mechanism, but they can still shape actor's behaviour significantly.

Ruggie (1998: 871) differentiates regulative from constitutive rules, exemplifying them with the correct side to drive on the road and the game of chess, respectively. Onuf (2014: 3-4) argues against the this distinction, making the case that

> *"all rules regulate conduct by definition and, in doing so, constitute the social arrangements within which they function. Thus the regulative function of rules—all rules—serves to connect agents to an ever-changing world, the structure of which is constantly being remade as those same rules simultaneously perform their constitutive function".*

According to him, rules can be directive-rules, instruction-rules and commitment-rules. Directive rules are "the kind of rules that we associate with law (and fear as a motivation for compliance with the law)" (2014: 3). These rules could be said to be in line with the coercion logic of IL discussed in section 3.1.3. On the topic of instruction-rules, he continues:

> *"If we do not follow instruction-rules, it will be harder for us to accomplish what we had hoped to. [...] If we ignore these rules, people will remind us of the value, indeed the need, to do what these rules tell us to; we are likely to feel ashamed when we are reminded" (2014:3).*

And finally,

> *"Commitment-rules are like contracts reciprocally undertaken to assure a mutually desired result. When generalized to the society as a whole, they create what we ordinarily call rights and duties. Rules of this sort reflect an emphasis on exchange; disputes over rights and duties are typically referred to third parties for resolution; failure to perform one's duties can elicit a feeling of guilt" (2014: 3).*

As mentioned in the quote, the mechanic of reciprocity – discussed in section 3.1.3 – seems the closest fit for commitment-type rules, although of course international agreements may also follow different logics. Treaties – or agreements – are an exemplification of that last kind of rule, and are of course what we are concerned with here.

Rules are central to our understanding of law, language and social cooperation and they underlie and weave through all of these concepts. One could argue that they are the central building block of all social understanding. It may not be too much to say that rules are the common denominator of all the key concepts discussed in this thesis.

3.1.5 International agreements in context

In the sections above, I have outlined several theoretical considerations which put international agreements into context. All of them interact to shape the drafting and implementation of international agreements, and they influence how their language is used and interpreted. I have discussed some of the implications of the fact that much of IL is organized in regimes, the dual dynamics of negotiation and implementation of law, both of which have implications for the language used in international agreements, the mechanisms and functions of IL and some characteristics of rules. On the one hand, all of these characteristics put in context the way language – including vague language – is used and understood in international agreements. On the other hand, it also provides the theoretical underpinnings for how academics conceptualize the causes and effects of vagueness. The theoretical considerations outlined here are therefore crucial to understand the second part of this chapter.

3.2 Debating the causes and effects of vagueness

In this part of the chapter, I will discuss the hypotheses made about the causes and effects of vagueness in international agreements. Most of the authors discussed below are not primarily concerned with vagueness or indeterminacy, but touch on the issues – more or less briefly – while developing broader ideas about the form and function of international agreements. Going forward, I will refer to their hypotheses on the causes and effects of vagueness as causal assumptions to differentiate from the more operationalized hypotheses

I will present in chapter 5. At first glance, some of the causal assumptions appear to contradict each other fundamentally. By laying them out against the background of different perspectives on IL outlined above, I will try to show that they simply follow different logics or emphasize different aspects, and could as such also be seen as complementary.

3.2.1 A note on causality

In the following, I will discuss a number of causal claims that have been made about the vagueness in international agreements. However, as Kratochwil (2006: 14) points out, when human agency is in play, causality is not as simple as it may first appear:

> *"I have used the two examples above to suggest that the recursivity problem arises particularly clearly in the social sciences, because the actors' understanding influences the world. Therefore, the causal arrows run from our (or the agent's) understanding to the world and not from 'the world' to our understanding or theory".*

It cannot simply be assumed that an event that occurred first is the cause of an event that occurs later. This will become clear when talking about some of the causal claims about vagueness. Sometimes, the fact that a negotiator assumes vagueness to have a certain effect, they will be more or less inclined to draw up a vague agreement. The temporal order is clear – but the cause that vagueness then appears – or not – may actually lie in the later event, or at least in the assumption on the part of the actor that it will occur. I will go into more detail of the phenomenon as it becomes relevant throughout the chapter.

3.2.2 The normativity of (theories on) vagueness

In the literature, vagueness is often perceived as 'less than'. This seems natural, as one possible definition of vagueness is the absence of precision, and as such the phenomenon may by its very nature be a negative.

> *"In the absence of transaction costs, the parties to an agreement would specify the precise conditions under which they would (or would not) perform. Agreements would list every possible state of the world and the obligations of the parties in each state" (Guzman 2002: 1856).*

Citing vagueness as one of the factors which making it so, Jackson (1992: 399) states that "like all human endeavor, and perhaps all national legislation, many treaty rules will not be perfect". Equating precision with perfection and vagueness with imperfection is a fairly common implication, which also underlies Guzman's point above.

In the previous chapter, my own conclusion was that vagueness is a very common phenomenon. For quite a lot of the following discussion, authors assume that it is an

exception, if not in frequency then in acceptance. The underlying question quite often seems to be 'Why on earth would treaty parties allow vagueness to appear in their agreements?' and proceed from there. It is from this vantage point that the question 'Why is there vagueness in treaties?' is most often asked – although this evidently only describes a trend, and by no means all the discussed literature. A different understanding of the functions of rules may lead to different value of vagueness. In so far as agreements between states are seen as creating new possibilities (of cooperation, communication etc.) the functioning of the rule may not rely as much on being obeyed, as by simply being there it gives the option to cooperate. If on the other hand rules are regarded as primarily prohibitive, the logic of coercion and possibilities of enforcement become much more immediately relevant, which also leads authors to regard vagueness as more negative. As a subjective impression, it also seems like authors who are the least concerned with vagueness – only mentioning it in passing in their paper – often have the most critical view of it, perceiving it as an imperfection at best. This is of course not to say that authors who are primarily concerned with the study of vagueness are not critical about the concept – far from it – but they do tend to at least adopt a more nuanced view, citing some positive aspects of the phenomenon in addition to what they view as problematic features.

3.2.3 Theorized causes of vagueness in international agreements

In the first part of this sub-chapter, I will describe the suggestions on causes for vagueness in international agreements found in the literature. This is section is shorter than the next one, because more causal assumptions have been made about the effects of vagueness in international agreements than their causes. Nevertheless, there are also notions in why vagueness appears in agreements in the first place. The propositions discussed in this part are mostly relevant in the context of the negotiation of IL (see section 3.1.2), because that is when the actual texts of agreements are written and thus vague or precise phrases are used. However, as discussed above in section 3.2.1, causal timelines can get inverted when human intentions come into play. Therefore, some of the causal assumptions also touch on the parties' or negotiators preconceptions as to what will happen in the implementation phase of these agreements.

One hypothesis found in the literature is that the conclusion of precise agreements has prohibitively high transaction costs[17] (see for example Abbott/Snidal 2000: 433). Therefore, vague treaties will be concluded even if they would otherwise be less optimal. Guzman (2002: 1856) sums this position up as follows:

> *"In practice, however, substantial transaction costs prevent international agreements from specifying every possible future contingency".*

Sunstein (1996: 16-7) makes a similar point when talking about the minimalism of courts when faced with decision costs. Taking a closer look, the assumption here is clearly that vagueness on the part of negotiators requires less effort – of communication, research, formulation – and thus less cost than precision, which is seen as an unattainable ideal. While not explicitly spelled out, this proposition fits well into the logic of the coercive or reciprocal functions of IL. It contends that vagueness is largely undesirable for all parties and thus likely to be unintentional, but accepted as an inevitability. It also suggests that key parts of treaties – the ones especially important to one, some or all parties to it will be more precise than those parts considered less important by the negotiators. Guzman (2002: 1855) goes on to specify four reasons for these high costs: the impossibility to predict every possible future state of the world, the even greater impossibility of knowing what their probability is at the time of treaty conclusion, the different interpretations varying parties may have of any given state of the world and the difficulties of ratification faced by treaties containing long lists of exceptions and conditions. This last reason will be discussed in more detail in section 3.2.4, as it pertains more to an (assumed) effect of vagueness.

The first two of these reasons are made out to be causal assumptions in their own right by Maley (1987: 41), writing:

> *"However, a just law may need to accommodate to changing times and unforeseen circumstances, so a certain amount of flexibility is desirable for some legislative rules. Hence rules with VAGUE meanings, using judgment words".*

Both Maley and Guzman thus agree that uncertainty of the future is a central cause of vagueness in international agreements. The normative implications of their respective propositions is very different, however: while for Guzman precision, where possible, is always and unquestionably preferable, Maley has a more nuanced view of the value of vagueness.

[17] According to Johnson (2005), transaction costs are "[t]he costs other than the money price that are incurred in trading goods or services". Dahlman (1979: 148) elaborates that "search and information costs, bargaining and decision costs, policing and enforcement costs" fall under the term.

Guzman's third reason relates to the hypothesis that the diversity of parties to the treaty leads to vaguer agreements. The more diverse the parties, the less common ground they will have – and the more diverse their interpretation of the world – leading to more varying interpretations, more difficult communication and thus vaguer treaties. With regards to vagueness in particular, Jackson (1992: 399) explains the connection in the following way:

> *"[C]hances are that treaties will be more ambiguous than national laws (because of the difficulty of reaching consensus), and that vague treaty norms will therefore be used to challenge national regulations".*

This causal assumption could be understood either as an unintentional consequence born out of more complexity due to larger differences between parties (see section 3.1.1), or as a more reflected approach by parties to effectively deal with their differences.

Continuing on the theme of differences, a smaller-scale hypothesis takes into account differences in opinion of diplomats and the countries they negotiate for. As in the previous example, the suggestion is that diversity and differences lead to vagueness. Hisschemöller and Gupta (2001: 166) reflect on the reasons for vague outcomes of negotiation processes and conclude that vague treaties result when negotiators have different views on a topic than the public in the country they represent.

A fourth hypothesis is that parties may push for vague treaties if they believe a vague treaty would be advantageous to them. The difficulties with causal directionality discussed in section 3.2.1 definitely apply in this case. At first glance, the observation that parties will do what is of benefit to them may seem very obvious, but the clue here is that it might pass more easily and frequently than usual interest-based modifications of treaties because both parties may think it is advantageous to them in different ways. This means that treaties will be vaguer than is rational (because all parties overestimate the advantages), and the hypothesis mostly works if the underlying assumption is that vagueness is an irrational feature in otherwise rational actors and circumstances and thus somewhat of a puzzle.

Karrer (2005: 97) elaborates that sometimes parties to treaties can opt for imprecise formulations because they both hope that the underspecified treaty text is closer to their own advantage, underestimating the opportunity costs. Incomplete agreements thus may feed the wishes of both contracting parties who tend to underestimate costs and risks and overestimate the probability of positive outcomes.

Attributing slightly different reasons to the negotiating parties, Schachter (1977: 298) supports the general causal assumption:

> *"[I]f the text or circumstances leave the intention uncertain, it is reasonable to consider vague language and mere declarations of purpose as indicative of an intention to avoid legal effect".*

This assumption will be discussed in more detail in its reversed form – that vagueness leads to less legal effect – in section 3.2.4. It is worth noting, however, that this causal assumption leans firmly on the side of attributing intentionality to vague formulations, as opposed to them being mere side effects or inevitable consequences. This stands in contrast – or complement – to the first causal claim discussed in this chapter, namely that prohibitively high transaction costs lead treaties to be vague whether the negotiators want them to be or not.

One hypothesis which breaks out of the parameters I set for them here, but nonetheless is important to mention is that conflicting rules lead to indeterminacy (Ranganathan 2014). The claim is important especially since the fragmented nature of IL lends itself to conflicting legal rules, which cannot readily be pressed into any one consistent hierarchical order. However, as can be taken from the formulation of the hypothesis, conflicting rules lead to *indeterminacy* of legal texts, not to vagueness in their formulation. Therefore, while extremely significant when studying the indeterminacy of international legal texts, the line of thought is only tangentially relevant when concerned with the way (textual) vagueness leads to indeterminacy in IL.

Summing up, four main causal assumptions have been made about the causes of vagueness in international agreements: high transaction costs, differences between treaty parties, differences in opinion between countries and the diplomats negotiating for them and finally the belief of parties or negotiators that vagueness will be advantageous to them are all hypothesized to lead to vaguer treaty texts. Out of the four, the first one appears to be the most common in the literature, while the last is most closely related to expectations about the effects of vagueness in international agreements, which will be discussed in the following section.

3.2.4 Theorized effects of vagueness in international agreements

This section is aimed at describing the causal assumptions about the effects of vagueness made in the literature. While the caveat that most of these claims are made in passing and not usually the focus of the articles they originate from still applies, the potential effects of vagueness in international agreement have been more widely studied than its causes. The lack of dedicated research on the topic, given its potential consequences, is surprising. Of course, both sets of claims are related in important ways, and often made by the same

authors. Most notably, if negotiators belief in the causal effect of one of the claims made here, that may lead her to push for more or less vagueness in an agreement, depending on her interests. Like the propositions discussed in the previous section, the causal assumptions explored here relate to the framework given in section 3.1 in various significant ways.

The most commonly asserted effect of vagueness in international legal agreements is that it reduces compliance and makes treaties less likely to be implemented (see for example Stokke 2001: 16, Franck 1990: 53, or Mitchell 2001: 228). According to Guzman (2002: 1875), "when possible, countries that wish to increase the level of commitment prefer more formal and detailed agreements", because such agreements are harder to violate. Werksman and Herbertson (2009: 2) also contend that

> *"Lessons from other treaties tell us that international agreements with binding, specific content backed by robust review procedures are generally more effective than those with vague content or limited review procedures."*

Unfortunately, they do not specify which other agreements they refer to. Staton and Vanberg (2008: 504) make the case that this proposed effect applies to court rulings, stating "[v]ague rulings decrease the likelihood of compliance". In fact, most of the proposed causal connections that follow can be drawn more or less directly from this main statement. And while we have seen in chapter 2 that not all legal philosophers support the claim that indeterminacy of rules leads to lower rates of compliance, this certainly appears to be the majority opinion of legal philosophy as well. According to Mitchell (2001: 228–9) Compliance with a regime corresponds to clarity in three ways:

> *"Compliance problems arise from failures of three types. The first type is a failure of obligational clarity. The regime needs to provide clarity with respect to "who must do what." This requires the regime to minimize ambiguities about what behaviors must be undertaken and what outcomes must be achieved, as well as about who is responsible for undertaking or achieving those standards and who is responsible if they are not achieved. A second type of failure involves performance clarity. The regime needs transparency—that is, knowledge about what behaviors relevant actors actually undertook and what environmental outcomes resulted. Reporting, monitoring, and verification provisions seek to address this potential source of implementational failure. A third type of failure involves response clarity. A regime's success depends on the expectations actors have about how other actors, both within and outside the regime, will respond if it fulfills or fails to fulfill the regime's rules".*

While the difference in terminology may lead to doubts if clarity is meant as the opposite of vagueness here, it seems likely that this should be at least partly the case. Others authors certainly do seem to interpret Mitchell in this fashion (see Stokke 2001: 16).

However, while most authors agree that vagueness leads to lower compliance, their explanations for why that is the case can differ substantially. In the following, I will expand on these arguments.

A first line of reasoning is that vague rules simply communicate less input about what behaviour is expected and therefore can lead even those actors who actually wish to comply with the rule to inadvertent non-compliance. Putting it another way,

> *"The determinacy of a rule affects its compliance pull because 'due process' requires a clear message about what is expected of those addressed by the rule" (Stokke 2001: 16).*

Another reason found in the literature focusses on those actors whose non-compliance is at least somewhat intentional. The reasoning is based on the lack of hierarchically organized enforcement mechanisms in the international sphere, and therefore main sanction for non-compliance with IL is considered to be the loss of reputation. Parties in this perspective have an interest in understanding and interpreting the contract in a way suitable to their own advantages (Morgenthau 1954: 255; Abbott et al. 2000). The following quote illustrates this explanation:

> *"Not only do indeterminate normative standards make it harder to know what is expected [...] indeterminacy also makes it easier to justify non-compliance. To put it conversely, the more determinate a standard, the more difficult it is to justify non-compliance. Since few persons or states wish to be perceived as acting in flagrant violation of a generally recognized rule of conduct, they may try to resolve a conflict between the demands of the rule and their desire for interest gratification by 'interpreting' the rule permissively. A determinate rule is less elastic and thus less amenable to this strategy than an opaque one" (Franck 1990: 53–4).*

Relatedly, vague agreements are seen as less likely to be implemented because they are understood to give parties more room for (diverging) interpretations. In the absence of a central ruling authority like a court this could lead to significant hindrances in the implementation of treaties. A similar argument to this has come up in the discussion of indeterminacy in legal philosophy, here enhanced by the fact that in the international arena there does not tend to be an impartial judge but rather the treaty parties themselves interpret their own agreements.

The causality of this claim is disputable, however, when considering the parallel proposition that diversity in negotiation parties leads to vaguer treaty provisions. Taken together, one might as well skip the intervening variable of vagueness and say that agreements concluded between parties with strongly diverging values will less likely be complied with.

Another reason given for this hypothesized lack of compliance again shows a reversal of the timed directionality of causality discussed in section 3.2.1 above: Vague rules are less likely to be complied with because it is assumed that the writers of the rule did not intend for the rule to be binding. Abbott et al. (2000: FN 36) spell this out:

> *"The State Department's Foreign Relations Manual states that undertakings couched in vague or very general terms with no criteria for performance frequently reflect an intent not to be legally bound".*

In this, the assumptions made about the causes and effects of vagueness may be said to have come full circle. Nevertheless, while it is important to be aware of the difficulties causality presents in cases where human intention comes into play, this does not mean that all of the assumptions made must necessarily be false or senseless. After all, the same people who draft agreements are often charged with interpreting them, and they are the ones knowing most about their own intentions during the drafting process.

The main reasoning why vague agreements lead to less complete implementation appears to be that enforcement mechanisms are less likely to be effective or put into place at all. This strand of thinking follows the coercion logic of the mechanism of IL. The reasoning is clear: if agreements spell out exactly what kinds of behaviour will be met with which consequences, this makes it more likely that those consequences will come into force, and thus increases the force of coercion of the agreement in question. Abbott and Snidal claim that legalization, which is comprised of the three components obligation, precision, and delegation, "increases the costs of violation through normative channels" (Abbott et al. 2000: 428).The question of enforcement is difficult in IL even under the most favourable circumstances, so much so that it has been used as a reason to doubt whether IL is law at all. When combined with another factor which calls into question the legality of legal rules (see void for vagueness doctrine), the proposition that vague international agreements will likely not be complied with does not appear to be far-fetched.

Franck discusses the implications of indeterminate rules on the legitimacy of treaties and regimes. For him, while vague rules have the advantage of allowing flexibility when reacting to new developments, "Indeterminate normative standards not only make it harder to know what conformity is expected, but also make it easier to justify noncompliance"(1988: 714). He goes on:

> *"The degree of determinacy of a rule directly affects the degree of its perceived legitimacy. A rule that prohibits the doing of "bad things" lacks legitimacy because it fails to communicate what is expected, except within a very small constituency in which "bad" has achieved a high degree of culturally induced*

> *specificity. To be legitimate, a rule must communicate what conduct is permitted and what conduct is out of bounds" (Franck 1988: 716).*

Legitimacy, in turn, increases the "compliance pull" (1988: 712) of a rule and thus supports the claim that vague rules are less likely to be implemented.

Building on Franck's theories, Abbott et al. (2000: 413, FN 26) take up precision as one indicator for the degree of legalization of a rule. While they hedge that vagueness does not necessarily have to come at the cost of impact of the rule,

> *"Nevertheless, for most rules requiring or prohibiting particular conduct–and in the absence of precise delegation–generality is likely to provide an opportunity for deliberate selfinterested interpretation, reducing the impact, or at least the potential for enforceable impact, on behavior".*

This reasoning relates to the reasons given for lowered implementation of vague rules discussed above. Reus-Smit (2003: 592) expands on the matter, calling into question the legality of a vague rule:

> *"A regime that incurs strong obligations, has precise rules and delegates authority is said to be highly legalized, and one that involves weak obligations, has vague rules and confines decision-making to the contracting parties is legal only in the most general sense of the word".*

Adding a twist to this, Michael Byers (2004) argues that vagueness or ambiguities of a document (in his case, of a Security Council Resolution), while potentially reducing the legitimacy of the legal rule itself, can actually help to preserve the legitimacy of an institution or regime. When the document can be legitimately interpreted in more than one way, it may enable one party to do what it wants in spite of objections by other parties, while the other parties can claim that they do not condone the action. All this can happen without either party explicitly violating the terms of the regime, which would otherwise lead to a significant decrease of its legitimacy.

In conjunction with the presumption that more precise treaties have higher transaction costs at the negotiation stage, this idea assumes that vague treaties are more likely to be concluded. However, in a sort of parallel to the lower transaction cost at the negotiation stage, the transaction cost for implementing vague agreements may be higher. Posner (2009: 110-1) calls these costs "ex-post costs". Both hypothesis follow the same logic: There is a cost of rendering legal clauses precise enough so that they may be transferred to material situations and allow their implementation which remains the same no matter when it is addressed and which must be paid at some point. If they are taken care of in the negotiation phase of an agreement, then that phase is more costly and/or difficult but the implementation phase becomes easier, while in the reverse case the cost of clarification

must be met in the implementation phase. “In short, precision (1) increases the difficulty of reaching agreement ex ante but also (2) promotes compliance ex post” (Goodman/Jinks 2004: 681). In fact, Abbott (2001: 142) even argues that those transaction costs are higher in the implementation phase than in the negotiation phase of an agreement:

> *“For several reasons, ex ante bargaining costs of reaching a precise agreement are likely to be lower than ex post bargaining costs over the terms of an imprecise agreement, particularly in the absence of delegation of decision making authority regarding implementation of the agreement to standing institutions”.*

Most of the propositions above follow some variant of the logics of coercion or reciprocity, where vagueness is seen as unfavourable for implementation. If following the acculturation logic, however, vague treaties may actually lead to better implementation. Goodman and Jinks (2004: 681) make the case that the causal claims discussed above only hold when viewed in the context of the coercion or persuasion mechanism of IL.

> *“Under the acculturation approach, however, these effects are potentially reversed: precision that outstrips existing preferences might propel agreement, and imprecision will sometimes help to produce behavioral conformity. In other words, precision is potentially beneficial ex ante and costly ex post”.*

They explain this in terms of the different role played by norms in the context of acculturation approaches: “Conformity depends less on the properties of the rule than on the properties of the actor’s relationship to the community. Because the Convention or norm is associated in general terms with the identity of the group, rules best foster conformity by “establishing] broad hortatory goals with few specific proscribed or prescribed activities” (Goodman/Jinks 2004: 683).

It is worth noting that this is a very minor avenue of literature, the vast majority of authors tends at least in large parts to the opposite causal claim.

From a slightly different angle of view, causal claims have also been made about the temporal progression of vagueness in precision. Predominantly, scholars argue that vague treaties are more a phenomenon of young regimes and interstate relationships, while more precision develops over time.

The assumption is that states’ interests diverge to a great extend at the beginning of negotiations on a certain topic, and when they get together for the first time, they can only agree on very broad and general terms, leaving all of them room to continue on basically the same lines as before. Then, however, by meeting regularly, discussing the matter and exchanging opinions, as well as more and more availability of scientific evidence and increased public opinion on the matter, more clear-cut treaties become possible, ultimately resulting in better outcomes for the community of states. Steffek (2005: 50) points out that

"in the short term, incomplete agreements may provide new, even innovative solutions to political problems". While according to him, ambiguous agreements can become a foundation for later, more specific and effective treaties, this does not have to be the case, and is likely to take a long time (2005: 50). Wettestad (1999: 237) implicitly agrees with this assessment when he writes about "the generally diffuse character of international environmental politics, with [...] quite vague and general regulatory yardsticks". And Goldstein et al. (2000: 386) state that

> *"Legally binding environmental treaties have proliferated in recent years. These agreements often trace their lineage to hortatory political pronouncements but often become closer to hard law over time".*

Staying in the context of the UNFCCC, Bode (2004: 82) proposes that the problems of vagueness only become clear in subsequent negotiations and therefore are not addressed at the outset, but rather in later agreements:

> *"After the United Framework Convention on Climate Change entered into force remarkably quick, it turned out at the first Conference of Parties that the nonbinding targets in the Convention for the year 2000 were too vague and inadequate to address the global and long-term problem of climate change".*

To fully appreciate the logic behind this temporal causal claim, it is necessary to view it in combination with some of the other propositions and frameworks discussed above. As we have seen, vague agreements are often considered more likely to be concluded. In section 3.1.1, I have also outlined that regimes, once established, foster communication and mutual understandings about issues. Taken together, these characteristics form a coherent picture: Great initial differences are glossed over by vague language in agreements, which are mainly aimed at fostering communications. This initial phase then follows the acculturation mechanism of IL. Over time, parties discover similarities or areas of compromise, and more precise agreements become possible and desirable, because when confronted with more specific issues the outlook then switches to a coercion approach.

While a certainly a majority opinion, this view is not shared by all International Relation scholars – Morgenthau (1954: 255) reasons the opposite:

> *"Thus the lack of precision, inherent in the decentralized nature of international law, is breeding ever more lack of precision, and the debilitating vice that was present at its birth continues to sap its strength".*

Similarly, Whalley/Walsh (2009: 279) could be interpreted in a way supporting this claim:

> *"As such, the imprecision of the mandate can be a major factor in preventing a successful conclusion to the negotiation simply because of the ambiguity of what is being negotiated on",*

because the mandate to further negotiations is usually given in the previous or constitutive agreement. So if a vague mandate is given, for example in the form of an indeterminate framework Convention, further negotiations would in this context prove unsuccessful.

This diversity in proposition shows both how important it is to acknowledge different perspectives as the backgrounds for all the arguments made here and a distinct need for empirical research.

In this section, I have attempted to show the range of causal claims made about the effects of vagueness in international agreements. The most common proposition is that vagueness hinders compliance. Reasons given for this include lack of communicated information, lack of enforcement opportunities, diverging interpretations, and lower legitimacy and legality of vague rules. However, this claim is not entirely uncontested – when following the logic of acculturation, vagueness may actually be helpful in the implementation of international agreements. Another claim about vagueness in international agreements is temporal in nature: the majority of authors proposes that vagueness in agreements generally decreases during the span of existence of a regime. Here, too, researchers arguing the opposite exist. In the end, empirical evidence is needed to confirm which hypothesis proves reliable in which circumstance.

In the previous two sections I have given an overview over the causal claims about vagueness in international agreements that exist in the literature. I have attempted to show that both theorized causes as well as effects of vagueness have to be seen in the context of the broader theoretical framework they originate in. Most prominently, scholars have argued that vagueness is caused by the high transaction costs of precision, that its presence in agreements hinders compliance with them and that its presence within a given regime decreases over time. However, none of these claims are uncontested.

3.3 Conclusion: The case for measuring vagueness in international agreements

It is interesting to note that very few of the causal assumptions try to get at differences of the scale of vagueness in international agreements. Most often, they are formulated in absolute terms: ‘vague treaties will be harder to implement than precise treaties’, for example. In chapter 1, I have attempted to show that all language – and therefore all law – is at least a little vague, and differences must necessarily be a matter of degree. It seems

curious that this is most often not the case. On a side-note, considering the place that vagueness has in the dismantling of binary logic, the dualistic nature of those formulations could appear somewhat ironic.

The main argument I make in this thesis is that while, when put in their theoretical contexts, all of the causal claims discussed above are theoretically plausible, none of them are yet hypotheses in the sense that they are operationalized enough to be tested empirically. This fact also explains their variety and how sometimes claims that are contradictory – at least in their unqualified states continue to coexist. The few hints authors give on what exactly they mean by vagueness and how they recognize it in an international agreement will be discussed in the methodology section (4.3) of this paper, but most of the authors seem to simply take for granted that everyone would recognize vague treaties in the same way, without elaborating on it. And one of the main challenges that such an empirical test would face is the question of an intersubjective method of measurement to detect the differences in vagueness in international agreements. I argue that the questions posed by varying degrees of vagueness – as exemplified by the causal claims discussed above – are important enough to warrant empirical research. In order to do so, a coherent method to recognize and compare the vagueness in international agreements is necessary. Chapters 5 and 6 of this thesis attempt to address just this research gap.

4 The linguistic perspective: existing methods to indicate vagueness

The goal of this chapter is to investigate how vagueness has been identified by other scholars. This is important for several reasons: first, it will show where research on the comparison and measurement of vagueness in language stands today, to form a point of departure for further studies. Second, the landscape of what has been done before will outline the possibilities and limits of identifying vagueness. Third, previous findings on how to recognize vagueness will give essential clues on what might be identified as vague in the current study. The limits of applicability of previous studies will also be discussed. The chapter proceeds in five parts: the first, and longest, will draw on the findings of linguistic scholars studying vague language. This is the area in which most research has taken place, and the most advanced techniques for identifying vagueness are found. However, there are strong limitations on which of these findings are applicable to studying IL. The reasons for these limitations will be specified in part 2. Part 3 is concerned with the few instances in which scholars of IL and IR have attempted to expand on how to identify vagueness. Part 4

discusses one central remaining problem: the importance of context. The last part draws on the findings of this chapter to prepare my own analysis of vagueness in international agreements.

4.1 Linguistic measures

There are two main ways to research vague words. The first is to start with a definition, and from there deduce which words are vague and which are not. The second is to do empirical research on which words participants of a specific speech community perceive as vague. Both methods have advantages and drawbacks which should be carefully weighed against each other according to the goal of the study. As discussed in chapter 1, the field of study concerned with the language component of vagueness – vague language – is linguistics. In the following, I will present first deductive approaches to identifying vagueness in the field of linguistics before turning to the few existing inductive studies.

4.1.1 Deductive approaches[18]

Indicators for vagueness can be directly deduced from definitions. For example, it should theoretically be possible to establish a list of vague words by looking at a number of nouns individually and asking whether or not they are susceptible to the Sorites Paradox. Unfortunately, the existence of higher order vagueness and the ways in which natural language is used makes it virtually impossible to complete such a task without additional, arbitrary criteria. In other words, it is virtually always possible to construe a borderline case for any given word or concept. Morzycky (2013: 4-5) illustrates why identifying vague language is a far from simple undertaking:

> *"Vagueness is ubiquitous. It's in obvious places, such as in the semantics of GRADABLE ADJECTIVES—that is, adjectives that admit degree modification or can occur in comparatives and related constructions. Accordingly, that's the spot linguists have most concentrated on. But as the classic form of the sorites reflects, it's also found in nouns as well. One can construct sorites sequences for PPs like across the quad. Verbs can be vague: love gives rise to borderline cases (for some people, apparently with alarming regularity), as can run (how fast must you go to count?)".*[19]

Therefore, most approaches here called deductive are a little less immediate. They identify sub-classes of vehicles or carriers of vague language and proceed to deduce – usually

[18] Brunner (2011: 64) gives a comprehensive overview over the various categorizations of vague words and phrases in French.

[19] PP means prepositional phrase in this case.

incomplete and exemplified – lists from there. Another characteristic that most of these approaches have in common is that they take their examples from spoken language.

The definitions given fall in three broad categories: Formal definitions, functional definitions, and exemplifications. Most useful for the project at hand are formal definitions, because they can be can be applied more or less directly to instances of text. Exemplifications are a little trickier, because while the examples themselves can be directly applied, the limits of what counts as being in the same category as the example isn't clear – a phenomenon that will incidentally be topical to the discussion below. Functional definitions are very common in linguistics, but they are very difficult in their application for two reasons. One, the function of the expression in question must be recognizable from the text. This is especially challenging in legal texts, as I will discuss in section 4.2.2. Two, this kind of definition may lead to a problem of circularity if we want to identify vagueness to then see in a second step what it is used for. Quite a few of the definitions proposed in the following combine two or even all three of these categories.

Hedge words

The first category of words I would like to look at is hedge words. The relationship between hedge words and vague language is not quite straightforward (see chapter 2.1.2 for a closer discussion). Nevertheless, hedge words are very good at fulfilling the criterion of being 'purposely and unabashedly vague' as postulated by Channell (1994: 19), because in the vast majority of cases, hedge words could be left out of the sentence with only very slight change in meaning, but making a sentence much more precise. The first person to name the class of words and study them in detail is Lakoff (1973). He gives a list of 63 words and expressions titled "Some hedges and related phenomena" (1973: 473). Interestingly, not only does he make no claim of completeness, he does not even attempt to clearly define hedges from similar expressions. What he is clear about, though, is that hedges are words and phrases that render bits of language vaguer, by reflecting the fuzziness of a truth value of a statement (Lakoff 1973: 471).

Building on his works, several different definitions of hedges have been proposed. One of the more prominent ones comes from Salager-Meyer (1994). Her definition of hedges is two-fold: she employs a formal and a functional approach. She sets out by defining hedges through there function:

> *"embraces a three-dimensional concept: (1) that of purposive fuzziness and vagueness (threat- minimizing strategy); (2) that which reflects the authors' modesty for their achievements and avoidance of personal involvement; and (3)*

> *that related to the impossibility or unwillingness of reaching absolute accuracy and of quantifying all the phenomena under observation" (1994: 153).*

Later, she goes on to give formal criteria on how to identify hedges. According to her, hedges include five separate linguistic phenomena. Some of these relate to vagueness in their own right, and will be discussed more extensively below[20]. Shields are "modal verbs expressing possibility" (154). Approximators are taken from Prince et al. (1982: 86) to mean "stereotyped 'adaptors' as well as 'rounders' of quantity, degree, frequency and time" (154). She also includes "expressions such as 'I believe', 'to our knowledge,' 'it is our view that . . .' which express the authors' personal doubt and direct involvement." (154). Here, and in the next category of 'emotionally-charged intensifiers" (154) she blurs the distinction between a functional and a formal approach of identifying vagueness. The last category is concerned with "compound hedges" (154), which describes instances when the above categories are combined by language users (155).

Crompton criticises this mixing of functional and formal definition. He states "It seems clear that hedging cannot, unfortunately, be pinned down and labelled as a closed set of lexical items." (Crompton 1997: 281) and then goes on to define a hedge as "an item of language which a speaker uses to explicitly qualify his/her lack of commitment to the truth of a proposition he/she utters." (1997: 281), departing entirely from the formal aspects of definition.

Cutting (2012: 285) elaborates on the functional aspects:

> *"A review of the literature of VL in general suggested that there are three hedging functions; in the study, these were labelled 'courtesy', 'modesty' and 'caution'. These are particularly relevant to the assumption that the level of informativeness [sic] would be low because of the community of practice's shared knowledge and Conventions".*

The problem with these sorts of functional approaches to hedges is that they again muddy the waters on which words and phrases may be counted as a hedge. To still be able to recognize hedges when they come up, Crompton (1997: 282) proposes the following test:

> *"Can the proposition be restated in such a way that it is not changed but that the author's commitment to it is greater than at present? If "yes" then the proposition is hedged. (The hedges are any language items in the original which would need to be changed to increase commitment)".*

It is interesting to note that unlike most studies of vague language, which are concerned with spoken language, the largest area of studying hedges has come to be academic writing, a specific sub-class of written language.

[20] Due to authors using differing vocabulary, such an overlap is unfortunately unavoidable.

General extenders/vague categories

Another marker for vague language is what I will call vague categories. Different authors have discussed the phenomenon under various names: general extenders (Overstreet 2006: 25), extension particles (DuBois 1992), or vague category markers (O'Keeffe 2004: 5). What they mean by this is a form of speech where a person makes a statement about something, and then, after the sentence could potentially be complete extends it by making it an example for a vague category. This of course presupposes a certain amount of shared knowledge between speaker and audience about what else could fall into this category. O'Keeffe (2004: 6) gives a functional definition along these lines:

> *"The vague category markers in the corpus will be seen as recognisable chunks of language that function in an expedient way as linguistic triggers employed by speakers and decoded by participants who draw on their store of shared knowledge".*

While O'Keeffe hints at some formal criteria that might be used to identify vague categories, she herself does not use them. She states: "The analysis focuses on any forms that make vague reference to sets or categories. Research tells us that vague category markers are found in clause-final positions and mostly comprise a conjunction and a noun phrase" these are admittedly thin criteria, and she immediately continues to say "however because a bottom up approach to identifying all vague categories in the data was used, there was no pre-selection criteria based on form".

Overstreet/Yule (1997: 250-1) give some thought to a formal definition:

> *"The forms in question can be characterized as a class of clause-final expressions that have the syntactic structure of conjunction plus noun phrase. […] They all have nonspecific or 'general' reference, and they 'extend' otherwise grammatically complete utterances. They can be divided into one subset called adjunctive general extenders (typically beginning with and) and another called disjunctive general extenders (beginning with or)".*

They go on to give some examples, making it clear that they do not envision a definite, closed list of general extenders

Channell (1994: 119) gives the examples of 'or something', 'and things', and 'whatnot'. She proceeds to give a list which, contrary to most instances, appears to be intended to be exhaustive. According to her, what she calls tags, or vague category identifiers can either start with and or with or. Special emphasis is laid on the phenomenon that vague categories require a certain amount of knowledge from the interlocutors to work: "Hearers and readers need to draw on pragmatic information in order to identify the intended vague category. They use in particular: (a) the surrounding linguistic context; (b) the purpose of the text or conversation; and (c) their world knowledge" (Channell 1994: 143).

Vague Quantifiers/approximators

One feature of language that is intimately related to understanding precision and vagueness are numbers. Usually, numbers of any kind evoke precision. However, there are also ways to refer to quantities that render language decidedly vaguer.

> *"We saw […] that there exist a number of ways of being vague about quantities in English. In particular, speakers have the option of either adding something to a precise number or numbers, or using a round number, or using a vague quantifier" (Channell 1994: 42).*

Although Channell dedicates three separate chapters to these phenomena, I will summarize them all as vague quantifiers, because they are all concerned with expressing vague quantities, and they seem to go together for other authors. This category of vagueness indicators is one where most clarifications come in the form of examples, so I will try to give an overview of what different authors consider to be approximators or vague quantifiers.

Examples for approximators in conjunction with numbers given by Channell are: 'five or six', 'about ten', 'around the twenty percent mark', 'eighty or so' (1994: 42-43). Some of these words can have a completely different meaning when used in a different context: in the sentence 'I am talking about this', the word 'about' is used very differently than in Channell's example above. This shows that one has to be careful when only looking at words, especially when the analysis is partly or completely automated. Using round numbers is another indicator Channell (1994: 78) uses for identifying vagueness. However, while this may work in general speech, it is very hard to identify which numbers are meant as precise, yet incidentally round, and which are meant as an approximation.

Examples for non-numerical vague quantifiers, according to Crystal/Davy (1975: 113) are 'oodles', 'bags of', 'heaps of', 'umpteen', 'touch of'. Channell (1994: 95) adds 'some', 'several', 'lots', 'a high rate of', 'masses of', 'a lower', 'a higher', 'and an extensive'. Talking about vague quantifiers in surveys, Bradburn/Miles (1979: 92) give the examples 'very', 'pretty' and 'not too' (often, strong, etc., see also Wänke 2002). Coventry et al. (2010: 222) give the examples of "few, some, many and lots of". Lin (2013: 76) speaks of "vague quantifiers such as a bit and a little bit".

In addition to these exemplifications, there are also attempts at defining vague quantifiers through their function. "One of the functions of proximative adverbs (around, about, almost, nearly, etc.) in discourse is to make reported events, times, distances, numbers, etc. vague." (Janney 2002: FN 9). But even in this case, exemplification is used, and the function cited is almost tautological. Although it dies seem to be the case that when it is not the numbers themselves that are used to convey vagueness, approximators as a group do seem to mostly consist of adverbs, with a few adjectives added for good measure. There

is a slight curiosity in that Channell (1994: 118) gives a definition of a category by exemplifying the exclusions rather than the inclusions:

> *"There exist in English a large number of ways of quantifying without using numbers. […] With the exception of the minority of terms which are at the ends of scales, for example, all, never, these terms are vague. […] Many of the vague quantifiers are restricted to spoken, informal contexts." (Channell 1994: 118)".*

A main takeaway from all these examples is that there is no definite list of words or expressions in the category of vague quantifiers, and that deductive methods always have to rely on the assessment of the researcher. Somewhat ironically, most of the definitions used take the form of vague category markings themselves.

Placeholder words

Channell (1994) identifies several English placeholder words: 'thingy' (158), 'thingummy', 'thingummyjig', 'thingumabob' (159), 'whatsisname', 'whatnot' (160), and 'whatsit' (161). Of course, these words exist in any language. The functional definition follows shortly after:

> *"Placeholders are used for two communicative reasons: 1 when speaker does not know or has forgotten a name or noun; 2 when speaker does not wish to use a name or noun" (Channell 1994: 164).*

Placeholder words are almost exclusively used in spoken language (164). It is also obvious that these words are also very informal, and it is perhaps the category that illustrates most clearly why findings from the field of linguistics cannot simply be applied to specialized languages like those used in IL – at least not without careful further investigation. Even though the examples above are not meant to be a closed list, it borders on the ridiculous to find even similar words in any kind of legal text.

Mass nouns

The concept of mass nouns is a little more complicated than most other indicators discussed here. Chierchia (2010: 100) informs us that "[w]hile every noun/concept may in a sense be vague, mass nouns/concepts are vague in a way that systematically impairs their use in counting". Accordingly, this category might come closest to the philosophical definition of vagueness as discussed in section 2.1.1. The opposite of a mass noun is a called a 'count noun' (Trask 2013: 67).

Comparing the two expressions 'more than half' and 'most', Solt (2011) try to get at the specific differences in meaning of mass concepts.

> *"The felicitous corpus example (37a) contrasts with the awkward (37b); the issue here is that racism is not something that can be quantitatively measured. On the other hand, more than half is acceptable in examples such as (38), the difference being that energy use can receive a numerical measure (expressed e.g. in*

> *kilowatt hours). (37) a. But black activists acknowledge that most racism is not so blatant. (Associated Press, 16/9/1991) b. ?But black activists acknowledge that more than half of racism is not so blatant. (38) More than half of home energy use goes to space heating and cooling. (Popular Mechanics, 184(6), p. 79, 2007) As the preceding discussion shows, most and more than half —despite their superficially similar meanings—are used very differently by speakers" (Solt 2011: 170).*[21]

Solt illustrates the inherent immensurability of racism: It is a concept that cannot be counted and therefore fulfils the criteria of being a mass noun.

Determiners

In conjunction with mass nouns, determiners are sometimes discussed as relating to vagueness. According to Trask (2013: 80), a determiner is

> *"A lexical category, or a member of this category, whose members typically occur within noun phrases and indicate the range of applicability of the noun phrases containing them. English determiners include the articles the and a, the demonstratives this and that, possessives like my and your, quantifiers like many, few, no and stressed some, and various other items like either, which, both and unstressed some. In most versions of the X-bar system, determiners are regarded as specifiers of nouns"*

They have a special relationship to mass and count nouns, as some determiners go only with mass nouns, while others can only be used in conjunction with count nouns. This was already hinted at in Solt's distinction above. Determiners could therefore potentially be used to distinguish mass nouns from count nouns in automatic language processing, especially since there can normally be only one determiner in a noun phrase (cf. Aarts et al. 2014: 119).

Shields

As the name suggests, shields are defined primarily through their function: distancing the speaker from what she is saying and thereby protecting her. Jucker et al (2003: 1761) tell us:

> *"First, speakers indicate in various ways their degree of commitment toward a proposition being expressed. Vague expressions are themselves an explicit Conventional device for conveying a lack of commitment. One way of doing this is to use a so-called 'shield'".*

They are also overlapping with many of the phenomena already discussed above, including adjectives like 'probably'. The reason why I include them as a separate category here is that when talking about shields, scholars lay the most emphasis on verbs, which are not at the centre of any of the other classes of vagueness indicators.

[21] The question mark in example 37b indicates a grammatically or otherwise wrong sentence in linguistic Convention.

> *"First of all, there exists a range of parenthetical phrases like I think, I believe or I guess […] Finally, different kinds of modal verbs, e.g. might, would and could, can indicate lack of certainty" (Jucker et al 2003: 1762).*

Jucker indicates that shields are one of the means which the relationship between uncertainty and vagueness, mentioned in chapter 2.1.2, can be expressed.

The indicators for vagueness identified in this section are various and diverse, and practically all of the definitions are vague themselves, in the sense that it is not clear which words and expressions are included or excluded. On the other hand, they do present a relatively detailed picture on what to look for when trying to find vagueness. They all have in common that they are useful when attempting to find vagueness within a sentence: all of them relate to either words or short phrases, and none rely on the entire sentence or the sentences surrounding it to apply.

4.1.2 Inductive approaches

Another way to approach the question of how to identify vague language is by proceeding inductively. This involves empirical studies, asking people how they perceive different characteristics of text. This sort of process is much more time-intensive, and much rarer, than the theoretical approach discussed above.

Caraballo/Charniak (1999) test various measures to automatically determine the degree of specificity of nouns in text, for instance factoring in how often they occur and how often they are modified by adjectives, verbs or other nouns. This method is interesting as it considers the words immediately around the noun in question, and also as it does not require extensive questionnaires. However, directly linked to this, the criteria they give to their program are actually arrived at from theoretic preconceptions and in this sense not empirical. On the other hand, the results were then re-evaluated and checked, and therefore this may still count as an attempt to go beyond mere positing of theoretical notions on how to recognize vague words. One interesting indicator that has not appeared above is that the authors suppose that

> *"One possible indicator of specificity is how often the noun is modified. It seems reasonable to suppose that very specific nouns are rarely modified, while very general nouns would usually be modified" (1999: 64).*

However, in this very idea encapsulates its main problem: would the idea expressed by the modified general noun still lack specificity? Their model does not address this, being aimed only at the specificity of the original, unmodified nouns.

Gilhooly/Logie (1980) use a questionnaire to determine (among other things) the ambiguity of 1944 words, but they simplified it by letting subjects first rate all words as

ambiguous or not ambiguous, and then only asked for interpretations of those words that were rated ambiguous. While this comes close to empirically determining the perceived vagueness of words, two points are worth keeping in mind. For one thing, their instructions regarding the rating of words according to their ambiguity reads as follows: "Words referring to objects, materials, or persons were to receive a high concreteness rating, and words referring to abstract concepts that could not be experienced by the senses were to receive a low concreteness rating" (1980: 396). This, while certainly in a way related, is quite different from the concepts of vagueness discussed here. Second, the fact that they look at single words, while advantageous in terms of research design and feasibility also bears certain limitations – as we have seen with regards to vague quantifiers above, words may have radically differing meanings depending on how they are used in a sentence. Nevertheless, the list of results on all tested words they provide (397-428) is maybe the most comprehensive assessment provided in this direction of study.

Chandra et al. (2003) develop an Ambiguity Index by empirically determining which different interpretations of a term actually are understood in practice, and how often they occur. Their definition of ambiguity, as something that "can be defined as the tendency for a term or concept to be understood in more than one way" (2003: 48) comes much closer to the concept of vagueness under discussion here. The way they measure this is by comparing the most frequent response option to the frequency of all other options. So, they effectively measure the probability of people interpreting a statement differently than the dominant interpretation. This is very interesting as it does not require respondents to realize any vagueness or ambiguity in a statement at all: they may all be convinced that their interpretation is the only plausible one.

In this and the previous section I have attempted to give an overview of the most common ways to identify vague language. It is not easy to identify clear categories, since the terminology is far from fixed and even when authors make reference to one another, they still usually modify the concepts at least a little. One thing that this overview makes clear is that vague language is not restricted to one or a few word classes. Instead, it is necessary to take a closer look at nouns, adjectives, adverbs, verbs and conjunctions alike. It is also noteworthy that most of the emphasis lies on vocabulary (even if it is sometimes compound phrases like in the case of general extenders) and not on grammatical features.

4.2 Legal language as specialized language

While the study of vague language and how to identify it is by far the most prevalent in the field of linguistics, there are a few problems with the applicability of the findings to the current study. Most importantly, the overwhelming majority of the studies is based on spoken, everyday language. The language and expressions used under such circumstances varies greatly from those used in formal legal agreements like the ones studied here. To give one obvious example, words like 'thingy', as discussed in section 4.1.1, while clearly vague, never appear in international agreements. But even when the differences are not quite so glaring, there is no indication that expressions that would be seen as vague in one context would appear so in another context as well. This point is important for all different categories of language, but perhaps especially so in the highly specific world of legal language.

Consequently, this section will first outline the arguments on why differences among speech communities and (professional) contexts are important to take into account when analysing language. In a second part, I will briefly touch on the specific peculiarities of legal language. This serves as a background to the description of how vagueness has been identified so far in legal/IR contexts. In both cases, I will use the phenomenon of vagueness to anchor and exemplify, but the importance of differences between speech communities are more widely applicable than that.

4.2.1 Specialized languages

The findings of the previous section cannot simply be applied to international agreements. As Easterbrook (1984: 87) states:

> *"Words have meaning only to the extent there is some agreement among a community of users of language. Unless both writers and readers understand the same thing by some construction of words, the writing either fails of any purpose or, as with literary interpretation, liberates the reader to supply his own meaning or story".*

Therefore, it is important to take into account the community of language users. There is not one clear and accepted definition of specialized languages. One of the most widely received ones comes from Castellví (1999: 59)

> *"The set of rules, units and restrictions that form part of the knowledge of most speakers of a language constitutes the common or general language. [...] In contrast, we speak of special or specialized languages to refer to a set of subcodes (that partially overlap with the subcodes of the general language), each of which can be 'specifically' characterized by certain particulars such as subject field, type of interlocutors, situation, speakers' intentions, the context in which a communicative exchange occurs, the type of exchange, etc.".*

She emphasizes that in addition to looking for units and rules, it must be considered how speakers actually use a language. As can be seen by the exemplifications, it is not actually

easy or clear to decide what counts as a specialized language in the end, and which features of the language will be different from general language. Maia (2003: 27) suggests that the differences go beyond specialized terminology:

> *"simply knowing the correct vocabulary for legal terms in another language […] is not sufficient to translate from the legal genre of one language into that of another".*

However, it is clear that specialized languages can be attributed to groups of people who speak a language. According to Gumperz (1971: 114), these communities can be constituted according to all kinds of criteria:

> *"Most groups of any permanence, be they small bands bounded by face-to-face contact, modern nations divisible into smaller subregions, or even occupational associations or neighborhood gangs, may be treated as speech communities, provided they show linguistic peculiarities that warrant special study".*

In the case of this paper, the speech community of interest is a professional one. In addition to allowing professional communities to communicate efficiently about very specific problems, they can also be a mark of status and rank, and act as a gate-keeping mechanism to prevent specialized knowledge from being obtained by the general public. Explicitly mentioning the legal and the medical field, Heath (1979: 103) elaborates on this point:

> *"Anthropologists and linguists, working in cultures other than our own, have described the secret and /or special languages of those members of societies who hold access to restricted information, and we have learned much about the role linguistic competence plays in socialization into these privileged positions. It is reasonable to attempt to gain corresponding knowledge about the communicative habits of those who serve somewhat similar purposes in today's modern and complex societies".*

While it is difficult to say in general how language is modified in specialized communities – precisely because different communities engender different (types of) modification – it is nevertheless clear that it is important for any analysis of language to know which kind of specialized language the discourse under study is part of. As Gumperz (1971: 115) illustrates, understanding at least the basics of the specific kind of speech community is fundamental to any kind of understanding:

> *"Just as intelligibility presupposes underlying grammatical rules, the communication of social information presupposes the existence of regular relationships between language usage and social structure. Before we can judge a speaker's social intent, we must know something about the norms defining the appropriateness of linguistically acceptable alternates for particular types of speakers; these norms vary among subgroups and among social settings".*

It is therefore necessary to look more closely at the properties of legal language, in order to understand how it differs from general language and avoid false equivalences.

4.2.2 Characteristics of legal language

Having established that specialized languages are a fundamental analytical category, I will now take a look at legal language in particular. The first question to ask is whether legal language is indeed a specialized language – and thus a subsection of, in our case, English, or if it might even be a different language altogether:

> *"Legal language is treated as a sub-type of natural language. It shares with natural language several relevant semantic and pragmatic features, such as fuzziness, contextuality of meaning and viability as an instrument of communication. The characteristics of a legal language are rather problematic and controversial, and even its very existence is discussed" (Wróblewski 1985: 240).*

Discussing the same doubts, Tiersma nevertheless agrees to treat legal language as a sub-category of general language rather than putting it in an entirely different category.

> *"Although some have suggested that legal English is a separate language, it seems best to regard it as a variety of English. For the most part, legal language follows the rules that govern English in general. At the same time, it diverges in many ways from ordinary speech, far more than the technical languages of most other professions" (Tiersma 1999: 49).*

I would agree with Tiersma in his assessment that legal English is indeed a variety of English, but his argument shows how large the differences can be perceived as.

The following subsections detail a number of features that distinguishes legal language from ordinary language.

Formality

Mellinkoff (1963: 19) writes: "There is a ceremonial quality to the language of the law achieved by the use of 'formal words' – words set apart from the language of street or locker room". Formal and ritualistic language is another trait of the legal lexicon. Such language tends to be archaic as well, although it need not be". Tiersma (1999: 100) notes that in texts as well as court proceedings, the language tends to be at its most ritualistic in the beginning and at the end. One of the reasons Tiersma gives for the formality of most legal documents is that it is used to mark the document as a legal act, with all the consequences that entails. However, he also adds that there is such a thing as 'lawyerly slang' (1999: 137), countering the notion that legal language is exclusively formal. Mattila (2006: 39) argues that one of the functions of legal language is to uphold the authority of the law, which entails a need for some formality.

Specialized vocabulary

Several of the characteristics Mellinkoff identifies in his seminal book 'The language of the Law' as distinguishing legal from everyday English have to do with vocabulary: "Frequent

use of common words with uncommon meanings", "Frequent use of Old English and Middle English words once common use, but now rare", "Frequent use of Latin words and phrases", "Use of Old French and Anglo-Norman words which have not been taken into the general vocabulary", and "Use of terms of art". (1963: 11) Of course, Mellinkoff is talking here specifically about English legal language, so even when international agreements are written in English, that is not to say that they have the same characteristics as the English legal discourse we are concerned with. For example, current international agreements appear to contain much less Latin words than Mellinkoff attests for the English legal system. Nevertheless, vocabulary has often been cited as the main difference between legal and everyday language (see critically Charrow et al 1982: 175). The existence of legal dictionaries and thesauruses stands testimony to the importance of legal vocabulary (see Garner/Black 2004 or Curzon 2007).

Grammatical structures

The grammar of the legal language is of course closely based on general grammar (Mattila 2006: 3). Nevertheless, Charrow et al (1982: 176) identify several grammatical peculiarities of the legal language. According to them, passive constructions, the interjection of phrases in the middle of a clause, phrases beginning with 'as to', nominalizations, word lists and multiple negatives, while unusual in everyday language, are common in the legal setting. On the other hand, it could of course be argued that since these constructions do exist in the general language but are used less frequently there, this does not amount to changes in grammar but simply to a legal writing 'style' (Mattila 2006:3). Mellinkoff (1963: 145) suggests that some of the grammar used in legal texts has been taken from Latin.

Exclusivity

All of the features mentioned above come together make it extremely difficult for laypeople to understand legal language. In accordance with Heath's point about the exclusiveness of specialized languages, a large part of the academic discourse on legal language has focussed on the difficulties for ordinary people to understand the language of the law. For example, Levi (1990: 25) writes on the so-called 'legalese':

> *"it is the quality of the written language in those documents which elicits intense reactions from many citizens, who feel that the language in those cases serves not as a means of, but rather as an obstacle to, effective communication".*

Indeed, when compared to other technical professional languages like that of engineers or even medical professionals, the question of accessibility to the general population seems

especially glaring in the legal field, because of the direct impact the language itself has on every person's life. After all, „language is the vehicle by means of which law is transmitted, interpreted and executed in all cultures" (Levi 1990: 4). In fact, the scope of the problem is such that a concerted effort – the plain English movement – is dedicated to rendering legal language more accessible to the general population.

However, there may also be positive aspects to this exclusivity:

> *The other side of the coin is that a specialist language strengthens group cohesion. This holds true, too, for the legal profession: an in-house language encourages among lawyers a feeling of solidarity towards their colleagues; it consolidates professional identity in legal circles, and expresses the commitment of lawyers to the values and traditions of their profession" (Mattila 2006: 52).*

The specialized nature of legal language excludes people untrained in law, but it does enhance a sense of community among those who speak it.

Performativity

Maley (1987: 27) takes a completely different approach to getting at the distinguishing characteristics of legal langue: "Controlling actions by words; here is the key to the understanding of legislative language." What he refers to here are the performative properties of the language of the law. Mattila (2006: 31) elaborates:

> *„Speech acts are of fundamental importance from the standpoint of the legal order. Given that the law is a metaphysical phenomenon that is only 'alive' in language, it is only by language means that it is possible to change legal relationships. The language of the law is thus an instrument of speech acts: it has a performative function".*

The above quotes may seem to imply that performativity itself is the factor that sets legal language apart from other types of language. This is not quite the case: After all, Austin (1962: 147) has famously shown that all language is, in some way, performative. However, legal language indeed stands out as a particularly explicit type of performative language. By their very definition, laws are prescriptive, and this is reflected in the language they are formulated in. While in ordinary life, illocutionary acts are often implied in the language (i.e. one would say 'do the dishes' instead of 'I order you to do the dishes'), legal language often expresses this illocutionary force directly in the provisions themselves. And a large part of this explicit illocution is expressed in verbs. Austin's argument gradually abstracts from these verbs, showing that it is not necessary to say 'I promise' for a promise to be made (1962: 32), but it is not at all a random choice that these verbs are his starting point. It follows that verbs which make an illocutionary act, deserve particular attention in the legal context. Based on Austin's (1962: 32) expression, I will call these verbs explicitly performative verbs, or explicit performatives. As a testament to how important verbs are to legal language, entire articles

have been written on the use of the modal auxiliaries 'shall' and 'should' in treaty text (cf. e.g. Williams 2005, D'Acquisto/D'Avanzo 2009).

4.2.3 Legal language and vagueness

At first glance, legal language seems to contain less vagueness than ordinary language. As Tiersma (1999: 71) puts it:

> *"Much of the linguistic behaviour of the legal profession is geared towards speaking and writing as clearly and precisely as possible".*

The notion that legal language is more precise than general language is almost as ubiquitous as the view that it is difficult to understand for laypeople.[22]

Interestingly, two of the characteristics Mellinkoff (1963: 11) ascribes to legal language are "deliberate use of words and expressions with flexible meanings", and "attempts at extreme precision of expression" They appear contradictory: How can legal language be characterized, at once, by use of deliberately flexible meanings and extreme precision? Mellinkoff seems to hint that the question of vagueness is particularly important in the legal discourse, and it can consequently tend to both extremes of the spectrum.

Quite apart from research discussing the nature of vagueness in the legal language, an argument can also be made deriving from the properties of vague language. As we have seen in section 4.1.1, indicators of vague language are often defined functionally. By way of summary, Drave (2001: 26-7) lists the functions of vagueness in everyday conversation: as filling lexical and knowledge gaps, (de-)emphasising or deliberately withholding certain information, conveying tentativeness, an evaluation of, or expectation about, or a proposition, or maintaining an atmosphere of friendliness, informality or deference. Having discussed the peculiarities of legal language above, it is clear that several of these functions do not apply in the legal field. Some of them are tied to spoken language, while other function may relate to entirely different forms in the legal field. Fjeld (2005: 157) sums it up, saying

> *"the precisification strategies used by law experts and by normal readers/laymen are of different kinds, which often leads to misunderstandings and misinterpretations".*

Another anchoring point ties back to the vagueness indicator of vague categories, discussed in section 4.1.1. O'Keeffe points out that in order for vague categories to function communicatively, a shared framework is needed:

> *"It is argued that the shared knowledge required in order to construct vague categories has a common core of socio culturally ratified 'understandings' and*

[22] See Mattila (2006: 3) or the discussion on so-called ‚legalese', e.g. Benson/Kessler 1986 or Hill 2001.

> *that the range of domains of reference of these categories is relative to the depth of shared knowledge of the participants and relative to their social relationship" (2004: 1).*

Additionally, vagueness is usually implicitly assumed to apply to the locutionary dimension of language. However, especially considering the very explicitly performative nature of the legal language, it is a fair question whether or not this assumption is actually true – vagueness may well be found in both the locutionary and the performative aspect of a legal text.

As this section has shown, the indicators for vague language as discussed above can only be useful as a first point of departure.

4.3 Vagueness indicators in IL

Having seen the importance of specialized languages, it stands to reason that in order to find useful indicators for the presence of vagueness in international agreements, it is necessary to turn to scholars of IL. Unfortunately, even though many researchers have hypotheses on the causes and effects of vagueness in IL (as shown in chapter 3), very few give any indication as to which criteria they use to identify and compare vagueness. This is not to say that scholars make no assertions on which agreements and provisions they believe to be vague: on the contrary. But those who do assert vagueness often stick to specific examples and rarely go into the details on why exactly they consider some phrases to be vague and not others. This makes generalizations extremely difficult.

Franck (1988: 713-714) gives some examples on what a comparison between sentences in a treaty would look like:

> *"To illustrate the point, let us compare two textual formulations defining the boundary of the underwater continental shelf. The 1958 Convention places the shelf at "a depth of 200 metres or, beyond that limit, to where the depth of the superjacent waters admits of the exploitation of the natural resources of the said areas." The 1982 Convention on the Law of the Sea, on the other hand, is far more detailed and specific. It defines the shelf as "the natural prolongation of . . . land territory to the outer edge of the continental margin, or to a distance of 200 nautical miles from the baselines from which the breadth of the territorial sea is measured," but takes into account such specific factors as "the thickness of sedimentary rocks" and imposes an outermost limit that "shall not exceed 100 nautical miles from the 2,500 metre isobath," which, in turn, is a line connecting the points where the waters are 2,500 meters deep".*

While he thus provides a somewhat more detailed description of what he means by vagueness than most other scholars of IL, he nonetheless fails to provide the empirical means to actually test his theory. In order to do so, we would need criteria allowing a classification of some treaty texts as vaguer than others, before then seeing if they are actually obeyed

more readily. The examples that Franck gives could be used to deduce such criteria, but the only one that seems obvious from his text is that determinate rules seem to contain more numbers. While his theoretical argument is compelling, the empirical foundation for it is still somewhat lacking.

Abbott et al. (2000) have theorized on vagueness in international agreements in the context of their study of the legalization of international regimes. They propose a three-partite conception of legalization, dividing it into obligation, precision and delegation.

> *"Obligation means that states or other actors are legally bound by a rule or commitment or by a set of rules and commitments […] in the sense that their behavior thereunder is subject to scrutiny under the general rules, procedures and discourse of international law, and often domestic law as well. Precision means that rules unambiguously define the conduct they require, authorize, or proscribe. Delegation means that third parties have been granted the authority to implement, interpret, and apply the rules; to resolve disputes; and (possibly) to make further rules" (Abbott et al. 2000: 401).*

If the level of each of these indicators is high, a rule or regime will then be highly legalized. While they embed their hypotheses about precision in legal documents in the broader framework of legalization, they also provide one of the further-reaching attempts to operationalize the concept, building on Franck's theories. According to them,

> *"A precise rule specifies clearly and unambiguously what is expected of a state or other actor (in terms of both the intended objective and the means of achieving it) in a particular set of circumstances. In other words, precision narrows the scope for reasonable interpretation. In Thomas Franck's terms, such rules are ''determinate.'' For a set of rules, precision implies not just that each rule in the set is unambiguous, but that the rules are related to one another in a noncontradictory way, creating a framework within which case-by-case interpretation can be coherently carried out. Precise sets of rules are often, though by no means always, highly elaborated or dense, detailing conditions of application, spelling out required or proscribed behavior in numerous situations, and so on" (Abbott et al. 2000: 413).*

However, while they do provide a few guidelines on how to distinguish vague treaties from precise ones – most notably 'rule-like' instead of 'standard-like' formulation, use of numbers and non-contradictory rules, most of their accounts still remain in the area of examples, not easily applicable criteria, as the following quote shows:

> *"Much of international law is in fact quite precise, and precision and elaboration appear to be increasing dramatically, as exemplified by the WTO trade agreements, environmental agreements like the Montreal (ozone) and Kyoto (climate change) Protocols, and the arms control treaties produced during the Strategic Arms Limitation Talks (SALT) and subsequent negotiations. […] Even many nonbinding instruments, like the Rio Declaration on Environment and Development and Agenda 21, are remarkably precise and dense, presumably because proponents believe that these characteristics enhance their normative and political value" (Abbott et al 2000: 414).*

While being critical of Abbott et al.'s approach, Flohr also offers some explanation on what more precise legal provisions look like:

> *"Rule-like provisions would narrow the room for interpretation and might, for example, provide that 'truckers shall drive no more than 50 mph and take breaks every 4 hours'. Compared to the principle of 'driving cautiously', precision would have increased, thereby reducing addressee discretion. However, addressee discretion, especially in the international legal system, also depends on the degree to which rules are prescriptive, meaning whether they impose duties as absolute requirements or as recommendations or options to consider. In practice, these criteria will often overlap and they may even appear indistinguishable, although they are analytically distinct" (Flohr 2014: 53-54).*

There are very few studies which discuss specific indicators for vagueness in international legal language. In an article comparing the English and Croatian versions of the Arbitration Agreement between the Government of the Republic of Croatia and the Government of the Republic of Slovenia, Šočanać (2009: 190) discusses determiners, modifiers, modal auxiliaries and passive constructions as to their effect on vague language. According to her,

> *"exact quantification tends to give a high degree of precision to the language of the document. On the other hand, determiners such as some, any, all, both, either, neither, each and every imply varying degrees of precision both in general and legal language" (2009: 199).*

Determiners seem to be important in order to determine whether a provision is vague or precise, but it remains unclear which determiners point in which direction, and why.

On the subject of modifiers[23], Šočanać also finds that while some modifiers enhance precision, others render the text vaguer.

> *"In general terms, quantification involves greater determinacy than qualification, which, in itself, implies different degrees of precision. Most of the examples [of modifiers] in our corpus are descriptors, involving different degrees of indeterminacy/precision" (Šočanać 2009: 199).*

Modal auxiliaries can also be indicators for vagueness or precision. In particular, 'may' and 'shall' are discussed by Šočanać:

> *"May can cause vagueness since it can express both deontic permission and epistemic possibility. […] Shall denotes 'mandatory intent' which means that noncompliance is punishable by sanction" (Šočanać 2009: 200-1).*

It is not entirely clear if she therefore also considers 'shall' to be more precise than other modal auxiliaries, but the way the data is presented seems to imply just that.

Fjeld (2005) considers the role of vague adjectives in normative text.[24] According to her,

> *"Most nouns are indefinite and need specification, either according to the situation, or according to linguistic specification. The main function of adjectives is normally to specify or identify vague or indefinite nouns, and this was also the case in the analysed acts. A certain set of adjectives proved to be extremely*

[23] According to Trask (2013: 174), a modifier is a member of „any category which serves to add semantic information to that provided by the head of the category within which it is contained, such as an adjective or a relative clause within an NP or an adverbial within a VP."

[24] Although Fjeld is mostly concerned with (Norwegian) domestic legal texts, I discuss her work here because it is the basis for some research focussed on the international sphere.

frequent in the acts examined. An analysis of their semantics show that they are in fact often used as deprecisification tools in normative texts" (Fjeld 2005: 157).

According to her, most adjectives serve to make a sentence more precise, but a certain subset of them actually makes legal sentences vaguer. She distinguishes dimensional and evaluative adjectives: The former "are interpreted according to the external properties of the referent of a noun and a metric scale of comparison to make up their quantity parameter" (2005: 159-60), while the latter "refer to internal and often prototypical properties, and must be interpreted according to an unpredictable and subjective scale of measurement related to their class of comparison" (2005: 160). They are therefore much more problematic in their interpretation (2005: 164).

Partly basing her work on Fjeld, Di Carlo (2011) refers to 'weasel words', which are very similar to hedge words, in her analysis of the indeterminacy of Security Council resolutions on the Second Gulf War. She differentiates consequence adjectives (di Carlo 2011: 48) and modal adjectives (di Carlo 2011: 49), both of which are sub-classes of the evaluative adjectives specified by Fjeld (2005: 164-5), as well as weasel nouns (di Carlo 2011: 49). Interestingly with regard to the discussion on the performativity of legal language, both classes of adjectives used by di Carlo mainly affect the performative dimension of the sentence. According to her, these three groups of words are indicative of intentional vagueness in legal texts (di Carlo 2011: 47). It is important to note the di Carlo uses deductive categories, which do not necessarily reflect the opinions of the community of legal professionals.

In this part of the dissertation, I have shown both that a reliable measurement tool is necessary for deeper insights into the fields of IR and IL and that such a tool does not as of yet exist. However, I have also provided some starting points for the following investigation: There are a few indicators of vagueness to be drawn from the literature. For example, it seems to be a common assumption that numbers increase the precision of a text. Similarly, words that more closely describe a rule (i.e. adjectives and adverbs) are seen as reducing vagueness, but some subclasses of adjectives may increase it. Flohr (2014: 54) and Šočanać (2009: 200) also add the dimension of prescriptiveness to the possible indicators. In the following, I will build on these thoughts to attempt the realization of the possibility to systematically measure vagueness in international legal documents.

All of the indicators used by political or legal scientists discussed here are deduced from theoretical definitions and not based on what people perceive as vague. They are, however, usually concerned with existing legal documents and as such provide more insight than those based on general language.

4.4 Conclusion: uses and limits of existing methods

It may seem like I have used a lot of space detailing ways to indicate vagueness in the first part of this chapter, only to then argue that they are not useful when looking at international agreements. This is not quite the case: While I do argue that existing linguistic indicators for vagueness cannot be applied to the context of IL without further thought, they do provide important starting points. It is quite possible that some or even all of the indicators are still valid in this context. But in order to know which ones remain applicable, and which ones do not carry over to the international legal context – and which new indicators may come into play that are not represented above – it is necessary to look directly at international agreements, and also to study how vague language is understood and interpreted by experts in the legal field. The next chapter is therefore aimed at developing new methods on the basis of those outlined here to specifically detect indicators of vagueness in international agreements.

5 Looking for indicators of vagueness: original methods

While the previous chapter discussed methods for discovering vagueness already existing in the literature, this chapter will lay out the method I used to find out features of international legal agreements that indicate vagueness. Section 5.1 will present the methods which were used to design a survey which was conducted among experts in IL in order to find out how to identify vagueness in legal texts. Section 5.2 goes to present the results of this survey in two parts: Section 5.2.1 presents the answers to a multiple choice section of the survey, wherein respondents were asked to rate sentences according to how vague they found them. Section 5.2.2 discusses the comments left by respondents in answer to open-ended questions. Section 5.3 describes the analysis of the data gained by the survey and ultimately presents a consolidated tool which can be used to measure the relative vagueness of international agreements.

5.1 Methods for the survey

In this section, I will explain the methods used to survey international legal professionals, and my reasons for using them. The main goal of the survey was to learn what professionals in the field of IL regard as vague in international agreements. This is primarily for two reasons. For one, it is vital to capture the opinions and impressions of people who belong to the language community. The language of IL differs in significant ways from general other types

of professional or regional languages (see chapter 4.2.2), and therefore what appears vague will likely differ as well. Therefore, people familiar with the language will be better able to recognize vagueness where it is present. Secondly, their perceptions matter because at least in part, their interpretation becomes reality. As we have seen in the section on causal assumptions, negotiators' beliefs on what vagueness may do can influence their behaviour around it. Similarly, how they believe vagueness to be expressed will change what formulations to use according to what they would like to achieve. In this way, even though the survey is limited to finding out what some people in the community of IL – speakers perceive to be vague, this limit is actually broader then it at first sounds, because the perceptions of the people surveyed are, albeit in a small way, part of what actually makes a sentence vague or not.

In general, while designing the survey I had several, sometimes conflicting concerns. Firstly, as with most surveys, I aimed at maximising the responses. This was necessary for three reasons. One, the survey was made to collect, at least in part, quantitative data, and for quantitative studies more data is – ceteris paribus – better. Two, most statistical evaluation methods commonly used to evaluate raw data points call for a certain lower limit of data points to be applicable. Furthermore, while the study was not meant to be representative, I did try to make the field of respondents (within the intended target group) as large and diverse as possible by sending invitations to participate to universities as well as research institutes and international organisations from different countries.

Secondly, the material I wanted to ask about also needed to be as close as possible to actual agreements. This is the case for several reasons enhancing the quality of the responses. Leaving the text as unaltered as possible would avoid the accidental exclusion of significant characteristics. As mentioned in chapter 4.2 above, it is also imperative to stay within the framework of the language of IL. If this was the only concern of the study, it would have been best to have people read complete, unaltered agreements and ask for their perceptions on the vagueness. However, there are other concerns prohibiting that approach.

The most obvious one is a respect for the time of busy professionals, and a commitment to user-friendliness in surveys. The reasons for this are both ethical and pragmatic. For one, making a survey as user-friendly as possible is an end in and of itself, being the bare minimum acknowledgement that people donate their time to fill it out. For the other, people are much more likely to continue filling out a survey, or even recommend it to others, if it is user-friendly and demands only a limited amount of time to fill out (see Dillmann 1993: 298). Therefore, respecting people's time constraints, and making their experience as pleasant as

possible also maximises the potential for usable answers, enhancing the quality of the data gained in accordance with the aim outlined above.

The second constraint on the necessity to stay as close as possible to the original materials is the aim of comparability. The more closely one stays to the original text and the more idiosyncrasies and details are taken up, the less likely it will be to apply to any other text. Therefore, the limits of the usefulness of the data would be the exact texts that would have been asked about in the survey.

As mentioned above, another aim of the survey was to collect expertise from a specific target group, namely people familiar with IL. This is because random people may have very different conceptions about language and, as mentioned in chapter 4.2, opinions of people in the language community shape the language. This aim conflicts with the goal of getting as many responses as possible, because it narrows the target group considerably.

Two more conflicting goals are the need to check for likely hypotheses, and the equally important need to stay open to new insights. The few hypotheses that were won by looking through relevant literature needed to be tested, because the sort of correlations drawn here depend on firm theoretical footing to be taken seriously to be even approximated to causal relationships. Therefore, hypotheses already taken up in more theoretical literature are more convincing in showing possible causalities, not simply correlations.

On the other hand, since they were so few and far between, it was just as necessary to be open to new insights. Since the people surveyed were best placed to know which signs point to vagueness in international agreements, it makes sense that they would add some insights not present in the literature so far. Relatedly, all surveys entail the inherent danger of priming the respondents with answers that are thought likely by the survey designer (see Tversky/Kahnemann 1973), thus rendering the quality of the data questionable.

Given the amount and the inherent contradiction between some of the aims, it is of course impossible to maximise them all at once. It follows that they need to be balanced against each other in some way. I have encapsulated the compromises in four solutions, which together map out the methodology behind the survey. I will now discuss each of the solutions in turn.

The first solution I settled on was to conduct the survey at sentence level. Sentences were taken in their original form from treaty texts, but no longer units of text were tested. This procedure has several consequences. The first one is that some measure of vagueness cannot be accounted for, because there is no reason why all vagueness should be on the sentence level. However, this vagueness is certainly the easiest to identify. Additionally, as

was shown in chapter 4.3, when vagueness in international agreements is discussed, it is most often talked about on the sentence level. Most importantly, in the drafting process of documents, the unit of discussion tends to be on the sentence level, as can be observed in the fascinating and informative textual history of the Kyoto Protocol (Depledge 2000). In order to mitigate some of the obscuring effects, I included a broad range of sentences, chosen partly at random, partly to include as many different characteristics as possible. I will expand on the procedure used to choose sentences for each part when getting to the sections of the survey. The fact that some of the sentences remained long and fairly complex, although a trade-off with user-friendliness, had one welcome side-effect: participants not used to the language would have found it hard or time consuming to complete the survey, helping out in the selection process.

The second solution settled on was to split the survey into three parts, of which I will discuss the two that are pertinent to this study below.[25] This served several aims: for one, it made it possible to combine relatively closed questions with open ones, thus accounting both for the test of indicators found in chapter 4 and the need to stay open enough in the research design to discover new ones. Additionally, it allowed to focus different parts better on different goals. Lastly, the diversity in questions enhanced user-friendliness by offering more variety and interest to participants (see Converse/Presser 1986: 63). In the next paragraphs, I will discuss the design of each of the survey parts in turn.

The first part consists of six sentences taken directly from treaty text. To select the sample sentences, I separated all agreement texts included in this study (just under 90.000 words in total) into separate sentences. I then used Microsoft Excel to count the words in each sentence to order each document's sentences by length. The choice of sample sentences was half-randomized. I aimed to achieve a sample representing a broad range of documents, sentence lengths, verbs and numbers of adjectives used etc. The sentences were equally distributed between those taken from UNFCCC and NPT documents, as well as from documents of differing legal statuses. Unfortunately, the constraints of an online survey make it impossible to fully randomize a large-scale selection of data, which is why some targeted selection needed to take place. For instance, the shortest chosen sentence is 14 words long, while the longest amounts to 129 words. They were taken both from preambles as well as operative clauses. The number of different sentences was again constrained by the need to not burden any one respondent with a too lengthy questionnaire, and by the need

[25] A third part was present in the survey, but since its results are not taken into account in the analysis, I will also refrain from discussing it here.

to repeat the questions often enough to gain a number of estimates for each, so as to minimize the effect of outlier opinions. In this way, I established a list of 30 sample sentences, which participants of the survey were then asked to rate according to their vagueness. Choices given were 'not at all vague', 'contains some vagueness', 'fairly vague', 'very vague' and 'no answer' (on the benefits of labelling each point of a scale see Krosnick 1999: 544). The scale remained the same for each of the questions (see Weisberg et al. 1996: 98). The choices were intentionally left fairly open to the interpretation of the participants, yet clearly mark different levels of vagueness. Participants were encouraged to use the whole range of possible answers. In order to maximize the data gained from the survey, but not overwhelm participants, the six sentences each contributor saw were drawn at random from a pool of 30 sentences (see Weisberg et al. 1996: 99 on limiting the number of survey questions). The order of the sentences was equally randomized (on the importance on varying the order of text in online surveys see Krosnick 1999: 545). This procedure has the added benefit of eliminating biases of comparison, where one sentence is rated a certain way simply because the previous sentence is perceived as comparably vague or precise. Respondents were given a box at the end of the page to add any comments or thoughts they would like.

A second part of the survey simply asked each respondent two questions: "In general, which features of a sentence make you consider it vague?" and "Which features of the sentences in this survey made you consider them precise?" This was aimed at giving them the chance to add their own understandings of vagueness, and thus generating more target-specific hypotheses (see Fink 2003: 17 on the importance of both open and closed questions in surveys). This part was beneficial both from a data collection and a user friendliness point of view. For one, it gave respondents a chance to interact with the survey relatively freely and give thoughts of their own apart from strict numbered scales. Since the answers to these questions were not mandatory to complete the survey, they did not place an extra burden on the experts. For two, the same freedom to express themselves enhanced data collection, since it minimized suggestiveness and opened up the chance to find points might otherwise have been overlooked.

The third solution hinged on making the survey as technically flawless as possible. To this end, the choice of a good hosting website was essential, which is why the survey was made on SoSciSurvey[26]. When a first draft of the questionnaire was completed, it was thoroughly tested in both face-to-face and online pretest sessions[27]. When possible, the Think Aloud

[26] Visit website at https://www.soscisurvey.de//.
[27] On the importance of pretests, see Schnell/Hill/Esser 2008: 347.

technique was used (see Prüfer/Rexroth 2005: 14), meaning that testers were encouraged to talk continuously while taking the survey, expressing everything coming to their minds while completing it. After the comments and suggestions from the first round had been incorporated into the design, a second round followed. In this way, I tried to maximize a good experience for participants as well as clarity in asking for the questions the survey was designed to answer.

After a welcoming page explaining the general intent of the survey, the next page contains two questions meant to ascertain the expertise status of participants: They were asked to rate their level of expertise in the field of IL, and also to supply information on whether they had any background in either environmental/climate law or the NPT. The content-oriented part of the survey is divided in three parts, only the first of which is relevant to the results discussed in this paper.

Solution four focussed on reaching the target group of the survey. The target group for this survey were people with a background in IL. This is because a very specific type of language is used in international agreements and interpretations and assigned meanings are likely to differ between members of the general population and experts. This presumed difference may be very important in general, but the purpose of this study was first to find out how experts interpreted the data. They are the ones drafting and working with them, after all. The invitations to participate were sent out exclusively via email. Both general email-lists as well as writing to specific persons and writing to departments of IL were used. Emails were sent out to mailing lists concerned with International Law and IR as well as faculties and institutes of IL. The number of people reached is difficult to estimate, as is the percentage of experts within the people reached (though through the targeted invitations, this last number was probably fairly high – see section 5.2). To screen participants further, two questions were introduced in the very first pages of the survey. They asked participants to rate both their level of expertise in IL, and whether they had a background in the fields of either environmental or nuclear non-proliferation law.

5.2 Results of the survey

All in all, the Survey received 341 clicks in the 90 days it was online. 131 questionnaires were started. Of those, 50 were aborted after the first page. This could in large part stem from participants realizing they do not have the expertise needed to fit in the target group. 68 questionnaires were filled out completely. These are the ones the results of this study draw from.

16 of the interviewees self-identified as students of IL, 18 as PhD students, 16 as having a PhD in IL, 17 as international legal professionals with 1-20 years of experience, and 15 people as 'other', most of whom indicated degrees in IR. The slight variation in the total comes from the fact that double mentions were possible, most often used to convey both professional and academic experience or a lower degree in law and a higher one in another field.

5.2.1 Sentences voted as vague

The first part of the survey asked respondents to rate 5 sentences according to their vagueness, as described in section 5.1. 68 people responding to 5 out of 30 questions each leads to a total of 340 responses to 30 questions. In the following, sentences used in the survey are referenced by their running number. The complete list of sentences, including full references, can be found in section 9.3.

The first result of note is that the sentences were rated very differently from each other. This clearly points to the fact that experts in IL do judge sentences to contain different amounts of vagueness, and that they are at least partly agreed in which sentences are vaguer than others. At the most precise end, 93.75% (15 out of 16) of respondents thought the sentence

> *"Each Party included in Annex I shall submit the information required under paragraph 1 above annually, beginning with the first inventory due under the Convention for the first year of the commitment period after this Protocol has entered into force for that Party" (VS 25)*

was either precise or mostly precise. At the most vague, 90.91% (10 out of 12, with one 'don't know') said the following sentence was either 'somewhat vague' or 'very vague':

> *"The Parties should cooperate to promote a supportive and open international economic system that would lead to sustainable economic growth and development in all Parties, particularly developing country Parties, thus enabling them better to address the problems of climate change" (VS 22).*

This very first juxtaposition of sentences could, to a certain extent, be used to understand what is considered vague or not by example. However, it would be extremely difficult to apply this to any other documents in a systematic way.

In between the two extremes, the sentences were relatively evenly distributed. Four sentences (VS 4, VS 16, VS 22, and VS 26) were rated vague by exactly 50% of respondents. The graph below illustrates the variance of perceived vagueness in all the sample sentences:

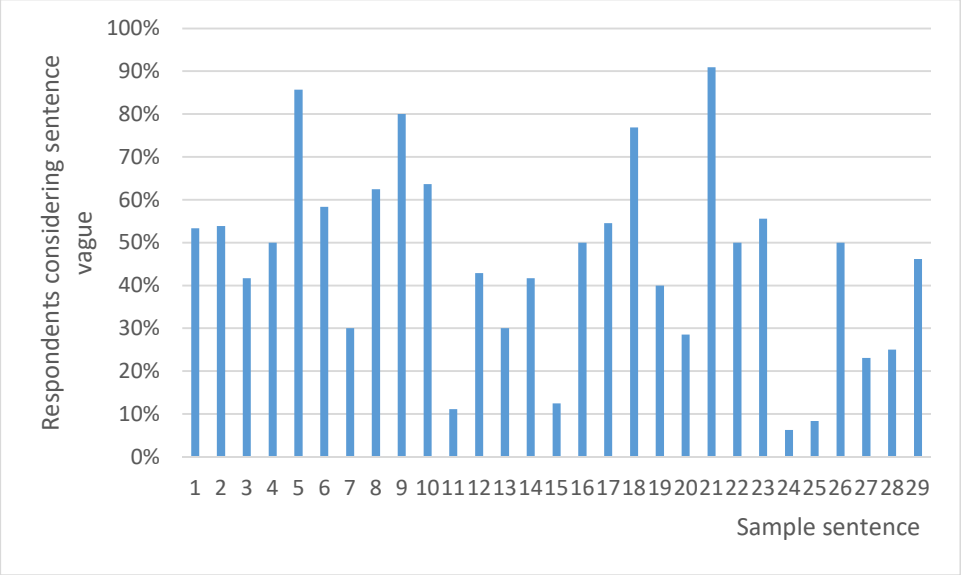


Figure 2: Variance of perceived vagueness

The graph above shows that over all participants, answers vary by up to 83 percentage points. While the number of people who were undecided on a sentence were usually low, one sentence eluded judgement: 4 out of 9 people said they couldn't decide whether the sentence "A clean development mechanism is hereby defined" was vague or not. Because this number is so high, it has been excluded from further analysis. One participant noted that this sentence in particular was dependent on context:

> *"what does "hereby" refer to? What is a clean development mechanism? If the Definition were vague, so would the sentence be" (Interview 49, 19.11.2014, 13:57h).*

The table below shows the differences between various groups of experts participating in the study.

Table 1: Perceived vagueness of sentences by profession of respondents

Sentence	**All respondents**	Students	PhD Students	PhD holders	Professional	Other
VS 25	**0.06**	0.00	0.00	0.25	0.00	0.00
VS 26	**0.10**	0.00	0.00	0.00	0.00	0.50
VS 12	**0.11**	0.33	0.00	0.00	0.00	0.00
VS 16	**0.13**	0.00	0.33	0.00	0.00	0.00
VS 28	**0.25**	0.00	0.00	0.00	0.17	0.67
VS 29	**0.25**	0.00	0.33	0.00	0.00	0.50
VS 21	**0.29**	0.00	0.00	0.67	0.33	N/A
VS 14	**0.30**	0.00	0.00	0.50	0.25	N/A

VS 8	**0.33**	0.33	0.50	0.33	0.00	0.00
VS 15	**0.42**	0.00	1.00	0.67	0.00	0.25
VS 20	**0.44**	0.00	0.40	0.00	1.00	0.67
VS 3	**0.45**	0.25	0.50	0.33	1.00	1.00
VS 30	**0.46**	0.33	0.20	1.00	0.00	1.00
VS 4	**0.50**	0.67	0.60	0.00	0.25	0.50
VS 13	**0.50**	0.50	N/A	0.50	1.00	N/A
VS 17	**0.50**	N/A	1.00	0.00	1.00	0.00
VS 23	**0.50**	1.00	0.00	1.00	0.80	1.00
VS 27	**0.50**	1.00	1.00	0.25	0.00	0.50
VS 2	**0.54**	0.75	0.67	0.00	0.50	0.33
VS 24	**0.56**	0.50	0.80	0.67	0.40	0.25
VS 1	**0.57**	0.67	0.25	0.60	0.50	1.00
VS 18	**0.60**	0.00	1.00	1.00	0.67	0.75
VS 6	**0.64**	0.33	0.50	1.00	0.33	0.50
VS 11	**0.64**	1.00	0.67	0.50	1.00	0.33
VS 9	**0.71**	N/A	0.60	1.00	1.00	1.00
VS 19	**0.77**	1.00	1.00	0.75	0.80	0.75
VS 10	**0.80**	N/A	0.83	1.00	1.00	0.50
VS 5	**0.86**	1.00	0.50	1.00	1.00	1.00
VS 22	**0.91**	1.00	0.75	1.00	1.00	1.00

Unfortunately, the number of participants does not allow a systematic analysis for each of the different groups. However, with few exceptions, they generally appear to be of similar opinions. The results are not definite – there are too many categories for the number of data points, so further research on this phenomenon is necessary. They do suggest, however, that going ahead with the aggregated data is not likely to lead to distorted results.

Excluding the above sentence VS 7 from further analysis, and not counting the 'I don't know'- type answers, this part of the survey totals 309 responses, which makes it very comfortably possible to analyse statistically. Due to the randomized nature of giving out sentences, varying amounts of responses have been collected for each sentence. As 30 sentences were included in the study, responses per sentence range from 7 at the lower end and 18 at the higher end, with a mean of just over 10 responses per sentence.

This section of the survey is the backbone of the study, as it serves as the basis for testing the hypothesis on vagueness. Before I come to that part, I will discuss the comments left by respondents.

5.2.2 Comments

Taken together, the survey gathered a total of 94 comments. Using qualitative content analysis based on Mayring (2010), I first ordered the comments by topic, then by their orientation in the topic. In the following, I will present the most common themes and the insights gained from each of them. This is a qualitative, not quantitative, analysis, aimed not at conclusive, comparable results but at gaining a better understanding of expert's opinions, and seeing which hypotheses might be worth pursuing.

The numbers of commenters are only included to provide rough impressions on general trends, and should not be read as inherently conclusive. Additionally, to present a detailed and in-depth picture, I often present similar amounts of comments on each side of a debate, even when the vast majority of commenters clearly favoured one particular argument.

The majority of the comments can be said to fit one of eight categories: verbs expressing obligations, references, definitions, hedge words, sentence length, measurability, the need for context and generality.

Most often, participants commented on the question whether or not verbs expressing a higher degree of obligation are also more precise. 28 comments in total were made on this topic. The vast majority of these were of the opinion that verbs associated with stronger obligations like 'shall' and 'will' are also indicative of a higher level of precision. One commenter stated that

> *"Those verbal expressions clearly ´demanding´ that a state follow that what is stated in the provision grant a more precise connotation to the reading of the text" (Interview 119, 03.12.2014, 11:33h).*

Another answered the question 'what makes sentences more vague' with the following:

> *"The use of certain expressions that diminish the level of obligation (or at least the sense of urgency of compliance) prescribed by the text, such as should instead of shall, recommend instead of urge etc." (Interview 180, 04.12.2014, 16:07h).*

Commenters also referred to specific verbs, stating for example that a sentence is vague

> *"[i]f only support for a certain goal is declared ("[re]affirm"/"note"/...) with no corresponding duties ("shall")" (Interview 359, 05.01.2015, 12:24h).*[28]

[28] The high numbers of certain interviews are due to the fact that all interactions with the survey have been counted continuously from the first testing stage, even those where someone only clicked on the link to the survey without answering any questions. This has no effect at all on the study but simplifies labeling.

or

> *"[w]hen they contain words implying that the sentence does not contain an obligation which is legally binding, e.g. "welcomes", "promotes", "should", "urges", "invites", etc..." (Interview 87, 02.12.2014, 16:11h).*

There were no comments proposing the opposite, i.e. that these verbs make sentences more vague, but four commenters disagreed with the premise entirely, arguing that binding force and precision are completely separate categories and thus these verbs do not influence the level of precision at all. One of them said

> *"In particular, the use of "shall" or "should" does not indicate a difference in precision, but a difference in obligation. They represent different things, not different degrees of the same thing" (Interview 123, 03.12.2014, 11:40h).*

Another participant made a very interesting distinction:

> *"adding further information makes sentences more precise; use of legally more determining vocabulary does not (it makes the legal consequences following from such phrase more precise, but not the sentence as such)" (Interview 99, 02.12.2014, 18:19h).*

Here, the distinction between legal indeterminacy and vagueness is clearly drawn. However, the legal indeterminacy in this case comes still from within the text, and not from conflicting legislation or other sources of contextual indeterminacy.

The second most comments were made on the terms generality and specifics. All but one of the fourteen commenters agreed that more general provision were vaguer, while more specific sentences were seen as more precise. One participant, when asked which features of a sentence they considered vague, replied "commitment to a general goal without precise steps to be taken in order to achieve it" (Interview 111, 03.12.2014, 9:30h). Other answers to the question were "general aims" (Interview 104, 02.12.2014, 18:42h), "general clauses", "too general phases" (Interview 336, 29.12.2014, 18:52h) and "broad statements with unclear policy implication" (Interview 179, 04.12.2014, 14:45h). On the other hand, when asked what made a sentence more precise, commenters answered with "use of specific terms" (Interview 309, 24.12.2014, 13:20h) or simply "specifications". Again, no participant stated that the reverse was the case, but one commenter argued that generality and specificity were entirely different categories from vagueness and precision. The following comment makes this point quite clearly:

> *"I couldn't decide which sentence is more precise quite often, because I had the impression that the difference between the two sentences was not vague vs. precise, but general vs. specific..." (Interview 118, 03.12.2014, 11:33h).*

This distinction relates back to the difference in use of the term vagueness in philosophical terms and everyday language discussed in chapter 2.1.1. It appears that most

respondents understand the term vagueness in the everyday sense of the word, while this last commenter uses the philosophical concept.

A third area eliciting comments concerned referencing of other legal texts. Most commenters were of the opinion that these kinds of references render a sentence more precise. For example, one commenter stated that

> *"The lack of reference to other – based upon more precise background, such as expert-based documents, resolutions, etc. – norms is also the sign or root of vagueness" (Interview 371, 14.01.2015, 16:39h).*

Another, asked about indicators for precision in sentences, cited "referrals to previous documents" (Interview 78, 02.12.2014, 15:05h). In this category, nobody made a point excluding references as indicators of vagueness or precision. Interestingly, however, three comments were made that indicate the exact opposite of the point made above, namely that references actually render a given legal sentence vaguer. One comment explained this position in the following way:

> *"Vagueness was due to [...] the constant ´references´ to other articles or preambles present in the Convention. This act of self referring [sic] may lead to a variety of interpretations, leaving the provision read [sic] very imprecise" (Interview 119, 03.12.2014, 11:33h).*

This comment draws the interesting parallel between references in the text and the point made in chapter 3.2.3, namely that fragmentation in the law leads to legal indeterminacy. However, this point of view remains a minority among respondents to the survey.

The inclusion and use of definitions was unanimously seen as enhancing precision by commenters. They said that "the use of clearly defined legal terminology" (Interview 94, 02.12.2014, 17:02h) makes sentences more precise, and that the "inclusion of terms which are understood differently in different contexts" (Interview 131, 03.12.2014, 12:30h) is what renders sentences vaguer. One commenter explained that

> *"Precision thus follows from [...] the inclusion of at least first order definitions of concepts (that is each concept is at least defined on a first level, that is the definitions includes necessary conditions of the concept; greater clarity could be reached by again finding necessary conditions for the necessary conditions of the concept" (Interview 126, 03.12.2014, 12:02h).*

Commenters also said that they considered shorter sentences to be more precise, while associating longer sentences with vagueness. Statements such as "Particularly compound or otherwise long sentences produce vagueness" (Interview 197, 11.12.2014, 13:39h) or "[precise sentences] went straight to the point, short, and relatively simple." (Interview 197, 11.12. 2014, 13:39h) For the commenters, vagueness via sentence length is strongly associated with complexity and readability, stating that "the shorter sentences the better for

understanding" when asked what made a sentence precise. Following this train of thought, grammatical clarity was seen as an indicator of precision as well:

> *"Vagueness seems to be related to [...] the grammatical setup of a sentence (in terms of unclear sentence structure)" (Interview 126, 03.12.2014, 12:02h).*

On the other hand, the inclusion of more information in a sentence – necessarily lengthening it – was mostly seen as reducing vagueness. As such, one participant commented that "more informative / explanatory text" (Interview 306, 24.12.2014, 9:01h) is indicative of precision, another said "adding further information makes sentences more precise" (Interview 99, 02.12.2014, 18:19h).

Commenters were not undivided on this issue, however. One respondent said that

> *"Sometimes, by introducing a description, one can actually make the initial statement less precise" (Interview 346, 03.01.2015, 11:54h).*

These related topics are a little less clearly biased toward one direction than the ones discussed above. However, the general trend is that shorter and less complex sentences are seen as more precise.

Commenters did agree on the topic of hedge words (although not usually calling them by this name). The seven comments on this topic are in agreement that hedge words render sentences vaguer. Examples included were "to the extent possible" (Interview 77, 02.12.2014, 14:52h), "as appropriate"(Interview 77, 02.12.2014, 14:52h), "equitable" (Interview 105, 02.12.2014, 19:10h), "timely" (Interview 105, 02.12.2014, 19:10h), "regularly" (Interview 257, 22.12.2014, 21:38h), "effective"(Interview 91, 02.12.2014, 16:49h) and "soon" (Interview 199, 18.12.2014, 22:52h). They referred to these words as "undetermined terms" (Interview 105, 02.12.2014, 19:10h) or "vague descriptions that are not really measurable".

This last quote leads to the next topic of comments: measurability. Five experts said that to be considered precise, sentences must include measurable concepts. The following comment illustrates the sentiment:

> *"Also vague are descriptions that are not really measureable such as 'as much as possible' or 'an environment that will promote...'. As an example, 'economic growth' is measureable. 'an environment that will promote economic growth' is not" (Interview 346, 03.01.2015, 11:54h).*

This discussion relates back to the philosophical conception of vague concepts as immensurable and incommensurable as discussed in chapter 2.1.1.

Three commenters called into question the entire premise of the survey by arguing that vagueness can only ever be identified from the context of a sentence, never from the sentence itself. I will provide all three comments in their entirety for two reasons: First, since they are critical of the work done in this thesis, I want to render them unaltered, so as to not

fudge interpretation to my advantage, but rather let the reader judge the value of each of the arguments for themselves. Second, I think that the debate is important and merits detailed attention and discussion. I also want to give them sufficient room because, in a way, the survey was biased against them – not intentionally, but simply by the fact that I was not explicitly inviting participants to comment on the lack of context, and instead implicitly made it clear that I hoped to find indications of vagueness in the sentences themselves through the design of the survey. Therefore, I consider that there position may be more relevant or widespread in the community of experts in IL than the low number (3 out of 68) commenters might suggest.

> *"My experience tells me that there is not any sentence which cannot be extremely vague. It depends on the context. Since no context was given, it is impossible to decide how vague a sentence is. Hence, the survey is based on a fundamental misunderstanding of legal practice and methodology. Only the context would allow for the identification of precision. Since no context was given, most questions could not be decided" (Interview 123, 03.12.2014, 11:40h).*
>
> *"It is not easy to rate the vagueness of the sentences without knowing the context, for example no. 6 'A clean development mechanism is hereby defined' – what does 'hereby' refer to? What is a clean development mechanism?" (Interview 49, 19.11.2014, 13:57h).*
>
> *"You seem to believe that vagueness could be evaluated by merely looking at singular sentences of a legal text. However, some of your cited sentences refer to other clauses within the legal text – its vagueness can therefore only judged by looking systematically at the sentence within its context (i.e. systematic interpretation). That is because the sentence receives its content also through these other sentences" (Interview 126, 03.12.2014, 12:02).*

The last two comments seem to me to be less absolute than the first, because they explicitly or implicitly state that the problem of context is present – or at least more pertinent – for some sentences than for others.

I agree with these commenters on a central point: context is very important for the study of vagueness. The choice not to include context in the data collection has been a methodological one, and while I remain convinced that it was necessary in the interest of gaining any usable data, I agree that it also makes the study incomplete: I will not be able to present a comprehensive list of factors influencing the vagueness of legal text. This is a hard limit of this study, and one I recognize. However, I am confident that context is not the only possible way to identify a vague sentence, but argue that there are also indicators at the sentence level. For one thing, this is supported by the other participants and commenters, who did find differences in vagueness between the sentences and identified indicators. Furthermore, most of the literature we saw earlier bases its examples on passages of one or two sentences, no more, clearly showing that however important context may be, there are also aspects of the sentences themselves which are relevant for the study of vagueness. My

third argument is epistemological: If there were absolutely no indication for vagueness in a single sentence, the context – which is also comprised of sentences – could not contain any vagueness, either.[29] There could be legal indeterminacy, because of contradictions amongst sentences or simply because the documents themselves could be unclear in their binding force, or because of characteristics of legal systems. But for there to be vagueness in the text, some traces of it must be found in at least some of the sentences, even if the presence of other sentences can then mitigate or exacerbate it.

5.3 What makes sentences vague?

In the previous sections, I described the survey design and its overall results. The next two chapters are concerned with analysing the data thus gained and ultimately making a tool which can provide an automated comparison of the vagueness of different international legal agreements. In a first step, I will test nine hypotheses – drawn both from the literature and comments discussed above – on the data gained by the survey. The next step is to take the hypotheses which proved to have a statistically significant relationship with the perceived vagueness of a sentence and combine them into a tool which can process a large amount of legal text quickly.

5.3.1 Testing hypotheses

Both the linguistic approaches in Chapter 4.1, as well as the comments by survey respondents shown in section 5.2.3 above can be used to identify several hypotheses on which features of a sentence make it vaguer and which make it more precise.

In the literature discussed in chapter 4.1, hedge words, general extenders, vague quantifiers, placeholder words, mass nouns, determiners, and shields were identified as potentially relevant indicators for vagueness in text. The sparse literature on identifying vagueness in international agreements (discussed in section 4.2.3) adds modifiers, modal auxiliaries, determiners, numbers and the use of passive voice as features to look out for. Survey respondents (see section 5.2.3) have claimed that verbs expressing obligations (here referred to as explicit performatives), hedge words, sentence length, referencing, generality and measurability affect the vagueness of a legal sentence. Not all of the mentioned indicators can be tested within the limits of this study. Some of them, like placeholder words,

[29] On the notion that the context is independent from the indeterminacy of the text, see also Peller (1985: 1173) as discussed in section 2.2.2.

are clearly not applicable to legal language. Others, like generality or referencing, are exceedingly hard to operationalize in a manner suitable to quantitative data analysis. Keeping these considerations in mind, eight indicators are included in this study.

Leaving out those criteria too vague to be operationalized and those that clearly do not apply to legal language leaves a number of hypotheses which will be discussed below. The first hypothesis, based on the comments of survey respondents, is that longer sentences are perceived as vaguer. The second hypothesis, which is supported by linguistic literature in general, by studies related to legal language specifically, as well as by survey respondents, is that the presence of numbers renders a sentence more precise than it would otherwise be. A third hypotheses, which is similarly widespread, is that that sentences containing hedge words are vaguer than those which do not. In order to check this hypothesis against the one that would ascribe an effect on vagueness by all modifiers (which has been suggested by literature on legal language), I also check the effect of the number of adjectives and adverbs on the perceived vagueness of a sentence. A fifth hypothesis, which is supported by literature on both general and legal language, but is not supported by survey respondents, is that determiners have an effect on the vagueness or precision of sentences. Two more hypotheses, which appear to be specific to legal language and have also been mentioned in the comments of survey respondents are that the presence of a weak explicitly performative verb renders a sentence vaguer, while the presence of a strong explicitly performative verb leads a sentences to be perceived as more precise. Additionally, I have included the hypothesis that the regime of origin of a sentence may influence its perceived vagueness. This hypothesis does not have its roots in the literature or survey comments, but it is necessary to test for in order to be able to estimate the limits of application for this study. In the following, the quantitative data from the survey is used to test these nine hypotheses.

Hypothesis 1: sentence length

One of the features mentioned in the comments of the survey was sentence length. As shown in section 5.2.3, most commenters agreed that shorter sentences are an indication of a sentence being more precise, while longer sentences indicate more vagueness.

The indicator I chose for the length of a sentence were the number of words it contained. Alternatives could have been the number of letters, but using word count as an indicator seemed adequate for this study because it is more intuitively understandable – it is easier to picture a sentence containing 10 words than one containing 65 letters. The first step to

testing this hypothesis, then, was to establish the number of words in each of the tested sentences. The longest sentence comprised 129 words,

> *"The Conference of the Parties [...] further recognizes that deep cuts in global greenhouse gas emissions are required according to science, and as documented in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, with a view to reducing global greenhouse gas emissions so as to hold the increase in global average temperature below 2 °C above pre-industrial levels, and that Parties should take urgent action to meet this long-term goal, consistent with science and on the basis of equity; also recognizes the need to consider, in the context of the first review, as referred to in paragraph 138 below, strengthening the long-term global goal on the basis of the best available scientific knowledge, including in relation to a global average temperature rise of 1. 5 °C" (VS 11),*

while the shortest was only 10 words long:

> *"The Conference expresses its strong support for effective IAEA safeguards" (VS 1).*

On average, the sentence length was around 45 words, with the median at 43.

The next task was to establish a dividing point. Sentences under this point are considered short for the purposes of this study, sentences with the same number or more words than that are considered long. In order to keep approximately the same number of sentences in each group, I chose a cut-off point at 41.

The resulting table is shown below:

Table 2: Answers of respondents by length of sentence

	Not at all vague	Contains some vagueness	Fairly vague	Very vague	Sum
1-39 Words	29	38	41	30	138
40+ Words	33	60	41	32	166
Sum	62	98	82	62	304

Combining the first two categories under the label 'not vague' and the last two categories under 'vague' turns out a simpler table for better overview:

Table 3: Number of respondents considering a sentence vague by length of the sentence

	Not vague	Vague	Sum
1-40 Words	67	71	138
41+ Words	93	73	166

Sum	160	144	304

This table foreshadows the results discussed below: While the number of people who consider a sentence precise rises slightly when considering longer sentences, the number of people who consider it vague stays approximately the same regardless of its length. This in itself is interesting, as it contradicts the impression gained from looking at the comments: If anything, longer sentences were considered slightly more precise. The following graph visualizes this result:

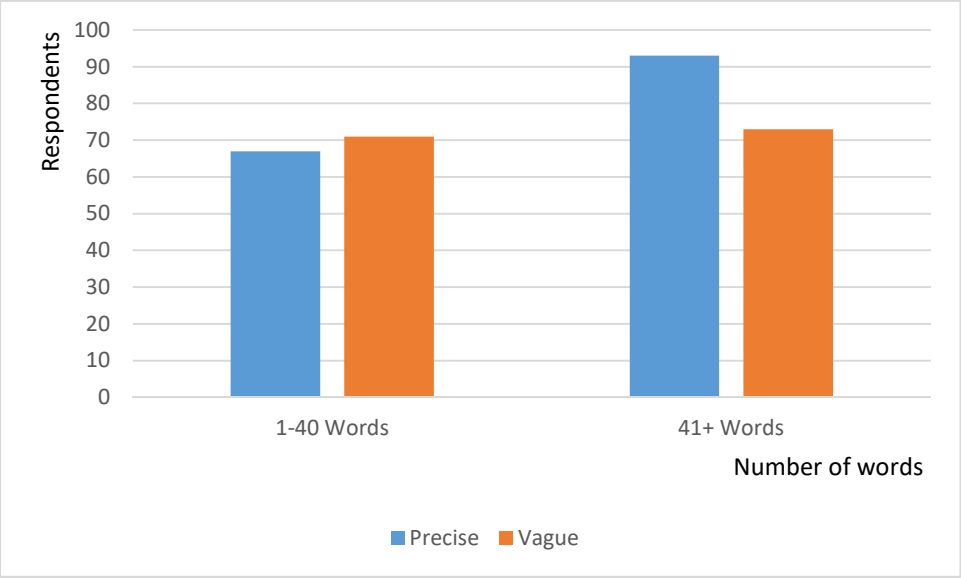


Figure 3: Respondents considering sentences precise or vague, by sentence length

Taking a closer look at the numbers, however, reveals that there is no statistically relevant correlation at all. When plotting all the combination of sentence length and assessment of vagueness on a scatter graph, this becomes intuitively apparent:

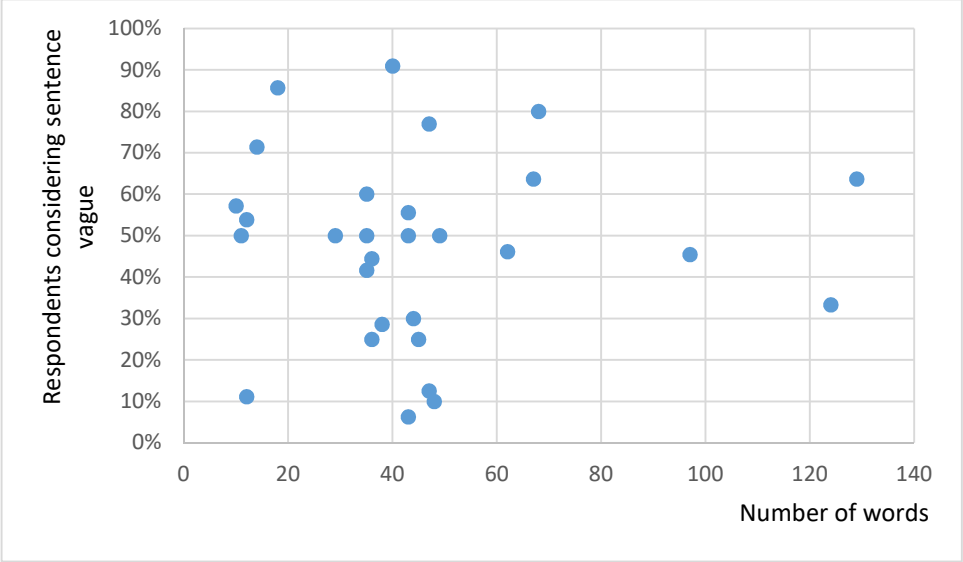


Figure 4: Perceived vagueness by sentence length

Spearman's rho[30] (which is the appropriate correlation coefficient for ordinal scales (Gingrich 1992: 796) is -0,032. This number is close to 0, which suggests that the two variables are independent of each other. The result does not change in any meaningful way if the data is expressed in the form of percentages of people who consider a sentence vague.

While the results vary slightly if the expertise status of the respondents is taken into consideration, none of the numbers suggest any statistical correlation or significance. Below is a table of Spearman's rho values for the different groups:

Table 4: Spearman's rho values of the relationship between vagueness and sentence length by respondent group

Respondent group	Spearman's rho values
Students of IL	-0.157
PhD Students in IL	-0.021
People with PhD in IL	0.083
Legal Professionals	-0.085
Other	0.119

In light of this, I draw the conclusion that the perceived vagueness of a sentence does not correlate with its length. This result is particularly interesting considering the high number

[30] "[Spearman's rho] produces a correlation coefficient which has a maximum value of 1, indicating a perfect positive association between the ranks, and a minimum value of -1, indicating a perfect negative association between ranks. A value of 0 indicates no association between the ranks" (Gingrich 1992: 820).

of survey respondents who explicitly commented that length is a factor in how vague they consider a sentence to be.

Hypothesis 2: numbers

Several commenters mentioned that they perceived a sentence as vague when its content was 'measurable'. Since it is not possible to measure this feature directly, I simplified it by testing whether or not a sentence contained numbers (both cardinal and ordinal numbers were counted). The presence of numbers also features heavily as a factor for vagueness in the relevant literature (see section 4.1.1).

Contrary to expectations, however, the participants of this survey did not perceive sentences containing numbers differently from those which did not. The complete data on the presence of numbers and the ratings of participants are summarized in the table below.

Table 5: Survey responses by number of numbers

	Not at all vague	Contains some vagueness	Fairly vague	Very vague	Sum
0 Numbers	29	54	48	41	172
1 Number	6	16	17	12	51
2 Numbers	6	12	6	2	26
3+Numbers	21	16	11	7	55
Sum	62	98	82	62	304

The value of the test used to calculate correlation is -0,210, which is not enough to reject the null hypothesis, both because it only shows a weak correlation and because it is not statistically significant. This result does not change significantly (to a value of -0.230) when the sentence length is taken into account.

Given that the emphasis on numbers in linguistic literature is almost equal to that of hedge words, this result is among the more surprising of this study. However, there are several possible explanations for numbers working in different ways in legal language than they do in everyday use. For one thing, numbers are often used to reference other provisions of the same agreement, or provisions in different agreements. Without knowing their content in turn, it is not possible to say whether they render the given sentence more or less vague and therefore there should not be any effect in the perceived vagueness. Another factor may be that numbers in legal agreements are frequently used in minor detail which

do not affect the main meaning of the provision. Consider the following sentence, which was also included in the survey:

> *"The Conference of the Parties serving as the meeting of the Parties to this Protocol shall, at its first session or as soon as practicable thereafter, decide upon modalities, rules and guidelines as to how, and which, additional human-induced activities related to changes in greenhouse gas emissions by sources and removals by sinks in the agricultural soils and the land-use change and forestry categories shall be added to, or subtracted from, the assigned amounts for Parties included in Annex I, taking into account uncertainties, transparency in reporting, verifiability, the methodological work of the Intergovernmental Panel on Climate Change, the advice provided by the Subsidiary Body for Scientific and Technological Advice in accordance with Article 5 and the decisions of the Conference of the Parties" (VS 8).*

The words in question are 'first', '(Annex) I', and '(Article) 5'. While the numbers certainly specify matters – to which session, which annex and which article the provision is referring to – these numbers hardly have any impact on the vagueness or precision of the main content of the provision in question. It is possible that the minority of numbers in legal agreements that are related to the actual content of the provision function similarly to numbers in everyday language, i.e. making the provision more precise. If that is the case, the effect is obscured by all the other numbers present in legal documents. It may also be that legal experts see numbers as less important than other indicators of vagueness (although this option contradicts respondents' statements in the comments to this survey). In any case, the matter requires more detailed investigation than is possible in the framework of this study. For the time being, the hypothesis that numbers lead to a sentence being perceived as more precise is not supported by the available data.

Hypothesis 3: hedge words

One widespread assumption in the literature is that hedge words make sentences vaguer (see chapter 4.1.1). The survey participants agreed in the comments (see section 5.2.3). The first step to accomplish here was to choose the words to be included in the category of hedge words. This was somewhat limited by the original material, since it is not possible to test for words that did not appear in the 30 sentences that were tested. I settled on the following words: 'relevant', 'appropriate', 'available', 'specific', 'particular', 'well', 'effective', 'good', 'possible', 'general', 'practicable', and 'regularly'. All of these words fulfil the general criterion of being hedge words in the sense that their addition makes the language of a sentence vaguer than it would be without it. Each of the 30 sentences contains either none, one or two of these words. Therefore, I established a table with these three categories.

Table 6: Survey responses by number of hedge words

	Not at all vague	Contains some vagueness	Fairly vague	Very vague	Sum
0 Hedge words	51	46	29	28	154
1 Hedge words	4	30	36	15	85
2 Hedge words	7	22	17	19	65
Sum	62	98	82	62	304

Again, for the better overview, here is the same table with the four options condensed into two:

Table 7: Respondents considering sentence vague or precise, by number of hedge words

	Precise	Vague	Sum
0 Hedge words	97	57	154
1 Hedge words	34	51	85
2 Hedge words	29	36	65
Sum	160	144	304

The difference to the graph on sentence length becomes clear immediately. While the number of people considering sentences containing zero hedge words precise is almost double the amount of people who consider them vague, this picture has changed significantly when sentences with two hedge words are considered. In this case, the number of people considering the sentence precise is lower than the number of respondents who see it as vague. Again, let's look at a visualization of the table to make the finding more accessible:

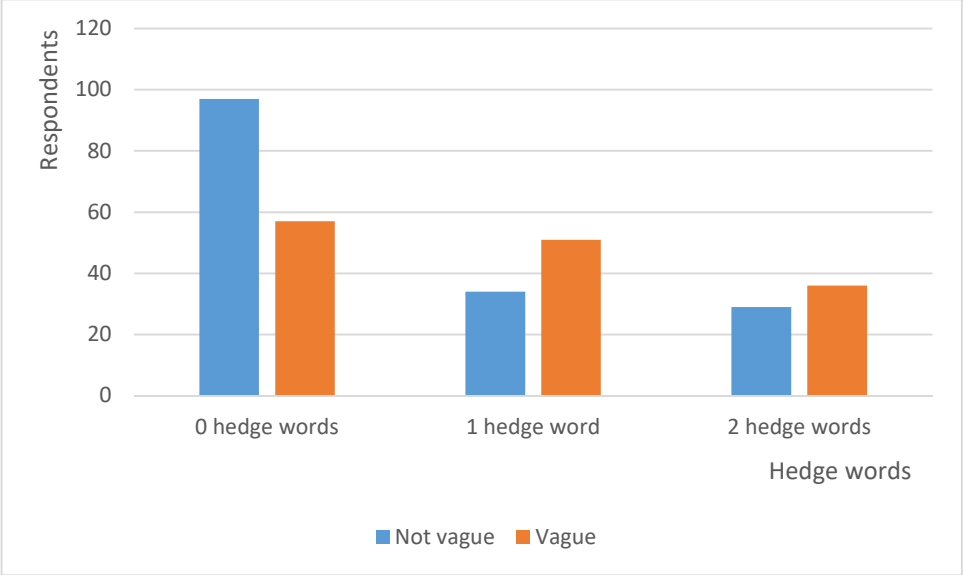


Figure 5: Respondents considering sentences precise and vague by number of hedge words

Interestingly, the ratio of people considering a sentence vague does not seem grow when moving from one hedge word to two – it even decreases slightly. This puzzling occurrence may be solved when we turn away from absolute number of hedge words and consider the percentage of hedge words in a sentence, when set in relationship to the total number of words. It is possible that sentences with two hedge words simple tend to be much longer, so that the presence of an extra hedge word does not alter the perception of the meaning that much given the amount of information present in the sentence.

However, when looking at the numbers, this explanation does not seem to hold, either. When taking the percentages of hedge words as the basis for analysis, the above graph looks as follows:

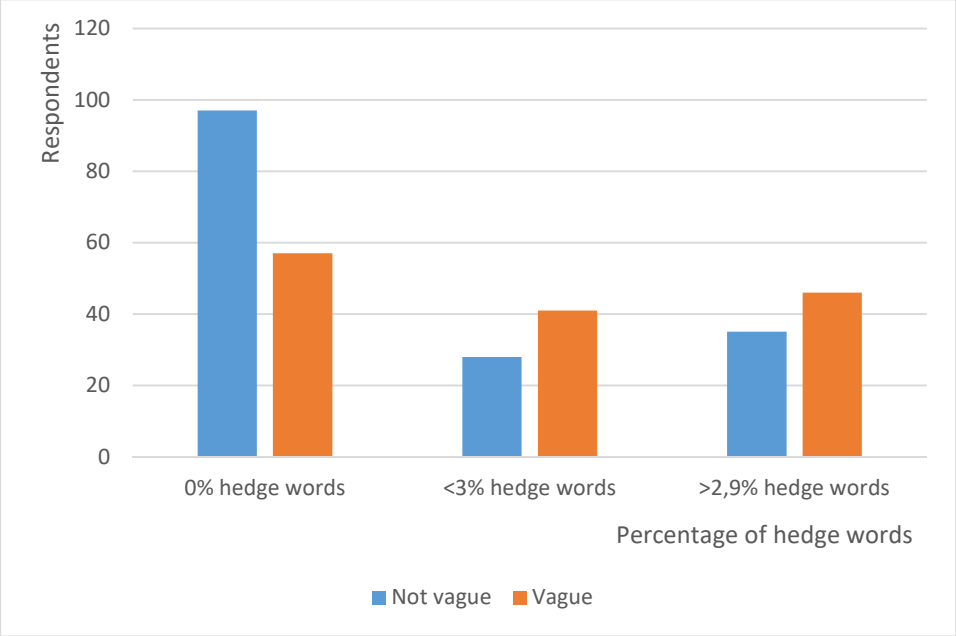


Figure 6: Respondents considering sentences precise and vague by percentage of hedge words

The difference does not grow in the last two categories, but rather declines very slightly. It looks like sentences without any hedge words are seen as more precise than usual, while sentences including at least one hedge word – regardless of their number or the length of the sentence – are seen as somewhat more vague.

Spearman's rank correlation coefficient is 0.442, which means that the number of hedge words (relative to the number of words) in a sentence is moderately positively correlated with the percentage of people who thought the sentence was vague. The relationship between the pairs of data is visualized in the following graph:

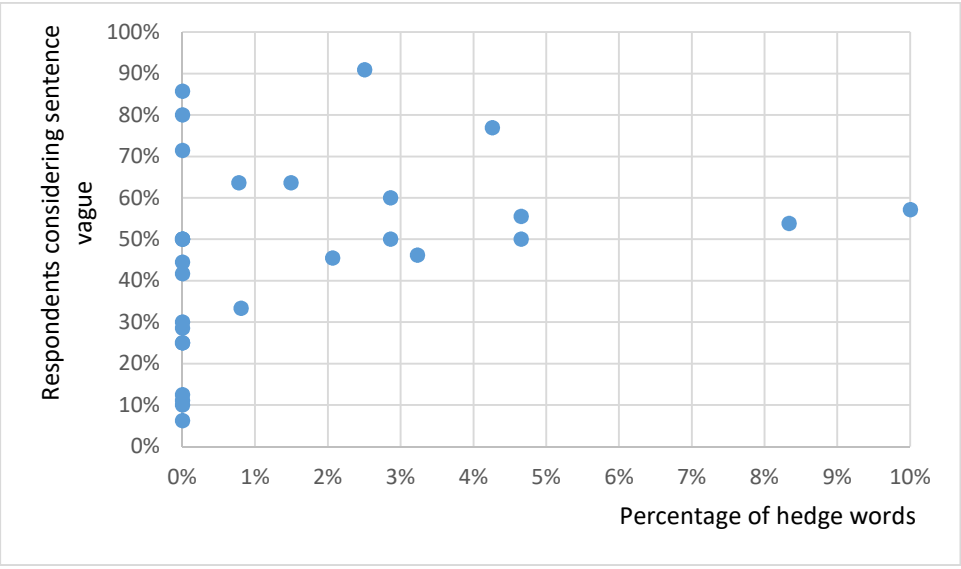


Figure 7: Perceived vagueness by percentage of hedge words

The more hedge words there are – or in this case, if there are hedge words at all – the vaguer experts in IL will perceive it as. The correlation is significant at $p<0.05$, which means that there is less than a 5% chance that this result has been obtained by chance.

In light of the heavy emphasis on hedge words in linguistic literature, this result is not particularly surprising. It seems that hedge words function similarly in the legal language than they do in everyday language use, at least when it concerns vagueness.

Hypothesis 4: adjectives and adverbs

At first glance, there appears to be no significant relationship between the number of adjectives and adverbs and the number of people who perceive a given sentence as vague. However, when modulating the relationship by sentence length, a moderate, but statistically significant positive correlation emerges – the more adjectives and adverbs there are per sentence, the vaguer it appears to be perceived. The resulting graph illustrate these results.

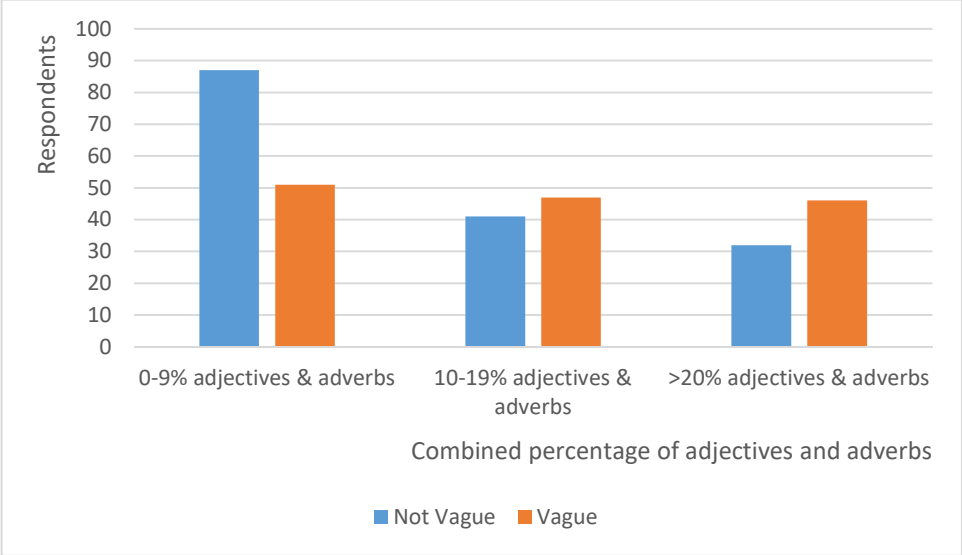


Figure 8: Respondents considering a sentence precise and vague by combined percentage of adjectives and adverbs

This is actually at odds with the hypotheses gained from studying the literature: adjectives in general are assumed to make a sentence more precise. Because the result is so surprising, I studied the numbers even more closely and found a different explanation. As shown above, the number of hedge words correlates positively with vagueness. Hedge words are primarily – and in this case, exclusively, adjectives and adverbs, so the two hypotheses are not actually independent of each other.

When removing hedge words from the number of adjectives, the weak positive correlation with vagueness, while still present, is no longer statistically significant. It is therefore not included in the list of vagueness indicators in this study.

Hypothesis 5: determiners

Determiners are words like your, their, these, some, whose or which[31]. These words put the other words of a sentence in relation to each other and provide some context. In the literature covered in chapter 4.2.3, determiners are seen as one possible factor for vagueness in legal documents, although whether they render a text more vague or more precise remains unclear. In this study, sentences with more determiners were seen as more precise than sentences that contained a lower number of them.

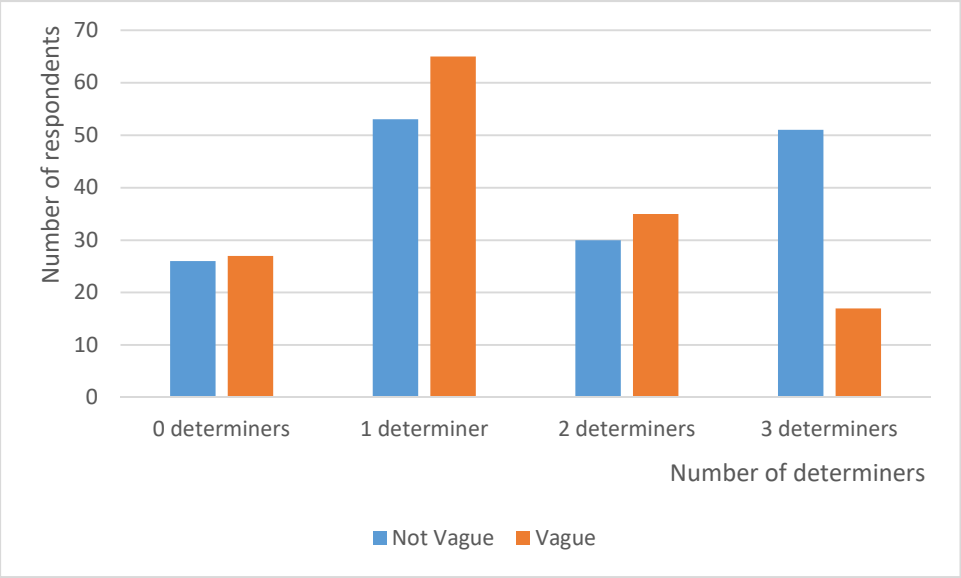


Figure 9: Respondents considering a sentence precise and vague by number of determiners

The graph shows that sentences with three determiners are much more often considered precise than is the case otherwise. Further tests indicate that this result is not dependent on the length of the sentence. There appears to be a moderate, but statistically significant negative correlation between the absolute number of determiners a sentence contains and its perceived vagueness.

[31] For a definition of the term ‚determiner' see section 4.1.1.

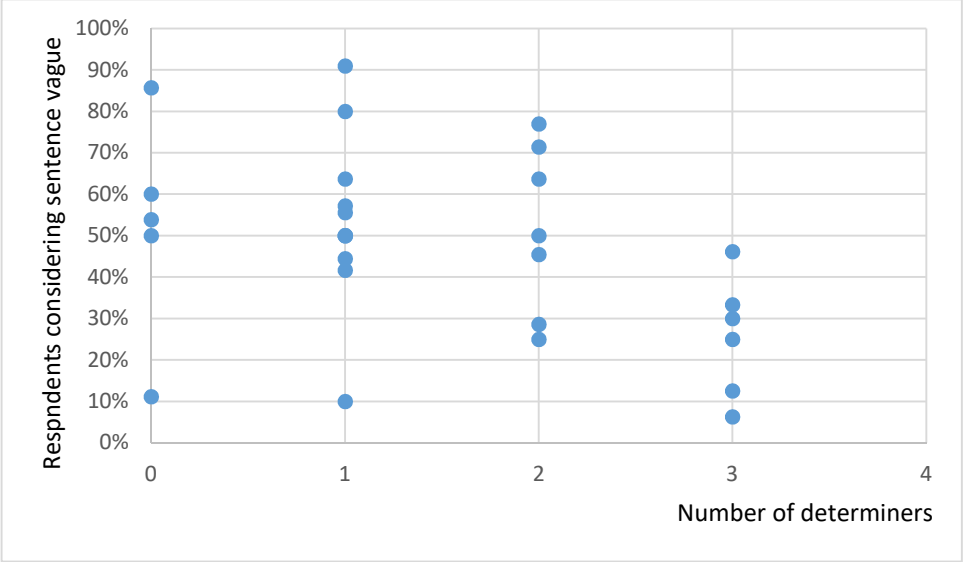


Figure 10: Perceived vagueness by number of determiners

Neither the literature nor the comments of survey respondents provide any real explanation for this relationship. It could be speculated that those determiners most commonly used in legal text specify the meaning of the nouns they are attached to, in a similar way as in the case of adjectives discussed in section 4.3. The significant drop in perceived vagueness when sentences reach three determiners could be explained by the fact that while all sentences use some determiners, specific legal language uses an unusually high number of them. In any case, the relationship between vagueness and the number of determiners in a legal sentence definitely merits further study.

Hypothesis 6: weak explicit performatives[32]

Lastly, I come to two of the most interesting results. Verbs might be the most important type of words in the legal language (see section 4.2.2 for a more detailed discussion), but they are not very prominent in linguistic literature on vagueness indicators (although they are sometimes mentioned, see chapter 4.1.1). There is therefore some incentive to check whether, and in what way, explicitly performative verbs influence the perceived vagueness of legal sentences.[33] When the operational verb of a sentence expresses only a weak intent to be legally binding – which is the case with verbs like 'should' or 'undertakes to'[34] – international legal practitioners appear to perceive the sentence as vaguer than average. The relationship is easily visible in a graph:

[32] For an explanation of what explicit performatives are, see chapter 4.2.1.

[33] This matter is closely related to the discussion on vagueness vs. bindingness outlined in section 5.2.3.

[34] The full list of explicitly performative verbs which tested as vague in this study were ‚would', ‚could', ‚should', ‚underline', ‚underscore', ‚recognize', ‚have the right', ‚call', and ‚undertake'.

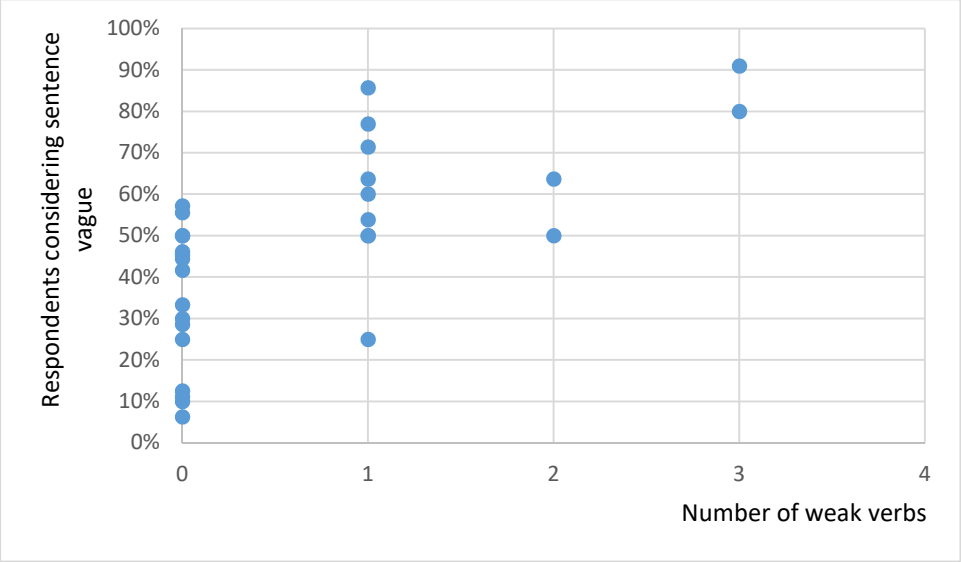


Figure 11: Perceived vagueness by number of weak explicitly performative verbs

The Spearman's rho of 0.698 suggests a strong positive correlation between vagueness and the presence of a weak operational verb. This is the highest result among all the hypotheses.

Hypothesis 7: strong explicit performatives

On the other end of the scale, verbs like 'shall' and 'will' express a strong binding force. In keeping with the previous hypothesis, sentences containing these words are seen as more precise than others. The correlation is moderate and negative.

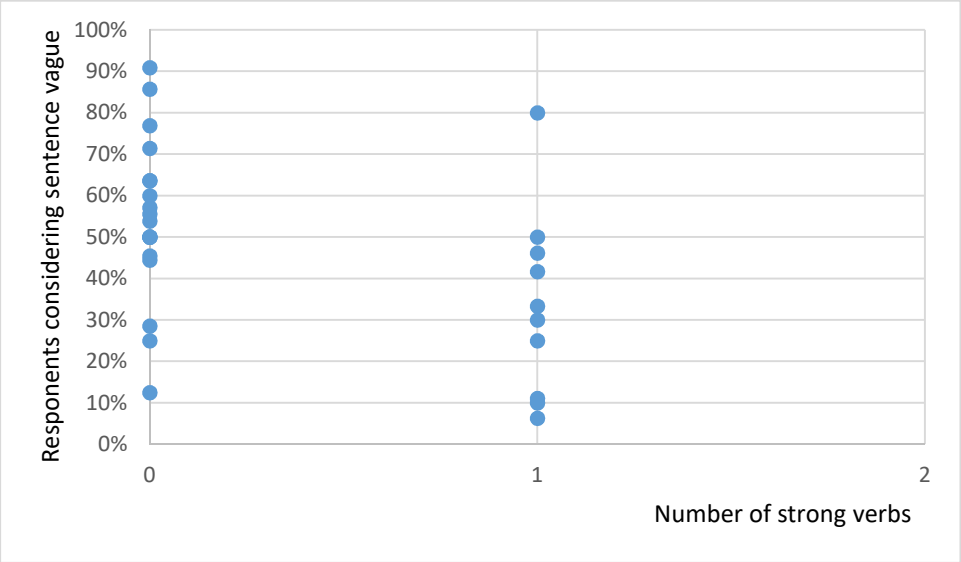


Figure 12: Perceived vagueness by presence of strong explicitly performative verbs

The above graph clearly shows that while sentences without strong explicitly performative verbs (represented in the first column of graph 14) have a wide range of vagueness, almost all sentences containing the verbs ‚shall' or ‚will' are seen as less vague

than average, with only one exception. These verbs tend to make sentences be seen as more precise. The outlier sentence which can be seen in the chart is the following:

> *"To achieve the ultimate objective of the Convention to stabilize greenhouse gas concentration in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system, we shall, recognizing the scientific view that the increase in global temperature should be below 2 degrees Celsius, on the basis of equity and in the context of sustainable development, enhance our long-term cooperative action to combat climate change".*

This exception notwithstanding, the negative correlation between strong explicit performatives and vagueness remains clear, and provides a counterpoint to the weak explicit performatives discussed above. Some explicitly performative verbs, like 'remain convinced', and ‚note', appear to have no relationship with a sentence's perceived vagueness.

Hypothesis 8: regime

The hypothesis that the regime from which the sentence is taken may influence its perceived vagueness stands out in two respects. For one, it is the only characteristic tested for which is not inherent in the sentence itself, but lies in its context. Secondly, it does not stem from the literature or the survey participants. Nevertheless, it is an important case for the interpretation of the following results. If there turns out to be a significant difference between sentences from each regime, it would be problematic to analyse them in aggregate form. Furthermore, it might reveal a bias by the people surveyed: a large difference in assessment could indicate that experts were influenced by the context of the sentence (because they believe one regime to be vaguer than the other) and did not judge the sentence on its own.

However, there does not appear to be a correlation between the regime a sentence is taken from and the number of people perceiving it as vague. While there is a very slight tendency for sentences from the UNFCCC to be perceived as vaguer than the NPT (the value of spearman's rho is 0.149 and -0.149, respectively – the mirroring results from the fact that all sentences were taken from one of the two regimes), these numbers are far from statistically significant and should not be taken as meaningful differences.

Summing up, five of the eight tested hypotheses are not supported by the data: this study cannot support a correlation between vagueness and the regime, sentence length, how many adjectives are used, the use of conjunctions and how many numbers are used in a sentence. Four hypotheses, however, revealed a statistically significant correlation: the percentage of hedge words, the vague and precise explicitly performative verbs and the

amount of determiners present in a sentence. Hedge words come as no surprise, given their prominence in studies on vague language (see chapter 4.1.1). Determiners, as well, have been hypothesized to be indicators for the vagueness of legal sentences (see chapter 4.2.3). Verbs are less prominent in ordinary language studies of vagueness. However, given the peculiarities of legal language (see chapter 4.2.2), and the prominent role verbs play therein, it is not surprising that verbs would also feature prominently in how vague a legal sentence is perceived to be.

5.3.2 Building the tool

In the previous section, I have shown which hypotheses on what leads sentences in international agreements to be perceived as vague are supported by the data. The goal of this study is to build a tool which makes it possible to quickly assess how vague the sentences of an international agreement are. To that end, the most diagnostic findings of the previous analysis are combined into a points system.

Upon closer inspection, four hypotheses appear to be supported by the data:

- Hedge words make sentences vaguer
- A low number of determiners make sentences vaguer
- Weak explicitly performative verbs make sentences vaguer
- Strong explicitly performative verbs make sentences more precise.

In order to facilitate later assessments and analyses, I have combined these four elements into a points system. If a sentence contains one or more hedge words[35], it receives 1 point. If it contains less than 3 determiners[36], it receives another point. If it contains any of the verbs 'would', 'could', 'should', 'underline(s)', 'underscore(s)', 'recognize(s)', 'has/ve the right', 'call(s) for' or 'undertake(s)', it receives 1 point, and if it contains either 'shall', 'will' or 'remain(s) convinced', it receives -1 point to indicate the enhanced precision.

When these points are added up, each sentence can now be rated with a score between -1 and 3. This allows for some differentiation between sentences, while also keeping the scoring system relatively straightforward and intuitively related to the results.

When comparing the scores of the sentences with the percentages of people who rated them as vague, the resulting scatter plot looks as follows:

[35] To repeat, the hedge words used in this study are 'relevant', 'appropriate', 'available', 'specific', 'particular', 'well', 'effective', 'good', 'possible', 'general', 'practicable', and 'regularly'.
[36] As identified by the CLAWS5 tagging method (http://ucrel.lancs.ac.uk/claws5tags.html).

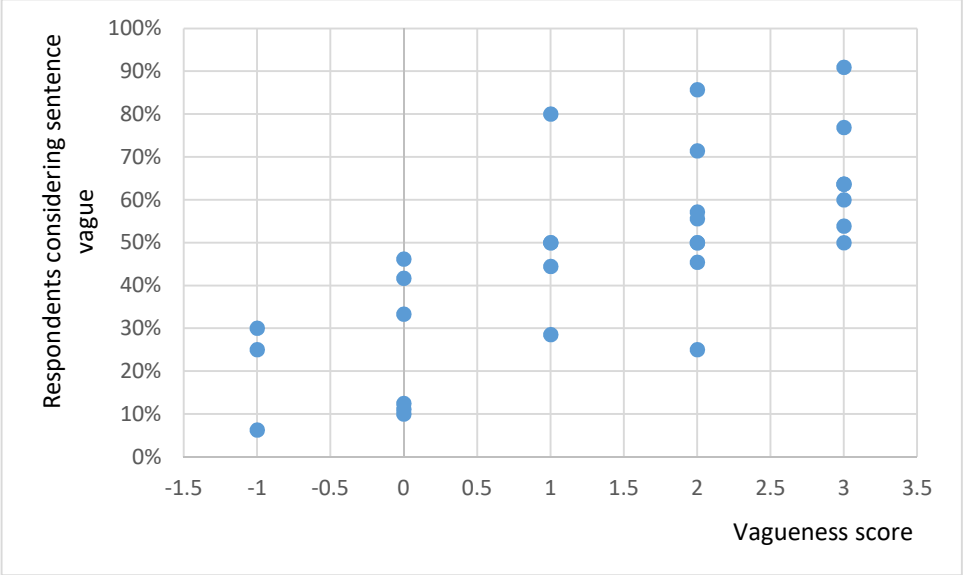


Figure 13: Vagueness score by perceived vagueness

The graph shows that this score and the perceived vagueness of a sentence are positively correlated. The higher the score a sentence receives, the likelier it becomes that survey participants rated it as vague. The corresponding spearman's rho correlation test is at 0.751, which means that there is a strong positive correlation between the combined features and the perceived vagueness of a sentence. The result is also statistically significant at $p<0.5$. By combining the indicators, a stronger correlation is achieved than any of the single indicators, enhancing the predictive value of the tool. In effect, if a sentence in an agreement has a score of three (i.e. it contains at least one hedge word, less than three determiners, and a verb of weak explicit performative, but no strong explicitly performative verb), this can be translated as saying 'It is likely that a person with expertise in IL would consider this sentence vague'.

5.4 Conclusion: indicators of vagueness in international agreements

In the course of this chapter, eight indicators were tested as to their relationship with perceived vagueness in international legal agreements. These were the regime of origin, sentence length, presence of hedge words, presence of numbers, number of adjectives and adverbs, number of determiners, weak explicitly performative verbs and strong explicitly performative verbs. Of these, four proved to have a statistically significant relationship with the perceived vagueness of a sentence: If a sentence contains at least one hedge word, if it contains less than three determiners, or if it contains a weak explicitly performative verb, the sentence is perceived as comparatively vaguer than other sentences by survey respondents. Conversely, if a sentence contains a strong explicitly performative verb, it is perceived as comparatively more precise. These four indicators are combined to constitute a vagueness

scale ranging from -1 (very precise) to 3 (very vague) with which to rate sentences in international legal agreements. Each sentence is assigned a binary score for each of the categories (1 or 0 for the first three categories, -1 or 0 for the last). Simple addition provides a combined vagueness score for each sentence. In a next step, the scores for each sentence in a legal document can be used to calculate an average vagueness score for the entire legal agreement. The automated nature of this tool allows the processing of large amounts of text in very little time.

The following chapter will apply this measure to 17 different international agreements. This will allow a test of the vagueness measuring tool in practice, explore a range of different comparisons, and a detailed study of what each component of the score actually signifies in legal sentences.

6 Finding vagueness in international agreements: Applying the vagueness score

In this chapter, I will apply the vagueness score developed in chapter 5.3.2 to a total of 17 documents stemming from two international regimes: the NPT regime and the UNFCCC. The first part of the chapter contains some clarifications on comparison groups. In a second part, I will discuss some high-level aggregate data, in order to point out general trends and compare some groups of documents. In section 6.3, I will look at the documents in more detail, discussing the particularities of each and also translating some of the quantitative data back into actual characteristics of legal sentences by giving examples of sentences scoring in each of the categories. In a fourth part of this chapter, I will compare the results of this study with previous assessments by researchers as to the vagueness of documents. While a comparison of vagueness has never been done systematically, there are some individual assessments. This enables me to discuss various aspects of the tool developed here and to see where it overlaps with these previous assessments as well as the cases in which it leads toward different conclusions. Section 6.5 takes up insights gained in the two previous sections as to the limitations and challenges of the tool and starts a discussion on further research.

6.1 Comparison groups and legal status

The analysed documents can be categorized in several ways. In this study, I concentrated on comparisons by regime (UNFCCC or NPT), text category (preamble, main text, annex,

titles, or final clauses), and legal status (decision, declaration or Convention/Protocol), though other possible axes for comparison certainly exist. In this section, I will briefly detail the different comparison groups, focussing especially on the legal status of documents.

The comparison of and within regimes is fairly straightforward. Separating documents by their regime of origin and only comparing documents has several advantages. For one, each regime may have its own idiosyncratic speech patterns which might make it harder to compare vagueness between regimes than within one. Additionally, the context and the content within a regime are much more fixed, so that the comparison is more likely to be focussed on actual differences in vagueness and not as susceptible to possible outside influences. On the other hand, it can also be beneficial to compare the vagueness of documents stemming from two or more regimes. This can help to identify language patterns which are in fact used in most IL, and in some cases the question whether documents from one regime are in general more vaguely formulated than from another may yield important insights into the dynamics and politics of the respective regimes. As with all the comparison groups discussed here, there is no one right way to compare international agreements, but it is very important to include information on which groups or documents are being compared and why.

I use a second set of characteristics to compare different legal texts is the part of the agreement in which a sentence features. In this study, I broadly follow Aust (2000: 332) in differentiating preambles, main body of text, annexes, titles, and final clauses. I focus on the main text, but occasionally other parts of the treaty are discussed when they are especially relevant. The differentiation is important for several reasons: One, different parts of agreements deal with different content. That may make different parts of the text more or less relevant for the analysis in question. Two, some parts of agreements like headings and some annexes are not actually composed in sentences, which presents a challenge for the tool (further discussion in section 6.5). Three, different sections of text are interpreted as having differing legal status. Talking about preambles, for example, Aust (2000: 337) writes:

> *"The preamble is part of the context of the treaty for the purposes of interpretation […] The more one burdens the preamble with unnecessary, but not always insubstantial, material, the greater the danger that it will be relied on to support a particular interpretation of the main text".*

In describing preambles as context to an agreement, Aust makes the distinction that the preambular paragraphs themselves are not included in the text that would require interpretation in the first instance. Headings are seen as even less integral to the text:

> *"As with the preamble, there should be no discussion of headings until the end of the negotiations: otherwise much time may be wasted arguing over the wording of what is no more than a label" (Aust 2000: 340).*

Final clauses, on the other hand, can be more important than one may assume.

> *"Final clauses can be a trap for the unwary. Compared with the main body of a treaty, the final clauses may appear to anyone who is not a foreign ministry lawyer to be less important, and in a sense they are. But [….] final clauses do play an essential, if different, role from the main text" (Aust 2000: 345).*

Aust refers to clauses on consent to be bound and entry into force here, which are often at the end of the agreement, and there is not usually much variance between agreements in how these clauses are formulated, but their content can nonetheless be crucial for the agreement in question. Even though most annexes form an integral part of the treaty they are part of, "it is prudent to distinguish their provisions from those in the treaty itself" (Aust 2000: 355). Each of these parts of text follows different Conventions and is constituted by slightly different language. It is of course possible to compare an entire agreement to another one, but the comparison will necessarily be slightly less detailed if different parts of text are not taken into account.

Lastly, the legal status of the document is of crucial importance. However, even the number of possible categories is controversial. According to the most intuitive, and Conventional, distinction there are only two types of legal documents: legally binding treaties on the one hand and any other document drawn up between two or more states on the other (see e.g. Aust 2000: 18). Unfortunately, whether an agreement is legally binging or not is contingent on whether or not the parties to the agreement intended it to be binding (Aust 2000: 17). This can of course create confusion. More pragmatically, legally binding documents are often identified by their need of ratification. Three documents in my corpus fulfil this criterion: NPT (1968), the UNFCCC (1992) and the Kyoto Protocol (UNFCCC 1997). Every other document then simply has the status of being not legally binding.

To make matters even more complicated, legal bindingness need not be considered a binary. One can, for example, classify decisions made by all parties within the framework of a given regime as a separate legal category in addition to legally binding Conventions and non-binding declarations. While coming at the cost of more fuzziness, this classification has the advantage of allowing more room for the perceived importance and bindingness of the different treaties. Certainly, such decisions do not usually require ratification. They are however adopted with consensus by an official body including all member states. As long as a country wishes to stay inside a given regime, it must also respect the decisions taken by the Conference of the Parties (COP). The legal nature of COP decisions in the climate regime, for example, has been the matter of some debate among scholars of IL. After having devoted an entire paper on the validity of COP-made law, Camenzuli (2007:25) comes to the conclusion that

> *"While the legally binding nature and characterization of CoP acts and decisions as 'law' will and should continue to be the subject of further research and discussion, the importance of this CoP made law in determining the effectiveness of MEAs [Multilateral Environmental Agreements] means that efforts should also be directed at ensuring its validity. Whether law making power of CoPs fall within the law of treaties or international institutional law, in either case, provided the interpretation adopted by the CoP is uncontested and not modified by further practice, it is likely to be regarded as authoritative".*

Considering also that the mechanisms of IL (discussed in chapter 3.1.3) are often based more on norm-changing, mutual acceptance and public recognition, these factors seem sufficient to allow for an extra category in-between the binding treaties and the non-binding declarations. In the following chapter, I do adopt this three-part split.

The question remains which documents to sort in which category. As stated above, the NPT and the Framework Convention as well as the Kyoto Protocol are unambiguously legally binding and constitute the group I call 'Convention/Protocol'. At the other end of the spectrum, Ministerial Declarations under the UNFCCC are clearly designed to be declarations of intent, without any legal consequence. CoP Decisions under the UNFCCC regime belong in the third category of decisions. This leaves RevCon Final Declarations. On the one hand, the fact that they are named 'declaration' is a hint that they should be categorized with the UNFCCC Ministerial declarations. On the other hand, they are universal decisions by all parties made on the basis of the NPT, reviewing and developing it over time. In this, they serve a similar purpose to CoP decisions. For these reasons, I categorize them as decisions for the purposes of this study.

In general, it is often useful to vary as few components as possible when comparing agreements: for example, when wanting to compare documents across regimes, it may be best to stick with documents of similar legal status and compare the same part of each text.

6.2 Overall results

The results of the application of the tool to documents of the NPT and UNFCCC regimes can be seen in the following table:

Table 8: Average vagueness points by document: all text

Regime	Year	Document Name	Legal Status	Sentences	Total Points	Average Points
NPT	1968	NPT 1968	Convention /Protocol	60	26	0.43
NPT	1975	NPT 1975	Decision	120	176	1.47
NPT	1985	NPT 1985	Decision	222	307	1.38
NPT	1995	NPT 1995	Decision	82	142	1.73
NPT	2000	NPT 2000	Decision	351	488	1.39

NPT	2010	NPT 2010	Decision	364	501	1.38
UNFCCC	1992	UNFCCC 1992	Convention /Protocol	282	216	0.77
UNFCCC	1995	UNFCCC 1995	Decision	28	42	1.50
UNFCCC	1996	UNFCCC 1996	Declaration	30	41	1.37
UNFCCC	1997	UNFCCC 1997	Convention /Protocol	297	104	0.35
UNFCCC	1998	UNFCCC 1998	Decision	460	540	1.17
UNFCCC	2000	UNFCCC 2000	Decision	227	192	0.85
UNFCCC	2001	UNFCCC 2001	Declaration	12	13	1.08
UNFCCC	2002	UNFCCC 2002	Declaration	31	54	1.74
UNFCCC	2007	UNFCCC 2007	Decision	54	59	1.09
UNFCCC	2009	UNFCCC 2009	Declaration	52	48	0.92
UNFCCC	2010	UNFCCC 2010	Decision	388	497	1.28

The first important finding, though it may seem basic, is that there indeed exists a diversity of vagueness among the different documents. Means range from 0.36 points in the Kyoto Protocol to 1.83 points in the Delhi Ministerial Declaration. Out of a maximum possible range of 4 (-1 to 3), average vagueness scores cover a range of 1.47. Considering that an average score of 3 means that all the sentences of the agreement would need to score as vague in each category, this is quite a large variation. It indicates that even though the basic unit under scrutiny are sentences, the vagueness scores still show differences at the level of entire agreements.

On average, there does not seem to be a significant difference between the two regimes: The NPT averages 2.89 points, while the UNFCCC sentences get 2.67 on average. Counting only the main text bodies of agreements does not make a difference in the distribution, although it does elevate the vagueness scores slightly. The next two graphs are intended to show the similarity of the data when using the entire text of each agreement (Figure 14) as opposed to restricting analysis to the main body of text, excluding preambles, headings, annexes, and final clauses (Figure 15).

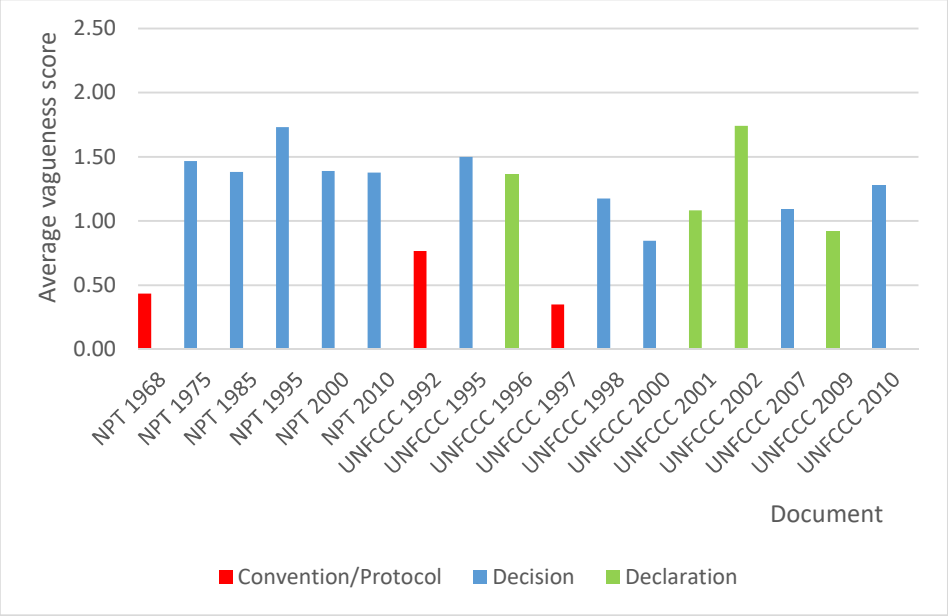


Figure 14: Average scores by documents, all text

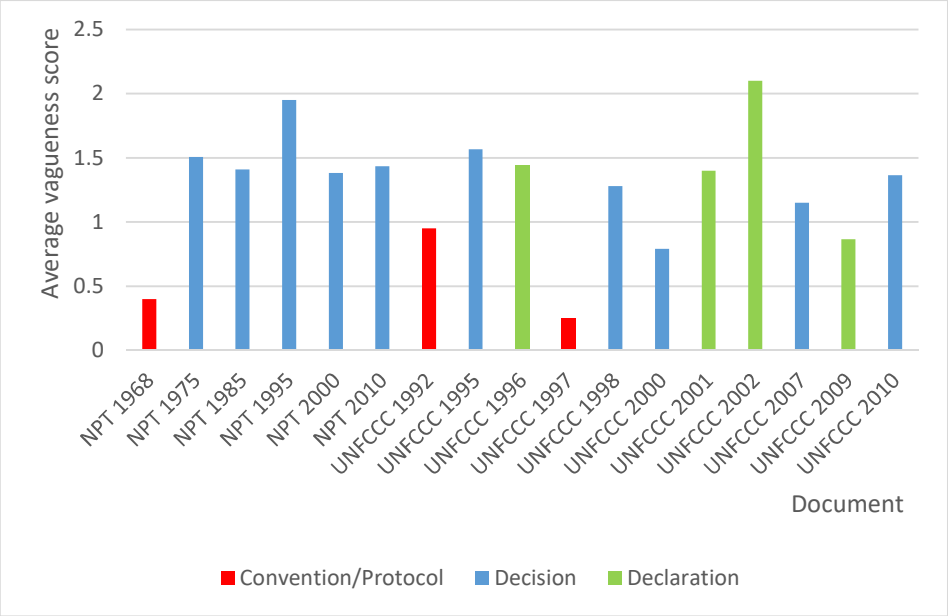


Figure 15: Average scores by documents, main body of text

The above graphs both show all documents, sorted by their respective regime and in order of their dates of adoption. Legally binding documents are coloured in red, Decisions made by a RevCon or CoP show up in blue, and Ministerial declarations are highlighted in green. What becomes apparent is that the three documents averaging the least points are the ones that are the most firmly counted as legally binding: The original NPT of 1968, the original UNFCCC (1992) and the Kyoto Protocol (1997).

The two documents scored by far as being least vague are the NPT and the Kyoto Protocol. A little vaguer, but still relatively precise, are the UNFCCC (1992), the Bonn agreement, and the Copenhagen Accord. On the vague end of the spectrum, the most vague document from

both regimes are, respectively, the NPT review conference (RevCon) closing document from 1995 and the Delhi declaration of 2002. While all of the red bars – representing the two Conventions and the Protocol – are clearly lower (meaning less vague) than all other documents, it cannot be said that the green bars are, on average, much higher than the others. When taking the averages of the three groups, this hunch becomes clear:

Table 9: Averages by document type, all text included

Document type	Average vagueness score
Protocol/Convention	0.52
Decision	1.32
Declaration	1.28

The difference between the first category and the next two is pronounced, while the latter categories are very close, the Declarations actually scoring slightly more precisely than the decisions. However, when looking only at the main text of the document – excluding preambles, annexes, headlines, and ending formula, this last fact changes. The table below shows the scores for the main text, aggregated by document group:

Table 10: Averages by document type, main body of text

Document type	Average vagueness score
Protocol/Convention	0.53
Decision	1.38
Declaration	1.45

While the numbers for all categories go up slightly, the average for declarations changes by 0.17, which is the most significant change and puts them above Decisions in terms of vagueness. This result is particularly interesting because in all other respects it does not seem to matter which parts of the text are included in the analysis. Most of the time, the differences are very slight and crucially they do not usually change the results of a comparison. Unless otherwise indicated, I will base the analysis that follows on the main body of text of documents for methodological reasons explained in section 6.1. As will be discussed in section 6.5, the group of comparison is crucial to obtain meaningful data. In this particular study, calculating the vagueness score of the entire documents as opposed to only using the main body of text does change the absolute results slightly, but does not change the relative results (i.e. which documents are vaguer than others), because all documents become slightly more precise. In order to provide a more solid comparison, I usually use only the main body of text in this analysis. In the rare cases where taking into account other parts of text proved necessary or useful, I have indicated the deviation.

All the scores shown above are averaged from all sentences in a given document being assigned a score from -1 to 3. Looking at the score distribution of all documents taken

together, the resulting graph approaches normal distribution, though skewed slightly to the right – vaguer – end of the spectrum:

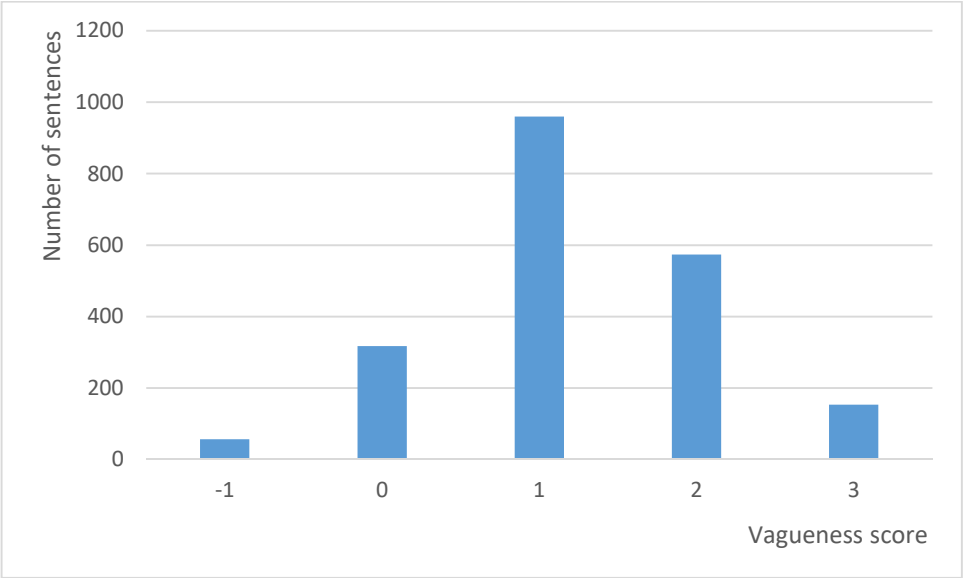


Figure 16: Overall score distribution

The majority of sentences score 1, meaning that most sentences contain some vague features and some that are precise. Just over half as many sentences score 2 points, meaning three out of the four measured criteria score as vague. The third largest category – again containing about half as many sentences as the previous one – is that of fairly, but not completely precise sentences, where usually one of the four categories remains vague, while three of them indicate precision. The two smallest categories are the most extreme ones. Again, the vague side contains about double the amount of sentences than the equivalent category on the left – precise side. If a sentence gets -1 points, all categories indicate a vague sentence, while if it receives 3 points, all categories indicate that the sentence is precise. It is important to note that the categories themselves are not scaled – there are only two values that can come out of one category. While the distribution definitely varies among the different documents, the basic pyramid shape mostly stays recognizable, if sometimes moved a step to the left or to the right.

Another way to analyse the data is to take a closer look at the categories themselves. As described in the previous chapter, there are four of them. The first one determines whether or not the sentence contains any hedge words. The second category is used to indicate if a sentence contains less than three determiners. The last two categories are concerned with verbs: they look at whether a weak or a strong performative verb is present, respectively. The category checking for strong performative verbs is the one responsible for the minus

points: here, the result is not either 1 (for vague) or 0 (for precise), but rather 0 as an indicator of vagueness when no such verb is present or -1 as an indicator for precision, when it is.

The following graph shows how many sentences overall score in each of the categories:

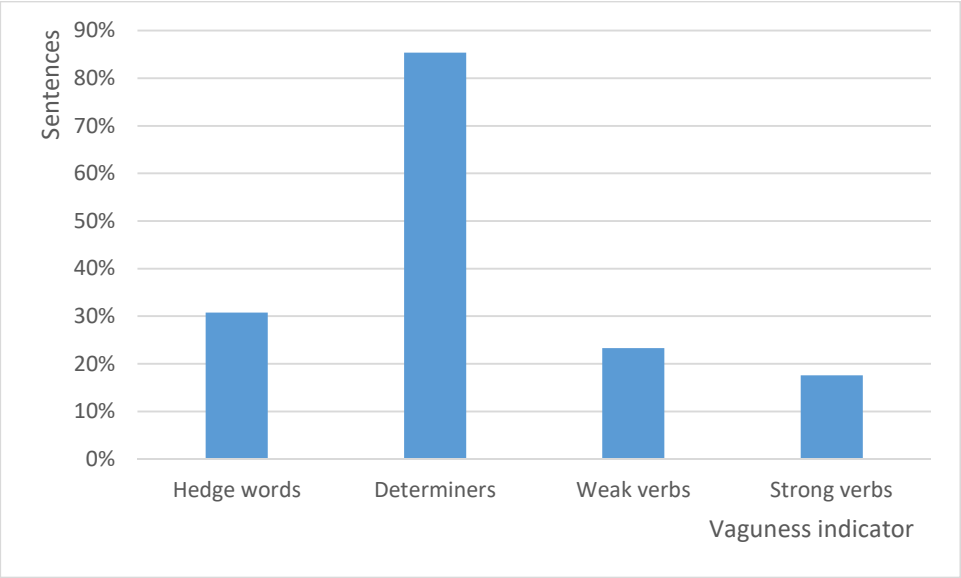


Figure 17: Percentage of sentences scoring in the respective category in all documents

As the graph shows, most of the sentences score in the category 'determiners' – 84.5% of them, in fact. This means that only around 15% of sentences contain three determiners or more, and thus come in at the precise end of this category. The next biggest category is hedge words. Just over 30% of sentences overall contain at least one hedge word. 23.3% of all analysed sentences feature weak performative verbs like 'should' or could, while strong performative verbs like 'will' and 'shall' are found least often, albeit still in 17.5% of the sentences. Of course, this data is aggregated, and the composition of scores varies widely when looking at the individual documents. The graphs below show overviews of these differences for each category.

The first graph shows the distribution of hedge words in each of the analysed agreements. Like in figure 16, the bars are ordered chronologically by regime, and Conventions and Protocols are marked in red, decisions in blue and declarations in red.

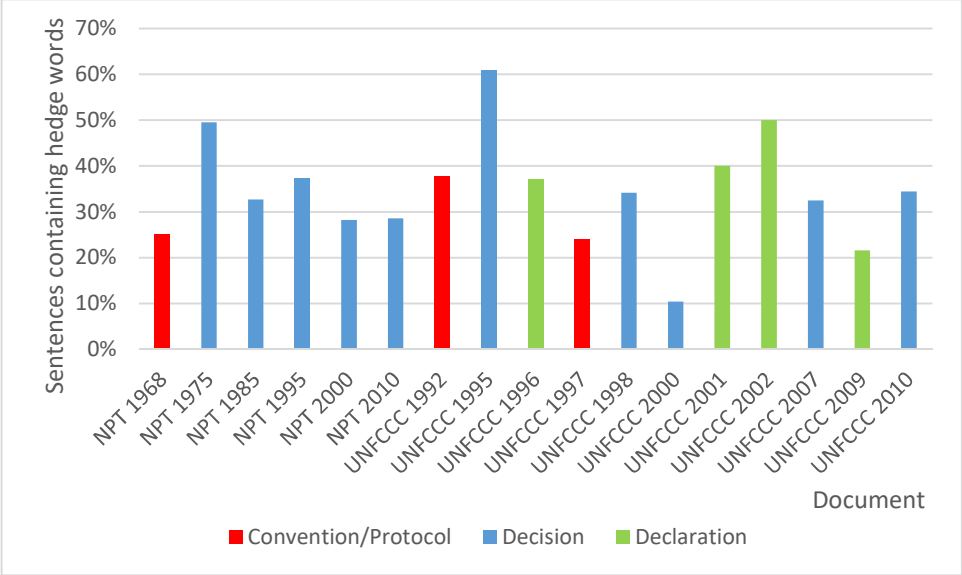


Figure 18: Percentage of sentences containing hedge words

In general, the percentage of sentences containing hedge words varies from 10.4% in the Bonn Agreements on the implementation of the Buenos Aires Plan of Action (or UNFCCC 2000) and 60.9% in the Final Declaration of the NPT RevCon of 1995. This is a variance of just over 50 percentage points, so it is a safe conclusion that not all international agreements use hedge words in equal measures. Interestingly, the legal form of the document does not seem to make a significant difference in this category at first glance. However, the average numbers by document type as shown in table 12 do show a difference, although it is not particularly large:

Table 11: Percentage of sentences containing hedge words by document type

Document type	Sentences containing hedge words
Protocol/Convention	29%
Decision	35%
Declaration	37%

The difference between the document type lowest in hedge words and the highest one stands at 8 percentage points.

In the category of determiners, the comparing graph looks differently.

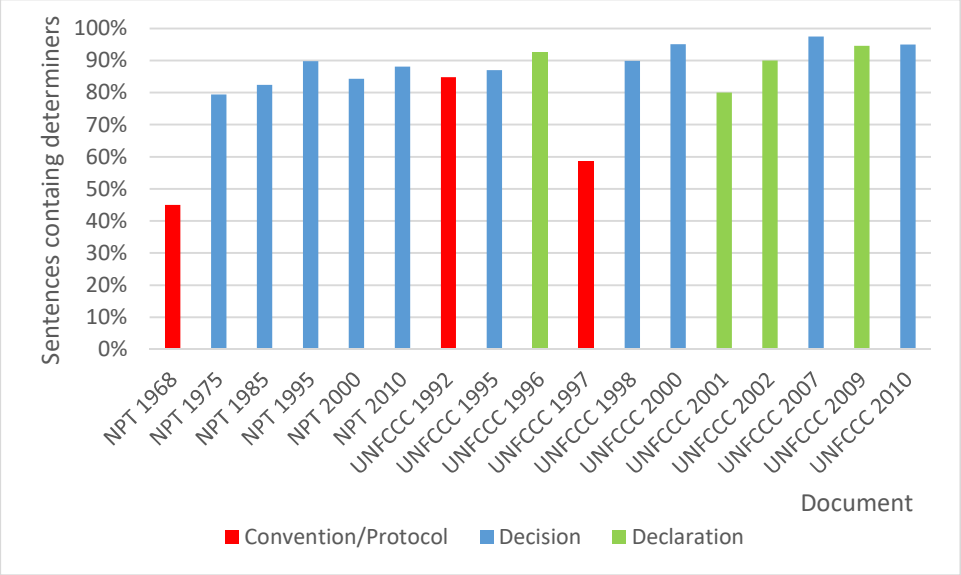


Figure 19: Percentage of sentences containing less than three determiners

The bars look more uniform in height, with the notable exceptions of the Nuclear Non-Proliferation Treaty and the Kyoto Protocol. Since both documents are legally binding, this result foreshadows the results of the grouped averages, where this type of document scores significantly lower. The NPT is the document that gets the least points for determiners meaning that it contains most of them, because sentences containing 3 or more determiners receive 0 points in this category, while those comprising two or less score 1. 45% of sentences in the NPT contain two or less determiners, while that number is 97.5% for the Delhi Ministerial Declaration (UNFCCC 2002). The variance is similar to the one for hedge words at just over 52 percentage points. Looking at the grouped averages, a far more distinct picture can be shown:

Table 12: Percentage of sentences containing less than three determiners by document type

Document type	Sentences containing determiners
Protocol/Convention	63%
Decision	89%
Declaration	89%

Decisions and declarations actually have the same percentage of sentences with two or less determiners, but the legally binding documents show a clear distinction from that number by 36 percentage points.

The next chart shows the prevalence of weak performative verbs in the documents. For the purposes of this study, weak performative verbs are ‘would’, ‘could’, ‘should’, ‘underline’, ‘underscore’, ‘undertake’, ‘recognize’, ‘have the right’, and ‘call’.

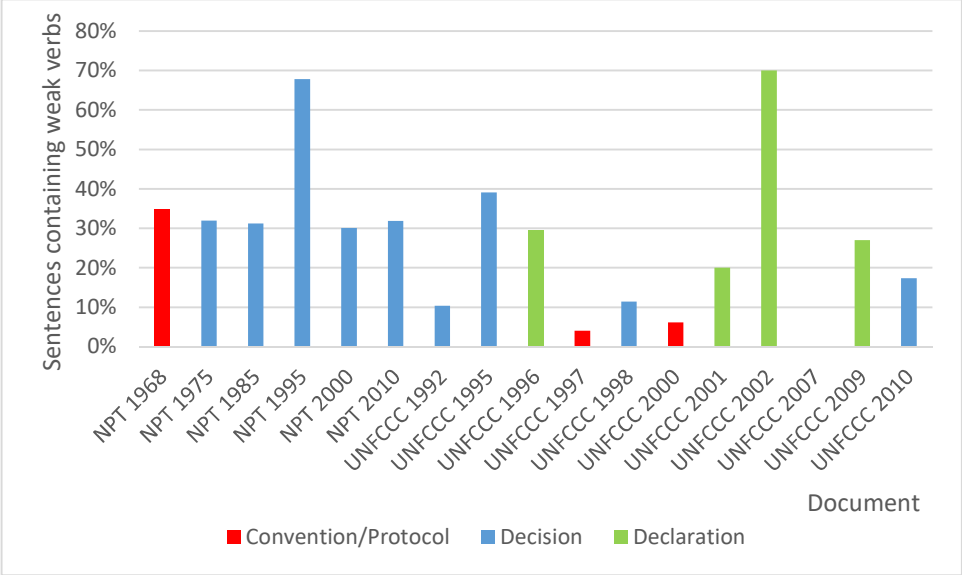


Figure 20: Percentage of sentences containing weak explicitly performative verbs

What is interesting here is that for the first time, there seems to be a notable difference among the regimes. While there is an outlier in the NPT regime – the Final Declaration of 1995 – most other documents score remarkably similarly, including the Non-Proliferation Treaty. In the UNFCCC, on the other hand, the percentage of sentences including weak performative verbs varies widely, and Ministerial Declarations appear to contain the vast majority of them. Both the UNFCCC (1992) and the Kyoto Protocol have very low percentages, though the Bali Action Plan (UNFCCC 2007) contains no weak performative verbs at all. The variation here is over 67 percentage points, which is extremely large.

When grouped by document type, the averages can be found in the following table:

Table 13: Percentage of sentences containing weak explicitly performative verbs by document type

Document type	Sentences containing weak verbs
Protocol/Convention	16%
Decision	27%
Declaration	37%

The differences here are noticeable and evenly spaced at 10-11 percentage points each. The scale follows the usual progression, with legally binding documents, NPT notwithstanding, featuring the least amount of weak explicitly performative verbs with 16% of sentences containing them and 37% of sentences in Declarations containing these verbs.

The last category is the opposite of the previous one, namely strong performative verbs. In the case of this study, this includes the verbs ‘shall’ and ‘will’.

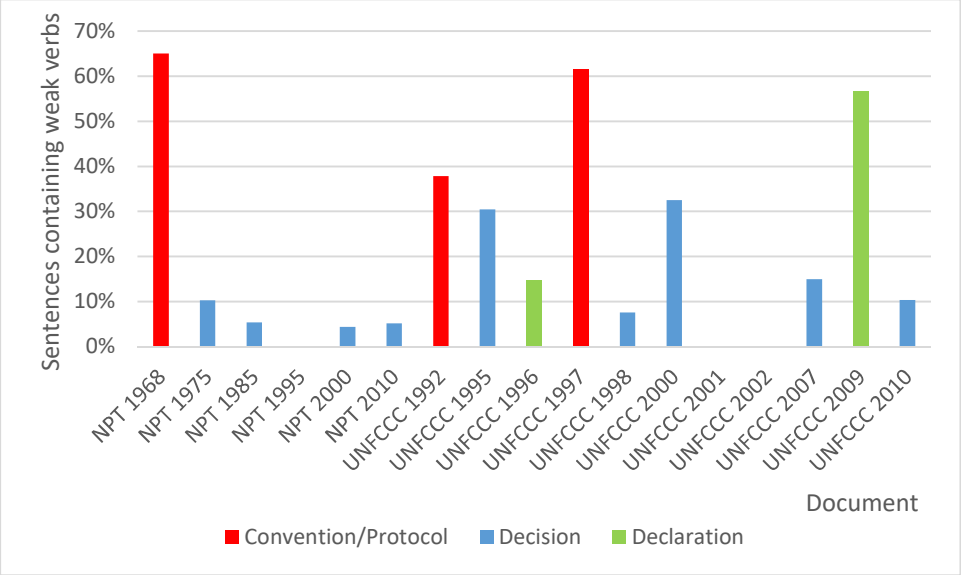


Figure 21: Percentage of sentences containing strong explicitly performative verbs

Here, the differences are even more distinct. Clearly the vast majority of times that 'will' or 'shall' is used in international agreements, it is in legally binding ones. Decisions sometimes feature a few instances, and Declarations do not contain any as often as they do – but there is the notable extreme outlier of the Copenhagen Accord, which will be discussed in detail later on. The variance comes in at 65 percentage points, which is comparable to the range of weak performative verbs, even though in total strong performative verbs are used much less often.

As expected, the differences among document types are stark: Legally binding documents feature strong performative verbs in over half of their sentences, while Decisions only employ them in 12% of theirs. Due to the extreme outlier status of the Copenhagen Accord, Declarations come in at 18%. In section 6.4, I will give some reasons for why I think this is an anomaly. The table below summarizes these results:

Table 14: Percentage of sentences containing strong explicitly performative verbs by document type

Document type	Sentences containing weak verbs
Protocol/Convention	55%
Decision	12%
Declaration	18%

In the following section, I will look at each individual document in more detail, and make an attempt to translate some of the quantitative data back to their qualitative meanings by giving examples of what sentences with specific vagueness scores or characteristics may look like.

6.3 Detailed result by agreement

The following two sections detail the application of the tool to individual agreements. In section 6.3.1, the nuclear non-proliferation treaty as well as five Final Documents of its RevCons are analysed. In section 6.3.2, 11 documents pertaining to the climate regime are studied, including the Convention, the Kyoto Protocol, four Ministerial Declarations and five Decisions. In each case, I show the detailed composition of the aggregate vagueness score that the specific treaty received and quote and discuss some especially interesting sentences.

The main goal of the chapter is to showcase how to apply the tool developed in this dissertation and what certain scores mean in practice, but in doing this I also provide some new information about the individual agreements. In some instances, limitations of the tool have been brought to light in the detailed analysis. I have made mention of it when this happened in this section, but the actual discussion of these issues takes place in section 6.5.

6.3.1 Documents under the NPT regime

The NPT regime has been called "one of the oldest, most ambitious, and most universal arms control treaties ever to enter into force" (Peloso 2011: 311). Its goal, as stated in its name, is to prevent the proliferation of nuclear weapons. In order to do that, the NPT strikes a "basic bargain", which

> *"allows countries to surrender their right to develop nuclear weapons in return for access to international assistance in civilian nuclear technology" (Way/Sasikumar 2004: 1).*

Every five years, the signatories meet at RevCons. RevCons are aimed at assessing the implementation status of each of the provisions in the NPT, providing recommendations to further the treaties' goals and, if possible, produce a Final Declaration on addressing these issues (Boon 2012: viii).

NPT 1968

The NPT, concluded in 1968, is the first of its kind and the starting point of the Non-Proliferation Regime. It can be divided into a preamble (ending with sentence 15), the operative clauses and ending formulas (starting at sentence 36). On the importance of distinguishing the parts of an agreement see section 6.1.

With a vagueness score of 0.43, the agreement is on the very precise side of the spectrum. Most parts of the text are fairly similar, the only exception being the preamble, which scores 1.36 vagueness points. The main body of text scores 0.4, and the concluding text scores -0.04 points. This is in line with preconceptions about these parts of text.

Looking only at the main body of text, the average score of 0.4 is combines four categories (hedge words, determiners, weak explicit performatives and strong explicit performatives), which can be seen in the graph below.

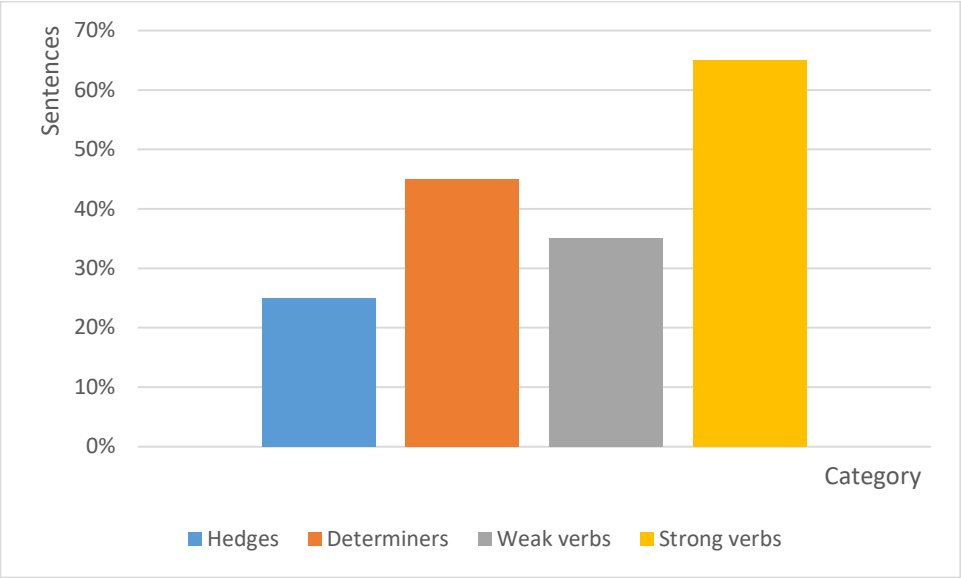
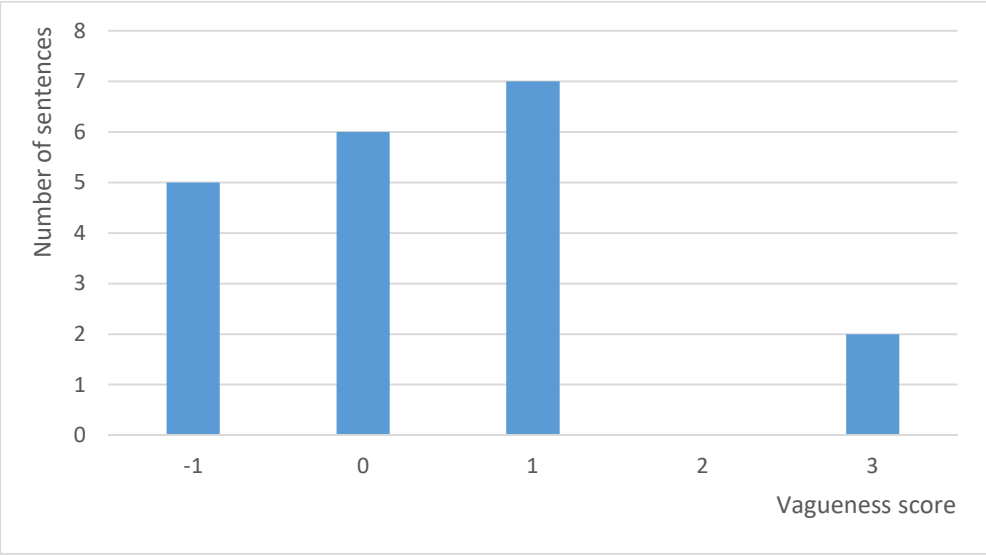


Figure 22: Percentage of vagueness indicators NPT 1968

The sub-score for hedge words is 0.25, which means that a quarter of the sentences in the main body of text (or more specifically, five out of 20) contain at least one hedge word[37]. Nine sentences, or 45% of sentences in the main body of text, contain three or more determiners. Seven sentences (or 35%) contain verbs of weak performative force, while 13 of them (65%) are based on verbs with strong performative force[38]. This last number is very high, and while it is not solely responsible for the low vagueness score, it is certainly a major factor. The distribution of sentences on scores looks as follows:



[37] See Section 5.3.1 for a list of the hedge words which were tested for.
[38] This last number is indicative of precision, not vagueness, so that the effect on the total is negative.

Figure 23: Score distribution NPT 1968

While most sentences score 1, meaning two indicators suggest precision and two suggest vagueness, it is clear that the distribution is heavily skewed towards the more precise end. In fact, only two sentences score more than 1 point, and both of them score 3, the highest number of points. They are Art. 4.2:

> *"All the Parties to the Treaty undertake to facilitate, and have the right to participate in, the fullest possible exchange of equipment, materials and scientific and technological information for the peaceful uses of nuclear energy" (NPT 1968: Art. IV.2)*

and Art. 6

> *"Each of the Parties to the Treaty undertakes to pursue negotiations in good faith on effective measures relating to cessation of the nuclear arms race at an early date and to nuclear disarmament, and on a treaty on general and complete disarmament under strict and effective international control" (NPT 1968: Art. VI).*

In the first case, 'the fullest possible' is a hedge, 'all' counts as a determiner, which means the sentence does score in this category, and 'undertake to facilitate' is a very weak verbal structure that indicates a low performative force, resulting in a weak obligation. In the second case, 'in good faith' and 'effective' count as hedge words or clauses, 'each' is the sole determiner, and the verb 'undertake' again indicates a low level of obligation.

The preamble also contains one sentence scoring 3. It is necessary to note, though, that the usual structuring of the preamble means that it consists mostly of half-sentences, held together by a clamp of subject ('the signatory states' etc.) in the beginning and a version of 'have agreed as follows' in the end. While they therefore cannot be analysed quite like complete sentences, the difference between them is still interesting. In this case, the sentence fragment in question reads

> *"Declaring their intention to achieve at the earliest possible date the cessation of the nuclear arms race and to undertake effective measures in the direction of nuclear disarmament" (NPT 1986: preamble).*

'Earliest possible date' is a hedge, undertake again takes the role of weakly performative verb and no determiners are present.

Five sentences in the main body of text of the agreement receive the lowest possible score of -1. I will not list all of them here, but simply give two examples. The first is a sentence under Art. 3.1:

> *"Procedures for the safeguards required by this Article shall be followed with respect to source or special fissionable material whether it is being produced, processed or used in any principal nuclear facility or is outside any such facility" (NPT 1968: Art. III.1).*

In this case, no hedge words are present. The operative verb is 'shall', which indicates a legal obligation and is a verb with strong performative force in this context. Determiners are

'any', and 'such', where 'any' is used twice and therefore reaching the threshold of 3. Similarly, this sentences contained in of Art. 3.4, is counted as precise:

> *"For States depositing their instruments of ratification or accession after the 180-day period, negotiation of such agreements shall commence not later than the date of such deposit" (NPT 1968: Art. III.4).*

The operative verb is again shall, there are no hedge words and 'their', 'such' (2) are determiners. Additionally, although this is not reflected in the scores, the sentence contains a reference to the clearly defined time period of 180 days. Furthermore, six sentences in the concluding text receive a score of -1. This is not surprising, since the closing paragraphs often contain clear-cut and uncontroversial arrangements concerning when to meet next or ratification procedures.

Content-wise, some of the most important parts of the agreement concern the actual provisions against nuclear proliferation, most notably articles 1 and 2 of the agreement. They are very similar in their sentence structure and read, respectively,

> *"Each nuclear-weapon State Party to the Treaty undertakes not to transfer to any recipient whatsoever nuclear weapons or other nuclear explosive devices or control over such weapons or explosive devices directly, or indirectly; and not in any way to assist, encourage, or induce any non-nuclear-weapon State to manufacture or otherwise acquire nuclear weapons or other nuclear explosive devices, or control over such weapons or explosive devices" (NPT 1986: Art.I).*

and

> *"Each non-nuclear-weapon State Party to the Treaty undertakes not to receive the transfer from any transferor whatsoever of nuclear weapons or other nuclear explosive devices or of control over such weapons or explosive devices directly, or indirectly; not to manufacture or otherwise acquire nuclear weapons or other nuclear explosive devices; and not to seek or receive any assistance in the manufacture of nuclear weapons or other nuclear explosive devices" (NPT 1986: Art.II).*

Both of these sentences receive a score of 1, bringing them to the middle of the vagueness scale. This is because even though they do not contain any hedge words and neither of them contains more than two determiners, the operative verb of the sentence is 'undertake' which indicates a relatively low level of obligation. Apart from this, though, the language that follows this operative word is unusually strong (which is not reflected in the results generated by the tool).

NPT 1975

Eight years after the conclusion of the original Nuclear Non-Proliferation Treaty and five years after its entry into force, the first RevCon ended with a Final Declaration. It is important

to note that the legal status between the original agreement and this document differ (see section 6.1). The Final Declaration goes through all the articles of the original agreement, noting the current status, progress and opinions on each. It has about double the length of the original agreement.

The total vagueness score of the entire document is 1.48, which is relatively vague but not yet among the vaguest of the tested documents. Considering only the main text does not change the relative position but heightens the score slightly to 1.52. In my analysis, the document consists of 120 sentences.[39]

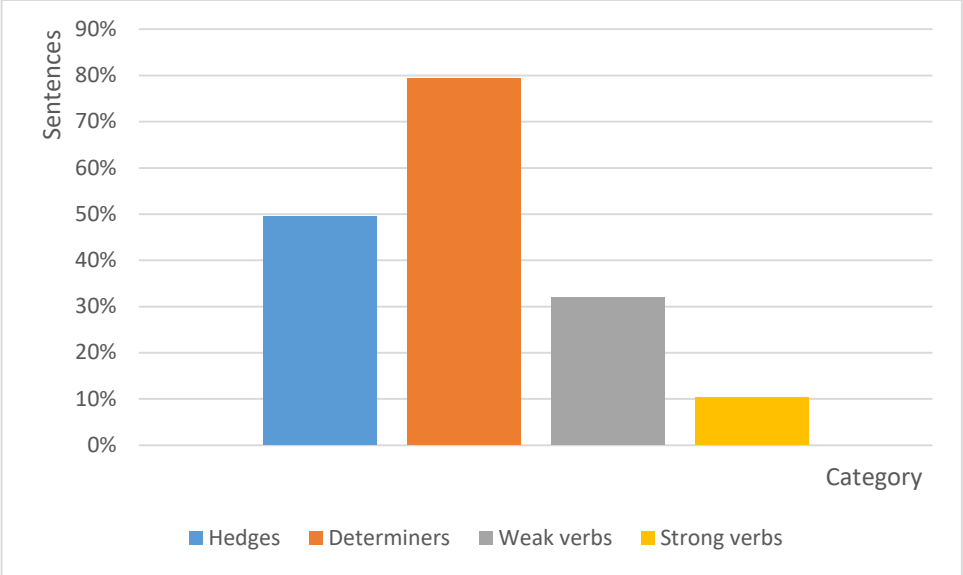


Figure 24: Percentage of vagueness indicators NPT 1975

As shown in graph 25 above, 50.5% of sentences in the 1975 NPT review document contain at least one hedge word. 79.4% score positively for containing less than three determiners. A third of the sentences contains a weak explicitly performative verb. In comparison to the previous document, it is notable that over half of the sentences contain neither strong nor weak explicitly performative verbs. These verbs are either neutral in their effect on the perceived vagueness, or have not been tested for in the original study and are therefore impossible to classify. Since the standards of which verbs to use when are more relaxed in less formal documents, it is natural that this would be the case for a larger amount of verbs in a review document than in the original agreement. Otherwise, scores in all sub-parameters are consistent in showing that the document is overall relatively vague. Breaking the scores down by the number of sentences that receive them creates the following graph:

[39] Each single provision was counted as a sentence if it concluded either with a dot or a line break.

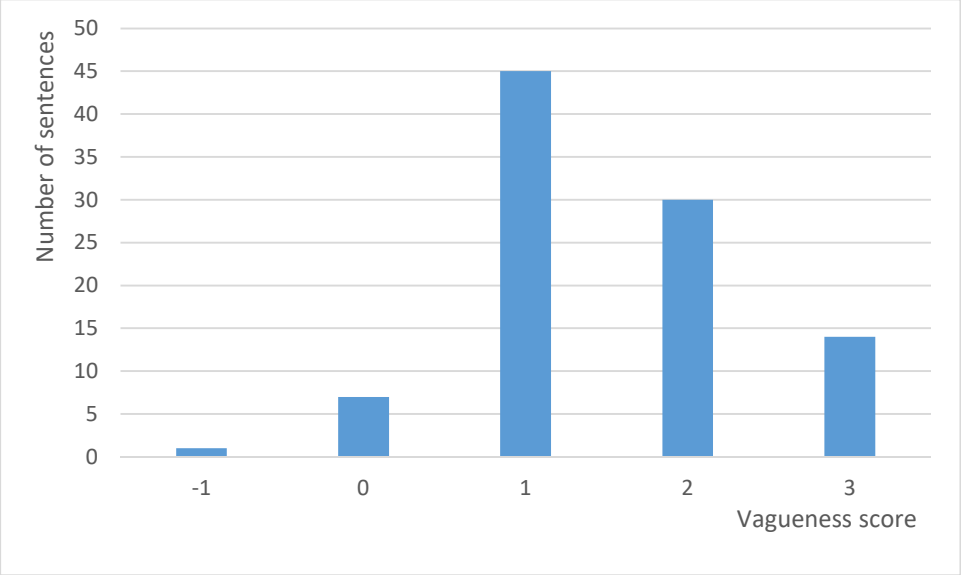


Figure 25: Score distribution NPT 1975

14 out of 97 sentences in the main text have a vagueness score of 3. One of them is exemplified here:

> *"Therefore, the Conference expresses the hope that States that have not already joined the Treaty should do so at the earliest possible date" (NPT 1975: 10).*

Even though it is relatively short, it manages to tick all the boxes for a vague sentence under this schematic: 'expressing hope' is a verb of very weak performative force, an 'earliest possible' date is a hedged expression and no determiners or strong performative verbs are present. Another one is as follows:

> *"The Conference underlines the importance of adherence to the Treaty by non-nuclear-weapon States as the best means of reassuring one another of their renunciation of nuclear weapons and as one of the effective means of strengthening their mutual security" (NPT 1975: 9).*

The verb in this sentence is 'underline', which has a weak performative force – in fact, it does not include any direct call for action at all. 'Their' is used twice, making it a total of two determiners in the sentence. Saying the adherence is the 'best' and 'one of the effective means' is a hedge, although less clearly so than in other cases.

Only one sentence in the main text (and in the whole document) has score of -1:

> *"The Conference takes note of the continued determination of the Depositary States to honour their statements, which were welcomed by the United Nations Security Council in resolution 255 (I968), that, to ensure the security of the non-nuclear-weapon States Party to the Treaty, they will provide or support immediate assistance, in accordance with the Charter, to any non-nuclear-weapon State Party to the Treaty which is a victim of an act or an object of a threat of aggression in which nuclear weapons are used" (NPT 1975: 9).*

Here, one problem with the tool becomes apparent: the word 'will', which counts as the verb with strong performative force in the sentence, is not the operative verb. That role falls to 'takes note', which was neutral as to its vagueness in the test cases and therefore does

not receive any positive or negative score. The implications of this issue for the tool itself will be discussed in section 6.5.

NPT 1985

While the NPT RevCon in 1980 did not produce a Final Declaration, the next one, five years later, did. Since it is a Final Document like its predecessor, the two documents are comparable in legal terms, as are the following three Final Declarations of NPT RevCons. At 222 sentences, the length of the document again almost doubles the previous one. Against the backdrop of heightened Cold-War tensions, it may be considered remarkable that the conference took place at all, much less produced a Final Declaration.

The 1985 Final Document receives a total vagueness score of 1.37. This is fairly similar to the previous document from 1975. While the number is still slightly above average for this study, it is the lowest score of all the Final Declarations of the NPT. Looking at the way this score is made up, it appears that 36.7% of the sentences contain hedge words, 82.4% receive a vagueness point because they contain less than three determiners, 31% contain weak explicitly performative verbs and only 6% include strong performative verbs:

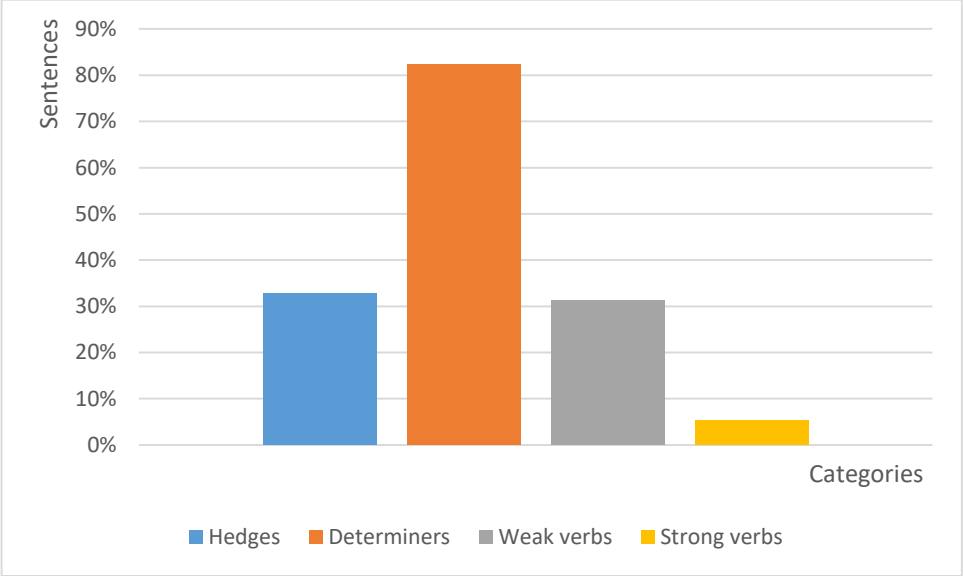


Figure 26: Percentage of vagueness indicators NPT 1985

This is a fairly typical distribution of the indicators. The graph below shows the distribution of vagueness points among sentences in the main body of text of the document:

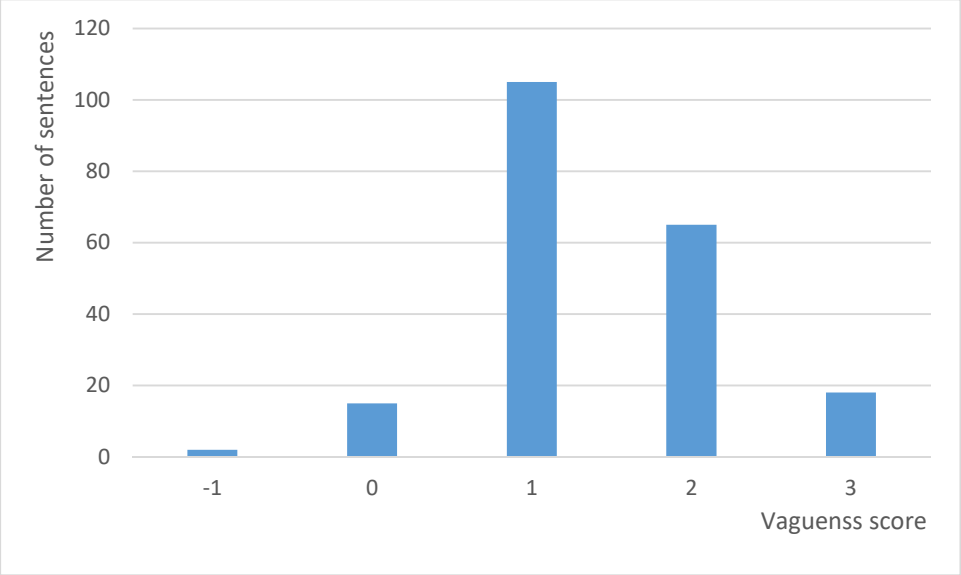


Figure 27: Score distribution NPT 1985

As should be expected, most sentences fall somewhere in the middle range of scores. 105 sentences score 1, that is to say that they include two features indicative of precise sentences and two which indicate vagueness. But while the middle categories contain by far the most sentences, the edge categories are interesting to analyse because they show the types of sentences that are counted as especially vague or precise by this method.

18 sentences in the main text score 3 vagueness points. I will take just two of them as an example here.

> *"The Conference underlines the need for IAEA to be provided with the necessary financial and human resources to ensure that the Agency is able to continue to meet effectively its safeguards responsibilities" (NPT 1985: 17).*

The first indicator is the verb 'underlines', which is hard to translate into any call or incitement for action. Effective – or in this case, the adverb effectively – is a hedge word because its presence makes it less, instead of more clear what financial resources are needed. The only determiner to be found in this sentence is again the word 'its', making it well under three determiners.

> *"The Conference further emphasizes the responsibilities of the Depositaries of NPT in their capacity as permanent members of the Security Council to endeavour, in consultation with the other members of the Security Council, to give full consideration to all appropriate measures to be undertaken by the Security Council to deal with the situation, including measures under Chapter VII of the United Nations Charter" (NPT 1985: 19).*

Only two sentences of the 205 in the main body of the text receive the lowest vagueness score of -1.

> *"The Conference takes note of the continued determination of the Depositary States to honour their statements, which were welcomed by the United Nations Security Council in resolution 255 (1968), that, to ensure the security of the non-nuclear-weapon States Parties to the Treaty, they will provide or support*

> *immediate assistance, in accordance with the Charter, to any non-nuclear-weapon State Party to the Treaty which is a victim of an act or an object of a threat of aggression in which nuclear weapons are used" (NPT 1985: 28).*
>
> *"The Conference commends the recent progress which the IAEA's Committee on Assurances of Supply (CAS) has made towards agreeing a set of principles related to this matter, and expresses the hope that the Committee will complete this work soon" (NPT 1985: 18).*

No hedge words are to be found in either case. Neither sentence contains a weak verb: 'take not of', which can be seen in the first sentence, was found to be neutral with regards to vagueness in the previous analysis, while 'commends' (as used in the second sentence) was not part of the original study and therefore cannot be assessed here. There are quite a few determiners: 'their', 'which', and 'any' in the first sentence and which, as well as twice 'this', in the second. As already discussed in the section above, in both of these cases, the strong verb 'will' is not actually the operative verb of the sentences. In contrast, here is a sentence where it is the operative verb:

> *"Such an extending of safeguards will enable the further development and application of an effective régime in both nuclear-weapon States and non-nuclear-weapon States" (NPT 1985: 16).*

However, sentences such as the above are extremely rare in the NPT regime. In fact, the likely conclusion is that in the NPT regime, words with a strong performative force or a specific legal meaning like will and shall are simply not used in Final Declarations. This problem will be discussed further in section 6.5.

Since there are so few sentences with strong performative verbs, I will also give an example of a sentence scoring zero, which is to say it gets no point in any of the categories.

> *"The Conference also invites the nuclear-weapon States, and all other States to render their assistance in the establishment of the zone and at the same time to refrain from any action that runs counter to the letter and spirit of United Nations General Assembly resolution 39/5413" (NPT 1985: 28).*

As can be seen, there is neither a strong nor a weak verb, nor any hedge word contained in this sentence. Determiners are 'their', 'any', and 'same'. Apart from the challenge that 'will' is only under certain circumstances indicative of precision, this sentence also shows another challenge for the method: references to other agreements, which are not themselves included in the tool, like the General Assembly resolution referred to above. This issue will be discussed further in section 6.5.

NPT 1995

Another ten years later, the 5th RevCon produced a Final Declaration. In the ten years between the two, the global climate shifted significantly. The fall of the Berlin wall and the subsequent restructuring of global relations also had significant effects on the non-

proliferation regime (see e.g. Simpson 1994). Additionally, 1995 was the year in which the Non-Proliferation Treaty was extended indefinitely.

At 82 sentences (2627 words) it is significantly shorter than the previous two Final Declarations, and just a little longer than the original agreement. This presents a break in the otherwise consistent production of ever longer documents over time:

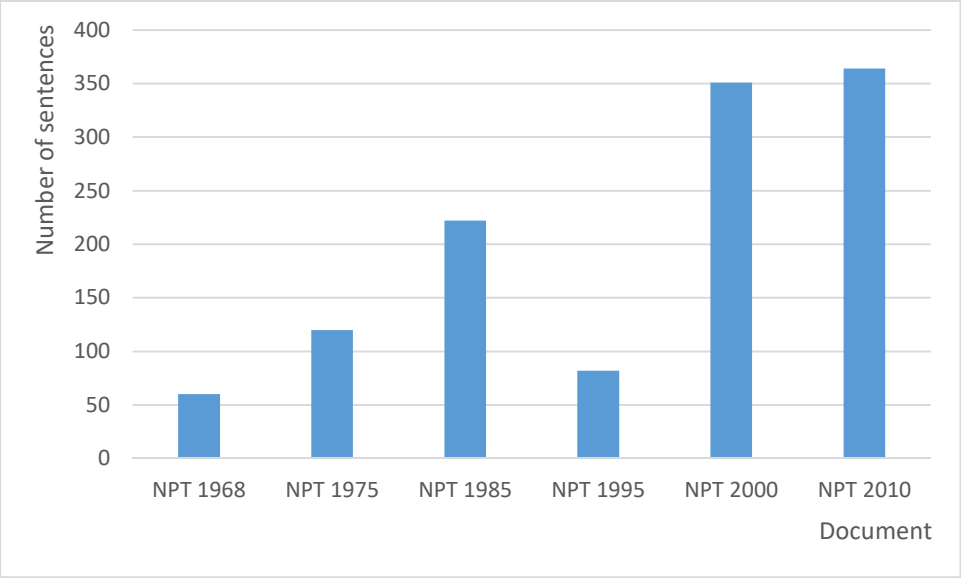


Figure 28: Number of sentences in different documents of the NPT Regime, all text included

At 1.79 vagueness points, this agreement is the second vaguest in this analysis and the vaguest Final Declaration produced by the NPT regime.

No sentence at all in the document receives a vagueness score of -1. This is due in part to the fact that there are no strong performative verbs in any of the sentences: nowhere have the drafters used either the word shall or will. This is consistent with the overall pattern of Final Declarations under the NPT regime. On the other hand, 67.8% of sentences in the main text of the document contain verbs with weak performative force, as shown in figure 30. That is by far the highest percentage in all analysed agreements (the second highest proportion can be found in the UNFCCC of 1995 and stands at just over 39%), and may be indicative of a wish to particularly weaken the legal force of the document. In this regard, the document is an outlier.

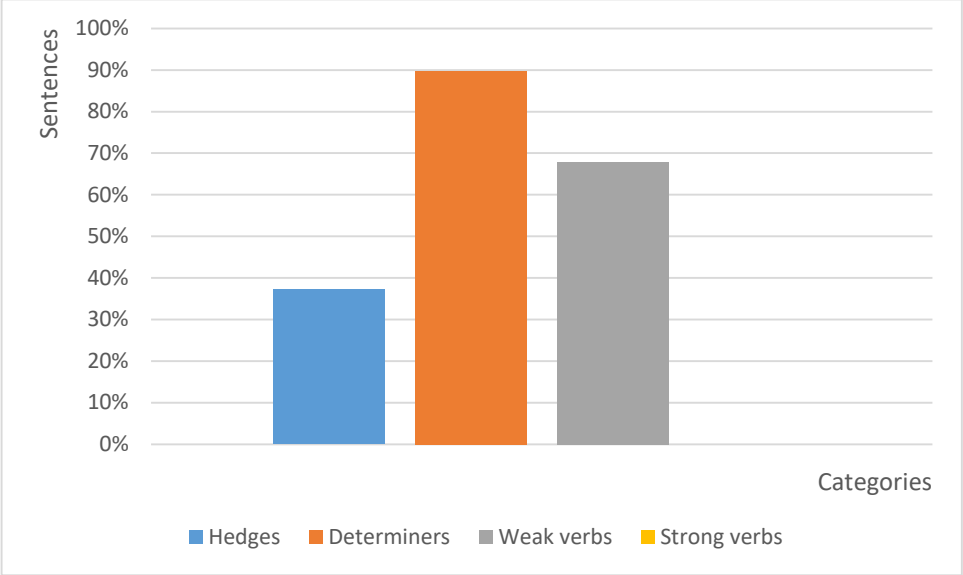


Figure 29: Percentage of vagueness indicators in the NPT 1995

The fraction of sentences containing hedge words or more than two determiners, while also on the higher end, are much more within the usual range of the documents studied here. The distribution of vagueness points is shown in the graph below:

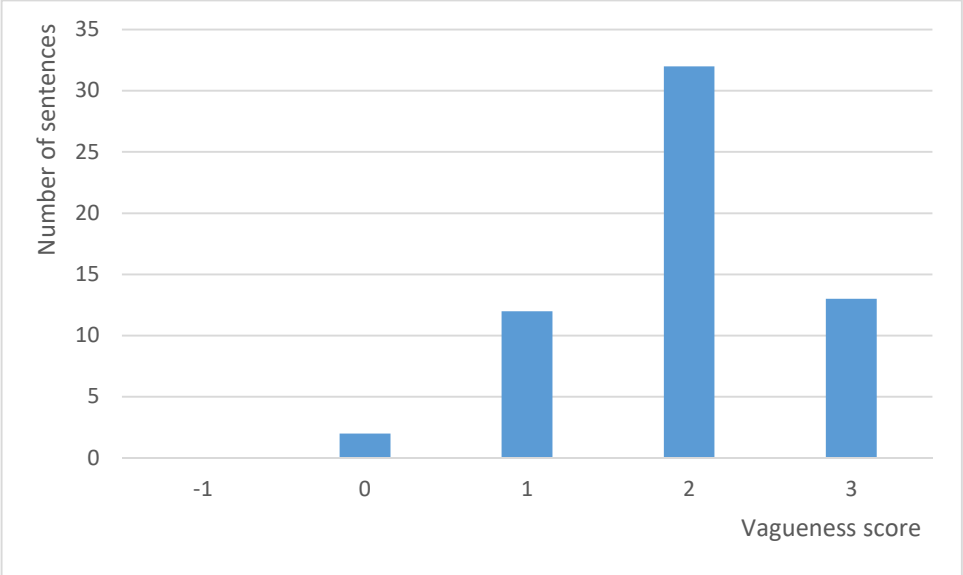


Figure 30: Score distribution NPT 1995

Compared to the previous graph, the whole distribution is skewed towards the left: not only are -1 and 0 much less prominent scores, but more sentences score 2 or even 3 as compared to 1 point. In the main text alone, 13 sentences score 3 points, the highest number on the scale. Two of them are shown as examples below:

> *"It was also agreed that subsidiary bodies could be established within the respective Main Committees for specific issues relevant to the Treaty, so as to provide for a focused consideration of such issues" (NPT 1995: 8).*
>
> *"Attacks or threats of attack on nuclear facilities devoted to peaceful purposes jeopardize nuclear safety and raise serious concerns regarding the application of*

> *international law on the use of force in such cases, which could warrant appropriate action in accordance with the provisions of the Charter of the United Nations" (NPT 1995: 12).*

The first sentence contains not one but two hedge words: both specific and relevant serve to obscure rather than clarify what issues are to be considered. In the second sentence, that role is filled by the word 'appropriate'. 'Could', used in both sentences is a verb of weak performative force. 'Such' is the only determiner to be found in either sentence.

Since there are no sentences containing verbs with a strong performative force, I will look instead at one of the two sentences that scored 0 points – that is, they do not score in any of the categories which indicate vagueness.

> *"The Conference requests that the President of the Conference bring the present decision, the decision on strengthening the review process for the Treaty and the decision on the extension of the Treaty on the Non-Proliferation of Nuclear Weapons, to the attention of the heads of State or Government of all States and seek their full cooperation on these documents and in the furtherance of the goals of the Treaty" (NPT 1995: 12).*

The operative verb here is 'requests'. It has not been part of the original analysis and is therefore not scored here, but it could be argued that it is likely to be one of the stronger performative verb choices when it comes to compelling action. In any case, the absence of any hedge words and the presence of the determiners 'all', 'their', and 'these' render the sentence relatively precise.

NPT 2000

Five years after the RevCon of 1995, the next RevCon also produced a Final Declaration. The parties agreed on 13 practical steps to further nuclear disarmament. The agreement's overall vagueness score of 1.41 is comparable to the other Final Declarations except for its immediate predecessor (discussed above). It is a little on the higher side of the scale, but not extremely so. Contrary to any other Final Declaration, the document contains concluding formula. At 351 sentences (12220 words) it is a long document, in the NPT regime second only to the Final Declaration of 2010.

The distribution of points over the four different categories seems pretty standard for NPT documents.

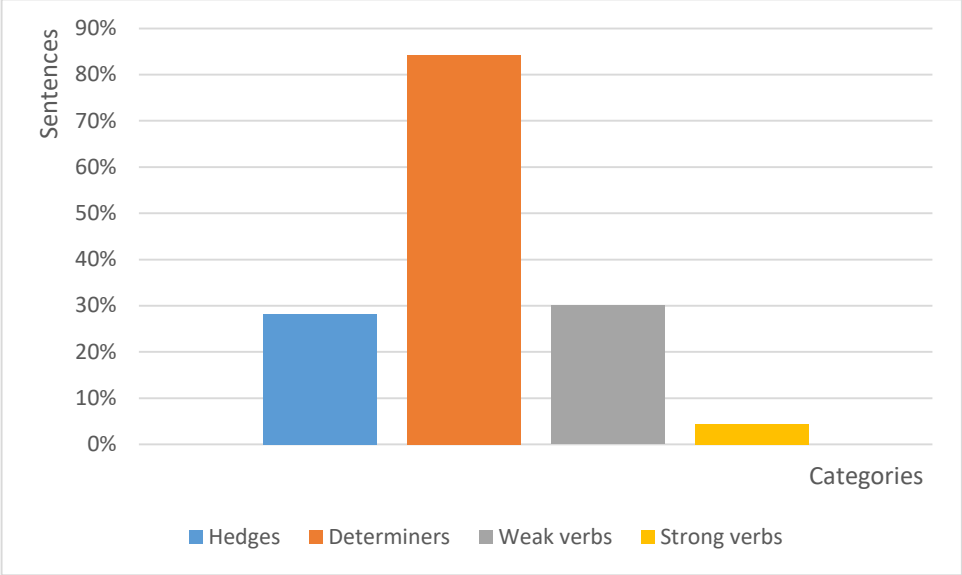


Figure 31: Percentage of vagueness indicators NPT 2000

30.1% of sentences in the main body of text score as vague in the hedge word category, 84.3% in the determiners category, again 30.1% or 96 sentences contain verbs with weak performative force and 5% those with strong performative force. The distribution of sentences over the vagueness points looks like this:

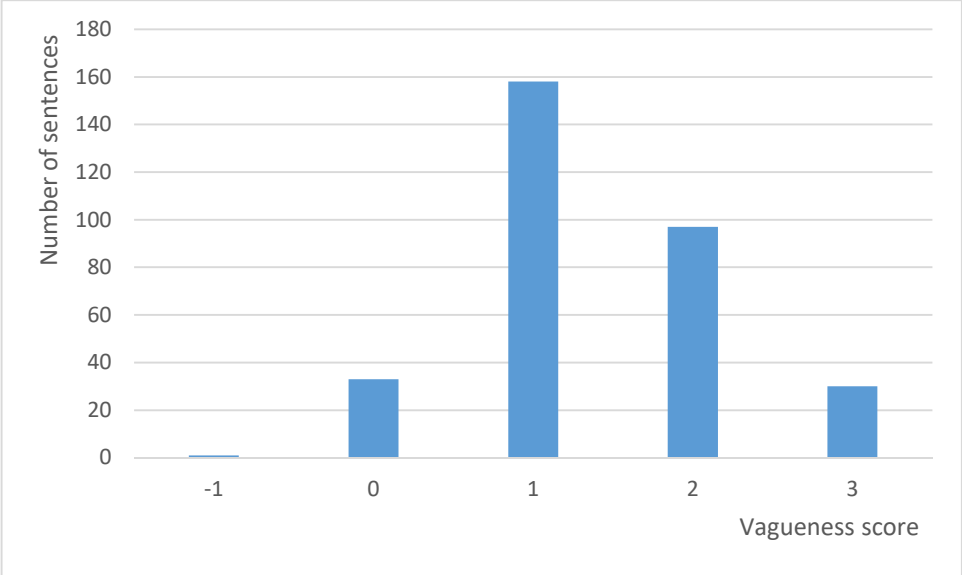


Figure 32: Score distribution NPT 2000

Once again, the bulk of sentences scores 1, as would be expected. For this document, the graph shows that more sentences are vague than precise.

As can be seen, only one sentence in the main part of the document scores -1 points:

> *"The Conference reaffirms that the cessation of all nuclear-weapon-test explosions or any other nuclear explosions will contribute to the non-proliferation of nuclear weapons in all its aspects, to the process of nuclear disarmament leading to the complete elimination of nuclear weapons and, therefore, to the further enhancement of international peace and security" (NPT 2000: 13).*

Here, while the operative word is 'reaffirms', the fact that the next verb is 'will' is important, and I would argue it is a valid indicator for precision, because it mirrors the formulation of the original sentence. As a further example, the following sentence receives a score of 0:

> *"The Conference remains convinced that universal adherence to the Treaty and full compliance of all parties with its provisions are the best way to prevent the spread of nuclear weapons and other nuclear explosive devices" (NPT 2000:2).*

33 sentences – the highest total number of sentences in any of the document (even though 3 documents are longer than this one) score 3 vagueness points. Two of them are given as examples below:

> *"The Conference recognizes that nuclear material supplied to the nuclear-weapon States for peaceful purposes should not be diverted for the production of nuclear weapons or other nuclear explosive devices, and should be, as appropriate, subject to IAEA safeguards agreements" (NPT 2000: 5).*
>
> *"Nuclear materials designated by each of the nuclear-weapon States as no longer required for military purposes should as soon as practicable be placed under IAEA or other relevant verification" (NPT 2000: 6).*

The hedge words used in these sentences are 'appropriate' and 'as soon as practicable' as well as 'relevant', respectively. Both sentences contain the weak explicit performative 'should'. There are no determiners in the first sentence, and the second sentence only contains the one determiner 'each'.

NPT 2010

The latest RevCon Final Document is, at 12958 words, also the longest produced so far. The action plan contained in the document sketches out steps toward nuclear disarmament as well as non-proliferation and the peaceful use of nuclear energy.

This document scores 1.47 on the vagueness scale, which is on the higher sider overall, but similar to most other NPT Final Declarations. The way the points are constituted is also fairly average for this type of document, as the graph below shows.

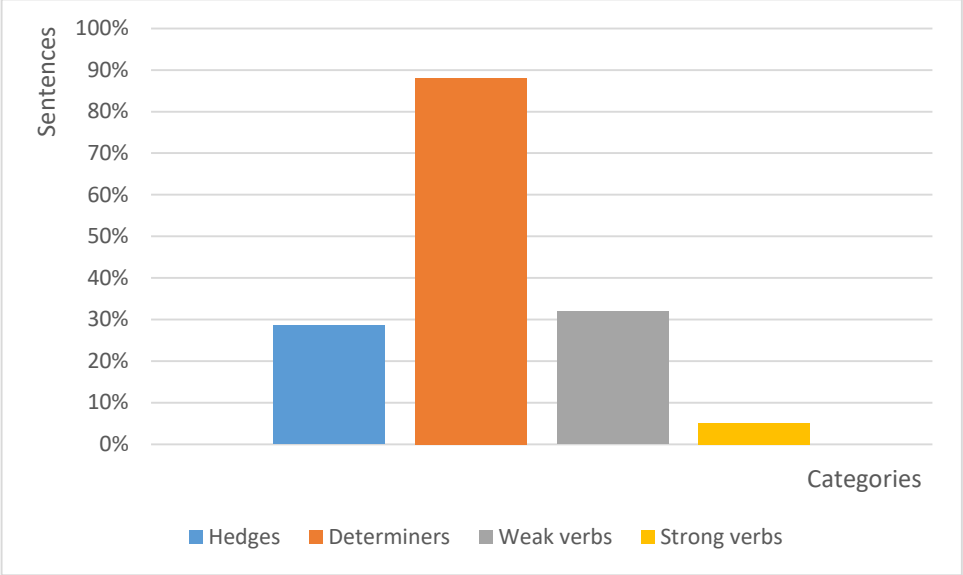


Figure 33: Percentage of vagueness indicators NPT 2010

32.2% of sentences include one or more hedge words, 88.1% of sentences contain two or less determiners, 32.2% contain words with weak performative force and 5.8% – or 19 sentences – contain verbs with strong performative force. Below is a graph showing the distribution of scores among the main sentences of the main body of text.

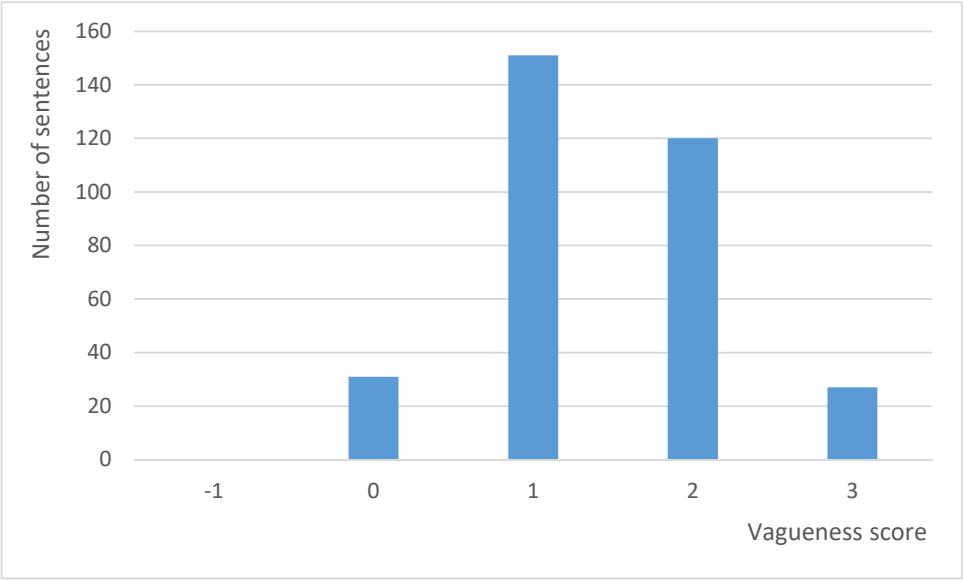


Figure 34: Score distribution NPT 2010

It reinforces the perception that the distribution of points in this document is fairly typical in its genre.

31 sentences (or 8.9%) receive the maximum number of vagueness points. Some examples are:

> *"The Conference recognizes that the development of an appropriate infrastructure to support the safe, secure and efficient use of nuclear power, in line with relevant IAEA standards and guidelines, is an issue of central*

> *importance, especially for countries that are planning for the introduction of nuclear power" (NPT 2010: 9).*

or

> *"The Conference recognizes the importance of having in place effective and coherent nuclear liability mechanisms at the national and global levels to provide compensation, if necessary, for damage inter alia to people, property and the environment due to a nuclear accident or incident, taking fully into account legal and technical considerations, and believing that the principle of strict liability should apply in the event of a nuclear accident or incident, including during the transport of radioactive material" (NPT 2010: 12).*

The operative verbs here is 'recognizes' in both cases. In the first case, there are two hedge words (appropriate and relevant). The second sentence includes the hedge word 'effective'. Neither of the sentences gets above the threshold of two determiners, even though they are both fairly long.

There is no sentence in the text that receives the minimum number of vagueness points, so instead I will look at three different sentences from the middle category to see what might be characteristic of these.

> *"In this regard, the Conference recalls that special efforts should be made and sustained to increase awareness in these fields, through participation of States parties, particularly those from developing countries, in training, workshops, seminars and capacity-building in a non-discriminatory manner" (NPT 2010: 9).*

This sentence scores a point in the category of weak performative verb, which is 'should'. It doesn't receive any points in the other three categories. There are neither hedge words nor a strong explicitly performative verb (which would result in a subtraction of a point), and there are three determiners: 'this', 'these', and 'those'. The second sentence scoring 1 point is

> *"The Conference encourages the efforts of IAEA, as well as of other relevant forums, in the promotion of safety in all its aspects, and encourages all States parties to take the appropriate national, regional and international steps to enhance and foster a safety culture" (NPT 2010: 10).*

This is an example of a sentence that includes a hedge word – relevant – and no strong performative verb, but also no weak performative verb and three determiners.

> *"The Conference supports national, bilateral and international efforts to train the skilled workforce necessary for developing peaceful uses of nuclear energy" (NPT 2010: 8) .*

Here, the only point is given because the sentence has less than three determiners. This is the most likely result for a sentence that receives one point, as could also be inferred from the fact that usually over 80% of sentences score in this category.

6.3.2 Documents under the UNFCCC regime

In 1992, the UNFCCC was adopted at the United Nations Conference on Environment and Development in Rio de Janeiro and now comprises 194 countries. Between 1995 and today, the COP has usually met once a year, with its subsidiary bodies additionally meeting once or twice in-between each conference. Included in this study are the UNFCCC (1992) itself and those documents produced by COP which have proper names – usually comprised of some indication of their legal status and the city in which they were concluded.

UNFCCC 1992

The UNFCCC officially started an international climate regime separate from other environmental concerns. The Convention entered into force in 1994. The objective of the agreement is to

> *"stabilize greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system" (UNFCCC 1992: Art.2).*

The UNFCCC scores 0.79 on the vagueness scale, making it the third most precise document after the original NPT and the Kyoto Protocol. The graph below shows the percentage of sentences scoring in each category:

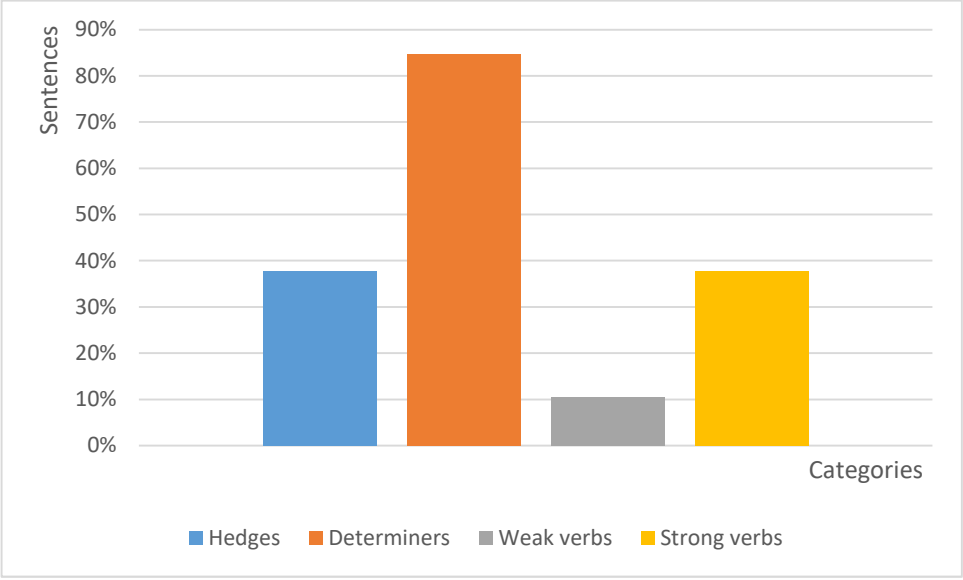


Figure 35: Percentage of vagueness indicators UNFCCC 1992

41.5% of sentences in the main body of text of this document contain hedge words. This is quite a lot of sentences compared to some other documents which score higher points overall. 84.8% of sentences contain two or less determiners. The big difference compared to other documents lies in the use of verbs. Only 10.4% of sentences contain weak explicitly performative verbs, while 37.9 of them contain strong explicitly performative verbs. When

putting the UNFCCC next to the other two documents of comparable status, however, this is also the point where the difference within this group shows most prominently: in both of the other two documents, over 60% of sentences contain strong explicitly performative verbs.

The number of sentences with the highest and lowest vagueness score is almost the same in this document: in the main text, 5 sentences score -1, while 7 sentences score 3. Out of a total of 164 sentences in the main body of text of the Convention, 70 score the exact median value of 1. The following graph provides an overview of the overall score distribution.

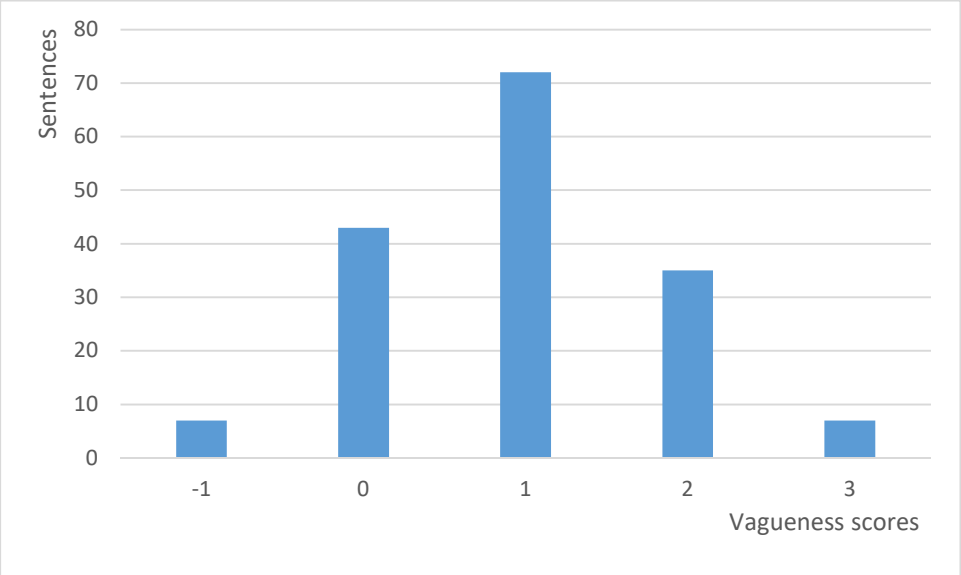


Figure 36: Score distribution UNFCCC 1992

As opposed to other documents, the distribution is almost perfectly symmetrical, with only a slight emphasis on sentences scoring 0 over those scoring 2 points.

One of the sentences receiving the highest vagueness score is the paragraph stating the objective of the Convention (and in effect the entire regime it establishes). For this reason, this article, already mentioned above, is often quoted in other documents of the regime.

> *"The ultimate objective of this Convention and any related legal instruments that the Conference of the Parties may adopt is to achieve, in accordance with the relevant provisions of the Convention, stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system" (UNFCCC 1992: Art.2).*

'Would' is counted as a verb of weak performative force. 'Relevant' clearly hedges the answer to the question what should actually be done to stabilize greenhouse gas concentration. The only determiner to be found in this sentence is 'any'.

In the next sentence, the first part in square brackets is an introductory sentence that gets counted individually. It is included to provide the subject.

> *"[All Parties, taking into account their common but differentiated responsibilities and their specific national and regional development priorities, objectives and circumstances, shall:] Take climate change considerations into account, to the*

> *extent feasible, in their relevant social, economic and environmental policies and actions, and employ appropriate methods, for example impact assessments, formulated and determined nationally, with a view to minimizing adverse effects on the economy, on public health and on the quality of the environment, of projects or measures undertaken by them to mitigate or adapt to climate change" (UNFCCC 1992: Art. 4.1(f)).*

Several hedge words can be found here. In addition to 'relevant' and 'appropriate', which are both counted at the moment, 'feasible' is another word that might be included in this category. Even though the sentence is not short, it only contains two determiners: 'their' and 'them'. The score in the category of weak performative verb is 'would'.

In contrast, here is a sentence that receives the lowest possible vagueness score:

> *"Each developed country Party and each other Party included in Annex I shall make its initial communication within six months of the entry into force of the Convention for that Party" (UNFCCC 1992: Art 12.5).*

It is not likely to be chance that this sentence is focussed on a much smaller and less contentious aspect of the regime. It provides a specific time frame in which to pass information to the regime. The operative verb is 'shall', indicating in this context of an international legal Convention a definite legal obligation. There are no hedge words to be found, and even though the sentence is short, there are four determiners: 'each' (twice), and 'its'. The contrast to the two sentences above should be clear.

Since this document is a Convention setting up a regime, the more administrative clauses at the end of the document are more relevant than usual. Below are two examples of the typically precise sentences:

> *"For each State or regional economic integration organization that ratifies, accepts or approves the Convention or accedes thereto after the deposit of the fiftieth instrument of ratification, acceptance, approval or accession, the Convention shall enter into force on the ninetieth day after the date of deposit by such State or regional economic integration organization of its instrument of ratification, acceptance, approval or accession" (UNFCCC 1992: Art. 23.2).*
>
> *"Any Party that withdraws from the Convention shall be considered as also having withdrawn from any protocol to which it is a Party" (UNFCCC 1992: Art 25.3).*

In both cases, the operative verb again is 'shall'. Determiners are 'each', 'such', and 'its' in the first case and 'any' (twice) as well as 'which' in the second case. Neither sentence contains any verb weakening their performative strength.

There are two core provisions of the Convention which deserve special attention because of their content. One paragraph that has been seen as the heart of the Convention is Article 4.2 (a) on the duties of Annex 1 parties.

> *"Each of these Parties shall adopt national policies and take corresponding measures on the mitigation of climate change, by limiting its anthropogenic emissions of greenhouse gases and protecting and enhancing its greenhouse gas sinks and reservoirs".*

This sentence scores 0 points on the vagueness scale. It receives a point for containing only one determiner ('its'), but receives -1 point because it contains the verb 'shall'. It is a relatively precisely formulated obligation to the parties listed in Annex 1 of the Convention.

Having analysed this sentence, it may be of interest to also take a look at the paragraphs which describe the responsibilities of all parties to the Convention. Those are outlined in Art. 4.1. Most of the sentences there receive 2 points, except for 4.1 (a), which concerns the establishment of national inventories and receives one point, and Art. 4.1(f):

> *"[All parties […] shall:] Take climate change considerations into account, to the extent feasible, in their relevant social, economic and environmental policies and actions, and employ appropriate methods, for example impact assessments, formulated and determined nationally, with a view to minimizing adverse effects on the economy, on public health and on the quality of the environment, of projects or measures undertaken by them to mitigate or adapt to climate change".*

Both 'take into account' and 'undertake' are weak explicitly performative verbs (although only 'undertake is counted with this method), 'relevant' is a hedge, as is 'appropriate'. The 'shall' in brackets is not counted for this sentence, but counts towards the introduction of the Article as a separate sentence. The only determiner is 'their'. These factors coming together mean that this particular part of Article four receives a vagueness score of 3.

UNFCCC 1995

The first CoP took place in 1995 – the same year as the 5th NPT RevCon – in Berlin and established the UNFCCC Secretariat, which has since set up its business in Bonn. It also agreed on the Berlin Mandate, which "establish[es] a process to negotiate strengthened commitments for developed countries" (UNFCCC 2014). The Mandate is only 3 pages long, but has been an important milestone in the negotiation process leading up to the Kyoto Protocol (cf. Andresen Agrawala 2002: 47). Its legal status is that of a COP Decision, which can roughly be compared in status to a Final Declaration under the NPT regime. The Berlin Mandate is very short (28 sentences, or 915 words), and it is the third vaguest of all the documents analysed with 1.5 points exactly.

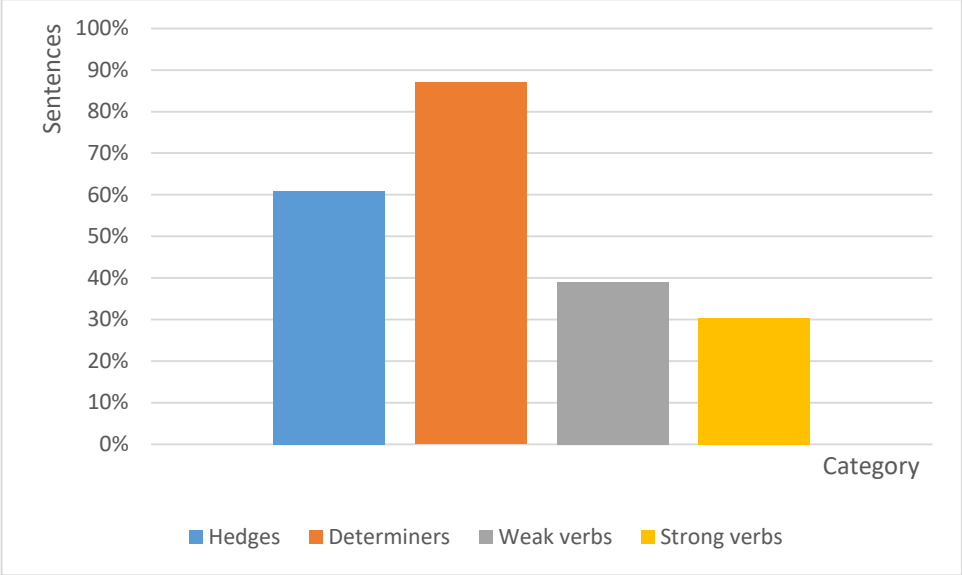


Figure 37: Percentage of vagueness indicators UNFCCC 1995

The breakdown of points in graph 38 gives some insights into why this is so: 60.8% of sentences contain hedge words, which is highest percentage of all documents. 87% contain less than three determiners, making this category fairly average. Similarly, 39% of sentences in the main text contain weak explicitly performative verbs but on the other hand, 30.4% contain strong explicitly performative verbs, which is a lot for the document's legal status and a clear difference to documents of similar status under the NPT regime. While only one of these cases represents a sentence with the word 'shall' and all other points are due to the use of the verb 'will', there is a crucial difference to the NPT regime: 'will' is here used as the operative verb of the sentences, not just in less crucial parts of the sentence, and not only within descriptions of facts.[40] A case in point:

> *"The process will be carried out in the light of the best available scientific information and assessment on climate change and its impacts, as well as relevant technical, social and economic information, including, inter alia, reports of the Intergovernmental Panel on Climate Change" (UNFCCC 1995: 6).*

This sentence clearly shows that 'will' is used in a different capacity under the climate regime than in the Final Declarations of the NPT regime.

The number of sentences receiving each score underline how vague this document is: the distribution is practically moved over by a point compared to other documents.

[40] Descriptions of facts in international legal agreements are of course also performative.

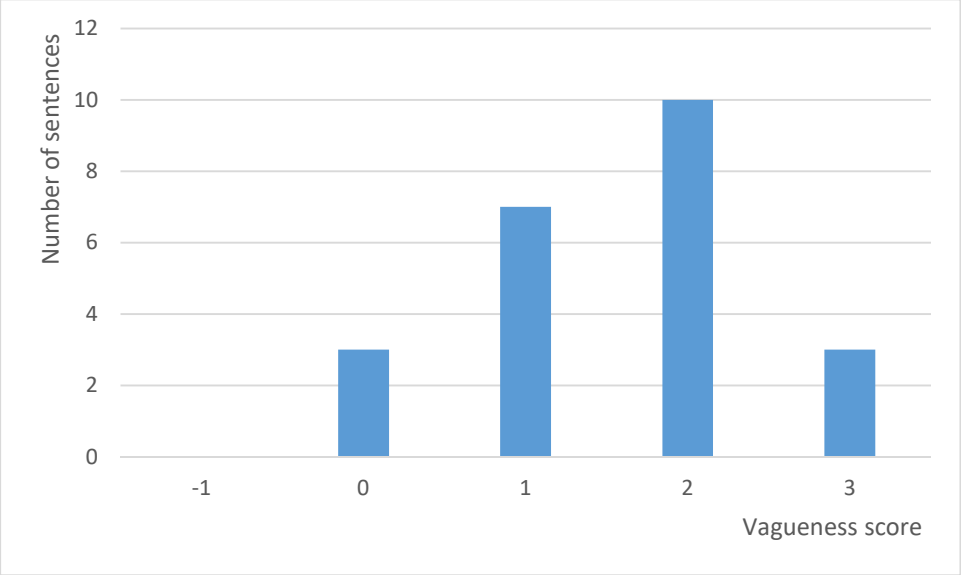


Figure 38: Score distribution UNFCCC 1995

There are three sentences scoring 3 points, one of which is mostly a quote from the Convention. The two other sentences read as follows:

> *"The protocol proposal of the Alliance of Small Island States (AOSIS), which contains specific reduction targets and was formally submitted in accordance with Article 17 of the Convention, along with other proposals and pertinent documents, should be included for consideration in the process" (UNFCCC 1995: 5).*

> *"The sessions of this [ad hoc] group should be scheduled to ensure completion of the work as early as possible in 1997, with a view to adopting the results at the third session of the Conference of the Parties" (UNFCCC 1995: 5).*

Uncharacteristically, both of these vague sentences are actually about relatively small-scope matters, with the latter sentence especially having an administrative character. The first sentence the consideration of a party proposal, and the second one concerns the scheduling for a newly established ad hoc group. Both sentences use the verb 'should', which is of course a weak performative verb. While in the first case, only 'specific' is counted as a hedge word, 'pertinent' could certainly also be counted. The second sentence includes the phrase 'as early as possible'. The first sentence contains one, the second sentence two determiners.

As can be seen in the graph above, no sentence in the main text of this document scores -1. There are two sentences scoring 0, two of which are actually introductory sentence fragments. Below, you can see one of these fragments in connection with the third relatively precise sentence of the document.

> *"The process shall be guided, inter alia, by the following: […] The fact that the largest share of historical and current global emissions of greenhouse gases has originated in developed countries, that the per capita emissions in developing countries are still relatively low and that the share of global emissions originating*

in developing countries will grow to meet their social and development needs" (UNFCCC 1995: 4).

The first part includes a strong performative verb – 'shall' – but it also scores a vagueness point for having no determiner, thus coming out at 0. The second part of the construct follows the more usual pattern of scoring in none of the categories.

Undoubtedly, the key sentence of this short document is the one including the actual Mandate it is named after.

> *"[The Conference of the Parties, at its first session] agrees to begin a process to enable it to take appropriate action for the period beyond 2000, including the strengthening of the commitments of the Parties included in Annex I to the Convention (Annex I Parties) in Article 4, paragraph 2(a) and (b), through the adoption of a protocol or another legal instrument" (UNFCCC 1995: 4).*

This paragraph receives a vagueness score of 2, because it contains the hedge 'appropriate' and only the determiner 'another'.

UNFCCC 1996

In its second session in Geneva, the COP noted – but did not adopt – the 4-page Geneva Ministerial Declaration. A large part of the conference was concerned with the preparation for what would later become the Kyoto Protocol, and in particular with the second IPCC[41] report. The fact that it is a Ministerial Declaration, rather than a set of decisions, means that the agreement is not actually legally binding. This type of document does not have an equivalent in the documents analysed under NPT.

The Geneva Ministerial Declaration has a vagueness score of 1.3, which is roughly in the middle of all the documents analysed.

Breaking this number down, 37% of sentences in the main text of the declaration contain at least one hedge word. 92.6% include less than three determiners, 29.6% include weak performative verbs, and 14.8% contain strong explicitly performative verbs as shown in figure 40.

[41] Second Assessment Report of the Intergovernmental Panel on Climate Change, 1995

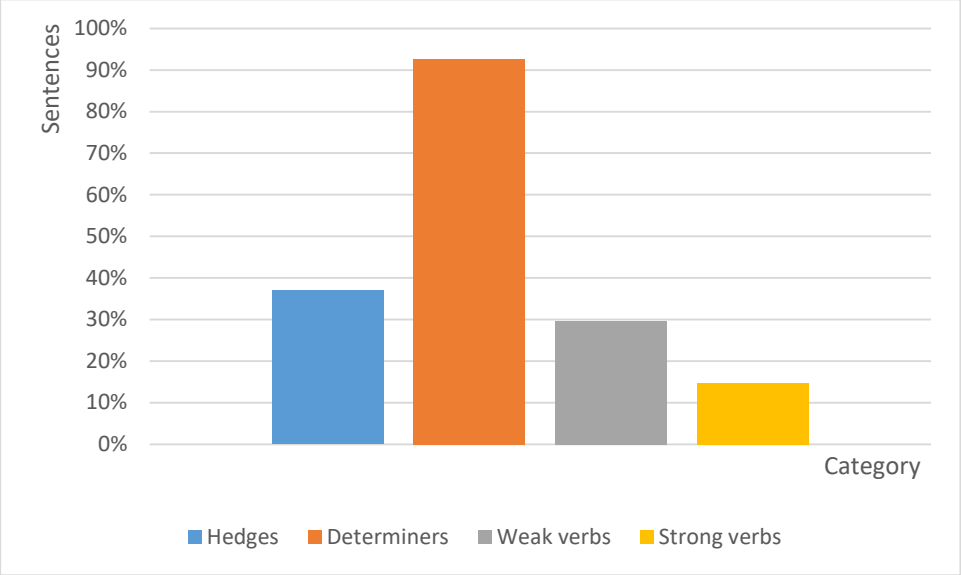


Figure 39: Percentage of vagueness indicators UNFCCC 1996

These last two figures are not unusual overall, but they are interesting considering the legal status of the document. When taking the non-binding nature of the document into account, one would expect that no strongly performative verbs would be used at all. It should be noted that in this text like in the Berlin Mandate, the verb used is 'will', while 'shall' is still avoided as an operative verb. I will explore this phenomenon further in section 6.5.

The following graph shows once more the distribution of scores within the main body of text of the declaration.

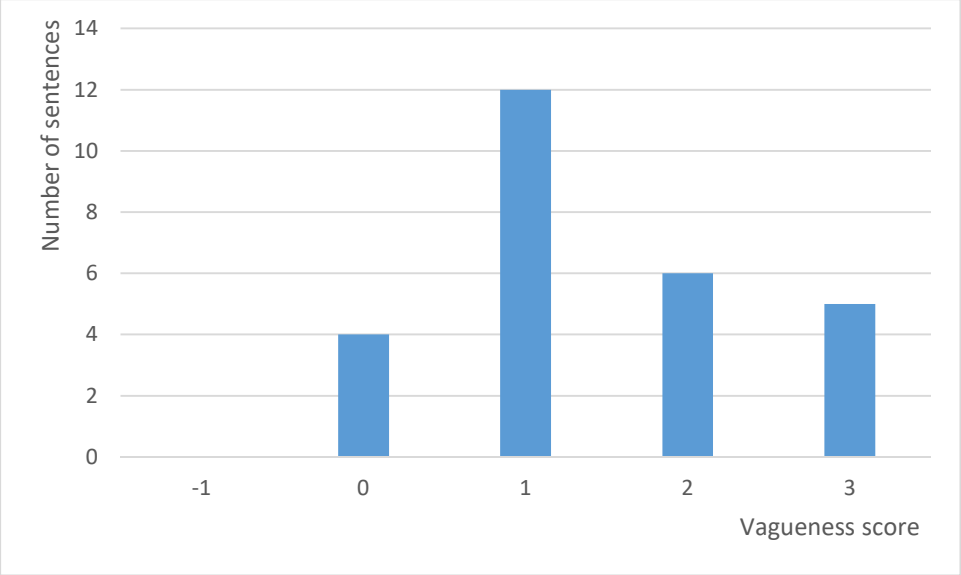


Figure 40: Score distribution UNFCCC 1996

In words: Out of 27 sentences in the main body of text, none gets a score of -1, four sentences score 0, 12 sentences score 1, six sentences received a score of 6 and five received the highest possible score of 3. A sentence from this latter category read as follows:

> *"Ministers believe that the Second Assessment Report should provide a scientific basis for urgently strengthening action at the global, regional and national levels, particularly action by Parties included in Annex I to the Convention (Annex I Parties) to limit and reduce emissions of greenhouse gases, and for all Parties to support the development of a Protocol or another legal instrument; and note the findings of the IPCC, in particular the following" (UNFCCC 1996: 1).*

This sentence is then followed by several resolutions on particular contents of the IPCC report. The trigger word for a weak performative verb is 'could', but 'believe' might also be in this category. The verb 'note' appears to be neutral as to its vagueness or precision in the results of the survey. 'Particularly' features as a hedge word. Even though the sentence is fairly long at 71 words, it contains only two determiners: 'all' and 'another'. On the other hand, as there were no sentences scoring -1, here are two of the four sentences scoring 0:

> *"The projected changes in climate will result in significant, often adverse, impacts on many ecological systems and socio-economic sectors, including food supply and water resources, and on human health" (UNFCCC 1996: 2).*

and

> *"Stabilization of atmospheric concentrations at twice pre-industrial levels will eventually require global emissions to be less than 50 per cent of current levels" (UNFCCC 1996: 1).*

In both cases, 'will' is the operative verb. As mentioned above, this is a difference to the Final Declarations under the NPT regime, where 'will', if used at all, only ever featured in later parts of the sentences. Even though the sentences therefore score -1 point each for including a strong performative verb, none of them include three or more determiners, so they score 0 points in total.

UNFCCC 1997

The third COP in Kyoto (1997) established the Kyoto Protocol, which, according to the UNFCCC website, is "billed the world's most ambitious environmental treaty". Indeed, the strong point of the Kyoto Protocol is that it constitutes a legally binding commitment to reduce their production of greenhouse gases for all developed country parties listed in Annex I of the Protocol (UNFCCC 1998). The treaty measures 25 pages in length and comprises two annexes. Especially Annex I is frequently cited, since it comprises the list of countries with quantified emission limitation commitments under the Protocol. The Kyoto Protocol is arguably the most important document in climate regime to date. Its legal status is that of a Protocol under Convention, so it is comparable in legal terms to the NPT and the Convention. At 297 sentences, or 8520 words, its length is comparable to the Framework Convention as well.

The Kyoto Protocol is the most precise document in this study. Its vagueness score is 0.36. When only looking at main body of text (excluding headlines, preamble, annexes and closing sentences), the difference to other relatively precise documents becomes even more pronounced: the score is then 0.28.

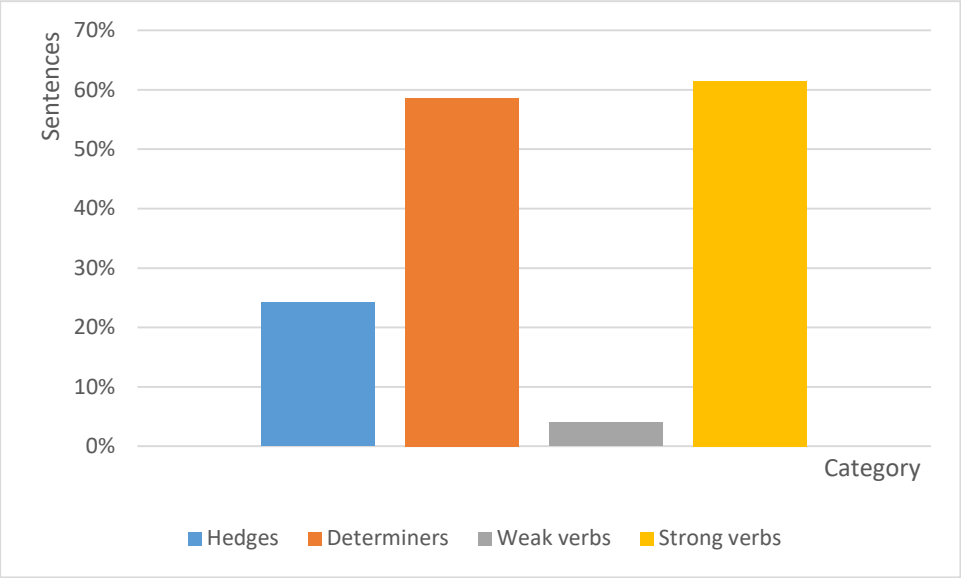


Figure 41: Percentage of vagueness indicators UNFCCC 1997

Looking more closely at the makeup of the score, some of the reasons for this become clearer. 26.4% of all the sentences contain hedge words, which is second only to the Copenhagen Accord. 58.6% of sentences contain less than three determiners, which is an exceptionally low score. There is one document in which it this score is lower, which is the original NPT at 45%. All other documents, including the Framework Convention, have a percentage of over 80, which marks an enormous gap in comparison to other differences. Only 4% of all sentences contain weak explicitly performative verbs, which is an extremely low number,[42] while 61.4% – the highest percentage in all analysed documents – contain strong explicitly performative verbs.

[42] It is important to keep in mind that, as explained in section 5.3.1, the analysis does not include all or maybe even most of weak explicit performative verbs that are possible to use in legal documents. It does however include the most common ones.

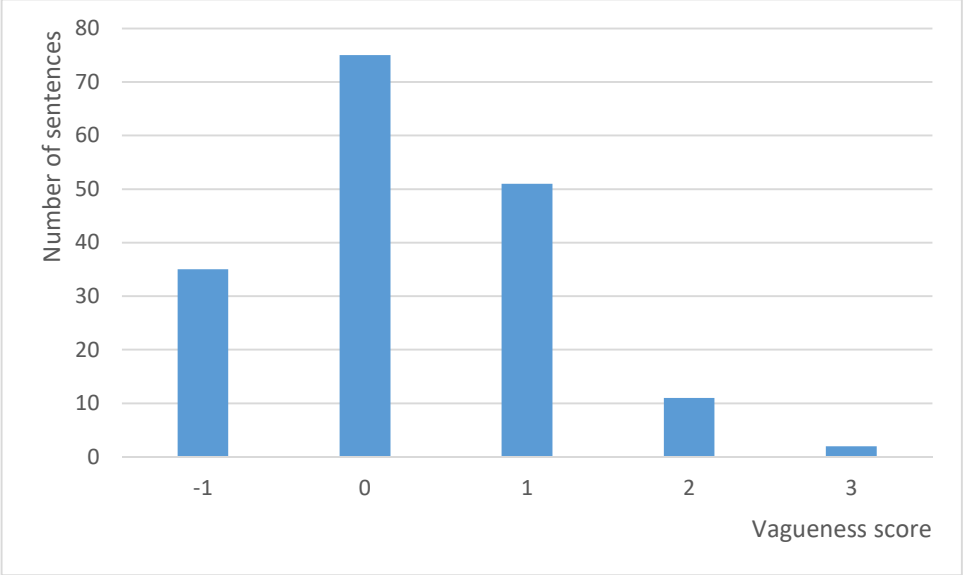


Figure 42: Score distribution UNFCCC 1997

As can be seen in the graph above, the distribution of sentences among points is shifted to the left of the usual distribution, as the majority of sentences score 0, not 1. In numbers, 35 sentences score -1, 75 sentences score 0, 51 sentences score 1, 11 sentences score 2 and only two sentences have a score of 3.

Here is an example of one of the two sentences scoring 3 (both of which are contained in art. 10 of the Kyoto Protocol):

> *"Suitable modalities should be developed to implement these activities through the relevant bodies of the Convention, taking into account Article 6 of the Convention" (UNFCCC 1997: Art. 10(e)).*

The operative verbs in these sentences are 'would' and 'could' respectively. The 'suitable' in the first sentences may well be a hedge word, but it is not reflected in the score. The word 'relevant', however, is.

On the precise side, there are many more sentences to choose from. The first example is on the topic of data provision – a matter that often makes for very precise sentences in the climate regime.

> *"Prior to the first session of the Conference of the Parties serving as the meeting of the Parties to this Protocol, each Party included in Annex I shall provide, for consideration by the Subsidiary Body for Scientific and Technological Advice, data to establish its level of carbon stocks in 1990 and to enable an estimate to be made of its changes in carbon stocks in subsequent years" (UNFCCC 1997: Art 3.4).*

The determiners used here are 'each', 'its' (twice) and 'this'. The operative verb is of course 'shall'. The impression of precision this sentences provides may be aided by the use of two relatively specific timeframes. However, this indicator did not prove statistically significant in the analysis of the survey data, so it is not included in the method used to score

sentences here. The next example of a sentence also concerns a major topic for the UNFCCC, namely Land Use, Land Use Change and Forestry (LULUCF).

> *"The net changes in greenhouse gas emissions by sources and removals by sinks resulting from direct human-induced land-use change and forestry activities, limited to afforestation, reforestation and deforestation since 1990, measured as verifiable changes in carbon stocks in each commitment period, shall be used to meet the commitments under this Article of each Party included in Annex I" (UNFCCC 1997: Art 3.3).*

Again, the operative verb of the sentence is 'shall'. The determiners used here are 'each' (used twice), as well as 'this'. One of the key provisions in the Kyoto Protocol describes the duties of Annex I parties. The sentence under Article 3

> *"The Parties included in Annex I shall, individually or jointly, ensure that their aggregate anthropogenic carbon dioxide equivalent emissions of the greenhouse gases listed in Annex A do not exceed their assigned amounts, calculated pursuant to their quantified emission limitation and reduction commitments inscribed in Annex B and in accordance with the provisions of this Article, with a view to reducing their overall emissions of such gases by at least 5 per cent below 1990 levels in the commitment period 2008 to 2012" (UNFCCC 1997: Art 3.1) .*

As there are no hedge words or weak performative verbs, but instead a strong performative verb ('shall') and no less than six instances of determiners – 'their' is used four times, 'such', and 'this'. This provision therefore also receives 3 points. This is important because, as we have seen with the NPT of 1968, a document can be very precise overall but still vague in its key provisions.

UNFCCC 1998

In the following years, the main concern lay in the retention of the Kyoto Protocol. The Buenos Aires Plan of Action was adopted by the CoP at its 4th session. Counting about 40 pages, it is lengthier than its predecessors, comprising seven individual decisions. It is mainly concerned with the financial and trading mechanisms called for in the Kyoto Protocol. A lot of the more technical – and more detailed – aspects of the decisions are dealt with in Annexes. At 460 sentences, it would be the longest document in this analysis. However, the majority of these sentences are annexes and tables, so that the main text is only 79 sentences, or 8670 words long. This fact pushes the limits of this method in several ways, which will be discussed in more detail in section 4.6.

In this specific case, it does not seem to make a big difference whether or not the annexes are included. While the vagueness score of the complete document – including annexes – is 1.17, while just the main text scores 1.28. The difference of 0.11 would not alter the order of the documents at all. Both versions of the vagueness scores are close to the overall average of 1.13.

The sources of this score is also fairly typical:

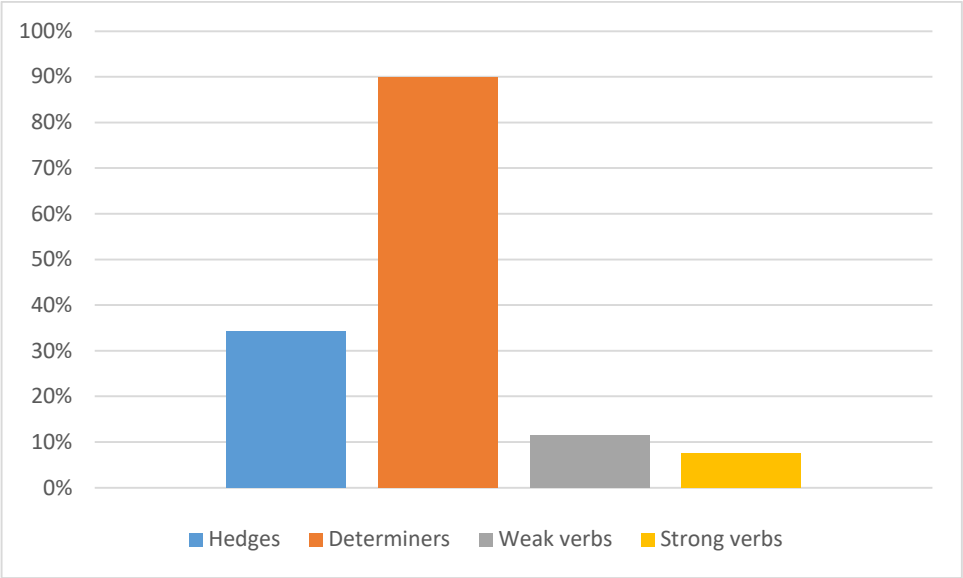


Figure 43: Percentage of vagueness indicators UNFCCC 1998

34.2% of the sentences in the main text of the regime contain hedge words, 89.9% contain less than three determiners, 11.4% of sentences feature weak explicitly performative verbs and 7.6% contain strong explicitly performative verbs. The following graph shows the distribution of scores among the sentences:

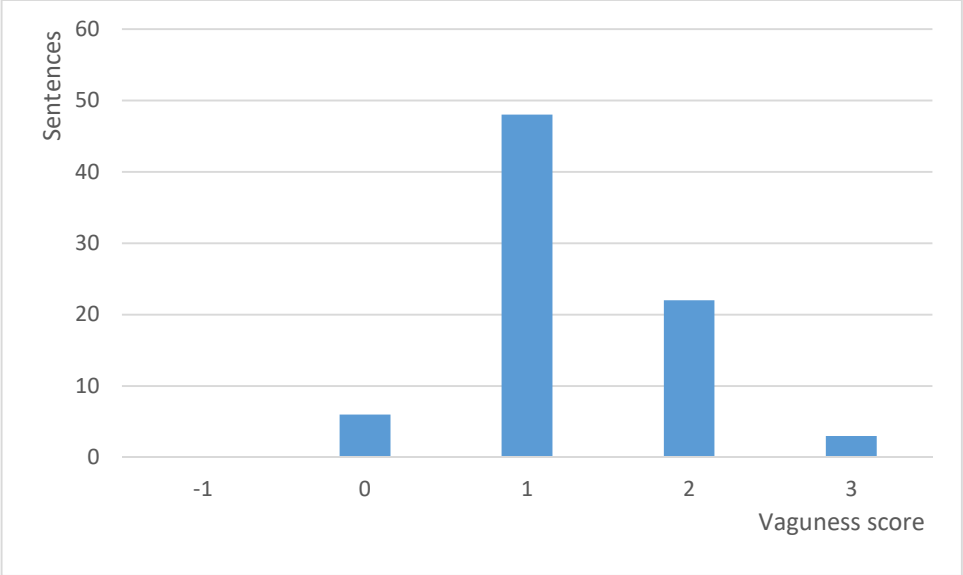


Figure 44: Score distribution UNFCCC 1998

While the graph shows at some more sentences score in the vague category than in the precise category, this distribution is still quite usual for this study.

Four sentences out of 79 in the main text score 3. One of them reads as follows:

> *"The consultative process could include, resources permitting, regional meetings, regional workshops and a SBSTA workshop process could include, resources permitting, regional meetings, regional workshops and a SBSTA and, as appropriate, experts engaged in the IPCC process" (UNFCCC 1998: 14).*

In this case, the operative verb is 'could' which has arguably even less performative force than 'should', which at least implies some level of incitement towards the described action. The hedge word is the most commonly used 'appropriate'. It actually doesn't contain any determiners at all, which is rare even with most sentences scoring in this category.

On the other end of the spectrum, there are actually no sentences scoring -1 (as can be seen in graph 45 above). Instead, here is a sentence scoring 0 points:

> *"[The Conference of the Parties [i]nvites all Parties and interested international and non-governmental organizations to identify projects and programmes incorporating cooperative approaches to the transfer of technologies which they believe can serve as models for improving the diffusion and implementation of clean technologies under the Convention, and to provide information thereon to the secretariat, by 15 March 1999, for compilation into a miscellaneous document to be considered by the Subsidiary Body for Scientific and Technological Advice (SBSTA) at its tenth session" (UNFCCC 1998: 13).*

This is a sentence which scores 0 and also contains neither a strong nor a weak performative verb.[43] It contains three determiners ('all', 'which', and 'its'), therefore not scoring in that category.

As annexes are so important – or at the very least voluminous – in this document, here is an example to illustrate the problems of including them in this analysis:

> *"What should be the objective of collaboration with relevant multilateral institutions to promote technology transfer and what practical steps should be taken? " (UNFCCC 1998: 15).*

Technically, this sentence would score as very vague, but it is not the kind of legal language which the subjects of the survey were asked about. For a further discussion of this issue, see section 4.6.

UNFCCC 2000

COP 6 in The Hague could not finish in the time allocated and was concluded at the so-called COP 6-bis in Bonn, Germany, producing the Bonn Agreement (Sari 2005: 34). The full title of the document ('Bonn Agreements on the implementation of the Buenos Aires Plan of Action') gives a fairly comprehensive summary of its contents. The short decision is followed by a 21-page annex structured along seven topics, all concerned with detailing the provisions of the Buenos Aires Plan of Action. The document has an overall very low score of 0.85. In fact, it is the lowest vagueness score of any analysed document that is not one of the three

[43] Some of these verbs have been teste das neutral in their effect on perceived vagueness by the study, but most of these verbs cannot be included in the tool because they were not part of the survey. Only the most common operative verbs were tested.

main legal documents of the study, which makes it by extension the most precise set of decisions under a Convention or Protocol.

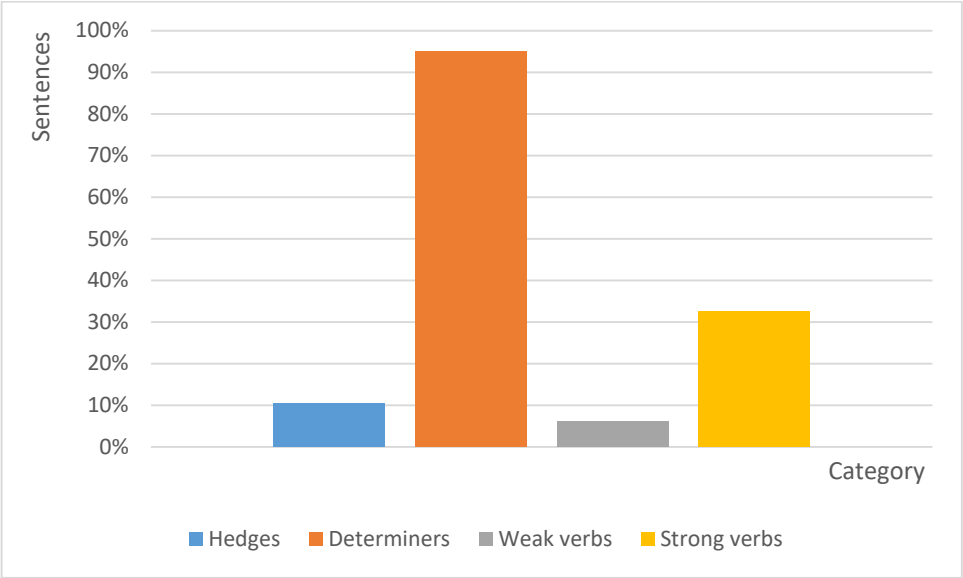


Figure 45: Percentage of vagueness indicators UNFCCC 2000

As the graph above shows, only 11% of sentences in the main body of text contain hedge words, which is the lowest score in any of the documents. On the other hand, the document is also exceptionally high in sentences that contain 2 or less determiners: they make up 95% of sentences in the main body of text. This category reduces the precision of the document to some degree. Only 6% of sentences contain weak performative verbs – this number again being lower than even the Kyoto Protocol, the overall most precise document in this study. 32.5% of sentences feature strong performative verbs.

The following graph shows the distribution of scores among sentences in the document. Two sentences score -1 point, 50 score 0, 92 sentences receive a score of 1, 18 sentences score 2 and 1 sentence appears in the vaguest category of 3 points.

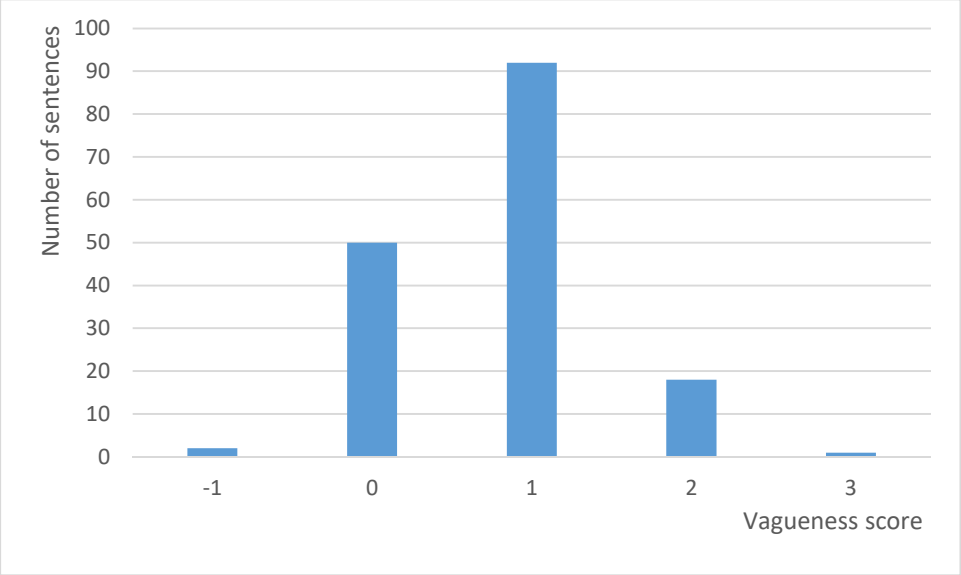


Figure 46: Score distribution UNFCCC 2000

The sentence scoring 3 concerns funding for non-Annex I parties.

> *"[The Conference of the Parties [r]ecognizes that funding should be made available to Parties not included in Annex I, which is new and additional to contributions under the Convention" (UNFCCC 2000: 3).*

Both 'recognizes' and 'should' actually count as weakly performative verbs. On the other hand, there is a rare mix-up in the hedge category, as 'available' is in this case used as a predicate, not an adjective, thus arguable losing its hedge word status. Since this happens only very rarely, it is negligible in the aggregate (see section 6.5 for a further discussion on this topic). There are no determiners and also no strong performative verbs in this sentence.

On the precise side, one of the sentences in the main body of text of this document scoring -1 is concerned with

> *"[The Conference of the Parties agrees] That the use of the mechanisms shall be supplemental to domestic action, and that domestic action shall thus constitute a significant element of the effort made by each Party included in Annex I to meet its quantified emission limitation and reduction commitments under Article 3, paragraph 1" (UNFCCC 2000: 7).*

As well as

> *"[The Conference of the Parties agrees] That, during the first commitment period, a Party that selects any or all of the activities mentioned in paragraph 4 above shall demonstrate that such activities have occurred since 1990, and are human-induced" (UNFCCC 2000: 10).*

Both sentences contain the word 'shall'. The determiners used are 'that', 'each', and 'it's' and 'any', 'all', and 'such', respectively.

UNFCCC 2001

The Marrakech Ministerial Declaration, as can be gathered from its name, is not a legally binding document. It shares this characteristic with the 1992 Geneva Ministerial Declaration, the 2002 Delhi Ministerial Declaration and the 2009 Copenhagen Accord. It is also shortest document in this study, containing only 12 sentences in total. Additionally, four of these are in the preamble, one is a closing formula, and two are headlines, so there are only five sentences in the main body of text. Since there are so few of them, I will look at them all in this section, because the small number of sentences might also mean that small inaccuracies in the results don't necessarily even out, so greater scrutiny is warranted.[44]

The vagueness score of this document is 1.08, which is relatively precise. However, due to the low number of sentences the difference between the entire text and the main body of text is actually quite large, with the latter scoring 1.4. For completeness' sake, below is a graph on the distribution of points, even though at 5 sentences it is perhaps not as informative as in other cases.

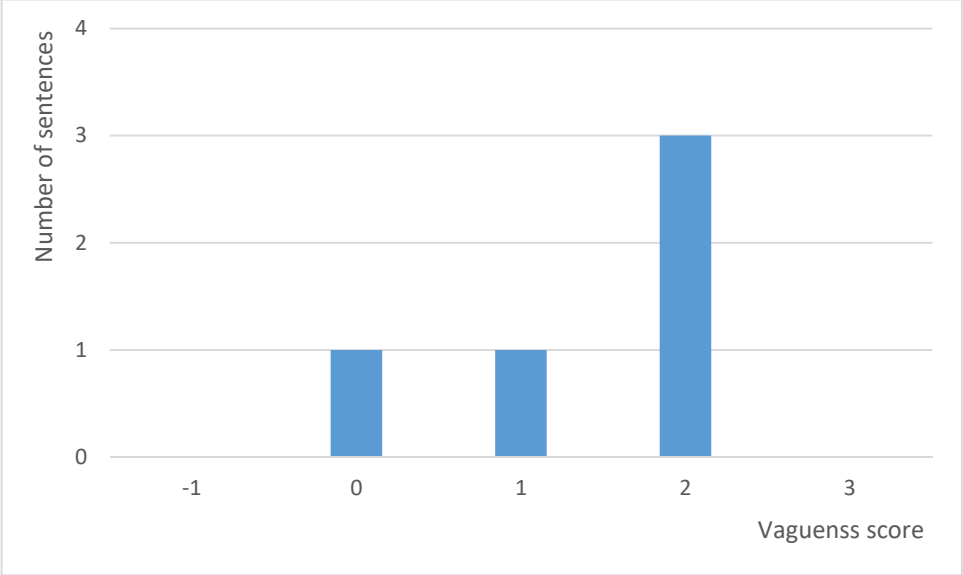


Figure 47: Score distribution UNFCCC 2001

All sentences start with this formula:

> *"The Ministers and other heads of delegation present at the seventh session of the Conference of the Parties to the United Nations Framework Convention on Climate Change" (UNFCCC 2001: 1),*

which as usual is not included in any of the points.

> *"Note the decisions adopted by the seventh session of the Conference of the Parties in Marrakesh, constituting the Marrakesh Accords, that pave the way for the timely entry into force of the Kyoto Protocol"* (UNFCCC 2001: 1).

[44] The question on how to deal with very short texts is further considered in section 6.5.

This sentence scores 1 point because it contains no determiners. 'Note' appeared neutral in its performativity in the results of the survey. The next sentence in the document scores 2:

> *"Remain deeply concerned that all countries, particularly developing countries, including the least developed countries and small island States, face increased risk of negative impacts of climate change"* (UNFCCC 2001: 1).

'Particularly' is a hedge word. The only determiner used is 'all'. Like 'note', 'remains concerned' is neutral in its performativity. The next sentence, even though it reads pretty differently, also scores 2:

> *"Recognize that, in this context, the problems of poverty, land degradation, access to water and food and human health remain at the centre of global attention; therefore, the synergies between the United Nations Framework Convention on Climate Change, the Convention on Biological Diversity, and the United Nations Convention to Combat Desertification in those Countries Experiencing Serious Drought and/or Desertification, Particularly in Africa, should continue to be explored through various channels, in order to achieve sustainable development"* (UNFCCC 2001: 1).

The score comes together in a different way, though. While this sentence also only contains one determiner ('those'), it does not include any hedge words but instead features the weak performative verb 'recognize'. The next sentence is back to the same features as the second one.

> *"Stress the importance of capacity-building, as well as of developing and disseminating innovative technologies in respect of key sectors of development, particularly energy, and of investment in this regard, including through private sector involvement, market-oriented approaches, as well as supportive public policies and international cooperation"* (UNFCCC 2001: 1).

While the determiner here is 'this', the hedge word is actually 'particularly', same as above. The last sentence in the main body of text of this document scores 0 in all categories, which is to say that it contains no hedge words, and no verbs of weak or strong performativity, but it does contain three hedge words: 'its' and twice 'all':

> *"Emphasize that climate change and its adverse impacts have to be addressed through cooperation at all levels, and welcome the efforts of all Parties to implement the Convention"* (UNFCCC 2001: 1).

Interestingly, neither 'emphasize' nor 'welcome' were rated as weak performative verbs by participants of the survey, though both verbs were tested for.

UNFCCC 2002

In 2002, COP 8 in New Delhi concluded with the Delhi Ministerial Declaration on Climate Change and Sustainable Development. Like most Ministerial Declarations, it is, at about 3 pages, relatively short.

With a vagueness score of 1.84, the declaration is the vaguest document in the entire study. As I have noted in talking about the Kyoto Protocol, the distance to other documents becomes even more pronounced when only the main text is taken into account: then score is 2.2. The graph below shows the percentage of sentences scoring in each of the four categories:

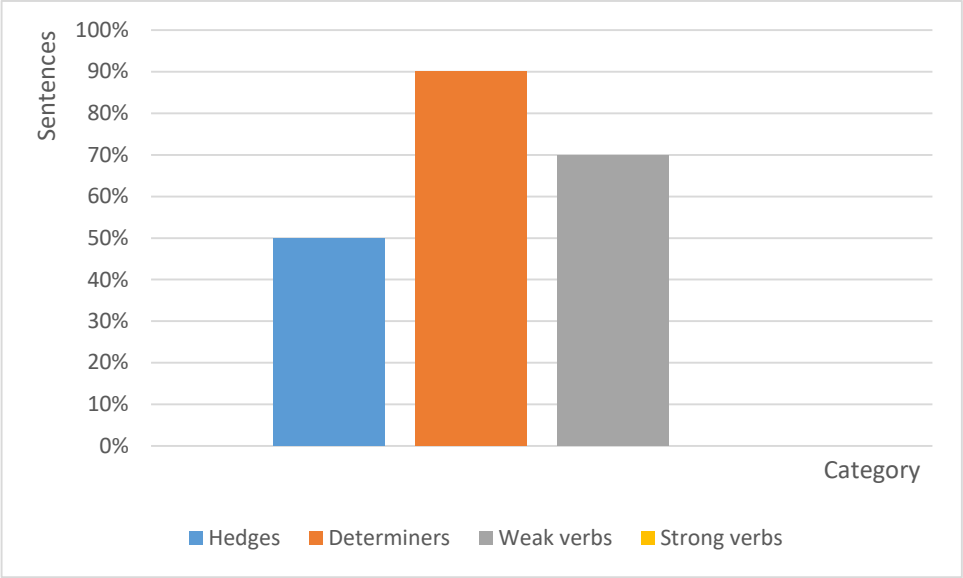


Figure 48: Percentage of vagueness indicators UNFCCC 2002

60% of the sentences in this document contain hedge words, 90% of them have less than 3 determiners, and a total of 70% of sentences feature weak explicitly performative verbs. Consistent with the non-legally binding nature of the document, no sentence contains a strong performative verb. As a result, the lowest category of -1 does not contain any sentences. One sentence scores 0, 3 sentences score 1, 8 score 2 and another 8 sentences score 3, as can be seen in the graph below:

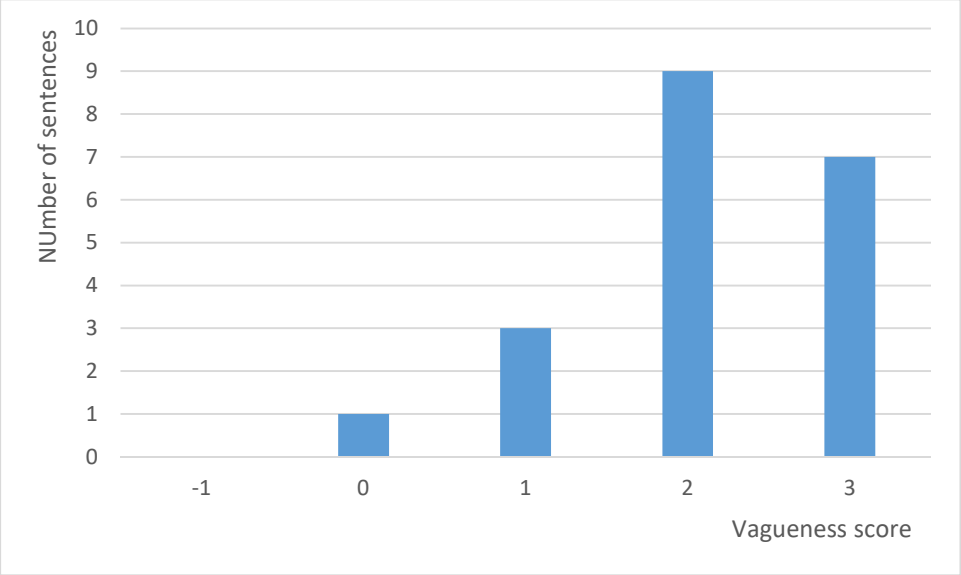


Figure 49: Score distribution UNFCCC 2002

The following sentence uses the verb 'should' not once, but twice:

> *"Policies and measures to protect the climate system against human-induced change should be appropriate for the specific conditions of each Party and should be integrated with national development programmes, taking into account that economic development is essential for adopting measures to address climate change" (UNFCCC 2002: 1).*

It also includes the hedge word 'appropriate'. Arguably, the content of the sentence – if it were used in a more legally binding setting – could relativize the entire goal of the UNFCCC by explicitly subordinating combatting climate change entirely to economic development. This is not the case of course both because of the wording of the sentence and the legal status of the document it is in. The second vague sentence is uncharacteristic as its topic is exchange of information, which usually makes for relatively uncontroversial and precise provisions.

> *"Parties should promote informal exchange of information on actions relating to mitigation and adaptation to assist Parties to continue to develop effective and appropriate responses to climate change" (UNFCCC 2002: 2).*

Two hedge words are used in this sentence: 'effective' and 'appropriate'. Incidentally, it is a good example of why these adjectives are counted as hedge words: instead of narrowing the range of possibilities of what responses to climate change should be taken, they actually serve to blur the lines further, making it less instead of more clear what parties should do. The operative verb being 'should' also cuts down the performative force of this sentence. There are no determiners either in this sentence or the next.

> *"Technological advances should be promoted through research and development, economic diversification and strengthening of relevant regional, national and local institutions for sustainable development" (UNFCCC 2002: 2).*

In this last example the verb is also 'should'. The hedge word used here is 'relevant'. Its present increases the arbitrariness of which institutions for sustainable development actually count under this proposal.

As there are no sentences with the score of -1, below is the one sentence of the document which scores 0:

> *"Actions are required at all levels, with a sense of urgency, to substantially increase the global share of renewable energy sources with the objective of increasing their contribution to total energy supply, recognizing the role of national and voluntary regional targets as well as initiatives, where they exist, and ensuring that energy policies are supportive to developing countries' efforts to eradicate poverty" (UNFCCC 2002: 2).*

The determiners used here are 'all', 'there', and 'that'. There are no hedge words included in this study in this sentence. The verbs are categorized neither as strong nor weak performative verbs.

UNFCCC 2007

The COP-13/CMP-3 2007 in Bali was perceived by negotiators as a breakthrough, since it saw the adoption of the 4-page Bali Action Plan, outlining the steps to the regime's future after 2012, when the first commitment period under the Kyoto Protocol ended. In spite of its shortness, the plan goes into relative detail on further steps to be taken, and its adoption is positively remembered by a lot of COP-participants and observers. At 54 sentences, it is a fairly short document.

With 1.09 points on the vagueness scale, which means the COP decision is on the precise end of the spectrum. However, the makeup of these points is interesting:

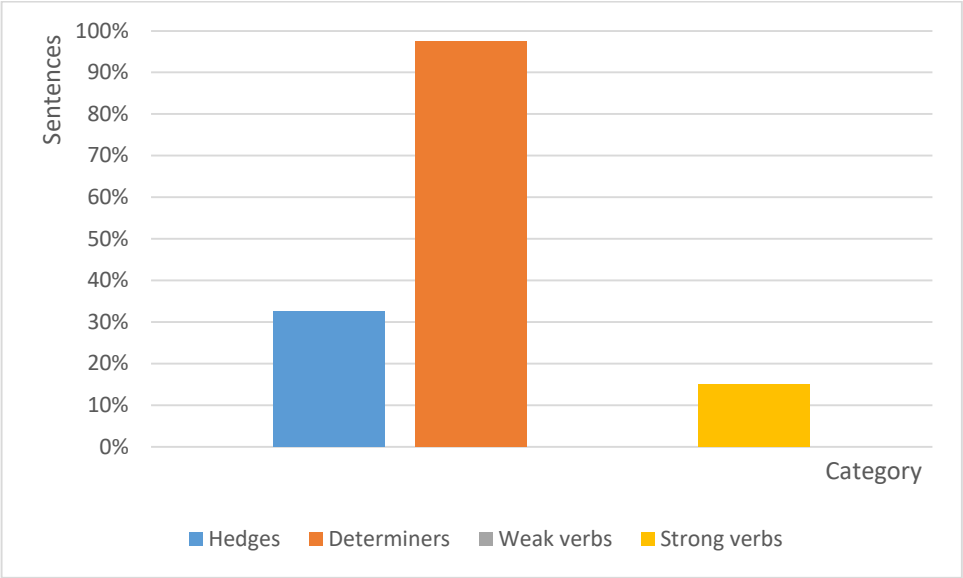


Figure 50: Percentage of vagueness indicators UNFCCC 2007

32.5% of sentences include a hedge word. 97.5% of sentences contain less than three determiners – this is actually the highest percentage of all studied documents in this category. The verb categories are what the relative precision come from. Interestingly, no sentences at all contain a weak verb, while 15% include a strong verb. In terms of the score distribution, one sentence receives a score of -1, three sentences score 0, 25 sentences get a score of 1, 11 score 2 and no sentence in the main body of text of this document scores 3 points:

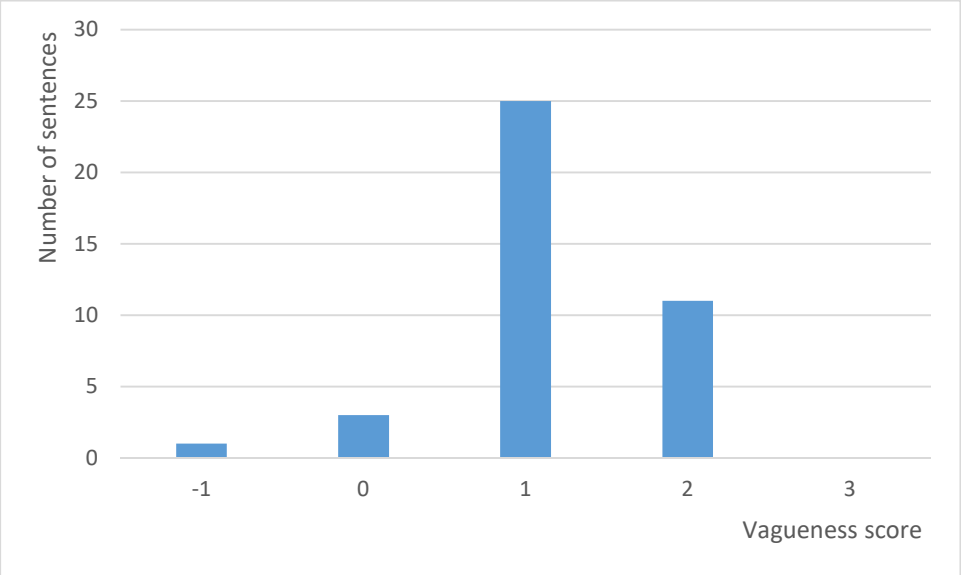


Figure 51: Score distribution UNFCCC 2007

As there are no sentences in the vaguest category, here are the first two sentences of the Bali roadmap, both of which score 2 vagueness points:

> *"The Conference of the Parties […] Decides to launch a comprehensive process to enable the full, effective and sustained implementation of the Convention through long-term cooperative action, now, up to and beyond 2012, in order to reach an agreed outcome and adopt a decision at its fifteenth session, by addressing, inter alia:*
>
> *A shared vision for long-term cooperative action, including a long-term global goal for emission reductions, to achieve the ultimate objective of the Convention, in accordance with the provisions and principles of the Convention, in particular the principle of common but differentiated responsibilities and respective capabilities, and taking into account social and economic conditions and other relevant factors" (UNFCCC 2007 3):.*

This is followed by several other issues the COP is deciding to address. The first sentence scores points in the categories of hedge word, because it contains the words 'effective' and determiners, because it only contains one instance of them: the word 'its'. The second sentence scores in the same categories, because it features the hedge words 'particular' and 'relevant' and actually does not contain any determiners at all.

In contrast, here is the sentence that scores -1 point:

> *"[The Conference of the Parties,] Decides that the process shall be conducted under a subsidiary body under the Convention, hereby established and known as the Ad Hoc Working Group on Long-term Cooperative Action under the Convention, that shall complete its work in 2009 and present the outcome of its work to the Conference of the Parties for adoption at its fifteenth session" (UNFCCC 2007: 5).*

It contains the strong performative verb 'shall' and three times the word 'its', which means it gets no points in the category of determiners.

UNFCCC 2009

The COP15/CMP5 in Copenhagen is surely the most publicized perceived failure of a climate summit. It has been noted for the enormous show of NGO/civil society protest and the failure to adopt the 'Copenhagen Accord', a proposal pushed by the United States of America which was however subsequently signed by 141 parties. The main objections by small developing states and many environmental activists to the accord were the fact that it is not legally binding and that it was tried to implement it without the consulting major developing country parties (Deitelhoff/Wallbott 2012: 18). For a declaration – and one that was not even adopted by the COP/CMP at that – it is unusual in a number of ways. First, it is almost 30 pages long, making it longer than the Convention itself or the Kyoto Protocol. Second, it lists from the outset the names of the participating states. Third, it departs from the extremely formal language common for declarations, using instead the more business – like language more widespread in decisions.

The Copenhagen Accord has a vagueness score of 0.92. That means it is the 4th most precise document out of 11 in UNFCCC Regime or the 3rd most precise for documents which are not Protocols or Conventions (14 of which are included in this study). As this might be a surprising result, I will discuss it further in section 4.3. The makeup of the score is charted in the graph below:

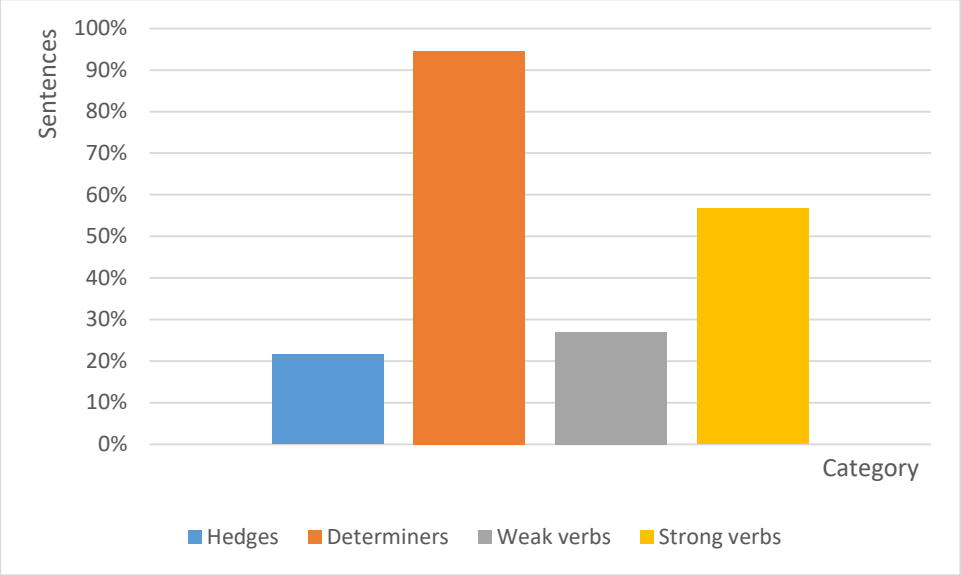


Figure 52: Percentage of vagueness indicators UNFCCC 2009

Only 21.6% of sentences contain hedge words, which is the second lowest percentage right after the Bonn Agreement. 94.6% score because the sentences contain a low number of determiners 27% contain weak explicitly performative verbs, which seems on the low end of average, and – which will be important – 56.8% of sentences contain strong operative verbs.

The score distribution of the main text of the document can be seen in the following chart:

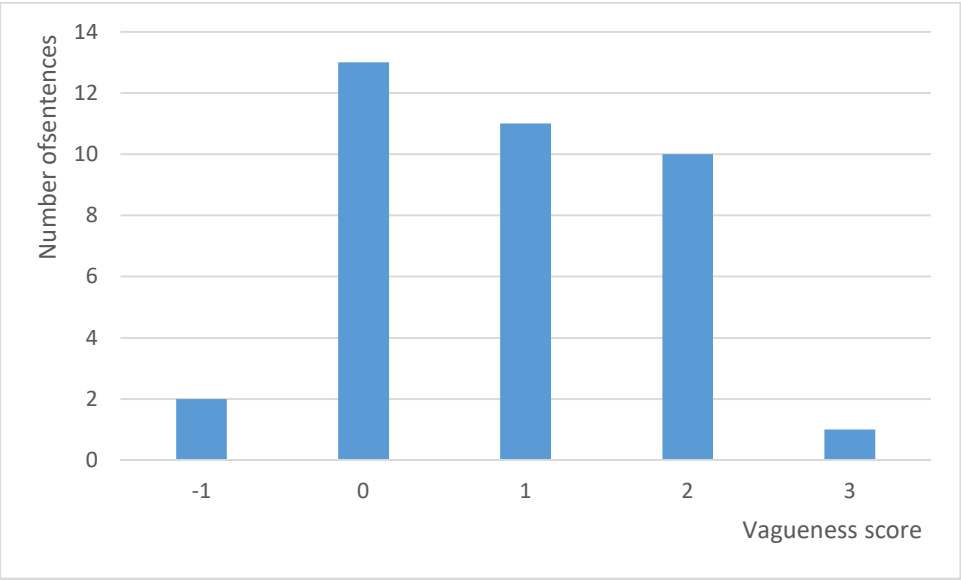


Figure 53: Score distribution UNFCCC 2009

As the graph shows, 2 sentences receive -1 points, 13 receive 0 points, 11 receive 1 point, ten sentences score 2 and one sentence scores 3 points.

Here is the vague sentence:

> *"We recognize the critical impacts of climate change and the potential impacts of response measures on countries particularly vulnerable to its adverse effects*

> *and stress the need to establish a comprehensive adaptation programme including international support" (UNFCCC 2009 5).*

The fact that the verb is 'recognize' makes this sentence score in the weak verb category. 'Particularly' counts as a hedge word. The only determiner is 'its', which means that it gets a vagueness point in that category as well.

On the other side of the scale, below are the two sentences scoring -1.

> *"Delivery of reductions and financing by developed countries will be measured, reported and verified in accordance with existing and any further guidelines adopted by the Conference of the Parties, and will ensure that accounting of such targets and finance is rigorous, robust and transparent" (UNFCCC 2009: 6).*
>
> *"Mitigation actions taken by Non-Annex I Parties will be subject to their domestic measurement, reporting and verification the result of which will be reported through their national communications every two years" (UNFCCC 2009: 6).*

Both of these sentences contain the verb 'will', and it is used in the prescriptive sense. In the first case, the determiners are 'any', 'such' and 'that'. In the second, they are 'their' (twice), and 'which'. Perhaps the most important and controversial part of the Copenhagen Accord are the paragraphs that talk directly about its goals and how to achieve them. The following sentence gets to the heart of that, and it scores 1 vagueness point:

> *"To achieve the ultimate objective of the Convention to stabilize greenhouse gas concentration in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system, we shall, recognizing the scientific view that the increase in global temperature should be below 2 degrees Celsius, on the basis of equity and in the context of sustainable development, enhance our long-term cooperative action to combat climate change" (UNFCCC 2009: 5).*

Here, we can see one of the rare cases when both a weak and strong verb: 'shall' on the one hand and 'would' as well as 'should' on the other. These two scores therefore cancel out. The sentence does not contain hedge words, and there is only one determiner ('our'), so the sentence scores one point in that category.

UNFCCC 2010

The COP16/ CMP6 2010 in Cancun concluded with the 30-page Cancun Agreement and was widely regarded as a measure to re-establish trust in the process and between the participants after what happened in Copenhagen. Indeed, the document itself contains only minor novelties. The Cancun Accord scores 1.28 points on the vagueness scale. That means it is in about the middle of the spectrum overall, but on the more precise side when it comes to regime decisions.

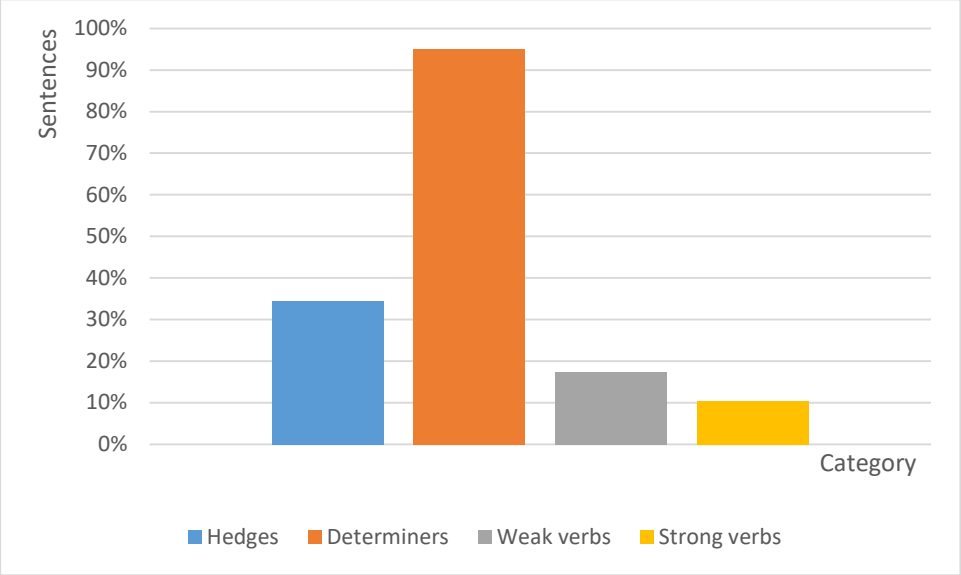


Figure 54: Percentage of vagueness indicators UNFCCC 2010

Overall, 37.1% of sentences in the main body of text in the Cancun Agreement contain hedge words. 94.9% of the sentences contain less than three determiners, 17.4% of them comprise weak performative verbs, while 10.4% feature strong performative verbs. All these numbers are fairly standard for a COP Decision.

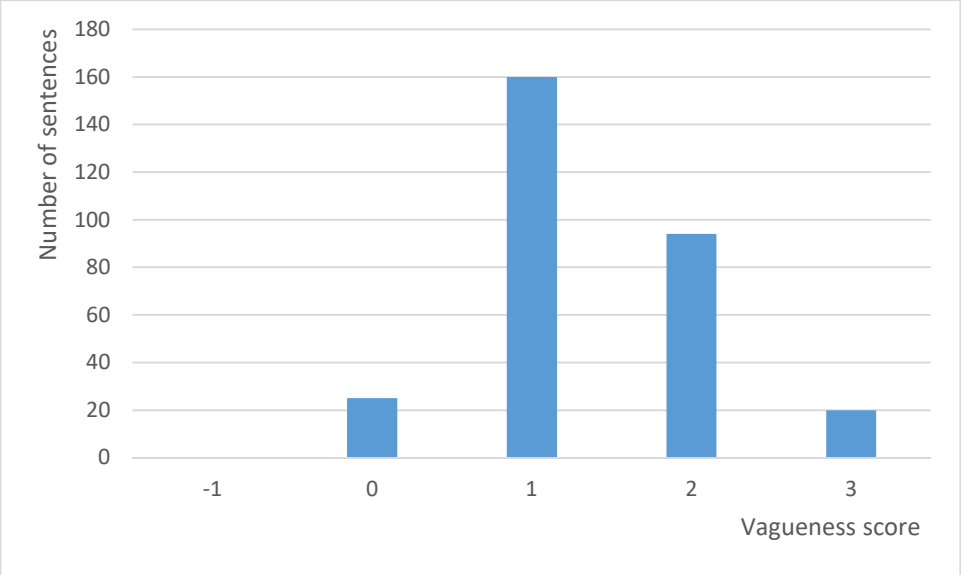


Figure 55: Score distribution UNFCCC 2010

The graph above shows that no sentence in main body of text scores -1, 25 sentences score 0, 160 score 1 point, 94 receive 2 points and 20 sentences score 3. This distribution more or less resembles the overall distribution of points in all documents.

One of the vaguest sentences in the agreement reads as follows:

> *"[The Conference of the Parties]Urges Parties, in particular developed country Parties, to support, through multilateral and bilateral channels, the development of national strategies or action plans, policies and measures and capacity-building, followed by the implementation of national policies and measures and*

> *national strategies or action plans that could involve further capacity-building, technology development and transfer and results-based demonstration activities, including consideration of the safeguards referred to in paragraph 2 of appendix I to this decision, taking into account the relevant provisions on finance including those relating to reporting on support" (UNFCCC 2010: 13).*

This sentences almost starts with the hedge word 'particular', and towards its end also contains 'relevant', another hedge word. In terms of verbs, the word 'could' counts as weakly performative. The length of the sentence notwithstanding, 'this' and 'those' are the only two determiners featured.

On the other hand here is a sentence which scores 0, so is one of the most precise sentences in the document.

> *"[The Conference of the Parties] Decides to hereby establish a process to enable least developed country Parties to formulate and implement national adaptation plans, building upon their experience in preparing and implementing national adaptation programmes of action, as a means of identifying medium- and long-term adaptation needs and developing and implementing strategies and programmes to address those needs" (UNFCCC 2010: 5).*

It does not contain any verbs that score either in a vague or in a precise direction. Neither is there any hedge word from the list included in this study. It does, however, feature 3 determiners: 'least', 'their', and 'those'.

The detailed analysis above shows the capacities and limits of the presented tool to compare the vagueness and precision of legal document. It should have outlined questions of comparability – both within and among documents as well as translated vagueness scores back to meaningful concepts by showing what parts of sentences they indicated. After comparing the results produced by the tool to independent analyses of other researchers in section 6.4, I will come back to some of the issues with the tool that came up in the detailed analyses above in section 6.5 in order to discuss them more thoroughly.

6.4 Comparison of the results

I will now undertake to compare the results of this tool with assessments of other researchers on the individual documents (and their comparison). There are several challenges to this: first, assessments on the vagueness or precision of agreements are heavily biased in favour of extremes (especially towards the vague end of the scale), and agreements that are seen as important or much-debated, which again means that more precise or neutral agreements are much less discussed. Therefore, I too will necessarily concentrate on the more extreme results of the study. Moreover, even in these cases direct assessments are few and far between, as the authors tend to focus on different issues and only mention an agreements' perceived vagueness or precision in passing. From this also follows that

agreements are seldom compared, it is much more likely to find a passing reference as to the vagueness or precision of an isolated agreement, and then it is much more difficult to compare it with the place that agreement has in this study. Nevertheless, the check is still necessary to be able to assess how well the method presented here works.

Starting with the original NPT of 1986, there is quite an interesting case to be made. As mentioned above, the agreement has the least vagueness points of the NPT regime, as well as the second least of all documents tested in the study. With an average of 0.433 (out of a possible range from -1 to 3), where the average of the tested documents is 1.14 (the average of NPT documents is 1.11). How does this compare to expert's assessment? For the most part, authors address specific provisions rather than assessing the document in its entirety. Majority opinion appears to be that provision VI and parts of provision IV, specifically on nuclear disarmament, are vague. Interestingly, as the following quotes will show, the authors use their assessment of the agreement as vague from very different perspectives.

> *"While [Article IV] is only a vague statement with no hard deadlines or timetables, it remains the only legally binding obligation on NWS Parties to ultimately eliminate their nuclear weapons" (Peloso 2011: 313).*

Peloso is arguing that, even though the article on disarmament is vague, elimination of nuclear weapons is still IL. Koplow (1993: 304) elaborates on why he thinks the article is vague, and while he in essence agrees with Peloso on the difference between vagueness and lack of obligation, his focus is more specifically the US:

> *"The Article is unable to be precise about exactly what behaviors are now required of the United States or about what specific timetable is mandated for negotiation or conclusion of the next step in the sequential evolution of test bans. But it is incumbent upon the United States to do something-a vague and aspirational obligation is not necessarily a void and meaningless one-and recent American policies, refusing even to talk about a comprehensive test ban agreement, fall short of our international responsibilities" (Koplow 1993: 304).*

Ford, on the other hand, argues the opposite:

> *"the language [in Article 6] about negotiations needing to be ''pursue[d] . . . in good faith'' clearly leaves open the possibility that such negotiations might not take place, let alone succeed" (Ford 2007: 404).*

For him, not even negotiations on the subject, let alone actual disarmament, are ultimately mandated by the wording of the article. In yet another take on the issue, Craig and Ruzicka focus more on why the provisions on nuclear abolition turned out to be vague than on its effects.

> *"By contrast, a continued focus upon isolated nonproliferation efforts, combined with vague talk of nuclear abolition, ran none of these risks" (Craig Ruzicka 2013: 334).*

On an incidental note, the above sentence also hints that Craig and Ruzicka would consider the rest of the original NPT as fairly precise. The risks they refer to are withdrawal

of funding and marginalization of non-nuclear weapon states, should they push for strong disarmament measures. But regardless of the direction of their argument, some of which are clearly contradictory, all these authors agree that the provisions in the NPT on nuclear disarmaments by states who currently are in possession of nuclear weapons are vague, as Fields and Enia sum up:

> *"The NPT is vague at best on how and when disarmament of the NWS should happen" (Fields Enia 2009: 190).*

At first glance, this seems a slight contradiction to the tool. However, when looking closer, this is not the case at all. First, these provisions appear to be singled out because the rest of the agreement is seen to be rather precise:

> *"Certainly, the drafters of the NPT knew perfectly well how to state legal requirements clearly. Articles I and II of the NPT *its core nonproliferation obligations* are quite unambiguous: nuclear weapon states ''undertake not to'' help others acquire nuclear weapons, and non-weapon states ''undertake not to'' acquire them. And Article III is quite clear that each non-nuclear weapon state ''undertakes to accept'' nuclear safeguards and that specific procedures and safeguards ''shall be'' accepted and followed" (Ford 2007: 403).*

This assessment is in line with more indirect assessments which take the NPT to be quite precise overall. Secondly, and most importantly, it is easily possible to only look at the provisions in question with the tool. Doing this shows that Article 6 has a score of 3 – the highest vagueness score possible on this scale – as does Article 4.2, which is generally the one that is described as vague by these researches (or at least is the part quoted in evidence of their proposal). This seems to be a very strong indicator that the tool is consistent with expert opinion.

The RevCons are much less discussed. Müller describes the development on article 6 of the original NPT in later RevCons:

> *"The major misgivings of non-nuclear weapon states parties to the NPT had always been that their undertakings were precise and controlled […], while those of the nuclear weapon states were vague with no strings attached. In 1995, when the NPT was extended indefinitely, nuclear weapon states had, for the first time, conceded a more precise interpretation of Article VI, which sets out their disarmament obligation: they had accepted the concluson [sic] of a CTBT and negotiations on a FMCT, and 'progressive and systematic steps' towards nuclear disarmament. In 2000, these 'steps' had been further refined and detailed" (Müller 2005: 35).*

As the authors mentioned above, Müller's focus is article IV. Instead of focussing on the original article, however, he traces its development through the RevCon, and notes a decline in vagueness over time. Looking at the section concerned with article VI in RevCon 1995, it includes 15 sentences, which is indeed more volume than the two sentences of article VI of the Original NPT. Additionally, the mean vagueness score of these sentences is 1.87 – significantly lower than the score of 3 of the original article, even though the treaty in its

entirety has a higher score than the NPT. In the Final Declaration of RevCon 2000, 27 sentences are directly concerned with art. VI, and their combined vagueness score is 1.04. Again, this low score is even more meaningful in light of the relatively higher score of the treaty overall. The tool matches Müller's analysis on article VI completely.

Stoiber gives another rare example of an actual comparison of the vagueness of different decisions in the same regime:

> *"Prior to the revelations of safeguard inadequacies in Iraq, RevCon documentation on strengthening safeguards tended to be vague and anodyne. For example, the 1985 Final Document records 'satisfaction with the improvements of IAEA safeguards, which has enabled it to continue to apply safeguards effectively during a period of rapid growth in the number of safeguarded facilities […] The 2000 Final Document reflects a much more rigorous evaluation of safeguards, giving support to the fundamental movement away from a narrow facility-based approach to a nation-wide approach using new authority under INFCIRC/540 to conduct short-notice inspections, utilize intelligence information, and conduct environmental sampling" (Stoiber 2003: 133).*

The overall scores of the Final Documents of RevCon 1985 and 2000 are virtually the same in my analysis (1.38 and 1.39, respectively). It is undoubtedly true that the final decision of Rev Con 2000 devotes a much larger part of its paragraphs to IAEA safeguards than the one of RevCon 1985 – while the latter only dedicates one paragraph (consisting of four sentences to the issue), a substantial part of the former document – 50 sentences – is concerned with safeguards. On the other hand, all of the four sentences the Final Document of 1985 does feature score 1 point, which is lower than the average of sentences concerned with safeguards in the RevCon 2000 Final Document of 1.72. As a matter of fact, this point is a very good reminder of the difficulties distinguishing between generality and vagueness (as discussed in chapter 2.1.1). Four sentences can address a subject much more generally than 50, but that does not necessarily mean that their language is vaguer. For now, the suggestion would be to take note of the amount of language dedicated to a specific issue in any analysis of its vagueness.

This last issue notwithstanding, however, the tool matches the receptions of the documents under the NPT regime quite precisely – even down to individual paragraphs.

I will now turn to similar assessments of documents under the UNFCCC regime.

One of the few comparisons that are made relatively routinely in the literature is that between the UNFCCC and the Kyoto Protocol. Opinion on this matter is consistent with the findings from my tool: researchers agree that the Kyoto Protocol is more precisely worded than the UNFCCC. For example, Roberts (2011: 778) states that the Convention was "carefully crafted but equally vague" and explains:

> *So the words were nice, but they were just that. The difficult parts – the details – were pushed off until later. The Framework Convention remains in effect, and making it more concrete has been the focus of great attention since it is the only global climate agreement the US has ratified. In efforts to put the UNFCCC into binding language requiring action of the Parties, consensus has repeatedly broken down into power politics. This happened in the Kyoto round of climate negotiations. In Japan in 1997, 129 nations agreed to a more concrete deal on how to begin the process of actually reducing greenhouse gas emissions". (2011: 778)*

Von Stein agrees with this assessment,

> *"In general, however, the [UNFCCC] is fairly imprecise. It does not specify quantified emissions targets. In contrast, the Kyoto Protocol is a substantially harder agreement—for the Annex 1 parties" (von Stein 2008: 247).*

Nanda (1999: 321) also argues that the convention "contains only vague commitments regarding stabilization of GHGs, and no commitments regarding reductions of GHGs", leaving more precise provisions for late discussion. In contrast, the Kyoto Protocol "sets specific targets and timetables for reducing overall global emissions of GHGs".

This assessment concurs with the tool, as the difference is clear: while the UNFCCCC receives a score of 0.77, the Kyoto Protocol scores 0.35, which is significantly lower. I can therefore conclude that the comparison made by most experts supports the results of this tool. When looking at what researchers say about each agreement individually, the assessments do not overlap quite so neatly, but it is important to keep in mind that the tool was specifically built to compare vagueness, not to assess single documents or paragraphs absolutely. I will go into more detail on this issue when discussing limitations and challenges of the tool in section 6.5. In any case, as the quotes above already show, the UNFCCC is considered as fairly vague. In my analysis, it is the third most precise document. However, as also shown in section 6.5, the categorization heavily depends on what the agreement is compared with – it is the least precise of the legally binding documents, and arguably that is the correct group for comparison. On the side of the Kyoto Protocol, it is rarely assessed as to its vagueness on its own.

The most challenging case for my analysis is undoubtedly the Copenhagen Agreement. Even though it was a mere Ministerial Declaration, which typically do not generate much interest, it is a much discussed and highly controversial document[45], both in academia and in the general news. And most assessments concur that it is vague[46]. For example, Dimitrov sums up what a lot of researchers imply:

> *"The Copenhagen Accord is an agreement among heads of governments vaguely alluding to limiting temperature rise to 2°C" (Dimitrov 2010: 21).*

[45] See for example Hofmann 2011: 5, or Saleh 2012: 24.
[46] See for example Michaelowa 2010: 2, and Bodansky 2010: 235.

In particular, provisions on adaptation funding are said to remain vague (Falkner et al. 2010: 253). When applying the vagueness score method presented here, however, with a score of 0.92 it is the most precise of all declarations and even quite precise when compared to COP or RevCon Decisions. This is quite a glaring discrepancy. However, there are several arguments on why that strengthens rather than weakens the case for the application of my tool. The first one is simply about intersubjective assessment. In addition to being perceived as vague, the Copenhagen agreement certainly was one of the agreements inciting the most disappointment in both civil society and the academic community.

> *"Rarely has as much anticipation accompanied an international meeting than swirled around the 15th Conference of the Parties of the United Nations Framework Convention on Climate Change" (Hunter 2010: 4).*

Because actors had hoped for a more binding and comprehensive agreement (see also Bodansky 2010: 230), it is easy to imagine why they would also perceive the accord as vague. It is the very fallacy this study sets out to address: Only because an agreement is (perceived to be) ineffective, that does not necessarily mean that it is vague in its wording. Especially when the agreement differs from expectations it is important to have a more objective tool for assessment at one's disposal.

Another reason for the discrepancy has its roots in the actual use of words, and highlights the importance of context for the application of the tool. The language used in the Copenhagen Agreement bears much more resemblance to legally binding documents like the UNFCCC or the Kyoto Protocol than it does to other Ministerial declarations. However, given its legal status, much of the language that indicates precision in legally binding documents (and is mostly reserved for these) may actually not indicate any such thing in declarations. The best example here are verbs like should and shall. There are several possible explanations for this: the conference it was concluded at was originally intended to produce a legally binding document, so it is not inconceivable that some of the language was kept the same even though it is unusual for a declaration. Given the careful thought that tends to be given to language in legal contexts it is also possible to doubt this explanation. There is another possibility: Knowing that a legally binding document would be expected, the diplomats may have wanted it to look like as much like such an agreement as possible and therefore kept the language similar to one, while of course being fully aware that the precision on one level would be undercut by the missing legal status. In general, it is astonishing how closely the overall linguistic structure of the Copenhagen accord matches that of legally binding documents such as the Convention and the Kyoto Protocol and how much it differs from that of other Ministerial Declarations. This is the case especially in the

categories of hedge words and strong explicitly performative verbs as shown in graphs 19 and 22. Going more in detail beyond assessments of the entire Copenhagen accord, Roberts specifically focusses on more specific provisions on finance:

> *"The accord fails on peaking, emissions reductions, process, and nearly every other justice principle. There seemed to be one slightly brighter spot: on adaptation finance, where the Copenhagen Accord included what seemed to be two clear and fairly ambitious promises. The finance offer was $US 30 billion ''Fast Start Finance'' over 2010-2012, ramping up to US $100 billion per year by 2002. However even these seemingly straightforward promises have led to major debates after Copenhagen because their language was so unclear" (Roberts 2011: 778-9).*

Even though Roberts is not exactly unambiguous himself, I interpret his statement to mean that the language on adaptation finance in the Copenhagen Accord is vague. The two sentences he is referring to both score 1, which is slightly vaguer than the overall score of the agreement of 0.86. This result may be seen as a similarity in trend, however it is necessary to tread carefully in this specific instance: it is not at all clear that Roberts is actually finding the sentences vaguer than the rest of the agreement – in fact, the opposite could be argued as well. In effect, Robert's analysis, and his reference to 'seemingly clear' language highlights the difficulties in differentiating lack of ambition and vagueness in legal agreements and the feedback loop that exists between agreements being debated and not adopted and them being understood as vague or unclear.

Even though there are some discrepancies, in general the tool matches other, more qualitative analyses quite closely. Even and especially at the level of individual articles and paragraphs, the tool can provide similar results to qualitative analyses. This alone is a point in its favour: the discrepancies which there are might well be worth the trade-off for comprehensive and quick data when compared to the detailed study of single documents, depending on what questions it is intended to answer. However, even the discrepancies give valuable insights: First, the results confirm the intuition that text-based comparisons of vagueness are most valuable when drawn between documents of the same legal status. Second, the tool highlights both the importance and practical difficulties in differentiating between linguistic vagueness – vague language – and legal bindingness. Furthermore, the fact that the largest discrepancies in results are found the context of the Copenhagen Accord, which has elicited much disappointment, emphasizes the value of having an intersubjective tool at one's disposal, because it can ensure that these kinds of emotions will not affect the assessment.

6.5 Assessing the tool: discussion

The tool as presented here has several advantages. It is specifically designed to be used on international legal agreements, taking into account much of the peculiarities of legal language and based on the considerations and opinions of experts in IL. It is designed to be able to process large amounts of legal text in little time, but also allows more detailed analysis of specific sections or provisions, even going down to the sentence level. It provides a comparison that is unbiased with personal opinion on any one agreement, but can easily be combined with a more qualitative analysis of the text. As the previous section (6.4) has shown, it has an overall good reliability when compared to expert opinion on the assessed agreements. I would therefore argue that the tool fulfils the purpose it was built for: providing a method to systematically compare the vagueness of international legal agreements.

Nevertheless, the previous chapter has also shown a number of weaknesses of the tool. Some of these are limitations inherent in the approach and simply have to be weighed against its advantages on a case-by-case basis, but quite a few of them arise due to the innovative character of the method and simply provide challenges that could be improved upon in further work with the tool. In this section, I elaborate on these issues, going from general to more specific and, when appropriate, provide some first thoughts for future improvement.

Small sample size

There are several considerations to be taken into account before actually using this tool to assess international agreements. The first and most obvious is the small scale and exploratory nature of the initial survey, which limits the predictive possibilities somewhat. While most of the independent assessment by other researchers confirms that the tool works well, it is likely that a more fine-tuned tool could be gained by conducting a larger scale survey and possibly conducting direct interviews with experts in order to profit even more from their comments.

The small sample size especially affects verbs, because many verbs used as operative verbs in international agreements simply could not be accounted for, thus scoring this category for the sentences as 0. For a lot of them, this may be the correct choice – quite a few verbs that were tested for in this study do not lead to a sentence being perceived as either vaguer or more precise (see section 5.3.1). But there may still be other, albeit less frequently used, verbs that do change the perception of vagueness of a sentence. The same

issue arises in the category of hedge words: not all hedge words used in the tool could be tested for in the survey, and the tool itself uses only the ten most frequently appearing hedge words. This allows a fair approximation, because all documents subjected to the tool are measured by the same imperfect standard, but the accuracy could certainly be further improved upon. To a lesser extent, this fine-tuning is also needed for determiners: not all could be tested for in the survey. The problem is less urgent only because this feature consists of a clearly defined group of verbs.

Application to other fields

One major concern is whether the results can be applied to regimes other than the ones which the sentences were taken from, because results could vary according to subject matter, or the language of specific regimes could be contained enough to feature their own specialized (sub)language. The fact that no such variance could be observed here suggests that a transfer to other areas of IL is possible, though the fact remains that this survey was conducted with sample sentences from only two regimes. Further study is needed before any definite claims on this issue can be made.

It is expected that adjustments would have to be made before results could be applied to fields that are not IL, because the indicators – especially the verbs – are dependent on this context and also the survey participants were experts only in this field. Nevertheless, the results shown here could probably be taken into account to facilitate locating the starting point of such research.

Comparison

Due to the complex nature of vagueness all the values should only be taken as comparative, not absolute. The tool assigns a vagueness score to each sentence or agreement, but this score by itself has little informative value. Only by comparing two or more sentences, sections, or agreements do the numbers become meaningful. In the previous section I have shown that in some instances, relative comparisons between agreements from other researchers are in accordance with the results of the tool, while individual assessments of agreements may not be quite as congruent (see discussion on Framework Convention and Kyoto Protocol in section 6.3). These differences may occur because the comparison group an author uses is not made explicit.

The scores produced by the tool should enable someone to determine whether one agreement is vaguer than another according to these criteria, not determine an absolute

vagueness score. It is for this reason that determining the appropriate comparison group is vital to achieve meaningful results. In this thesis, I have used several comparison groups to test the tool on: documents of one regime can be compared to one another, or the entirety of documents under one regime compared to another regime. Likewise, documents types of differing legal status can be compared, or, probably more usefully in most cases, several documents of the same legal type can be compared to each other. Another important consideration for comparing documents is which part(s) of the text to include. Preambles are probably best compared to other preambles. This concern is a very general one, applying to all research which includes comparisons. It is nevertheless crucial to keep in mind when applying this tool.

Annexes, titles and questions

Tying in to the general issue of comparison groups above, there is a special consideration to be made in this particular study about annexes. Under the Climate Convention,

> *"Annexes to the Convention shall form an integral part thereof and, unless otherwise expressly provided, a reference to the Convention constitutes at the same time a reference to any annexes thereto" (UNFCCC 1992: Art. 16).*

Nevertheless, annexes do constitute a separate textual category from the main body of text of an agreement, and have been treated as such in this study (see comments on categorizing above). More challenging is the fact that a big component of the annexes are tables with no traditional sentence structures at all or questions, which also cannot be measured by the same yardstick as legal sentences. Since the unit of investigation used in this study is the sentence, it follows that language not arranged in sentences, as is often the case in tables or titles, cannot be analysed by this tool. Technically, most questions would score as very vague, but it is not the kind of legal language which the subjects of the survey were asked about. Also vagueness and performativity of questions is a very different matter from those of a statement. For example, the performative force of 'should' when used in a question is much less clear than when used in a declarative sentence. This is usually not a problem when studying vagueness in international agreements, because international agreements do not normally contain questions of any kind. The sole exception, as pointed out here are annexes, and even then questions only feature in certain cases. It is mainly for this reason that annexes, and with them questions and tables, have been excluded from the analysis in this study. If it was truly necessary to include them, further research on the vagueness of questions or on single words/parts of sentences would certainly be necessary.

References

Referencing is a very important device in legal texts, yet the tool can give no indication as to whether the paragraph referred to is vague or not – which could significantly change the vagueness of the sentence itself. However, this is not so much a shortcoming of the method as a systematically excluded category. If a specific reference were of particular importance to an agreement, it would be perfectly possible to analyse it with the same method and include the results.

Explicitly performative verbs

One problem that became apparent thanks to the detailed analysis above is that the word 'will', which counts as a verb with strong performative force, is not always the operative verb in a sentence. It can be used to simply indicate an assumed or hoped for future, not to impose any kind of obligation on parties. This leads to a slight overemphasis of the precision of this sentence in this tool. In this study, this is only the case in a small minority of sentences, and the level of aggregation means there are only extremely small changes in overall results. This is an area where there is an interesting difference between the two regimes: Under the NPT regime, the only document featuring 'will' as an operative verb is the original document from 1968. As declarations under this regime do not feature any strong performative verbs at all, the only instances in which the verb 'will' appears is as an indicator of future tense, not in the legal sense of the word. This is not the case under the UNFCCC: here, strong explicitly performative words are used in all document types, albeit in different numbers.

Even though it only concerns a small quantity of cases, the issue should still be taken seriously, with a view to improving the accuracy in studies with lower quantities of text. One remedy would be to only count the first verb used in each sentence. This would entail quite complicated operations of computational linguistics, and in all likelihood would not change the results all that much. Another, more easily feasible solution to achieve more accuracy would be to fine-tune the method so that the verb 'will' is excluded from the analysis when it is concerned with Final Declarations, because it yields too many false positives to be ultimately useful. Since the total number of times the word 'will' appeared in the 90,000-word corpus is 111, it is of course also an option to check for incorrect uses manually.

Errors in automated word detection

In rare cases, the automated detection of words to determine vagueness points counts words that do not actually belong in that category, which can cause it to attribute points

wrongly. I have shown one of these cases in the context of the Bonn Agreement, where the word 'available' was used as a predicate ('make funds available') instead of as an adjective ('any available funds'). While the word is the same, and thus is picked up by the automated tool, it is only a hedge word in the second case. There are a few possibilities on how to remove these already rare instances, but there is not yet a perfect solution. Some quantitative text analysis tools allow users to take into account surrounding words. In this case, one could experiment with counting 'available' as a hedge word only if there it is not accompanied by 'make' or 'made' at a distance of, for example, ten words or less. However, if this would lower the error occurrence or overcorrect and create false negatives instead is an open question.

Very short texts

Many of the issues presented here become less important when assessing large quantities of text, because the results tend to average out. With very short documents, what would otherwise be small inaccuracies can lead to significant errors in the results. Fortunately, short texts are also the least in need of the kind of quantitative analysis presented here, because their very shortness allows them to be read in detail. A dual approach is therefore recommended: using the tool only as a first approximation, and additionally checking the results of the tool on whether or not they make sense for each of the sentences involved.

7 Conclusion and outlook

IR scholars have made a number of assertions as to the provenance and effects of vagueness in international agreements. They have contended that vague agreements are less likely to be implemented, less legalised, a result of high transaction cost, party interest, increased complexity, or insurmountable differences among parties, and that vagueness declines over time in any given regime. I have argued that to be able to know whether and which of these claims are true, it is necessary to be able to systematically measure and compare vagueness in international agreements.

The main objective of this thesis was therefore to develop and test a tool that would do so. This required establishing indicators of vagueness applicable to international agreements, which did not previously exist. As a first step, I assembled possible indicators by studying literature on vagueness in general language as well as the few sources concerned with how to detect vagueness in international legal text. I then conducted a survey among

experts in IL in order to find out which characteristics of a sentence would make them more likely to consider it vague. Based on the data generated by this survey, I conducted several statistical tests to find out which possible indicators of vagueness are relevant in the field of IL. Having thus established four indicators of vagueness in international agreements, I devised an automated tool which aggregates the sentences' vagueness over larger amounts of text. As an explorative case study, seventeen international agreements from the UNFCCC and the NPT regimes. In a last step, I compared the results provided by the newly developed tool to independent, qualitative assessments of researchers.

This process led to four sets of results: those of the initial survey, the ones gained from of the data analysis, the results of the application of the tool, and of the comparison of the tool with other assessments.

The survey provided a list of thirty sentences ordered by their perceived vagueness (see table 1), as well as additional insights by experts in IL into what features of the sentences made them see them as vague. Survey respondents claimed that verbs expressing obligations, hedge words, sentence length, referencing, generality, and measurability affected how vague they perceived a sentence to be.

The results of the data analysis showed that four out of eight tested hypotheses proved to be statistically significant. Sentences in international agreements are perceived by experts to be vaguer when they contain hedge words, a low number of determiners, or weak explicitly performative verbs, such as 'would', 'could', or 'recognizes', and more precise when they contain a strong explicitly performative verbs such as 'shall' or 'will'.

Applying the tool to seventeen international agreements generated the most extensive set of results. I list just some of the most pertinent results here: The tool allows for aggregate analysis – up to the level of entire treaty regimes – and can also provide insights to sentence-by-sentence analysis. The range of vagueness in the analysed agreements is diverse, but the overall vagueness scores of the two regimes are very similar. While decisions under the NPT regime have very similar vagueness scores to each other, this is not the case for decisions under the UNFCCC. The NPT and the Kyoto Protocol have the lowest vagueness scores, while the NPT RevCon decisions of 1995 and the Delhi Ministerial Declaration (2002) have the highest. Vagueness does not appear to develop in any consistent way over time. Legally binding documents are generally more precise than decisions or declarations.

Finally, the last set of findings supports the usefulness of the tool, as the results provided by the tool concur with the overall assessments experts have made of these documents.

The answer to my research question *'How can the vagueness of international agreements be systematically compared?* is therefore: It is possible to model the vagueness in international agreements with reasonable accuracy by using sentences as the basic scoring unit, and hedge words, determiners and explicitly performative verbs as indicators for vagueness in sentences.

While the focus of this thesis was certainly on the gathering and analysis of empirical data, the results do have some theoretical implications, which are worth noting.

In the first chapter, I discussed the indeterminacy debate – i.e. the question whether or not the law was indeterminate, and if so, whether or not this meant that it had any chance at legitimacy, or even exists at all. I argued for a middle ground: There is some vagueness in the law, but vagueness is not absolute. Legal texts differ in how vague they are, and containing vagueness does not automatically render a law meaningless. This premise is one of the underlying assumptions of this thesis. Looking for indicators of vagueness postulates at the outset that these indicators can be more or less present, therefore vagueness in legal text must vary. Nevertheless, actually finding these indicators, and showing that they indeed are present to varying degrees, does support the claim that there is a variance of vagueness (see Figure 16). Even though the study was set up under this assumption, and may therefore have been biased in its favour, the results could well have shown no significant differences at all.

In section 3.2.4, I briefly discussed the relationship between legal and vagueness. The performativity of legal language (see section 4.2.2) suggests that vagueness does not necessarily only affect the locutionary content of a legal text, but may also be found in its performative force. This in turn is closely related to legal bindingness, where the latter it is expressed directly in the text. The respondents of the survey confirmed this in their comments: most tied vagueness or precision to explicitly performative verbs (see chapter 5.2.2). This relationship is far from exhaustively researched, but this thesis shows that the two dimensions are related and not easy to separate in practice, in part because they can be tied to the same words: for example, 'shall' is used to indicate a certain legal status, and IL experts also consider sentences containing this verbs to be more precise than others.

In chapter 4, I have asked if some of the indicators used to identify vagueness in general language also apply to international legal agreements. According to my findings, hedge words are an indicator that is pertinent to both types of language. However, in the legal context much more importance has to be given to verbs than is otherwise the case. The

relationship between vagueness and determiners requires further research in both IL and linguistics.

The results of this study also provide a point of departure for scrutinizing the numerous hypotheses made in the literature concerning the causes and effects of vagueness in IR. In a first instance, the claim that regimes become more precise over time does not appear to be the case in the two regimes studied here. Even when only taking into account agreements of similar legal status, the agreements which are part of the NPT regime remain remarkably steady in their vagueness scores over time, while the agreements under the UNFCCC vary with no apparent connection to the time at which they were negotiated. It has to be noted, however, that the text of agreements under the NPT regime got longer in a fairly steady fashion over time. This in itself does not, of course, indicate precision, but it may signify that matters have been dealt with in more detail.[47]

The specific conditions under which each agreement was negotiated might also shed some light on the reasons some of the vagueness or precision is present. At the very least, the variance present in this analysis supports the claim that at least some of the vagueness in agreements is intentional and could be avoided if that was wanted by the negotiators. This finding could be the basis for further study into why negotiators use vague language in specific cases.

I will end with some suggestions for further research based on these findings.

First, as discussed in section 6.5, the tool itself as well as the automated word detection used by it should be refined further to generate even more accurate results, overcoming challenges such as the fairly small sample size of the original survey, inaccuracies in very short texts and the application of vagueness scores to headings or questions.

Second, the tool should be tested further. To do so, it should be applied to other legal regimes and international agreements. If the results of these additional tests do not fit expert opinion, or if the detailed analysis reveals other problems (such as a significantly different vocabulary used in international treaties depending on the subject matter they regulate), research should be conducted to adjust the tool accordingly.

A third area for further study would be using the results presented in chapter 6.3 in order to study both the NPT regime and the UNFCCC regime from a different angle. For example, combining an in-depth study of the overall reception of the Copenhagen Accord with the

[47] Of course, the other explanation is simply an expansion of the scope of the regime. It is likely that both explanations are true to some extent.

vagueness scores of the sentences may provide deeper insights into the controversy surrounding this agreement. Because my analysis was focussed on the tool itself, implications of vagueness for the efficacy of the actual documents under analysis were necessarily studied in less detail.

A fourth area for further research is in the area of linguistics: research could be done on whether or not the results of this study in a specialized language are potentially applicable to different specialized languages, or if the results could even refine the study of vagueness in everyday language.

Finally, the tool should of course be used for its intended goal: To systematically assess whether the causes and effects of vagueness hypothesized by IR scholars are true.

8 Bibliography

8.1 Corpus

NPT 1968: The Treaty on the Non-Proliferation of Nuclear Weapons. 21 U.S.T. 483, 729 U.N.T.S. 161.

NPT 1975: Final Declaration of the Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons. In: Review Conference of the Parties to the Treaty on the Non-Proliferation of nuclear weapons, Final Document, Annex 1.

NPT 1985: Final Declaration. In: Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons, NPT/CONF.UI/64/l. ANNEX I: FINAL DECLARATION.

NPT 1995: Decisions and Resolution Adopted by the Conference. In: 1995 Review and Extension Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons, NPT/CONF.1995/32 (Part I).

NPT 2000: Review of the Operation of the Treaty, Taking into Account the Decisions and the Resolution Adopted by the 1995 Review and Extension Conference. In: 2000 Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons Final Document, NPT/CONF.2000/28 (Parts I and II).

NPT 2010: Review of the operation of the Treaty, as provided for in its article VIII (3), taking into account the decisions and the resolution adopted by the 1995 Review and Extension Conference and the Final Document of the 2000 Review Conference Conclusions and recommendations for follow-on actions. In: 2010 Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons Final Document, NPT/CONF.2010/50 (Vol. I).

UNFCCC 1992: United Nations Framework Convention on Climate Change, S. Treaty Doc No. 102-38, 1771 U.N.T.S. 107.

UNFCCC 1995: The Berlin Mandate: Review of the adequacy of Article 4, paragraph 2 (a) and (b), of the Convention, including proposals related to a protocol and decisions on follow-up. In: FCCC/CP/1995/7/Add.1, 4-6.

UNFCCC 1996: The Geneva Ministerial Declaration. In: FCCC/CP/1996/15/Add.1, Annex.

UNFCCC 1997: Kyoto Protocol to the United Nations Framework Convention on Climate Change, U.N. Doc FCCC/CP/1997/7/Add.1, 37 I.L.M. 22.

UNFCCC 1998: The Buenos Aires Plan of Action. In: FCCC/CP/1998/16/Add.1, 4-40.

UNFCCC 2000: The Bonn Agreements on the implementation of the Buenos Aires Plan of Action. In: FCCC/CP/2001/5, 36-50.

UNFCCC 2001: The Marrakesh Ministerial Declaration. In: FCCC/CP/2001/13/Add.1, 4-5.

UNFCCC 2002: The Delhi Ministerial Declaration on Climate Change and Sustainable Development. In: FCCC/CP/2002/7/Add.1, 3-5.

UNFCCC 2007: Bali Action Plan. In: FCCC/CP/2007/6/Add.1*, 3-8.

UNFCCC 2009: Copenhagen Accord. In: FCCC/CP/2009/11/Add.1, 4-9.

UNFCCC 2010: Cancun Agreement. In: FCCC/CP/2010/7/Add.1, 2-31.

8.2 Literature

Aarts, Bas/Chalker, Sylvia/Weiner, Edmund 2014: The Oxford Dictionary of English Grammar. Oxford, Oxford University Press.

Abbott, Frederick M. 2000: NAFTA and the Legalization of World Politics: a Case Study. In: International Organization 54:03, 519-547.

Abbott, Kenneth W./Keohane, Robert O./Moravcsik, Andrew/Slaughter, Anne-Marie/Snidal, Duncan 2000: The Concept of Legalization. In: International Organization 54:3, 401-419.

Abbott, Kenneth W./Snidal, Duncan 2000: Hard and Soft Law in International Governance. In: International Organization 54:3, 421-456.

Abizadeh, Arash 2005: Does Collective Identity Presuppose an Other? On the Alleged Incoherence of Global Solidarity. In: American Political Science Review 99:01, 45-60.

Alter, Karen J./Meunier, Sophie 2009: The Politics of International Regime Complexity. In: Perspectives on politics 7:01, 13-24.

Alvarez, José E. 2006: International Organizations: Then and Now. In: American Journal of International Law 100:2, 324-347.

Andresen, Steinar/Agrawala, Shardul 2002: Leaders, Pushers and Laggards in the Making of the Climate Regime. In: Global Environmental Change 12:1, 41-51.

Aust, Anthony 2000: Modem Treaty Law and Practice. Cambridge, Cambridge University Press.

Austin, John L. 1962: How to Do Things with Words. Cambridge, Harvard University Press.

Beaulac, Stéphane 2004: The Power of Language in the Making of International Law: the Word Sovereignty in Bodin and Vattel and the Myth of Westphalia. Leiden, Brill.

Benford, Robert D./Snow, David A. 2000: Framing Processes and Social Movements: An Overview and Assessment. In: Annual Review of Sociology 26, 611-639.

Benson, Robert W./Kessler, Joan B. 1986: Legalese v. Plain English: An Empirical Study of Persuasion and Credibility in Appellate Brief Writing. In: Loyola of Los Angeles Law Review 20:2, 301-322.

Bernauer, Thomas 1995: The Effect of International Environmental Institutions: How We Might Learn More. In: International Organization 49:2, 351-377.

Besson, Samantha 2005: The Morality of Conflict: Reasonable Disagreement and the Law. Oxford, Hart.

Bhatia, Vijay Kumar/Candlin, Christopher N./Engberg, Jan/Trosborg, Anna, Eds.) 2003: Multilingual and Multicultural Contexts of Legislation: an International Perspective. Bern, Peter Lang.

Bodansky, Daniel 2010: The Copenhagen Climate Change Conference: A Postmortem. In: The American Journal of International Law 104:2, 230-240.

Bosch, Peter 1983: Vagueness is Context-Dependence. A Solution to the Sorites Paradox. In: Ballmer, T. T./Pinkal, M.: Approaching Vagueness. Amsterdam, North Holland: 189-210.

Bradburn, Norman M./Miles, Carrie 1979: Vague Quantifiers. In: Public Opinion Quarterly 43:1, 92-101.

Brunner, Pascale 2011: La notion du 'vague' en langue et en discours: une étude comparative entre le francais et l'allemand, Paris 3.

Burns, Linda Claire 1991: Vagueness: an Investigation into Natural Languages and the Sorites Paradox. Dordrecht, Kluwer Academic Publishers.

Byers, Michael 2004: Agreeing to Disagree: Security Council Resolution 1441 and Intentional Ambiguity. In: Global Governance: A Review of Multilateralism and International Organizations 10:2, 165-186.

Camp Keith, Linda 1999: The United Nations International Covenant on Civil and Political Rights: Does It Make a Difference in Human Rights Behavior? In: Journal of Peace Research 36:1, 95-118.

Caraballo, Sharon A/Charniak, Eugene 1999: Determining the Specificity of Nouns from Text. Proceedings of the 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora. Providence.

Castellví, M Teresa Cabré 1999: Terminology: Theory, Methods and Applications. Amsterdam, John Benjamins.

Chandra, Ashish/Holt, Gary A./Arora, Gurdial 2003: A New Measure to Determine the Ambiguity Level of Qualitative Statements. In: Clinical Research and Regulatory Affairs 20:1, 47-57.

Channell, Joanna 1994: Vague Language. Oxford, Oxford University Press.

Charnovitz, Steve 2003: Trade and Climate: Potential Conflicts and Synergies. Beyond Kyoto: Advancing the International Effort against Climate Change, Pew Center on Global Climate Change.

Charrow, Veda R./Crandall, Jo Ann/Charrow, Robert P. 1982: Characteristics and Functions of Legal Language. In: Kittredge, R./Lehrberger, J.: Sublanguage: Studies of Language in Restricted Semantic Domains. Berlin, de Gruyter: 175-190.

Chayes, Abram/Chayes, Antonia Handler 1995: The New Sovereignty: Compliance with International Regulatory Agreements. Cambridge, Harvard University Press.

Cheng, Winnie/Warren, Martin 2003: Indirectness, Inexplicitness and Vagueness Made Clearer. In: Pragmatics 13:3/4, 381-400.

Chierchia, Gennaro 2010: Mass Nouns, Vagueness and Semantic Variation. In: Synthese 174:1, 99-149.

CITES 1973: Convention on International Trade in Endangered Species of Wild Fauna and Flora, 27 UST 1087; TIAS 8249; 993 UNTS 243.

Converse, Jean M./Presser, Stanley 1986: Survey Questions: Handcrafting the Standardized Questionnaire. Thousand Oaks, Sage.

Coventry, Kenny R./Cangelosi, Angelo/Newstead, Stephen E./Bugmann, Davi 2010: Talking about Quantities in Space: Vague Quantifiers, Context and Similarity. In: Language and Cognition 2:2, 221-241.

Craig, Campbell/Ruzicka, Jan 2013: The Nonproliferation Complex. In: Ethics & International Affairs 27:03, 329-348.

Crompton, Peter 1997: Hedging in Academic Writing: Some Theoretical Problems. In: English for Specific Purposes 16:4, 271-287.

Crystal, David/Davy, Derek 1975: Advanced Conversational English. London, Longman

Curzon, Leslie Basil/Richards, Paul 2007: The Longman Dictionary of Law. Upper Saddle River, Pearson Education.

Cutting, Joan 2007a: Introduction to Vague Language Explored. In: Cutting, J.: Vague Language Explored. Basingstoke, Palgrave MacMillan.

Cutting, Joan, (Ed.) 2007b: Vague Language Explored. Basingstoke, Palgrave MacMillan.

Cutting, Joan 2012: Vague Language in Conference Abstracts. In: Journal of English for Academic Purposes 11:4, 283-293.

D'Acquisto, Germana/D'Avanzo, Stefania 2009: The Role of "Shall" and "Should" in Two International Treaties. In: Critical Approaches to Discourse Analysis across Disciplines 3:1, 36-45.

Dahlman, Carl J. 1979: The Problem of Externality. In: The Journal of Law & Economics 22:1, 141-162.

D'Amato, Anthony 1984: Is International Law Really Law. In: Northwestern University Law Review 79, 1293.

D'Amato, Anthony 1990: Aspects of Deconstruction: Refuting Indeterminacy with One Bold Thought. In: Northwestern University Law Review 85:113, 10-35.

Deitelhoff, Nicole 2006: Überzeugung in der Politik. Grundzüge einer Diskurstheorie internationalen Regierens. Frankfurt am Main, Suhrkamp.

Deitelhoff, Nicole/Wallbott, Linda 2012: Beyond Soft Balancing: Small States and coalition-Building in the ICC and Climate Negotiations. In: Cambridge Review of International Affairs 25:3, 345-366.

Della, Porta D./Diani, Mario 2006: Social Movements: An introduction. Malden, Blackwell.

Depledge, Joanna 2000: Tracing the Origins of the Kyoto Protocol: An Article-by-Article Textual History. UNFCCC Technical Papers. New York, UNFCCC.

Di Carlo, Giuseppina S. 2011: Indeterminacy and Vagueness in UN Resolutions Relating to the Second Gulf War. In: International Journal of Humanities and Social Sciences 1:21, 46-57.

Dillman, Don A./Sinclair, Michael D./Clark, Jon R. 1993: Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys. In: Public Opinion Quarterly 57:3, 289-304.

Dimitrov, Radoslav S. 2010: Inside UN Climate Change Negotiations: The Copenhagen Conference. In: Review of Policy Research 27:6, 795-821.

Drave, Neil 2001: Vaguely Speaking: a Corpus Approach to Vague Language in Intercultural Conversations. In: Language and Computers 36:1, 25-40.

Dubois, Sylvie 1992: Extension Particles, etc. In: Language Variation and Change 4:02, 179-203.

Dupont, Christophe 1996: Negotiation as Coalition Building. In: International Negotiation 1:1, 47-64.

Dupuy, Pierre-Marie 1998: The Danger of Fragmentation or Unification of the International Legal System and the International Court of Justice. In: New York University Journal International Law and Politics 31, 791-807.

Dworkin, Ronald 1963: Judicial Discretion. In: The Journal of Philosophy 60:21, 624-638.

Dworkin, Ronald 1975: Hard Cases. In: Harvard Law Review 88:6, 1057-1109.

Easterbrook, Frank H. 1984: Legal Interpretation and the Power of the Judiciary. In: Harvard Journal of Law and Public Policy 7, 87-99.

Eklund, Matti 2005: What Vagueness Consists in. In: Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 125:1, 27-60.

Elsig, Manfred 2011: Principal–Agent Theory and the World Trade Organization: Complex Agency and 'Missing Delegation'. In: European Journal of International Relations 17:3, 495-517.

Endicott, Timothy A. O. 2001: Vagueness in Law. Oxford, Oxford University Press.

Endicott, Timothy A. O. 2005: The Value of Vagueness. In: Bhatia, V. K./Engberg, J./Gotti, M./Heller, D.: Vagueness in Normative Texts. Bern, Peter Lang: 27-48.

Falkner, Robert/Stephan, Hannes/Vogler, John 2010: International Climate Policy after Copenhagen: Towards a 'Building Blocks' Approach. In: Global Policy 1:3, 252-262.

Fearon, James D. 1998: Bargaining, Enforcement, and International Cooperation. In: International Organization 52:02, 269-305.

FIELD, Foundation for International Environmental Law and Development 2006: Guide for Negotiators of Multilateral Environmental Agreements. Nairobi, UNEP/Earthprint.

Fields, Jeffrey/Enia, Jason S. 2009: The Health of the Nuclear Nonproliferation Regime. In: The Nonproliferation Review 16:2, 173-196.

Fine, Kit 1975: Vagueness, Truth and Logic. In: Synthese 30:3/4, 265-300.

Fink, Arlene 2003: The Survey Handbook: 2nd Edition. Thousand Oaks, Sage.

Finnemore, Martha/Sikkink, Kathryn 1998: International Norm Dynamics and Political Change. In: International Organization 52:4, 887-917.

Fischer-Lescano, Andreas/Teubner, Gunther 2004: Regime-Collisions: the Vain Search for Legal Unity in the Fragmentation of Global Law. In: Michigan Journal of International Law 25:4, 999-1046.

Fiss, Peer C./Hirsch, Paul M. 2005: The Discourse of Globalization: Framing and Sensemaking of an Emerging Concept. In: American Sociological Review 70:1, 29-52.

Fjeld, Ruth Vatvedt 2005: The Lexical Semantics of Vague Adjectives in Normative Texts. In: Bhatia, V. K./Engberg, J./Gotti, M./Heller, D.: Vagueness in Normative Texts. Bern, Peter Lang: 159-172.

Flohr, Annegret 2014: Self-Regulation and Legalization: Making Global Rules for Banks and Corporations. Basingstoke, Palgrave Macmillan.

Ford, Christopher A. 2007: Debating Disarmament. In: The Nonproliferation Review 14:3, 401-428.

Franck, Thomas M. 1988: Legitimacy in the International System. In: The American Journal of International Law 82:4, 705-759.

Franck, Thomas M. 1990: The Power of Legitimacy among Nations. Oxford, Oxford University Press.

Gamson, William A/Fireman, Bruce/Rytina, Steven 1982: Encounters with Unjust Authority. Belmont, Dorsey Press.

Garner, Bryan A./Black, Henry Campbell 2004: Black's Law Dictionary. St. Paul, Thomson/West.

GATT 1994: General Agreement on Tariffs and Trade, 1867 UNTS 187; 33 ILM 1153.

Gilhooly, Ken J./Logie, Robert H. 1980: Age-of-Acquisition, Imagery, Concreteness, Familiarity, and Ambiguity Measures for 1,944 Words. In: Behavior Research Methods & Instrumentation 12:4, 395-427.

Gingrich, Paul 1992: Introductory Statistics for the Social Sciences. Department of Sociology and Social Sciences, University of Regina.

Ginsburg, Tom/Shaffer, Gregory 2010: How Does Inter National Law Work? In: Cane, P./Kritzer, H. M.: The Oxford Handbook of Empirical Legal Research. New York, Oxford University Press.

Goffman, Erving 1974: Frame Analysis: An Essay on the Organization of Experience. Cambridge, Harvard University Press.

Goguen, J. A. 1969: The Logic of Inexact Concepts. In: Synthese 19:3-4, 325-373.

Goldsmith, Jack L./Posner, Eric A. 1999: A Theory of Customary International Law. In: The University of Chicago Law Review 66:4, 1113-1177.

Goldstein, Judith/Kahler, Miles/Keohane, Robert O./Slaughter, Anne-Marie 2000: Introduction: Legalization and World Politics. In: International Organization 54:03, 385-399.

Goodman, Ryan/Jinks, Derek 2004: How to Influence States: Socialization and International Human Rights Law. In: Duke Law Journal 54:3, 621-703.

Greenawalt, Kent 1992: Law and Objectivity. New York, Oxford University Press.

Greenawalt, Kent 2001: Vagueness and Judicial Responses to Legal Indeterminacy. In: Legal Theory 7:04, 433-445.

Greenhill, Brian 2008: Recognition and Collective Identity Formation in International Politics. In: European Journal of International Relations 14:2, 343-368.

Gumperz, John Joseph 1971: The Speech Community. In: Dil, A. S.: Language in Social Groups. Stanford, Stanford University Press. 3: 114-128.

Guzman, Andrew T. 2002: A Compliance-Based Theory of International Law. In: California Law Review 90:6, 1823-1887.

Haas, Peter M. 1989: Do Regimes Matter? Epistemic Communities and Mediterranean Pollution Control. In: International Organization 43:03, 377-403.

Hart, Herbert L. A. 1961: The Concept of Law. Oxford, Clarendon Press.

Hasenclever, Andreas/Mayer, Peter/Rittberger, Volker 1996: Interests, Power, Knowledge: The Study of International Regimes. In: Mershon International Studies Review 40:2, 177-228.

Hasenclever, Andreas/Mayer, Peter/Rittberger, Volker 1997: Theories of International Regimes. Cambridge, Cambridge University Press.

Heath, Jeffrey 1979: Diffusional Linguistics in Australia: Problems and Prospects. In: Wurm, S. A.: Australian Linguistics Studies. Canberra, Department of Linguistics, Australian National University 395-418.

Hilgartner, Stephen/Bell, Richard C./O'Connor, Rory 1982: Nukespeak: Nuclear Language, Visions, and Mindset. San Francisco, Sierra Club Books.

Hill, Claire A. 2001: Why Contracts Are Written in Legalese. In: Chicago-Kent Law Review 77:1, 59-85.

Hisschemöller, Matthijs/Gupta, Joyeeta 1999: Problem-Solving through International Environmental Agreements: The Issue of Regime Effectiveness. In: International Political Science Review 20:2, 151-173.

Hoffmann, Matthew J. 2011: Climate Governance at the Crossroads: Experimenting with a Global Response after Kyoto. New York, Oxford University Press.

Hunter, David 2010: Implications of the Copenhagen Accord for Global Climate Governance. In: Sustainable Development Law & Policy 10:2, 5.

Hyde, Dominic 2003: Higher-Orders of Vagueness Reinstated. In: Mind 112:446, 301-305.

Hyde, Dominic. 2014, 13.11.2016: "Sorites Paradox." The Stanford Encyclopedia of Philosophy, from https://plato.stanford.edu/entries/sorites-paradox/.

Jackson, John H. 1992: Status of Treaties in Domestic Legal Systems: A Policy Analysis. In: The American Journal of International Law 86:2, 310-340.

Janney, Richard W. 2002: Cotext as Context: Vague Answers in Court. In: Language & Communication 22:4, 457-475.

Johnson, Paul M. 2005, 29.11.2016: "Transaction Costs." A Glossary of Political Economy Terms, from http://www.auburn.edu/~johnspm/gloss/transaction_costs.

Jönsson, Christer 2002: Diplomacy, Bargaining and Negotiation. In: Carlsnaes, W./Risse, T./Simmons, B. A.: Handbook of International Relations. London, Sage Publications: 212-234.

Jucker, Andreas H./Smith, Sara W./Lüdge, Tanja 2003: Interactive Aspects of Vagueness in Conversation. In: Journal of Pragmatics 35:12, 1737-1769.

Kairys, David 1983: Law and Politics. In: George Washington Law Review 52:2, 243-262.

Karrer, Pierre A. 2005: Unbestimmtheit und Unvollständigkeit in Vereinbarungstexten und ihre Überwindung durch die Internationale Schiedsgerischtsbarkeit. In: Bhatia, V. K./Engberg, J./Gotti, M./Heller, D.: Vagueness in Normative Texts. Bern, Peter Lang: 97-133

Keefe, Rosanna 2000: Theories of Vagueness. Cambridge, Cambridge University Press.

Kennedy, Duncan 1997: A Critique of Adjudication: fin de siècle. Cambridge, Harvard University Press.

Keohane, Robert O. 1982: The Demand for International Regimes. In: International Organization 36:2, 325-355.

Keohane, Robert O. 1993: The Analysis of International Regimes: Towards a European-American Research Programme. In: Rittberger, V./Mayer, P.: Regime Theory and International Relations. Oxford, Clarendon Press: 23-45.

Kingsbury, Benedict 1998: The Concept of Compliance as a Function of Competing Conceptions of International Law. In: Michigan Journal of International Law 19, 345-372

Koh, Harold Hongju 1999: How Is International Human Rights Law Enforced? In: Indiana Law Journal 74:4, 1398-1417.

Koplow, David A. 1993: Parsing Good Faith: Has the United States Violated Article VI of the Nuclear Non-Proliferation Treaty? In: Wisconsin Law Review, 301.

Koskenniemi, Martti 1990: The Politics of International Law. In: European Journal of International Law 1, 4-32.

Krasner, Stephen D. 1982: Structural Causes and Regime Consequences: Regimes as Intervening Variables. In: International Organization 36:2, 185-205.

Kratochwil, Friedrich V. 2006: History, Action and Identity: Revisiting the 'Second' Great Debate and Assessing its Importance for Social Theory. In: European Journal of International Relations 12:1, 5-29.

Kress, Ken 1989: Legal Indeterminacy. In: California Law Review 77:2, 283-337.

Krosnick, Jon A. 1999: Survey Research. In: Annual Review of Psychology 50:1, 537-567.

Kumm, Mattias 2004: The Legitimacy of International Law: A Constitutionalist Framework of Analysis. In: The European Journal of International Law 15:5, 907-931.

Lakoff, George 1973: Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. In: Journal of Philosophical Logic 2:4, 458-508.

Leiter, Brian 2010: Legal Formalism and Legal Realism: What Is the Issue? In: Legal Theory 16:2, 111-133.

Levi, Judith N. 1990: The Study of Language in the Judicial Process. In: Levi, J. N./Walker, A. G.: Language in the judicial process. New York, Springer, 3-35.

Levy, Marc A./Young, Oran R./Zürn, Michael 1995: The Study of International Regimes. In: European Journal of International Relations 1:3, 267-330.

Liebwald, Doris 2013: Law's Capacity for Vagueness. In: International Journal for the Semiotics of Law - Revue internationale de Sémiotique juridique 26:2, 391-423.

Lin, Yen-Liang 2013: Vague Language and Interpersonal Communication: An Analysis of Adolescent Intercultural Conversation. In: International Journal of Society, Culture & Language 1:2, 69-81.

Lutz, Ellen L./Sikkink, Kathryn 2000: International Human Rights Law and Practice in Latin America. In: International Organization 54:3, 633-659.

Madry, Alan R. 1998: Legal Indeterminacy and the Bivalence of Legal Truth. In: Marquette Law Review 82, 581.

Maia, Belinda 2003: Ontology, Ontologies, General Language and Specialised Languages. In: CLUP, C. D. d.: Língua Portuguesa: Estruturas, Usos, e Contrastes. Porto, Centro de Linguìstica da Universidade do Porto.

Maley, Yon 1987: The Language of Legislation. In: Language in Society 16:1, 25-48.

Manor, Ruth 2006: Solving the Heap. In: Synthese 153:2, 171-186.

Mastenbroek, Willem F. G. 2002: Development of Negotiating Skills. In: Kremenyuk, V. A.: International Negotiation: Analysis, Approaches, Issues. San Francisco, Jossey-Bass: 433-454.

Mattila, Heikki E. S. 2006: Comparative Legal Linguistics. Hampshire, Ashgate.

Mayring, Philipp 2010: Qualitative Inhaltsanalyse: Grundlagen und Techniken. Weinheim, Beltz.

Mellinkoff, David 1963: The Language of the Law. Boston, Little, Brown & Co.

Michaelowa, Axel 2010: Copenhagen and the Consequences. In: Intereconomics 45:1, 2-3.

Mitchell, Ronald B. 2001: Institutional Aspects of Implementation, Compliance, and Effectiveness. In: Luterbacher, U./Sprinz, D. F.: International Relations and Global Climate Change. Cambridge, MIT Press: 221-244.

Morgenthau, Hans J. 1954: Politics among Nations. New York, Alfred A. Knopf.

Morzycki, Marcin 2015: Modification. Cambridge, Cambridge University Press.

Müller, Harald 2004: Arguing, Bargaining and All That: Communicative Action, Rationalist Theory and the Logic of Appropriateness in International Relations. In: European Journal of International Relations 10:3, 395-435.

Müller, Harald 2005: The 2005 NPT Review Conference: Reasons and Consequences of Failure and Options for Repair. Stockholm, Weapons of Mass Destruction Commission 31.

Nanda, Ved P. 1999: The Kyoto Protocol on Climate Change and the Challenges to Its Implementation: A Commentary. In: Colorado Journal of International Environmental Law and Policy 10, 319.

O'Keeffe, Anne 2004: 'Like the Wise Virgins and All that Jazz': Using a Corpus to Examine Vague Categorisation and Shared Knowledge. Applied Corpus Linguistics. Connor, U./Upton, T. A. Amsterdam, Brill: 1-20.

Oberheim, Eric /Hoyningen-Huene, Paul. 2016 (03.12.2016): "The Incommensurability of Scientific Theories." The Stanford Encyclopedia of Philosophy, from <http://plato.stanford.edu/archives/win2016/entries/incommensurability/>.

Olsen, Frances/Lorz, Alexander/Stein, Dieter, Eds.) 2009: Translation Issues in Language and Law. Basingstoke, Palgrave Macmillan.

Onuf, Nicholas 2010: Old Mistakes: Bourdieu, Derrida, and the "Force of Law", IPS Forum Contribution. In: International Political Sociology 4:3, 315-318.

Onuf, Nicholas 2014: Rule and Rules in International Relations. Erik Castrén Institute of International Law and Human Rights. University of Helsinki.

Onuma, Yasuaki 2003: International Law in and with International Politics: The Functions of International Law in International Society. In: European Journal of International Law 14:1, 105-139.

Oppenheim, Lassa F. L. 2005: International Law: a Treatise. Clark, The Lawbook Exchange.

Ostrom, Elinor 1990: Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge, Cambridge University Press.

Overstreet, Maryann/Tran, Jennie/Zietze, Sylvia 2006: Increasing Pragmatic Awareness: Die Vagheit der Sprache und so. In: Die Unterrichtspraxis / Teaching German 39:1/2, 24-29.

Overstreet, Maryann/Yule, George 1997: On Being Inexplicit and Stuff in Contemporary American English. In: Journal of English Linguistics 25:3, 250-258.

Pěchouček, Michal/Marík, Vladimír/Bárta, Jaroslav 2002: Knowledge Based Approach to Ootw Coalition Formation. Second International Conference on Knowledge Systems for Coalition Operations. Prague, Defense Technical Information Center (DTIC).

Peller, Gary 1985: The Metaphysics of American Law. In: California Law Review 73:4, 1151-1290.

Peloso, Chris 2011: Crafting an Updated Nuclear Non-Proliferation Treaty: Applying the Lessons Learned from the Success of Similar International Treaties to the Nuclear

Arms Problem. In: Santa Clara Journal of International Law 9, 309.

Pfister, Jannik/Rühlig, Tim/Weismann, Hanne 2011: Virtuous by Weakness? The Role of Small Developing States in the 2010 Cancun Climate Negotiations. Frankfurt, J. W. G. U., unpublished manuscript.

Posner, Eric A. 2009: The Perils of Global Legalism. Chicago, University of Chicago Press.

Postema, Gerald J. 2011: A Treatise of Legal Philosophy and General Jurisprudence, Springer Netherlands.

Prince, Ellen F./Frader, Joel/Bosk, Charles 1982: On Hedging in Physician-Physician Discourse. In: Pietro, R. J. D.: Linguistics and the Professions: Proceedings of the Second Annual Delaware Symposium on Language Studies. Norwood, Ablex. 8: 83-97.

Prüfer, Peter/Rexroth, Margrit 2005: Kognitive Interviews. In: ZUMA How-To 15, 3-21.

Putnam, Robert D. 1988: Diplomacy and Domestic Politics: The Logic of Two-Level Games. In: International Organization 42:3, 427-460.

Ranganathan, Surabhi 2014: Strategically Created Treaty Conflicts and the Politics of International Law. Cambridge, Cambridge University Press.

Reus-Smit, Christian 2003: Politics and International Legal Obligation. In: European Journal of International Relations 9:4, 591-625.

Roberts, J. Timmons 2011: Multipolarity and the New World (Dis)order: US Hegemonic Decline and the Fragmentation of the Global Climate Regime. In: Global Environmental Change 21:3, 776-784.

Ruzaitė, Jūratė 2007: Vague References to Quantities as a Face-Saving Strategy in Teacher-Student Interaction. Lodz Papers in Pragmatics. 3: 157.

Sainsbury, Mark 1991: Is There Higher-Order Vagueness? In: The Philosophical Quarterly 41:163, 167-182.

Salager-Meyer, Françoise 1994: Hedges and Textual Communicative Function in Medical English Written Discourse. In: English for Specific Purposes 13:2, 149-170.

Salleh, Ariel 2011: Cancun and After: A Sociology of Climate Change. In: Arena Magazin 110, 24-31.

Schachter, Oscar 1977: The Twilight Existence of Nonbinding International Agreements. In: The American Journal of International Law 71:2, 296-304.

Schauer, Frederick 2011: On Open Texture of Law. In: Virginia Public Law and Legal Theory Research Paper: 35.

Schnell, Rainer/Hill, Paul B. /Esser, Elke 2008: Methoden der Empirischen Sozialforschung. München, Oldenbourg.

Senter, RJ/Smith, Edgar A. 1967: Automated Readability Index, DTIC Document.

Setear, John K. 1996: An Iterative Perspective on Treaties: A Synthesis of International Relations Theory and International Law. In: Harvard International Law Journal 37, 139.

Simmons, Beth A. 2000: International Law and State Behavior: Commitment and Compliance in International Monetary Affairs. In: American Political Science Review 94:4, 819-835.

Simpson, John 1994: Nuclear Non-Proliferation in the Post-Cold War Era. In: International Affairs 70:1, 17-39.

Singer, Joseph William 1984: The Player and the Cards: Nihilism and Legal Theory. In: The Yale Law Journal 94:1, 1-70.

Snow, David A./Rochford, E. Burke/Worden, Steven K./Benford, Robert D. 1986: Frame Alignment Processes, Micromobilization, and Movement Participation. In: American Sociological Review 51:4, 464-481.

Šočanać, Lelija 2011: Indeterminacy vs. Precision in the Language of International Arbitration. In: Lapland Law Review: Special Issue on Law and Language in Partnership and Conflict 1:1, 190-205.

Solt, Stephanie 2011: Vagueness in Quantity: Two Case Studies from a Linguistic Perspective. In: Cintula, P./Fermueller, C. G./Godo, L./Hajek, P.: Understanding Vagueness. Logical, Philosophical and Linguistic Perspectives. London, College Publications: 157-174.

Solum, Lawrence B. 2010: The Interpretation-Construction Distinction. In: Constitutional Commentary 27, 95-118.

Staton, Jeffrey K./Vanberg, Georg 2008: The Value of Vagueness: Delegation, Defiance, and Judicial Opinions. In: American Journal of Political Science 52:3, 504-519.

Steffek, Jens 2005: Incomplete Agreements and the Limits of Persuasion in International Politics. In: Journal of International Relations and Development 8:3, 229-256.

Stoiber, Carlton 2003: The Evolution of NPT Review Conference Final Documents, 1975–2000. In: The Nonproliferation Review 10:3, 126-166.

Stokke, Olav Schram 2001: The Interplay of International Regimes: Putting Effectiveness Theory to Work. In: FNI report 14, 27.

Stratton, Jane 2009: International Law: Overview. In: Hot Topics: Legal Issues in Plain Language 69.

Sunstein, Cass R. 1996: Legal Reasoning and Political Conflict. New York, Oxford University Press.

Tallberg, Jonas 2010: The Power of the Chair: Formal Leadership in International

Cooperation. In: International Studies Quarterly 54:1, 241-265.

Tamiotti, Ludivine/Finger, Matthias 2001: Environmental Organizations: Changing Roles and Functions in Global Politics. In: Global Environmental Politics 1:1, 56-76.

Tiersma, Peter 1999: Legal Language. Chicago, University of Chicago Press.

Trask, Robert Lawrence 2013: A Dictionary of Grammatical Terms in Linguistics. London, Routledge.

Tussie, Diana 2005: Developmental Opposition in International Trade Regimes: Regional Groupings, State and Civil Society Coalitions. Pittsburgh, Matthew B. Ridgway Center for International Security Studies.

Tversky, Amos/Kahneman, Daniel 1973: Availability: A Heuristic for Judging Frequency and Probability. In: Cognitive psychology 5:2, 207-232.

Ullmann, Stephen 1962: Semantics: an Introduction to the Science of Meaning. New York, Barnes & Noble.

Underdal, Arild 1983: Causes of Negotiation 'Failure'. In: European Journal of Political Research 11:2, 183-195.

UNCLOS 1995: United Nations Agreement for the Implementation of the Provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 relating to the Conservation and Management of Straddling Fish Stocks and Highly Migratory Fish Stocks. A/CONF.164/37.

UNFCCC 2014 (05.01.2017): A Brief Overview of Decisions. from http://unfccc.int/documentation/decisions/items/2964.php.

Van der Burg, Wibren 2001: The Expressive and Communicative Functions of Law, Especially with Regard to Moral Issues. In: Law and Philosophy 20, 31-59.

Von Stein, Jana 2008: The International Law and Politics of Climate Change: Ratification of the United Nations Framework Convention and the Kyoto Protocol. In: Journal of Conflict Resolution 52:2, 243-268.

Waldron, Jeremy 1994: Vagueness in Law and Language: Some Philosophical Issues. In: California Law Review 82:3, 509-540.

Wänke, Michaela 2002: Conversational Norms and the Interpretation of Vague Quantifiers. In: Applied Cognitive Psychology 16:3, 301-307.

Way, Christopher/Sasikumar, Karthika 2004: Leaders and Laggards: When and Why Do Countries Sign the NPT? REGIS Working Paper. Montreal, Research Group in International Security.

Weisberg, Herbert/Krosnick, Jon A./Bowen, Bruce D. 1996: An Introduction to Survey Research, Polling, and Data Analysis. Thousand Oaks, Sage.

Werksman, Jacob/Herbertson, Kirk 2009: The Legal Character of National Actions and Commitments in a Copenhagen Agreement: Options and Implications. WRI Working Paper. Washington, World Resources Institute.

Wessel, Jared 2010: International Law as Language-Towards a "Neo" New Haven School. In: International Journal for the Semiotics of Law 23, 123-144.

Wettestad, Jørgen 1999: Designing Effective Environmental Regimes: the Key Conditions. Cheltenham, Edward Elgar Publishing.

Whalley, John/Walsh, Sean 2009: Bringing the Copenhagen Global Climate Change Negotiations to Conclusion. In: Cesifo Economic Studies 55:2, 255-285.

Williams, Christopher 2006: Fuzziness in Legal English: What Shall We Do with Shall. In: Wagner, A./Cacciaguidi-Fahi, S.: Legal Language and the Search for Clarity. Bern, Peter Lang: 237-263.

Williamson, Timothy 1999: On the Structure of Higher-Order Vagueness. In: Mind 108:429, 127-143.

Xi, Fang 2013: A Study on the Vagueness in English Language Teaching from the Pragmatic Perspective. In: Theory and Practice in Language Studies 3:9, 1596-1602.

Young, Oran R. 1979: Compliance & Public Authority: A Theory with International Applications. Baltimore, Johns Hopkins University Press.

Young, Oran R. 1982: Regime Dynamics: the Rise and Fall of International Regimes. In: International Organization 36:2, 277-297.

Yule, George 2014: The Study of Language. Cambridge, Cambridge University Press.

Zangl, Bernhard 1994: Politik auf zwei Ebenen. Hypothesen zur Bildung internationaler Regime. In: Zeitschrift für Internationale Beziehungen 1:2, 279-312.

Zartman, William 1994: The Elephant and the Holograph: Toward a Theoretical Synthesis and a Paradigm. In: Zartman, W.: International Multilateral Negotiation: Approaches to the Management of Complexity. San Francisco, Jossey-Bass.

Zartman, William, (Ed.) 1994: International Multilateral Negotiation: Approaches to the Management of Complexity. The Jossey-Bass Conflict Resolution Series. San Francisco, Jossey-Bass.

Zhang, Grace Q. 2015: Elastic Language: How and Why we Stretch our Words. Cambridge, Cambridge University Press.

Zürn, Michael/Binder, Martin/Ecker-Ehrhardt, Matthias/Radtke, Katrin 2006: Politische Ordnungsbildung wider Willen: ein Forschungsprogramm zu transnationalen Konflikten und Institutionen. Discussion Paper SP IV.

9 Appendix

9.1 Table of graphs

9.2 Table of tables

9.3 List of sentences used in the survey

Number	Sentence	Reference
VS 1	The Conference expresses its strong support for effective IAEA safeguards.	NPT 1975: 14
VS 2	The Conference reaffirms that IAEA safeguards should be assessed and evaluated regularly.	NPT 2000: 10
VS 3	The Conference calls upon States parties to [...] continue to discuss further, in a non-discriminatory and transparent manner under the auspices of IAEA or regional forums, the development of multilateral approaches to the nuclear fuel cycle, including the possibilities of creating mechanisms for assurance of nuclear fuel supply, as well as possible schemes dealing with the back-end of the fuel cycle without affecting rights under the Treaty and without prejudice to national fuel cycle policies, while tackling the technical, legal and economic complexities surrounding these issues, including, in this regard, the requirement of IAEA full scope safeguards.	NPT 2010: 34

VS 4	The Parties have a right to, and should, promote sustainable development.	UNFCCC 1992: Art. 3.4
VS 5	Accordingly, the developed country Parties should take the lead in combating climate change and the adverse effects thereof.	UNFCCC 1992: Art. 3.1
VS 6	Ministers believe that the Second Assessment Report should provide a scientific basis for urgently strengthening action at the global, regional and national levels, particularly action by Parties included in Annex I to the Convention (Annex I Parties) to limit and reduce emissions of greenhouse gases, and for all Parties to support the development of a Protocol or another legal instrument; and note the findings of the IPCC.	UNFCCC 1996: 1
VS 8	The Conference of the Parties serving as the meeting of the Parties to this Protocol shall, at its first session or as soon as practicable thereafter, decide upon modalities, rules and guidelines as to how, and which, additional human-induced activities related to changes in greenhouse gas emissions by sources and removals by sinks in the agricultural soils and the land-use change and forestry categories shall be added to, or subtracted from, the assigned amounts for Parties included in Annex I, taking into account uncertainties, transparency in reporting, verifiability, the methodological work of the Intergovernmental Panel on Climate Change, the advice provided by the Subsidiary Body for Scientific and Technological Advice in accordance with Article 5 and the decisions of the Conference of the Parties.	UNFCCC 1997: Art. 3.4
VS 9	We underline that climate change is one of the greatest challenges of our time.	UNFCCC 2009: 5
VS 10	To achieve the ultimate objective of the Convention to stabilize greenhouse gas concentration in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system, we shall, recognizing the scientific view that the increase in global temperature should be below 2 degrees Celsius, on the basis of equity and in the context of sustainable development, enhance our long-term cooperative action to combat climate change.	UNFCCC 2009: 5
VS 11	The Conference of the Parties [...] further recognizes that deep cuts in global greenhouse gas emissions are required according to science, and as	UNFCCC 2010: 3

	documented in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, with a view to reducing global greenhouse gas emissions so as to hold the increase in global average temperature below 2 °C above pre-industrial levels, and that Parties should take urgent action to meet this long-term goal, consistent with science and on the basis of equity; also recognizes the need to consider, in the context of the first review, as referred to in paragraph 138 below, strengthening the long-term global goal on the basis of the best available scientific knowledge, including in relation to a global average temperature rise of 1. 5 °C.	
VS 12	Australia will not purchase AAUs carried over from the first commitment period.	UNFCCC 2012: 13
VS 13	All the Parties to the Treaty undertake to facilitate, and have the right to participate in, the fullest possible exchange of equipment, materials and scientific and technological information for the peaceful uses of nuclear energy.	NPT 1968: Art. 4.2
VS 14	Nothing in this Treaty shall be interpreted as affecting the inalienable right of all the Parties to the Treaty to develop research, production and use of nuclear energy for peaceful purposes without discrimination and in conformity with Articles I and II of this Treaty.	NPT 1986: Art 4.1
VS 15	The Parties remain convinced that universal adherence to the Non-Proliferation Treaty is the best way to strengthen the barriers against proliferation and they urge all States not party to the Treaty to accede to it.	NPT 1985: 14
VS 16	The Conference further noted the calls on all States for the total and complete prohibition of the transfer of all nuclear facilities, resources or devices to South Africa and Israel and to stop all exploitation of Namibian uranium, natural or enriched, until the attainment of Namibian independence.	NPT 1985: 14
VS 17	The Conference reaffirms the determination expressed in the preamble of the 1963 Partial Test Ban Treaty, confirmed in Article I (b) of the said Treaty and reiterated in preambular paragraph 10 of the Non-Proliferation Treaty, to achieve the discontinuance of all test explosions of nuclear weapons for all time.	NPT 1985: 25
VS 18	Nuclear fissile material transferred from military use to peaceful nuclear activities should, as soon as practicable, be placed under Agency safeguards	NPT 1995: 17

	in the framework of the voluntary safeguards agreements in place with the nuclear-weapon States.	
VS 19	The development of nuclear-weapon-free zones, especially in regions of tension, such as in the Middle East, as well as the establishment of zones free of all weapons of mass destruction, should be encouraged as a matter of priority, taking into account the specific characteristics of each region.	NPT 1995: 16
VS 20	The Conference welcomes the fact that since May 1997, the IAEA Board of Governors has approved additional protocols to comprehensive safeguards agreements with 43 States and that 12 of those additional protocols are currently being implemented.	NPT 2000: 13
VS 21	Climate change means a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods.	UNFCCC 1992: Art. 1.2
VS 22	The Parties should cooperate to promote a supportive and open international economic system that would lead to sustainable economic growth and development in all Parties, particularly developing country Parties, thus enabling them better to address the problems of climate change.	UNFCCCC 1992: Art. 3.5
VS 23	The Conference of the Parties shall review, not later than 31 December 1998, available information with a view to taking decisions regarding such amendments to the lists in Annexes I and II as may be appropriate, with the approval of the Party concerned.	UNFCCC 1992: Art. 4.2(f)
VS 24	Significant reductions in net greenhouse gas emissions are technically possible and economically feasible by utilizing an array of technology policy measures that accelerate technology development, diffusion and transfer; and significant no-regrets opportunities are available in most countries to reduce net greenhouse gas emissions.	UNFCCC 1996: 1
VS 25	Each Party included in Annex I shall submit the information required under paragraph 1 above annually, beginning with the first inventory due under the Convention for the first year of the commitment period after this Protocol has entered into force for that Party.	UNFCCC 1997: Art. 3.4
VS 26	These consequences [of non-compliance] shall be the following: For the first commitment period, deduction at a rate of 1.3; [...] Development of a	UNFCCC 2000: 13

	compliance action plan to be submitted to the enforcement branch for review and assessment; [...] Suspension of the eligibility to make transfers under Article 17.	
VS 27	The Conference underscores the importance in complying with the non-proliferation obligations, addressing all compliance matters in order to uphold the Treaty's integrity and the authority of the safeguards system.	NPT 2010: 31
VS 28	The Conference remains convinced that universal adherence to the Treaty and full compliance of all parties with all its provisions are the best way to prevent the spread of nuclear weapons and other nuclear explosive devices.	NPT 2000: 8
VS 29	States parties that have concerns regarding non-compliance with the safeguards agreements of the Treaty by the States parties should direct such concerns, along with supporting evidence and information, to IAEA to consider, investigate, draw conclusions and decide on necessary actions in accordance with its mandate.	NPT 2010: 9
VS 30	The Conference of the Parties serving as the meeting of the Parties to this Protocol shall, at its first session, approve appropriate and effective procedures and mechanisms to determine and to address cases of non-compliance with the provisions of this Protocol, including through the development of an indicative list of consequences, taking into account the cause, type, degree and frequency of non-compliance.	UNFCCC 1997: Art. 18