



## Articles

## Linguistic variation in functional types of statutory law

Margaret Wood

Los Alamos National Laboratory, Los Alamos, New Mexico USA 87544

## ARTICLE INFO

## Keywords:

Statutory law  
Statutory interpretation  
Lexico-grammatical variation  
Corpus linguistics  
Key feature analysis

## ABSTRACT

When the meaning of an ambiguous word, phrase or grammatical structure in a statutory provision is disputed, courts are tasked with identifying the best meaning of the contested language. A common method of resolving linguistic ambiguities is to investigate the meaning of the contested word or structure in statutory provisions with similar subject matter. While the subject matter of a text has a demonstrated effect on language use, register variation research shows that the function of a text is also highly influential in predicting linguistic variation. Thus far, the function of a statutory provision (e.g., obligation to act, authorization to act) has not been considered in legal interpretative research. In the present study, I investigate the extent to which function influences the lexico-grammatical characteristics of statutory texts. 2,573 statutory provisions from the Arizona State Code are individually assigned to one of seven categories representing their function: Duties, Permissions, Impersonal Rules, Operational Definitions, Prohibitions, Procedural Guidelines, and Criminal Offenses. Key feature analysis is used to identify and describe patterns of lexico-grammatical variation between the seven functional types. Results reveal a great deal of lexico-grammatical variation associated with function in the register of statutory law. Furthermore, some functional types of statutory provisions are more linguistically distinct than others. These findings suggest that it may be beneficial to consider communicative function when investigating legal interpretative questions.

## 1. Introduction and background

Statutory law is written to directly alter our norms of social behavior in order to ensure that citizens act in a way that is deemed acceptable by society (Gunnarsson, 1984). As such, the language in statutory texts carries a considerable amount of power, explicitly creating, modifying, and terminating the legal rights and obligations of everyday individuals (Tiersma, 1999). Statutory texts accomplish this in a variety of ways. For example, while some statutory texts prohibit actions, others may authorize or obligate actions. In register variation research, these are considered different *communicative functions*<sup>1</sup> of the texts (i.e., to prohibit, to authorize, to obligate).

Variation in the communicative function of statutory texts is recognized in both linguistic and legal literature. While the term ‘communicative function’ has not been used in this body of literature, a variety of frameworks have been introduced to capture this variation. One of the earliest frameworks used to describe legislative texts was introduced by Gunnarsson (1984), who named three “classes of rules” in Swedish Acts of Parliament: (1) action rules, which include rights, duties,

prohibitions, exemptions, and recommendations, (2) stipulation rules, and (3) definition rules (p.84). Since this time, a variety of terms have been used to describe similar variation in legislative texts. For example, Kurzon (1986) named permissions, ordering, and prohibitions as three types of illocutionary forces characteristic of statutory language (also see Hiltunen 2012). Šarčević similarly (1997) named commands, permissions, and authorizations as three types of “regulatory instruments”. Cao (2007) later named permissions, rights, obligations, and prohibitions as four types of statutory texts (referred to as “legislative speech acts”), and Williams (2007) named definitions, commands, and prohibitions as four types of principal provisions in legal prescriptivist texts (e.g., codes) (pp. 45–50). Most recently, Fan (2020) categorized laws into two types: (1) authoritative guidance, which grants rights and powers, and (2) obligatory requirements, which include imperative requirements and prohibitive requirements. In this body of literature, recurring types of legislative texts include permissions, obligations, prohibitions, rights, and definitions. While there is some variation in terminology (e.g., commands, authorizations, ordering), there appears to be general consensus about these five pervasive types.

The results presented in this study are part of a large-scale study carried out for a doctoral dissertation (Wood, 2023).

E-mail address: [mkwood@lanl.gov](mailto:mkwood@lanl.gov).

<sup>1</sup> Commonly referred to as *communicative purpose* (see Biber 2019).

Definitions have garnered particular interest as precise and exact terminology is crucial to the integrity of written legal documents. Sub-categories have been proposed to reflect the way that definitions define terms and phrases. For example, Macagno (2010) identified three ways that terms are defined in legislative texts: (1) providing examples, (2) showing fundamental characteristics of the concept defined, and (3) listing constituent parts of a denotatum (p.200). Bowers (1989) distinguishes explanatory definitions from stipulative definitions; the former defining terms without altering conventional significations, and the latter altering the meaning of words by narrowing or enlarging their sense, or by creating a wholly new meaning for them (pp. 173–177).

In select circumstances, this body of literature has described the ways that these differences are reflected in language use. Bowers (1989) observed that while explanatory definitions use words such as *means* and *includes*, stipulative definitions favor language such as *is deemed* or *is*. Hiltunen (2012) observed that permissions, ordering, and prohibitions make use of the words *may*, *shall*, and *shall not*, respectively. The relationship between modal verbs and legislative texts has also been documented by Gunnarsson (1984), Foley (2022), Ni and Sin (2011), and Biel (2017). However, much of the existing research has focused solely on the pervasive linguistic features of legislative language, treating it as a linguistically homogeneous domain (e.g., Go'zd'z-Roszkowski 2011; O'zyildirim 2011; Sun and Chen, 2017; Fan 2020; Wood, 2022). Legislative language has commonly been described as reliant upon modal auxiliary verbs (Cooper, 2011; Gibova', 2011; Ni and Sin, 2011; Biel, 2014, 2017; Yankova, 2017), dependent clausal language (Go'zd'z-Roszkowski, 2011; Hiltunen, 2012; Wood, 2022), multi-word units such as lexical bundles, binomials and complex prepositions (Go'zd'z-Roszkowski, 2011; Williams, 2013; Biel, 2015; Giczela-Pastwa, 2019), and the passive voice (Williams, 2004; Bulatovi'c, 2013). While this literature has been instrumental in our understanding of legislative language, there has been little regard for the influence that the communicative function of individual statutory texts may have on the preferred linguistic features.

A rich history of register variation research has demonstrated that there is a three-way relationship between situational characteristics (e.g., communicative function), frequent and pervasive linguistic features, and the discourse function of those features at the text level<sup>2</sup> (see Biber and Egbert 2023). This is particularly true of communicative function, which is generally regarded as the most influential characteristic in describing the underlying situational-linguistic relationship. As a register comprising texts that serve multiple different communicative functions (e.g., to authorize, to prohibit, to define), it follows that statutory law is likely to exhibit distinct, systematic patterns of lexico-grammatical variation. This is potentially significant in the context of statutory interpretation, as legal interpretive research frequently relies upon pervasive phrasal and clausal patterns to identify the meaning of contested language. This suggests that it may be important to consider communicative function when attempting to identify pervasive linguistic patterns in statutes.

With this in mind, the aim of the present study is to investigate the extent to which functional types of statutory law display systematic patterns of lexico-grammatical variation. To do this, the study will identify and describe the discourse function of lexico-grammatical features associated with particular functional types. This means that unlike past literature, statutory law will be treated as a situationally heterogeneous register. The following section will describe the methodology

employed in the present study including a description of the corpus, the process of coding statutory provisions for function, and the method of linguistic analysis.

## 2. Method

The present study was carried out in five phases: (1) building a corpus of state-level statutory provisions, (2) developing a framework representing the communicative functions that statutory provisions serve, (3) categorizing provisions by their function in reference to the framework, (4) tagging the corpus for lexico-grammatical features, checking the accuracy of the tags, and correcting tags where necessary, and (5) conducting lexico-grammatical analysis to identify patterns of linguistic variation and provide functional interpretations of the characteristic features. These steps will be further detailed in Sections 2.1–2.4.

### 2.1. Compiling the corpus

The present study makes use of a corpus comprising statutes sampled from the Arizona State Code, which is a part of the larger Corpus of U.S. State Statutes (CorUSSS) (Egbert and Wood, 2023). In the present study, a 'statute' is defined by its unique universal citation, containing the exact title, article, or section that the text appears in (e.g., AZ Rev. Stat. §41–105; AZ Rev. Stat. §41–151.14). A statute often comprises multiple subsections (e.g., A., B., C., D.), as seen in Text Sample 1, below. These subsections will serve as the unit of analysis in the present study.

#### Text Sample 1: AZ Rev Stat § 5–332 (2019)

- A. All watercraft, unless exempted by the commission, carrying as fuel any volatile liquid having a flash point of one hundred ten degrees Fahrenheit or less shall have aboard a readily accessible United States coast guard approved fire extinguisher in a condition available for immediate and effective use.
- B. All watercraft over twenty-six feet in length and carrying as fuel any volatile liquid having a flash point of one hundred ten degrees Fahrenheit or less shall have aboard such fire extinguishers as may be prescribed or approved by the regulations of the United States coast guard.

To create the corpus, 1,000 statutes were sampled proportionally from the forty-seven titles of the Arizona State Code<sup>3</sup>. These titles represent a range of topical content, such as Environment, Businesses and Institutions, and Information Technology. The subsections of a statute tend to be self-contained and serve independent communicative functions. For this reason, the statutes were split at the subsection level to form independent texts. These texts are referred to as 'statutory provisions' in the present study.

Slight modifications were made to the corpus in preparation for linguistic analysis. Short texts are generally considered less than ideal for lexico-grammatical analyses that use normed frequency counts. Means and standard deviations rely upon repeated occurrences of lexico-grammatical features, which cannot be counted upon to appear frequently in short texts (Clarke, 2022; Liimatta, 2022). For this reason, texts shorter than twenty-five words in length were removed from the sample. 88 % of the statutory provisions in the original sample met this length requirement, resulting in a corpus comprising 2,573 statutory provisions and 218,675 words.

<sup>2</sup> The three aspects of this relationship are traditionally referred to as *situational characteristics* (e.g., *communicative purpose*), *linguistics forms*, and the *communicative function* of those linguistic forms (see Biber 2019). To avoid implying judgment of intent or reason in the context of statutory texts, the term *communicative function* has replaced the term *communicative purpose*. Accordingly, the original use of the term *communicative function* by Biber (2019) has been revised to *discourse function* for the purposes of this study.

<sup>3</sup> The 2019 version of the code was the most recent version available at the time that CorUSSS was compiled.

**Table 1**

Functional types of statutory law, definitions, and example provisions.

Description	Example	Additional Description
<b>Duties</b> Prescribes a mandatory duty, responsibility, or requirement to a human agent(s) or government entity(s).	<b>AZ Rev. Stat. §9–471 (2019)</b> G. The county assessor and the department of revenue, respectively, shall furnish to the city or town proposing an annexation, within thirty days after a request, a statement in writing showing the owner, the address of each owner and the appraisal and assessment of all such property.	<ul style="list-style-type: none"> <li>Agent may be unspecified.</li> <li>Includes mandatory contents/information in legal documents, forms, applications, etc.</li> </ul>
<b>Permissions</b> Authorizes human agent(s) or government entity(s) to act. Allows for inaction.	<b>AZ Rev. Stat. §10–2259 (2019)</b> C. A corporation formed under this chapter may amend its articles of incorporation provided that prior to the filing of articles of amendment with the corporation commission the same shall have been approved by the superintendent of financial institutions (. . .)	<ul style="list-style-type: none"> <li>Statements or a set of statements authorizing actions for individual(s), entity(s), or institution(s)</li> <li>Agent may be unspecified.</li> </ul>
<b>Impersonal Rules</b> States a general rule or provision that does not directly regulate a human agent or government entity's behavior.	<b>AZ Rev. Stat. §10–701 (2019)</b> C. The failure to hold an annual meeting at the time stated in or fixed in accordance with a corporation's bylaws does not affect the validity of any corporate action.	<ul style="list-style-type: none"> <li>Rules, descriptions, statements of fact.</li> <li>Usually concerns inanimate objects or abstract concepts/ideas.</li> <li>Includes describes the composition of entities, dates, etc.</li> <li>Defines terms or phrases through lists of examples, descriptions, or inclusion/exclusion criteria.</li> </ul>
<b>Operational Definitions</b> Defines a term or phrase from another statutory provision.	<b>AZ Rev. Stat. §34–603 (2019) L.</b> For the purposes of this section, "professional services" includes architect services, engineer services, landscape architect services, assayer services, (. . .)	<ul style="list-style-type: none"> <li>May include conditions or exceptions.</li> <li>Agent may be unspecified.</li> </ul>
<b>Prohibitions</b> Expressly prohibits human agent(s) or government entity(s) from taking a specified action.	<b>AZ Rev. Stat. §23–373 (2019)</b> I. An employer may not require that documentation under subsection G explain the nature of the health condition or the details of the domestic violence, sexual violence, abuse or stalking.	<ul style="list-style-type: none"> <li>May include a combination of different functional types.</li> <li>May concern the actions of a single agent or multiple agents.</li> <li>Actions are presented in a sequential fashion.</li> <li>Often discusses contingencies.</li> <li>Often includes temporal references.</li> </ul>
<b>Procedural Guidelines</b> Prescribes a set of sequential steps required to carry out a task in a specific situation or context, including description of the role(s) of one or more human(s) or government entity(s)	<b>AZ Rev. Stat. §3–239 (2019) B.</b> If the court finds the seed is in violation of this article and orders condemnation of the seed, it shall be denatured, processed, destroyed, relabeled or otherwise disposed of in compliance with the laws of this state. Before the court orders such disposition of the seed, it shall first give the claimant an opportunity to apply to the court for release of the seed or permission to process or relabel it to bring it into compliance with this article.	<ul style="list-style-type: none"> <li>Includes felonies, misdemeanors, and being held in contempt.</li> </ul>
<b>Criminal Offenses</b> Explicitly states that a prohibited action rises to the level of a criminal offense (i.e., felony, misdemeanor)	<b>AZ Rev. Stat. §13–2322 (2019)</b> A. A person or company that owns, sells, leases or brokers a transaction involving property or real property that the person or company knows will be used as a drop house is guilty of a class 4 felony.	<ul style="list-style-type: none"> <li>Includes multiple different functions that can be implemented independently from one another and are non-sequential.</li> </ul>
<b>Other</b> Subsections that do not fit clearly into one of the above seven functions.	<b>AZ Rev. Stat. §12–113 (2019)</b> C. The supreme court shall administer the fund and may expend monies in the fund, subject to legislative appropriation, for local, regional or statewide projects. The supreme court may directly provide or contract for services consistent with the purposes of the fund (. . .)	or <ul style="list-style-type: none"> <li>Serves a function not represented by the seven above.</li> </ul>

## 2.2. Development of the framework and coding rubric, piloting, and coding

To categorize statutory provisions by communicative function, a framework was developed to represent functional types of statutory law. The framework was developed with reference to past frameworks and assistance from expert informants from both legal and linguistic backgrounds. Seven functional types of statutory provisions are represented in the final framework: Duties, Permissions, Impersonal Rules, Operational Definitions, Prohibitions, Procedural Guidelines, and Criminal Offenses. An 'Other' category was also established to hold provisions that did not serve one of the seven functions appearing in the framework.

The framework was then adapted to create a coding rubric that could be used to categorize provisions by their functional type. This coding rubric includes: (1) the name of the functional type and its definition, (2) several example provisions, and (3) additional descriptions, including detailed inclusion and exclusion criteria to assist with coding decisions that were particularly challenging. There are inevitable instances of overlap between select functional types of provisions (e.g., Procedural Guidelines and Duties) making it necessary to establish additional

coding guidelines for potentially ambiguous cases. This issue of overlap will be revisited in [Section 4.1](#).

The coding rubric went through several iterations of revisions and pilot coding before reaching its final form. This process was facilitated with the help of several outside coders who were recruited to code small random samples of provisions based on the rubric and provide subsequent feedback. Once the coding rubric was completed, a random sample of one hundred statutory provisions were coded by two independent coders to test the reliability of the rubric. Agreement was high between the two coders with a Cohen's kappa value of .97.

A snapshot of the final rubric used to code the seven functional types of statutory provisions is presented in [Table 1](#). The complete coding rubric, including all example provisions and inclusion/exclusion criteria, can be found in [Wood \(2023, pp.65–69\)](#).

The coding process resulted in the corpus of statutory provisions presented in [Table 2](#). These groups constituted the sub-corpora used for analysis in the present study.

## 2.3. Tagging, tag-checking, feature selection

Once the statutory provisions were categorized by functional type,

**Table 2**  
Corpus coded for functional type.

	No. of texts	No. of words
Duties	726	51,218
Permissions	423	31,348
Impersonal Rules	284	17,203
Procedural Guidelines	167	23,277
Prohibitions	164	10,678
Operational Definitions	141	22,224
Criminal Offenses	68	4,252
Other	600	58,475
TOTAL	2573	218,675

the Biber Tagger was used to annotate texts for their lexico-grammatical features. A tag count was generated that provided normed frequency counts per 1000 words for over 150 lexico-grammatical features. A Python script was written to count tags for eight additional features (denoted with an asterisk in Table 3), including *to* and *that* noun complement clauses, and non-finite *-ing* complement clauses preceded by a preposition, among others. Fifty-six features were selected for analysis, informed primarily by past studies of legal language (Table 3). For example, features such as modal auxiliary verbs, a wide range of dependent clauses, and various passive voice constructions were selected, while *wh*- questions, color adjectives, and first and second person pronouns were not considered.

Following feature selection, forty-two texts representing all seven functional types of statutory provisions were randomly sampled to check the accuracy of the tags. The tag for each individual feature was checked for precision (the rate at which the tagger assigned a correct tag) and recall (the rate at which the tagger identified all target features present in the text). Based on results, a Python code was written to correct tags with low precision or recall, and accuracy was re-tested. Following this process, the tags in the corpus averaged 95 % precision and 85 % recall.

2.4. Linguistic analysis

A set of key feature analyses (see Biber and Egbert 2018; Egbert and Biber 2023) were used to identify patterns of lexico-grammatical variation between functional types of statutory law. A key feature analysis indicates the degree to which a linguistic feature is used with a markedly higher or lower frequency in a particular text variety (the ‘target corpus’) relative to a larger group of texts (the ‘reference corpus’) (Biber and Egbert 2018, p. 22).

In the present study, key feature analyses are used to identify the linguistic features that are most characteristic of each type of statutory provision relative to the others. Accordingly, seven individual key feature analyses were conducted. For each key feature analysis, one functional type of statutory provision served as the target corpus and was the focus of linguistic analysis. Texts from the remaining functional types of statutory provisions were combined together to serve as the reference corpus. A visual representation of each key feature analysis is presented in Table 4.

Note that a key feature analysis was not conducted to target the ‘Other’ category. This category does not represent a homogeneous function, meaning that examining the characteristic linguistic features of this category is not informative for the purposes of this study.

In each of the seven analyses, Cohen’s *d* effect sizes were calculated for each lexico-grammatical feature based on its mean rate of occurrence and standard deviation in both the target and the reference corpus. The size of the Cohen’s *d* value determines the ‘keyness’ of the lexico-grammatical feature. A large positive *d* value indicates that the feature is markedly more common in the target corpus compared to the reference corpus. A large negative *d* value indicates that the feature is markedly less common in the target corpus. *d* values approaching 0 indicate similar frequency of use. In accordance with Cohen (1977),

the strength of the effect size is interpreted as such:

<i>d</i> value	Interpretation
> ±.80	Large
> ±.50	Medium
> ±.20	Small

Once Cohen’s *d* values were generated for key features in each of the seven analyses, qualitative analysis was used to describe the underlying functional relationship between pervasive features and communicative function for each type of statutory provision.

Additional analyses were conducted to determine the extent to which functional types of statutory provisions are lexico-grammatically distinct from one another. This can be inferred through both the number of key features and the strength of the effect sizes. A large number of positive key features indicates that the functional type has a considerable number of lexico-grammatical features that are uniquely characteristic of it compared to the other types of provisions. A larger number of negative key features indicates the relative absence of many lexico-grammatical features that are shared by the other functional types. The following section will provide the results from each of the key feature analyses. Detailed functional interpretations of key features will be provided for select functional types.

3. Results and analysis

3.1. Key grammatical features in functional types of statutory provisions

Tables 5-11 present the key features for each of the seven functional types of statutory provisions. Cohen’s *d* values < ±.20 imply similar frequency of use and were not considered in linguistic analyses. Results are consistent with past register variation research examining the relationship between communicative function and language use, revealing that there are indeed systematic patterns of lexico-grammatical variation associated with function. Results also show that the extent to which a functional type is lexico-grammatically distinct from the other functional types is variable.

Tables 5-8 show that statutory provisions authorizing actions, mandating actions, prohibiting actions, and prescribing impersonal rules are relatively less distinct from the remainder of the statutory provisions in the corpus. While Permissions and Duties each have a modal verb with a large effect size, there are few positive or negative key features in either functional type (Permissions, *n* = 3; Duties, *n* = 6).

Prohibitions and Impersonal Rules are also relatively non-distinct. These two functional types have few positive key features (Prohibitions, *n* = 2; Impersonal Rules, *n* = 3), all with small effect sizes. Note that if a higher effect size cutoff had been implemented (e.g., > ±.30, see Biber and Egbert 2018), Impersonal Rules would only have a single positive key feature, and Prohibitions would have none. However, in contrast to Duties and Permissions, Prohibitions and Impersonal Rules have a reasonably large number of negative key features. This is particularly true of Impersonal Rules (*n* = 10). As negative effect sizes indicate the relative absence of features shared by the other functional types, this suggests that Prohibitions and Impersonal Rules rely on a smaller set of specific lexico-grammatical features to accomplish their intended function.

In contrast to the former four functional types, results indicate that Operational Definitions, Criminal Offenses, and Procedural Guidelines are relatively lexico-grammatically distinct (see Tables 9-11). Each of these three functional types have a relatively large number of positive and negative features, many with medium and large effect sizes. Operational Definitions and Criminal Offenses are particularly distinctive. The former has eleven positive key features and thirteen negative key features ranging from small to large effect sizes, while the latter has eleven positive key features and a substantial twenty-five negative key features.

**Table 3**  
Lexico-grammatical features for key feature analyses.

Feature category	Feature	Example feature
Verbs	modals of possibility	<i>may, can, might, could</i>
	modals of prediction	<i>shall, will, would</i>
	modals of necessity	<i>must, ought, should</i>
	present tense verb	<i>testify, bribe, accuse</i>
	past tense verb	<i>adjudicated, filed</i>
	perfect aspect	<i>has been, have been</i>
	progressive aspect	<i>trying, accusing</i>
	agentless passive verb	<i>the amendment was passed</i>
	<i>by</i> passive	<i>the bill was signed by the governor</i>
	infinitive verb	<i>to certify, to file</i>
	definition connectors	<i>means, includes, involves</i>
	verbs of causation	<i>let, permit</i>
	public verb	<i>assert, complain, testify</i>
	suasive verbs	<i>command, insist</i>
	activity verbs	<i>go, take, move</i>
	'be' verb	<i>be, am, is, was, being</i>
	'have' verb	<i>have, having, had</i>
Nouns	human nouns	<i>governor, judge, plaintiff</i>
	proper nouns	<i>Secretary of State, Arizona</i>
	process nouns	<i>meeting, trial</i>
	technical nouns	<i>arraignment, acquittal, dissent</i>
	common nouns	<i>affidavit, contract, agency, district, (all nouns except proper)</i>
	abstract nouns	<i>information, guilt, dispute, fact</i>
	place nouns	<i>courtroom, property</i>
	pre-modifying nouns	<i>school district, state treasurer</i>
	nominalization	<i>knowledge, application</i>
	<i>that</i> verb complement clause*	<i>If court [determines] that the company violated the terms of the agreement</i>
Dependent clauses	<i>that</i> noun complement clause*	<i>The document shall include [a statement] that the known place of business is also discontinued</i>
	<i>to</i> verb complement clause*	<i>The director may [deny, revoke or refuse] to renew the license</i>
	<i>to</i> noun complement clause*	<i>After notice on a petition for [permission] to resign</i>
	adjective complement clause*	<i>It is [unlawful] to participate in a financial transaction for...</i>
	<i>that</i> relative clause	<i>the court may make [any further order] that may be appropriate</i>
	<i>wh-</i> relative clause	<i>any [person] who contracts directly with members of the general public</i>
	<i>wh-</i> pied-piping relative clause	<i>"Agency placement adoption" means [an adoption proceeding] in which one or more of the requisite consents...</i>
	<i>-ing</i> post-nominal modifiers*	<i>the person injured shall submit [a complaint] to the board outlining the circumstances</i>
	<i>-ing</i> complement preceded by a preposition*	<i>On finding that a situation described in the testimony exists for any taxable year or years</i>
	<i>-ed</i> passive post-nominal modifier*	<i>the bill signed by the governor...</i>
Adverbials	adverbs *does not include downtoner/time/place	<i>knowingly, intentionally</i>
	adverbs	
	adverbs of time	<i>annually, quarterly, later</i>
	adverbs of place	<i>everywhere, around</i>
	conditional subordinating conj.	<i>if, unless</i>
	causative subordinating conj.	<i>because</i>
	subordinating conjunction - other	<i>as, except, until</i>
	adverb - downtoner	<i>nearly, only</i>
	prepositions	<i>in, on, by, at</i>
	clausal coordinating conjunction	<i>Any person who coerces or offers a bribe to manipulate...</i>
Other	phrasal coordinating conjunction	<i>The judge considered the documentary evidence and spoken testimony.</i>
	attributive adjectives	<i>legislative body, federal court</i>
	predicative adjectives	<i>The agreement is void.</i>
	pronoun <i>it</i>	<i>the holder of that security interest must exercise its right against...</i>
	3rd person pronoun	<i>he, she</i>
	nominal pronoun	<i>someone, everything</i>
	definite articles	<i>the</i>
	indefinite articles	<i>a/an</i>
	word count	—

Because these three functional types are considerably more distinct, the following sections will present the results of the three key feature analyses in detail, including effect sizes and functional interpretations of the key features.

### 3.2. Functional interpretation of Operational Definition key features

Operational Definitions have eleven positive key features. The feature with the highest keyness value is definition connectors (e.g., *means, includes*) ( $d = 1.44$ ). This is a narrow category of words that

connect the target term to its definition or description; for example, *[X] means [Y]*; *[X] includes [Y]* and *[Z]*; *[X] involves [Y]* and *[Z]*. Word count has a medium effect size in Operational Definitions ( $d = .52$ ) indicating that these provisions are markedly longer in length relative to the other functional types of provisions. Several semantic domains of nouns are key, ranging from medium to small effect sizes (common nouns,  $d = .71$ ; abstract nouns,  $d = .32$ ; proper nouns,  $d = .32$ ) as well as phrasal coordinating conjunctions ( $d = .39$ ), indicating the use of noun series. This functional type also has numerous key features associated with descriptive language, including pre-modifying nouns ( $d = .31$ ),

**Table 4**

Key feature analyses.

Key Feature Analysis 1	Key Feature Analysis 2	Key Feature Analysis 3	Key Feature Analysis 4
<b>Target corpus:</b> Duties	<b>Target corpus:</b> Permissions	<b>Target corpus:</b> Prohibitions	<b>Target corpus:</b> Impersonal Rules
<b>Reference corpus:</b> Permissions	<b>Reference corpus:</b> Duties	<b>Reference corpus:</b> Duties	<b>Reference corpus:</b> Duties
Prohibitions	Prohibitions	Permissions	Permissions
Impersonal Rules	Impersonal Rules	Impersonal Rules	Prohibitions
Procedural	Procedural	Procedural	Procedural
Guidelines	Guidelines	Guidelines	Guidelines
Operational	Operational	Operational	Operational
Definitions	Definitions	Definitions	Definitions
Criminal	Criminal Offenses	Criminal Offenses	Criminal Offenses
Offenses	Other	Other	Other
Other			
<b>Key Feature Analysis 5</b>	<b>Key Feature Analysis 6</b>	<b>Key Feature Analysis 7</b>	
<b>Target corpus:</b> Procedural	<b>Target corpus:</b> Operational	<b>Target corpus:</b> Criminal Offenses	
Guidelines	Definitions	<b>Reference corpus:</b> Duties	
<b>Reference corpus:</b> Duties	<b>Reference corpus:</b> Duties	Permissions	
Permissions	Permissions	Prohibitions	
Prohibitions	Prohibitions	Impersonal Rules	
Impersonal Rules	Impersonal Rules	Operational	
Operational	Procedural	Definitions	
Definitions	Guidelines	Procedural	
Criminal	Criminal Offenses	Guidelines	
Offenses	Other	Other	
Other			

attributive adjectives ( $d = .28$ ), and *that* and *wh*-relative clauses ( $d = .30$ ;  $d = .23$ ). Finally, Operational Definitions make markedly higher use of the present tense compared to other functional types of statutory provisions ( $d = .25$ ).

Definition connectors serve as the connecting word between a term and its definition in nearly all Operational Definitions (95 % of the sub-corpus). *Include* defines a term by providing a list of items that the term encompasses, while *means* is used to connect a term to a description, or to equate the term to other terms. *Include* is often negated (see Text Sample 2). The word *involved* is also used as a definition connector, though less frequently (definition connectors bolded).

#### Text Sample 2: AZ Rev. Stat. §44–1565 (2019)

In this article, unless the context otherwise requires:

1. “Beer” **means** any beverage obtained by the alcoholic fermentation, infusion or decoction of barley malt, hops or other ingredients not drinkable, or any combination of them. Beer **does not include** spirituous liquor.
2. “Franchise” **means** a commercial relationship between a supplier and a wholesaler which **includes** all of the following:
  - a. A commercial relationship of definite duration or continuing indefinite duration is **involved**. (. . .)

The markedly higher use of nouns in Operational Definitions reflects two characteristics of the texts: (1) the use of inclusive series of nouns in favor of single terms (resulting in lists), and (2) repetition of these series. These two characteristics are demonstrated in Text Sample 3. The text uses inclusive lists of nouns to ensure that the scope of the definition is clear and cannot be misconstrued (e.g., *a rule, regulation, practice or policy; an occupation, trade or profession*). This often results in a series of nouns that are typically regarded as synonyms or near synonyms by the lay person.

Series are often repeated in single provisions. Every time the text refers back to a previously mentioned series of nouns, all nouns in the series are re-listed. This is done to avoid the implication that a provision

**Table 5**

Permissions Key Feature Results.

Feature	Cohen’s $d$
modals of possibility	1.30
adjective complement clause	.20
modals of prediction	–1.28

**Table 6**

Duties Key Feature Results.

Feature	Cohen’s $d$
modals of prediction	.93
definite articles	.26
pre-modifying nouns	.21
indefinite articles	–.25
word count	–.25
modals of possibility	–.80

no longer applies to all nouns named in the original series. For example, if a provision first uses the series *occupation, trade, or profession*, and in a later clause only uses the term *occupation*, one must assume that this particular clause no longer applies to a trade or profession. In Text Sample 3, these two characteristics (use of noun series and repetition) result in the use of the series *occupation, trade or profession* three times in a single statutory provision. The regular appearance of nouns in Operational Definitions go hand-in-hand with the pervasive use of phrasal coordinating conjunctions (noun series italicized; phrasal conjunctions bolded).

#### Text Sample 3: AZ Rev. Stat. §9–842 (2019)

- I. For the purposes of this section, “occupational licensing requirement”:
  - A. Means:
    - a. A *rule, regulation, practice or policy* that is adopted by a *city or town* and that allows an individual to use an occupational title or work in a lawful *occupation, trade or profession* or a regulatory requirement that prevents an individual from using an occupational title or working in a lawful *occupation, trade or profession*.
    - b. A *fee or tax* that a city or town imposes on an individual to use an occupational title or work in a lawful *occupation, trade or profession*. (. . .)

Operational Definitions rely on pre-modifying nouns, attributive adjectives, *that* relative clauses, and *wh*-relative clauses to provide qualifying details about a term or to identify referents. Pre-modifying nouns and attributive adjectives are typically associated with literate discourse and condensed language. In Operational Definitions, they are used to provide precise, exact information in limited space. Pre-modifying nouns and attributive adjectives are markedly more frequent in Operational Definitions as more general nouns (e.g., school, court) often come in many types (e.g., *private* school, *public* school, *charter* school; district court, state court, municipal court) (see Text Sample 4). Many head nouns in Text Sample 4 take both an attributive adjective and a pre-modifying noun (pre-modifying nouns bolded, attributive adjectives italicized).

#### Text Sample 4: AZ Rev. Stat. §38–448 (2019)

D. For the purposes of this section:

1. “Agency” means: a. All offices, agencies, departments, boards, councils or commissions of this state. b. All **state** universities. c. All **community college** districts. d. All *legislative* agencies. e. All departments or agencies of the *state supreme* court or the court of appeals.



**Table 7**  
Impersonal Rules Key Feature Results.

Feature	Cohen's <i>d</i>
prepositions	.39
predicative adjectives	.30
verbs of occurrence	.20
non-finite <i>-ing</i> clause (preceded by prep)	-.20
pre-modifying nouns	-.21
<i>that</i> verb complement clause	-.21
activity verbs	-.22
suasive verbs	-.23
human nouns	-.26
modals of prediction	-.28
present tense	-.31
word count	-.38
modals of possibility	-.71

**Table 8**  
Prohibitions Key Feature Results.

Feature	Cohen's <i>d</i>
indefinite articles	.27
phrasal coordinating conjunctions	.24
present tense	-.21
modals of necessity	-.22
common nouns	-.28
modals of possibility	-.30
word count	-.31
definite articles	-.53

**Table 9**  
Operational Definitions Key Feature Results.

Feature	Cohen's <i>d</i>
definition connectors	1.44
common nouns	.71
word count	.52
phrasal coordinating conjunctions	.39
abstract nouns	.32
pre-modifying nouns	.31
<i>that</i> relative clause	.30
proper nouns	.29
attributive adjectives	.28
present tense	.25
<i>wh-</i> relative clause	.23
<i>that</i> verb complement clauses	-.20
modals of necessity	-.22
pronoun <i>it</i>	-.24
<i>to</i> verb complement clause	-.25
suasive verbs	-.27
infinitive	-.29
prepositions	-.30
'be' verb	-.31
activity verbs	-.36
predicative adjectives	-.42
definite articles	-.54
modals of possibility	-.65
modals of prediction	-1.17

2. “**Information** infrastructure” means telecommunications, **cable** and **computer** networks and includes the internet, the **world wide** web, Usenet, **bulletin board** systems, *on-line* systems and **telephone** networks.

In Operational Definitions, relative clauses often carry the crucial identifying details about a term. For example, the identifying information for the word *testimony* in Text Sample 5 appears in the *that* clause that modifies the head noun phrase *oral or written statements, documents or any other material*. *Wh-* relative clauses most often appear with the relativizer *who*, which is followed by a description of a human (see Text Sample 5) (head nouns bolded, relative clauses italicized).

**Table 10**  
Criminal Offenses Key Feature Results.

Feature	Cohen's <i>d</i>
predicative adjectives	1.63
<i>wh-</i> relative clauses	1.53
indefinite articles	1.38
human nouns	.80
present tense	.66
clausal coordinating conjunctions	.65
adverbs	.51
infinitive	.50
<i>to</i> verb complement clause	.35
3rd person pronoun	.31
<i>wh-</i> relative clause, prepositional fronting (pied-piping)	.26
verbs of occurrence	-.21
'be' verb	-.21
proper nouns	-.21
place adverbs	-.22
modals of necessity	-.22
subordinating conjunction - other	-.22
<i>that</i> relative clause	-.24
adjective complement clause	-.26
past tense	-.26
adverb (downtoners)	-.26
suasive verbs	-.26
attributive adjectives	-.28
<i>that</i> verb complement	-.30
word count	-.32
agentless passive	-.36
process nouns	-.36
subordinating conjunctions - conditional	-.37
prepositions	-.41
definitions verbs	-.49
non-finite <i>-ing</i> clause (post-nominal modifier)	-.52
non-finite <i>-ed</i> clause (passive post-nominal modifier)	-.52
nominalization	-.65
pre-modifying nouns	-.76
modals of possibility	-.77
definite articles	-1.10
modals of prediction	-1.11

**Table 11**  
Procedural Guidelines Key Feature Results.

Feature	Cohen's <i>d</i>
definite articles	.97
modals of prediction	.69
word count	.65
subordinating conjunctions - conditional	.46
agentless passives	.29
<i>that</i> verb complement clause	.25
perfect aspect	.20
<i>that</i> relative clauses	-.20
predicative adjectives	-.21
process nouns	-.22
proper nouns	-.23
abstract nouns	-.36

**Text Sample 5:** AZ Rev. Stat. §13–2801 (2019)

In this chapter, unless the context otherwise requires:

1. “Juror” means **any person** *who is a member of any impaneled jury or grand jury*, and includes **any person** *who has been drawn or summoned to attend as a prospective juror*.
2. “Official proceeding” means a proceeding heard before any legislative, judicial, administrative or other governmental agency or official authorized to hear evidence under oath.
3. “Physical evidence” means any article, object, document, record or other thing of physical substance.
4. “Testimony” means **oral or written statements, documents or any other material** *that may be offered by a witness in an official proceeding* ( . . )

Finally, the present tense is key in Operational Definitions as one of the primary functions of the present tense is to state a general truth or permanent state of being.

There are thirteen features with negative effect sizes, primarily representing verb-associated language (*to-* and *that-* verb complement clauses, the infinitive, *be* verbs, activity verbs, *suasive* verbs). All three semantic domains of modal verbs are also markedly less frequent in Operational Definitions: necessity modals ( $d = -.22$ ), possibility modals ( $d = -.65$ ), and prediction modals ( $d = -1.17$ ). The lack of verb-associated language in Operational Definitions can be attributed to the fact that definitions do not directly regulate human *actions*. Instead, they are concerned with objects, people, and ideas.

### 3.3. Functional interpretation of Criminal Offense key features

There are eleven positive key features and twenty-five negative key features in Criminal Offenses, many with medium and large effect sizes. This makes Criminal Offenses the most lexico-grammatically distinct functional type of statutory provision. Of the positive key features, four have large effect sizes, four have medium effect sizes, and three have small effect sizes. This analysis will focus on the key features with medium and large effect sizes, including predicative adjectives ( $d = 1.63$ ), *wh-* relative clauses ( $d = 1.53$ ), indefinite articles ( $d = 1.38$ ), human nouns ( $d = .80$ ), present tense ( $d = .66$ ), clausal coordinating conjunctions ( $d = .65$ ), adverbs ( $d = .51$ ), and the infinitive ( $d = .50$ ).

Predicative adjectives have the largest effect size due to the pervasive formulaic sequences *[X] is guilty of* and *[X] is unlawful* (see Text Sample 6). *[X] is guilty of* is used in a substantial majority of Criminal Offense provisions (89 %) to introduce the level of the offense (a misdemeanor, a felony, etc.). About 22 % of the Criminal Offense provisions use the sequence *[X] is unlawful*, which is used to explicitly indicate that an action is prohibited before indicating that the offense rises to the criminal level (predicative adjectives bolded).

#### Text Sample 6: AZ Rev. Stat. §16–1005 (B)

B. It is **unlawful** to offer or provide any consideration to acquire a voted or unvoted early ballot. A person who violates this subsection is **guilty** of a class 5 felony.

*Wh-* relative clauses are primarily used to describe a criminal action or to identify a human agent. In Criminal Offense provisions, this most often appears as *a person who [action/description]*. In Text Sample 7, two of the *who* clauses carry a description of the criminal action, and one provides identifying details about a human referent. Of these uses, the description of a criminal act is most common (head noun bolded, relative clauses italicized).

#### Text Sample 7: AZ Rev. Stat. §13–3623 (2019) (D)

D. A **person** *who intentionally or knowingly engages in emotional abuse of a vulnerable adult who is a patient or resident in any setting in which health care, health-related services or assistance with one or more of the activities of daily living is provided* or, having the care or custody of a vulnerable adult, *who intentionally or knowingly subjects or permits the vulnerable adult to be subjected to emotional abuse* is guilty of a class 6 felony.

Criminal Offenses use indefinite articles *a/an* to communicate broad application. They indicate that the offense is considered criminal for all citizens (e.g., *a person who is under the legal drinking age...*). Note also that indefinite article *the*, which is used to reference specific agents, has the second largest negative effect size in the sub-corpus ( $d = -1.10$ ).

Clausal coordinating conjunctions are associated with the long-winded, multi-clausal nature of statutory writing (Tiersma, 1999). In Criminal Offenses, they are most often used to connect two or more clauses together that contain an independent offenses related to the same subject matter (Text Sample 8; clausal coordinating conjunctions

bolded).

#### Text Sample 8: AZ Rev. Stat. §5–115 (2019) (D)

D. Any person who *holds or conducts* any racing meeting **or operates** an additional wagering facility without first complying with the provisions of this article, **or** any person who *fails to submit* to a drug test as directed by stewards **or** who *violates any other provision* of this article for which no other penalty is prescribed, is guilty of a class 2 misdemeanor.

Criminal Offenses rely on adverbs to describe the manner in which an action is carried out. This often relates to knowledge or intent. For example, Text Sample 9 states multiple times that the perpetrator must *knowingly* commit the act for it to be criminal. This is significant as it indicates that if the act was *not* done knowingly, the offense is not criminal. The occasional archaic adverb that remains from historical drafting (Williams, 2007) also appears in the text sample below (e.g., *therein, thereof*).

#### Text Sample 9: AZ Rev. Stat. §16–1016 (2019)

A person is guilty of a class 5 felony who: (. . .)

7. *Knowingly* adds a ballot to those *legally* cast at any election, by *fraudulently* introducing the ballot into the ballot box either before or after the ballots *therein* have been counted.
8. *Knowingly* adds to or mixes with ballots *lawfully* cast, other ballots, while they are being canvassed or counted, with intent to affect the result of the election, or to exhibit the ballots as evidence on the trial of an election contest.
9. *Knowingly* and *unlawfully* carries away, conceals or removes a poll list, ballot or ballot box from the polling place, or from possession of the person authorized by law to have custody *thereof*.

Criminal Offenses make use of the present tense to state general truths. These provisions use series of present tense verbs to list criminal acts, which safeguards against potential debate about sense meanings (Text Sample 10). As noted in Section 3.2, this sometimes results in the use of words that are typically regarded as synonyms (e.g., *change, alter*).

#### Text Sample 10: AZ Rev. Stat. §5–391 (2019) (G)

G. A person is guilty of a class 6 felony who knowingly *removes, defaces, obliterates, changes, alters* or *causes* to be removed, defaced, obliterated, changed or altered a factory, engine, serial, outdrive, lower unit, power trim or hull identification number or mark on a watercraft.

Criminal Offenses use infinitive clauses to describe criminal acts, similar to the way that relative clauses are used in this functional type (Text Sample 11; infinitive clauses italicized).

#### Text Sample 11: AZ Rev. Stat. §17–339 (2019) (B)

B. The failure *to make such return within thirty days* shall automatically cancel the license dealer's license, and intentional failure *to comply with the provisions of this section* is a class 1 misdemeanor. Any license dealer who makes a false or fraudulent return or report or who fails *to submit returns, reports, or all due and unremitted fees* as required under this section with the intent of defrauding the department is guilty of a class 6 felony.

Human nouns appear in Criminal Offenses when the guilty party is referenced. The frequency of use can largely be attributed to the word *person*.

There are twenty-five negative key features, seven of which have medium or large effect sizes. These features include modal verbs of prediction and possibility ( $d = -1.11$ ,  $d = -.77$ ), definite articles ( $d = -1.10$ ), pre-modifying nouns ( $d = -.76$ ), nominalization ( $d = -.65$ ), and non-finite *-ed* ( $d = -.52$ ) and *-ing* clauses ( $d = -.52$ ). Modal verbs



are comparatively infrequent in Criminal Offenses because communicating obligation (denoted by *shall*) and authorization (denoted by *may*) are not goals of a Criminal Offense provision. The infrequent use of definite articles reflects the absence of reference to specific individuals in Criminal Offenses, which instead aim for inclusivity of referents. It is notable that there is such a high number of negative key features, as this indicates that Criminal Offenses rely on a relatively narrow number of features to fulfill their function, relative to the other functional types.

### 3.4. Functional interpretation of Procedural Guideline key features

Of the three grammatically distinct functional types, Procedural Guidelines are the least distinct. However, due to the comparatively larger number of key features and higher effect sizes in relation to Permissions, Duties, Prohibitions, and Impersonal Rules, this functional type will be briefly discussed here.

There are seven positive key features in Procedural Guidelines. Definite articles have the largest effect size ( $d = .97$ ), followed by prediction modal verbs ( $d = .69$ ) and word count ( $d = .65$ ). The remaining four features have small effect sizes, including conditional subordinating conjunctions ( $d = .46$ ), *that* verb complement clauses ( $d = .25$ ), agentless passive verbs ( $d = .29$ ), and perfect aspect ( $d = .20$ ). The former two features indicate that the texts are highly clausal in nature.

Definite articles are markedly more common in Procedural Guidelines due to the highly specific contexts in which guidelines are prescribed. Procedural Guidelines are generally only relevant under specific circumstances, and only pertain to certain individuals and entities. In Text Sample 12, definite articles are used to identify the agents and objects referred to in the provision (e.g., *the person, the department of health services, the test results, the last known name and address*) (definite articles bolded, head nouns italicized).

#### Text Sample 12: AZ Rev. Stat. §13–1415 (2019) (E)

E. **The submitting entity** or **the department of health services** shall notify **the person** tested of **the results of the test** conducted pursuant to subsection B of this section and shall counsel **the person** regarding **the health implications** of the results. If **the submitting entity** does not notify **the person** tested of **the test results**, **the submitting entity** shall provide both **the name and last known address** of record of **the person** tested and **the test results** to **the department of health services** or a local health department for notification purposes.

Modals of prediction are key due to the pervasive use of *shall*. In these provisions, *shall* is used to communicate obligation of an action. This modal verb appears in a procedural guideline whenever a task or action is assigned to an agent in the provision. Prediction modal verbs appear multiple times in a single procedural guideline for two reasons: (1) there are often multiple actors that serve different roles, and (2) there are multiple steps in a procedure (an inclusion criterion for Procedural Guidelines). This is demonstrated in Text Sample 13, which references three independent agents with different roles (*the mental health expert, the clerk of the court, and the defense attorney*), and includes four distinct steps.

#### Text Sample 13: AZ Rev. Stat. §8–291.10 (2019) (A)

A. The mental health expert who consults with the restoration program **shall submit a written report** to the court before any hearing that is held pursuant to this section. The clerk of the court **shall seal and file the original report**. The mental health expert **shall provide a copy of the report** to the defense attorney for redaction. Within twenty-four hours after receiving a copy of the report, the defense attorney **shall provide copies of the redacted report** to the state and the court (. . .).

Conditional subordinating conjunctions (*if, unless*) are used to

describe the circumstances under which a procedure or a step in a procedure take place. For example, in Text Sample 14, *if* is used to communicate that the use of a container should be withheld pending a hearing – but *only* if the division directs so. Conditional subordinating conjunctions are also used to introduce contingencies.

The tendency to embed multiple conditional clauses that are dependent upon one another also likely contributes to the keyness of word count in Procedural Guidelines. This clausal embedding is demonstrated in Text Sample 14 (conditional clauses underlined).

#### Text Sample 14: AZ Rev. Stat. §3–2155 (2019) (C)

C. (. . .) *If the person using or proposing to use the marking, labeling or container does not accept the determination of the division*, such person may request a hearing, but the use of the marking, labeling or container shall, *if the division so directs*, be withheld pending a hearing and final determination by the director. Any such determination by the director shall be conclusive *unless*, within thirty days after receipt of notice of the final determination, *the person adversely affected appeals to the superior court*.

Agentless passive constructions are used when an agent is either unknown or unimportant. In Procedural Guidelines, agentless passives are used when the provision references events that occurred at a time prior to the procedure. It generally does not matter who the agents were in these events. This is demonstrated in Text Sample 15, which prescribes a procedure for obtaining and transmitting a blood sample from a detainee. The provision uses several agentless passive verb constructions to describe past events that necessitate the procedure described (agentless passive constructions italicized).

#### Text Sample 15: AZ Rev. Stat. §13–610 (2019) (B)

B. Within thirty days after a person *is placed on probation and sentenced to a term of incarceration in a county jail detention facility or is detained in a county juvenile detention facility*, the county detention facility shall secure a sufficient sample of blood or other bodily substances for deoxyribonucleic acid testing and extraction from the person if the person *was convicted of or adjudicated delinquent for an offense listed in this section*. (. . .).

*That* verb complement clauses complete meaning relationships with a preceding dependent clause. In Procedural Guidelines, they most often describe actions that must be carried out in the provision. In Text Sample 16, the verbs *recommends, direct, and order* are reliant upon the complement clauses that they control to communicate exactly what actions must be carried out (e.g., what action is recommended, what action is directed, etc.) (verb complement clauses italicized).

#### Text Sample 16: AZ Rev. Stat. §13–4092 (2019) (C)

C. If the certificate recommends *that the witness be taken into immediate custody* and delivered to an officer of the requesting state to assure his attendance in the requesting state, such judge may, in lieu of notification of the hearing, *direct that such witness be forthwith brought before him for the hearing*; and the judge at the hearing being satisfied of the desirability of such custody and delivery, for which determination the certificate shall be prima facie proof of such desirability may, in lieu of issuing subpoena or summons, *order that the witness be forthwith taken into custody* and delivered to an officer of the requesting state.

Procedural Guidelines rely on the perfect aspect to refer back to events that took place at any time leading up to the present moment. They generally reference past events that must occur prior to the procedural steps prescribed (Text Sample 17) (perfect aspect italicized).

#### Text Sample 17: AZ Rev. Stat. §27–1296 (2019) (B)

B. (. . .) After the reclamation measures *have been performed*, the remaining financial assurance shall be released, except that ten per cent shall be retained for the costs of care, monitoring and one reseeding, if necessary, for areas that *have been revegetated*. The inspector shall release the retained monies after a period of not more than three growing seasons after the supplemental management or other man-induced inputs *have been finally removed* or as otherwise provided in Sections 27–1292, subsection B.

Procedural Guidelines have five negative key features, including three semantic domains of nouns (abstract nouns,  $d = -.36$ ; proper nouns,  $d = -.22$ ); process nouns,  $d = -.23$ ), followed by predicative adjectives ( $d = -.21$ ) and *that* relative clauses ( $d = -.20$ ). It is unsurprising that nouns are generally less frequent relative to the remainder of the corpus as Procedural Guidelines principally concern human actions, rather than objects, concepts, or ideas. An abundance of verb-associated language is typically associated with the decreased use of nominal language.

The other two negative key features, predicative adjectives and *that* relative clauses, are less common in Procedural Guidelines because they are primarily used for description and identification, neither of which play a role in prescribing procedures.

To recap the findings of the present study and look towards future research, the following section will cover two topics of discussion: (1) the identification of the functional types of statutory provisions and decisions that went into functional classification, and (2) the implications of the findings presented above. Based on these two discussions, avenues of future research are proposed.

## 4. Discussion and conclusion

### 4.1. Reflections on the functional type framework

This study represents the first attempt to empirically classify statutory provisions by their function; however, this is not the primary aim of the study. While I am highly confident in the functional types identified in the present study and the soundness of the coding rubric, there is an opportunity for future researchers to bring a fresh perspective to certain aspects of the framework. First, there are inevitable instances of overlap between select functional types. Notably, Procedural Guidelines are conglomerates of other functions, frequently containing both duties and permissions within them. Consider the two text samples below. Text Sample 18 (coded as a Procedural Guideline) describes a set of sequential procedural steps applicable in a specific context. Text Sample 19 (coded as a Duty) describes a responsibility of the supreme court that is associated with their role in government.

**Text Sample 18:** AZ Rev. Stat. §49–403 (2019) (A)

A. A person may petition the director or control officer for a determination that a particular class or category of sources should be subject to a general permit instead of an individual permit that is issued under this chapter. The petition shall state the grounds for the determination that is the subject of the petition, including how the class or category meets the criteria prescribed in the applicable statute or rule for a general permit. The director or control officer shall either grant or deny the petition within sixty days after its receipt. If the petition is granted, the director or control officers shall initiate the formal process for issuing the general permit within six months. If the petition is denied, the denial is an appealable agency action pursuant to title 41, chapter 6, article 10.

**Text Sample 19:** AZ Rev Stat § 12–102 (2019) (A)

A. The supreme court shall discharge the duties imposed and exercise the jurisdiction conferred by the constitution and by law.

While a duty is prescribed in both text samples, it would be difficult

to argue that Text Sample 18 should be considered a simple duty, or that Sample 19 should be considered a set of procedural steps. They serve distinct functions, both of which are pervasive in the corpus. While this overlap likely decreased the effect sizes of select key features in these two functional types, this was accepted as a limitation in order to be true to the aim of identifying distinct functions. With this in mind, there is room for future researchers to approach this issue of overlap from a different perspective.

Next, researchers may chose to re-evaluate the ‘Other’ category, which represents ~21 % of the total provisions in the corpus. While ‘Other’ provisions accounted for a larger proportion of the corpus than predicted, this is not viewed as a limitation of the study. Allowing provisions to be classified as ‘Other’ is an important methodological aspect of the study as they serve distinct functions not represented in the framework. It follows that future iterations of this framework may include functional types of provisions that are presently classified as ‘Other’.

Moving forward, I believe that it would be advantageous for legal experts take on a larger role in developing future iterations of this framework.

### 4.2. Implications and future research

The findings of this study have the potential to impact the way that we approach legal interpretive questions. Results reveal that a great deal of lexico-grammatical variation exists within the register of statutory law based on the communicative function of the provision. This suggests that it may be beneficial to consider statutory provisions affiliated by function when investigating the meaning of contested words, phrases, or clauses. In application, a researcher using a corpus of statutory law to answer a legal interpretive question may simply narrow their search to a specific functional type. This would allow the researcher to analyze texts that more are situationally consistent with the target provision, which is generally regarded as a highly important part of sound corpus analysis. This would also increase confidence that the linguistic patterns identified are consistent with those in the provision containing the contested language. Consideration of function will also minimize the risk of drawing conclusions about linguistic patterns in statutory provisions that are linguistically dissimilar.

That said, I do not intend to suggest that the subject matter of a provision is in any way invalid, unreliable, or unimportant in statutory interpretation. Likewise, I do not make the claim that the functional type of the statute is in any way a superior measure of contextually similar provisions. There are two reasons for this. The first is a legal one: courts have historically relied upon topically related (or directly cited) provisions to resolve contested meaning, to great effect. For this reason, it would be illogical to ignore topic in the consideration of context. The second reason is a linguistic one: the subject matter of a text is also regarded as a distinct situational characteristic that may influence linguistic preference in a text. From a linguistic perspective, entirely ignoring the subject matter of a statutory provision may be equally as problematic as ignoring the function. In light of this, I suggest that a useful next step is to investigate the relationship between the topic of a statutory provision and the typical lexico-grammatical characteristics. As many state codes are organized by subject matter (e.g., Elections, Property, Marital and Domestic Relations, etc.) this would remove the time-consuming step of developing a framework and coding the provisions. One could also examine the interplay between function and topic in statutory law. Wood (2023) briefly examines this relationship, analyzing the proportional breakdown of functional types of statutory provisions in the forty-seven titles of the Arizona State Code. Notably, this was a qualitative, *non-linguistic* analysis, which leaves room for future research exploring the linguistic relationship between function and topic in statutory law, and its effect on language use.

## CRediT authorship contribution statement

**Margaret Wood:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

I would like to thank Jesse Egbert, Douglas Biber, Randi Reppen, and Thomas R. Lee for their feedback and guidance throughout this study. I gratefully acknowledge the support of the U.S. Department of Energy through the LANL/LDRD Program and the Center for Non-Linear Studies for this work.

## References

- Biber, D., Conrad, S., 2019. *Register, Genre and Style*, 2nd ed. Cambridge University Press.
- Biber, D., Egbert, J., 2018. Register Variation Online. Cambridge University Press.
- Biber, D., Egbert, J., 2023. What is a register?: accounting for linguistic and situational variation within—and outside of—textual varieties. *Regist. Stud.* 5 (1), 1–22.
- Biel, L., 2014. The textual fit of translated EU law: a corpus-based study of deontic modality. *Translator* 20 (3), 332–355.
- Biel, L., 2015. Phraseological profiles of legislative genres: complex prepositions as a special case of legal phrasemes in EU law and national law. *Int. J. Spec. Commun.* 37 (3–4), 139–160.
- Biel, L., 2017. Lexical bundles in EU law: the impact of translation process on the patterning of legal language. *Phraseology in Legal and Institutional Settings. A Corpus-Based Interdisciplinary Perspective*. Routledge, London/New York, pp. 10–26.
- Bowers, F., 1989. *Linguistic Aspects of Legislative Expression*. University of British Columbia Press, Vancouver.
- Bulatović, V., 2013. Legal language: the passive voice myth. *ESP Today* 1 (1), 93–112.
- Cao, D., 2007. Legal speech acts as intersubjective communicative action. Wagner, A., Werner, W., & Cao, D. (Eds.). *Interpretation, Law and the Construction of Meaning*. Springer, Dordrecht, pp. 65–82.
- Clarke, I., 2022. A multi-dimensional analysis of English tweets. *Lang. Lit.* 31 (2), 124–149.
- Cohen, J., 1977. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, New York.
- Cooper, P.K. (2011). Is there a case for the abolition of 'shall' from EU legislation. Riga: riga Graduate School of Law.
- Egbert, J., Biber, D., 2023. Key feature analysis: a simple, yet powerful method for comparing text varieties. *Corpora* 18 (1).
- Egbert, J., Wood, M., 2023. The corpus of united states state statutes—design, construction and use. *Appl. Corpus Linguist.* 3 (2), 100047.
- Fan, H., 2020. A corpus-based study on the English translation of modal verbs in Chinese legislative texts. In: *Proceedings of the Conference on Education, Language and Inter-Cultural Communication (ELIC 2020)*. Atlantis Press, pp. 461–465.
- Foley, R., 2022. Legislative language in the EU: the crucible. *Int. J. Semiot. Law* 15 (4), 361–371.
- Gibova', K. (2011). On modality in EU institutional-legal documents. *English Matters II: A Collection of Papers by the Institute of British and American Studies Faculty*, 6–12.
- Giczela-Pastwa, J., 2019. Inverse legal translation: a corpus-driven study of multi-word units related to the structure of translated statutory provisions. Biel, L., Engberg, J., Ruano, R.M., Sosoni, V. (Eds.). *Research Methods in Legal Translation and Interpreting*. Routledge, pp. 48–65.
- Go'zd'z-Roszkowski, S., 2011. Patterns of Linguistic Variation in American Legal English: A Corpus-Based Study. Peter Lang, Frankfurt am Main.
- Gunnarsson, B.L., 1984. Functional comprehensibility of legislative texts: experiments with a Swedish act of parliament. *Text Interdiscip. J. Study Discourse* 4 (1–3), 71–106.
- Hiltunen, R., 2012. The grammar and structure of legal texts. Solan, L. & Tiersma, P. (Eds.). *The Oxford Handbook of Language and Law*. Oxford University Press, pp. 39–51. <https://doi.org/10.1093/oxfordhb/9780199572120.013.0004>.
- Kurzon, D., 1986. It is Hereby Performed. Explorations in Legal Speech Acts. John Benjamins. *Labor Laws and Issues*. USA.GOV. Retrieved on March 22, 2023 from. <https://www.usa.gov/labor-laws>.
- Liimatta, A., 2022. Register variation across text lengths: evidence from social media. *Int. J. Corpus Linguist.* 28 (2), 202–231.
- Macagno, F., 2010. Definitions in law. *Bull. Suisse Linguist. Appl.* 2, 199–217.
- Ni, S., Sin, K.K., 2011. A matrix of legislative speech acts for Chinese and British statutes. *J. Pragmat.* 43 (1), 375–384.
- Özyildirim, I., 2011. A comparative register perspective on Turkish legislative language. In *law and language in partnership and conflict*. *Lapl. Law Rev.* 1 (1), 79–94.
- Šarčević, S., 1997. *New Approach to Legal Translation*. Kluwer Law International BV.
- Sun, Y., Cheng, L., 2017. Linguistic variation and legal representation in legislative discourse: a corpus-based multi-dimensional study. *Int. J. Leg. Discourse* 2 (2), 315–339.
- Tiersma, P.M., 1999. *Legal Language*. The University of Chicago Press, Chicago.
- Williams, C., 2004. Pragmatic and cross-cultural considerations in translating verbal constructions in prescriptive legal texts in English and Italian. *Textus* 17, 217–245.
- Williams, C., 2007. Tradition and Change in Legal English: Verbal constructions in Prescriptive Texts. Peter Lang. Vol. 20.
- Williams, C., 2013. 14 Changes in the verb phrase in legislative language in English. *The Verb Phrase in English: Investigating Recent Language Change with Corpora*, pp. 353–371.
- Wood, M., 2022. A lexico-grammatical comparison of statutory law and popular written language. *Res. Corpus Linguist.* 10 (2), 16–45.
- Wood, M., 2023. *Communicative Function and Linguistic Variation in State Statutory Law*. Northern Arizona University (Doctoral dissertation).
- Yankova, D., 2017. An attempt at redefining legal English contexts. *Int. J. Lang. Stud.* 11 (4), 173–191.