

Corpus-based Research on Variation in English Legal Discourse

Teresa Fanego and Paula Rodríguez-Puente, eds (2019)

John Benjamins viii + 294 pp

Reviewed by Zhiying Xin and Jiawei Wang

This volume, an addition to the *Studies in Corpus Linguistics* series, collates insightful and innovative articles and examines linguistic variation across a comprehensive range of legal discourses using advanced corpus tools to reveal synchronic and diachronic changes in legal language. *Corpus-based Research on Variation in English Legal Discourse* focuses on the *Corpus of Historical English Law Reports 1535–1999 (CHELAR)* and distinguishes itself in the way that it characterises linguistic representations in discipline-specific corpora while exploring a multitude of sub-genres of forensic discourse.

The volume encompasses an introductory chapter, summarising the past and present of English language forensic linguistics, and ten chapters divided into two parts, covering cross-linguistic and diachronic perspectives of comparative studies in the domain of legal discourse.

The first chapter commences with a brief review of the Fiction à Substrat Spécialisé and Plain Language movements as two cornerstones of forensic linguistics, then discusses how Biber's (1988) factor analysis and Multi-dimensional Model, together with Swales's (1981) move analysis and Bhatia's (1987) classification of text types paved the way for the enormous expansion of register and genre approaches to legal discourse. At the end of Chapter 1, the authors concur that

Affiliation

Xiamen University, PRC

email: xinzhiying@xmu.edu.cn wangjiawei@stu.xmu.edu.cn

large-scale corpora can be deployed as a plausible way of assessing the validity of direct observations.

Chapters 2 to 11 fall into two parts: ‘Cross-genre and cross-linguistic variation’ and ‘Diachronic variation’. However, some chapters go beyond this scope, excavating underlying mechanisms of linguistic phenomena inside and outside the language or across time and space, a categorisation that attempts to reveal the comprehensiveness and multiple facets of the compilation. A more unified system of categorisation is proposed by Biel (2010), classifying the domain of variation studies into four trajectories based on the nature of corpora. Trajectories 1 and 2 are external and internal variations that emphasise how legal language is distinct from other genres while proposing a diverse range of legal text types. Cross-linguistic and temporal variation in Trajectories 3 and 4 define themselves in revealing spatial and temporal discrepancies. Following this trail of thought, chapters with similar themes will be grouped together and reviewed within these trajectories.

Trajectory 1: Cross-linguistic variation in legal discourse

Using two self-compiled corpora of English and Italian land contracts collected from websites, Diani in Chapter 2 presents a relatively holistic picture of the similarities and differences between English and Italian legal instruments. Land contracts in both English and Italian share parallel macrostructures such as titles, main textual bodies, signatures and dates, but differ in other components such as introductions, recitals, operative provisions and punctuation. The use of cohesive devices and definite articles, abundance of compound adverbs, formulaic collocations, subordinate clauses, nominalisations, passive voice and overuse of some modal verbs appears to be shared, but Italian contracts are less formulaic than English, and variance between Italian and English contracts is manifested in both textual organisation and linguistic structure.

The cross-linguistic study in Chapter 3 compares the use of *if*-conditionals in spoken courtroom and parliament discourses in English, French and Spanish. Data extracted from the International Corpus of English, the Integrated Reference Corpora for Spoken Romance Languages and the Hansard Corpus are processed to observe the frequencies, metafunctions, semantic types, modal verbs and markedness in courtroom and parliament discourses. In contrast to textual and interpersonal conditionals, ideational conditionals predominate in both discourses across the three languages to convey a condition under which an event is enabled. Modal verbs in *if*-conditionals prevail in English, with alternative grammatical mechanisms in place for Spanish and French legal discourses. Marking of apodoses is absent in the parliamentary discourse of all three languages, suggesting the tone of parliamentary discourse is less categorical than that of the courtroom discourse.

Trajectory 2: Internal variation in legal discourse

In search of grammatical categories, Breeze in Chapter 4 mandates a part-of-speech analysis for four sub-genres of business law (academic articles, case law, commercial contracts and legislative commercial law). The distribution of grammatical categories brings to light overarching differences between the four genres in the following respects: the universal use of possessives; relative clauses, *that*-clauses and labelling markers as cohesive devices; the formulaic nature of legislative texts; and vagueness and inclusiveness surrounding *and* and *or*. Concurring with earlier results in Biber and Gray's (2016) study, the prolific use of relative clauses and tenses explains the discursive requirements of producing complex argumentations in legal academic articles and case laws. Irregular patterns of cohesion and modality in commercial contracts and legislative laws are characterised by the functional disciplinarity of both sub-genres.

Chapter 5 investigates the 50 most frequent three-word lexical bundles in spoken courtroom corpora obtained from the O. J. Simpson Trial, Early Modern British English and Present-Day British English. Five shared bundles across the three corpora are found to serve the same discursive purpose of offering information and specification, as well as addressing the audience. Bundles exclusively deployed in the Trial corpus are 'specific reference' and 'other-directed' bundles. The presence of such bundles can be attributed to the outcome of the specific content of the Trial. Based on these results, Reppen and Chen further disassemble Trial discourse into four sub-genres (opening statement, direct examination, cross-examination and closing argument) and discuss the frequency of bundles shared across the sub-genres. This study showcases how situational and functional goals of communication may influence lexical and grammatical choices.

Trajectory 3: External variation in legal discourse

In Chapter 6, nouns followed by *that*-clause complements are scrutinised to explore evaluation and stance in three corpora, i.e., academic journals, US Supreme Court opinions and British Law Reports. Processed by *Wordsmith Tools* (Scott 2012) and *Sketch Engine* (2019), the data analysis confirms previous assumptions that the majority of noun *that*-patterns in judicial genres are epistemic nouns, either expressing certainty or likelihood. The frequencies of epistemic use of *fact that*, *conclusion that*, *notions that* and *assumptions that* are similar in academic and judicial texts, because both discourses require a high level of intertextuality to maintain effective communications in discussing legal decisions. Compared to the construction of stances in judicial discourse, the demand of establishing neutrality and objectivity in academic discourse is greater and can be met with the employment of noun *that*-clauses.

Considering polarisation and economy as competing factors influencing language patterns in the corpus of law reports, fiction, newspapers and scientific prose, Biber and Gray in Chapter 7 track historical changes in linguistic features (nouns, verbs, adjectives, adverbs, passive verbs, nominalisations, prepositional phrases and average word length), colloquial features (modal and semi-modal verbs, contractions, aspect markings, personal pronouns, *get* passives and phrasal verbs), clausal complexity features (dependent clauses) and in phrasal complexity (attributive adjectives, nouns as nominal pre-modifiers and prepositional phrases as nominal post-modifiers). Quantitative research shows that nominalisations, semi-modals, noun *that*-complement clauses, and nouns as nominal pre-modifiers are dominant in *CHELAR*. Qualitative analyses explicating the balance between popularisation and economy hinge on the communicative and situational needs of law reports. Ample evidence gathered from quantitative and qualitative investigations shows legal reports to be resistant to diachronic change.

Trajectory 4: Temporal variation in legal discourse

Chapter 8 delves into a historical analysis of the use of personal nouns in *CHELAR* from 1535 to nowadays. Following Biber's (1988) taxonomy, Rodríguez-Puente delineates the distribution of personal nouns according to type, number and gender in *Wordsmith Tools* (Scott 2012), and tests the significance of the above results in *R* software. The use of anaphoric masculine third person singular references is widely attested in law reports, as compared with usage in parliamentary acts, proclamations and statutes. And the descriptive and narrative nature of law reports calls for both prescriptive and descriptive expressions to convey objectivity and subjectivity. The temporal patterns of first person, second person, and third person pronouns suggest that legal discourse has become more interpersonal and subjective over the course of time.

In Chapter 9, Groom and Grieve examine the types and sequences of rhetorical moves of 130 British patent specification texts between 1711 and 1860. Their hypothesis is that moves utilised by successful applicants are more likely to be re-used in future applications, though the over-popularity of these moves undermines their potential to be selected afterwards. The move type analysis denotes 17 moves across the years, and identifies 5 obligatory moves (Royal Grant, Condition Statement, Disclosure Statement, Invention Description and Witness) for their pervasiveness in the corpus. A core pattern of move sequences is found, i.e., Salutation–Royal Grant–Condition Statement–Disclosure Statement–Invention Description–Witness–Enrolment Confirmation. The authors generalise these empirical findings in legal texts, and argue that synchronic variations, diachronic reproductions and move selections are pillars of the architecture of language evolution.

The purport of Chapter 10 is to compare the social identities of citizens and monarchy in Acts of Parliament from 1800 to 2000. Being categorised by topics, the collocations of *Person, majesty, king, queen, crown and sovereign* are retrieved by *AntConc* (Anthony 2014) and further divided into 13 classes (legal actors, legal languages/processes, actions, crimes and punishments, works, possessions, welfares, relations to groups of people, types of person, numbers, times, locations and collocates with two or more of the topics). The semantic preference and semantic prosody of these words illustrate that linguistic portrayals of citizens as troublemakers gradually give way to constructions focusing on their social welfare, whereas the attitudinal meaning of the monarchy shifts from power and authority to nation and state. Hence negative semantic construction of citizens is replaced by positive lexical profiles, while the construction of monarchy changes from positive to neutral. Possible explanations include a desire for a more discreet legal writing style amid the launching of National Insurance Acts at the beginning of the twentieth century.

In a similar vein, Claridge in Chapter 11 investigates the historical changes of words meaning 'drunk' in courtroom proceedings from 1720 to 1913. The overlap of intoxication expressions among victims, defendants, witnesses, judges and lawyers are attributed to the effects of lexical priming and repetition. Generally speaking, law professionals and witnesses use more amplifiers and downtoners to refer to third and second person targets. Conversely, victims and defendants tend to use intensification and downplaying of drunkenness in referring to themselves. The purpose of negotiations between legal experts and laypersons in courtrooms is to downplay the degree of intoxication and to reduce responsibility for the crime.

As the leitmotif of the current collection is variation in legal discourse, each chapter applies its own lens to specific linguistic items in a diverse spectrum of legal genres either cross-linguistically or diachronically. To this end, these endeavours not only successfully interpret the historicist dynamics of the individual legal genres, but also prompt social accountability towards law enforcement. In addition, the studies of temporal-spatial variations incorporate morphology, syntax, semantics and pragmatics, representing a holistic landscape of language evolution. In this sense, how legal language evolves is compatible with the evolution of language in general and to some extent can be generalised to variations in other genres. Last but not least, the excellent introduction and application of 'ever more sophisticated computational tools' has transcended the 'original scope to encompass a much broader range of topics and methodologies' (p. 17). Throughout the book, a variety of auto-annotated corpus tools, language processors and statistical computer software are deployed to cater for the objectives of each study. Readability is enhanced with clearly illustrated tables, diagrams and figures, as

well as with updates on technical innovations in corpus linguistics.

The book nonetheless has its own weaknesses in aspects such as a lack of phonetic, phonological, multimodal or multilingual analysis, an imbalance between inter-lingual and intra-lingual studies, and confusion between the concepts of 'genre' and 'register'. Possible future research on variation in legal language might include analyses of sounds and gestures across languages, studies of variation between L1 and L2 speakers, and mechanisms of language evolution.

Corpus-based Research on Variation in English Legal Discourse may be of particular benefit to students and researchers in the fields of discourse analysis, forensic linguistics, historical linguistics and the sociology of law.

References

- Anthony, L. (2014) *AntConc*. Computer software. Tokyo: Waseda University. <https://www.laurenceanthony.net>
- Bhatia, V. K. (1987) Language of the law. *Language Teaching* 20(4): 227–234. <https://doi.org/10.1017/S026144480000464X>
- Biber, D. (1988) *Variation across Speech and Writing*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Biber, D. and Gray, B. (2016) *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511920776>
- Biel, Ł. (2010) Corpus-based studies of legal language for translation purposes: methodological and practical potential. In C. Heine and J. Engberg (eds) *Reconceptualizing LSP: Online Proceedings of the XVII LSP Symposium 2009*. Aarhus: Aarhus University.
- Scott, M. (2012) *WordSmith Tools*, version 6. Oxford: Oxford University Press.
- Sketch Engine*. (2019) <http://www.sketchengine.co.uk/>
- Swales, J. M. (1981) *Aspects of Article Introductions*. Birmingham: University of Aston.