Review

Patterns of Linguistic Variation in American Legal English: A Corpus-based Study
Stanislaw Goźdź-Roszkowski (2011)
Peter Lang 277 pages

Reviewed by Tammy Gales

Stanislaw Goźdź-Roszkowski's *Patterns of Linguistic Variation in American Legal English: A Corpus-based Study* is one of the newest volumes (22) in the Łódź Studies in Language series, edited by Barbara Lewandowska-Tomaszczyk. This book, which contains seven chapters, a bibliography, two appendices, and an index, presents a very well-written, thorough examination of variation in legal language. Goźdź-Roszkowski's primary goal is to "demonstrate that the universe of legal texts involves not only different situational characteristics of legal genres, such as different modes (speech, writing) and different production circumstances in which legal genres are created, different participants and the relations among them, or different communicative purposes, but that legal texts differ dramatically in terms of their linguistic characteristics" (p.11). By comparing legal language through a variety of corpus-based approaches, Goźdź-Roszkowski successfully demonstrates that what has traditionally been treated "as a largely monolithic phenomenon" (p.15), is, in reality, composed of a variety of genres that each contain highly systematic patterns of language use.

In the Introduction (Chapter 1), Goźdź-Roszkowski explores previous research on legal language, defines existing perspectives on categorizing text varieties, and outlines his own research within these traditions. First, while previous research has acknowledged that legal language is heterogeneous, scholars have primarily focused on describing lexico-grammatical features found across all legal categories (e.g., the use of passive voice, *shall*, archaic adverbs) or on describing one particular type of legal text (e.g., testament language or prescriptive legal texts). Thus, Goźdź-Roszkowski aims to fill this void by

describing both the "linguistic variation *within* legal language" and the "variation between legal language and other specialized languages" (p.16).

Second, in defining his method of categorizing texts, Goźdź-Roszkowski draws on Biber and Conrad's (2009) framework, wherein the terms *genre* and *register* refer to different approaches to the study of variation rather than different text types. For example, when analyzing a power of attorney document from a genre perspective, an outline of language characteristics emerges that includes the date the document was created, the name of the person who created it, and the designation of a person appointed to have power of attorney. From a register perspective, an analysis reveals that power of attorney documents pervasively contain lexical and grammatical features such as *to*-infinitive clauses, nominalizations, and prepositional phrases. Goźdź-Roszkowski utilizes both perspectives by using whole texts to examine lexico-grammatical features and specialized expressions and to identify the location of clusters of linguistic features within each text variety.

Finally, through a comprehensive linguistic description and functional analysis of variation among legal genres, Goźdź-Roszkowski proposes, after Bhatia's (2004) notion of disciplinary genres, that "what is commonly referred to as 'legal English' should be more accurately described as a system of related domain-specific genres, which vary widely in terms of patterning, understood here as recurring lexical and lexico-grammatical combinations discernible in large collections of authentic texts by means of quantitative and qualitative analytical techniques" (p.24). To this end, described below, he successfully achieves his goal.

Chapter 2, The Methods and the Corpus, offers a comprehensive description of the texts that comprise "the world of law" (p.27) and describes the methods used in their analysis. The American Law Corpus (ALC) contains more than 5,500,000 words from 687 texts across 7 legal genres: academic journals, briefs, contracts, legislation, opinions, professional articles, textbooks. From each category (except textbooks), Goźdź-Roszkowski used complete texts rather than excerpts and he aimed for representativeness by including at least 1 million words in each general category (e.g., legislation and contracts)

and 200,000-500,000 words in more narrowly-defined genres (e.g., professional articles and textbooks); all texts were randomly sampled from a range of written activities within American legal contexts.

The corpus was tagged using the Biber tagger and counted with Biber's tag-count program, which provides overall rates of frequencies for more than 150 linguistic features that occur in a corpus (Biber, 2006). WordSmith 5.0 was used to extract data from the corpus and all results were normed per 1000 words, except for the lexical bundles which were normed per 1,000,000 words.

Goźdź-Roszkowski used three corpus-based approaches in his study. First, he performed a keyword analysis (utilizing frequency cutoff points, statistical significance tests, and dispersion plots) to determine which words were significantly more common in one legal genre as compared to the others; second, he performed an analysis of lexical bundles, wherein he adopts the corpus-driven methods of Biber *et al.* (1999) and Biber *et al.* (2003) to uncover multi-word expressions that occur with a statistically high rate of frequency in legal genres; and third, he performed a Multi-Dimensional (MD) analysis (Biber, 1988, 1995; Conrad and Biber, 2001) to identify linguistic features that are statistically correlated and reveal the functions shared by those clusters of co-occurring features. In the final approach, Goźdź-Roszkowski both employs the dimensions from Biber's (1988) model and also performs a new MD analysis in order to "understand the specialized legal genres relative to a number of spoken and written genres or registers in English" and "establish the dimensions of variation valid for this domain of language use" (p.51), respectively.

Chapter 3 presents an innovative analysis of Vocabulary Use across Different Legal Genres. While the majority of research on vocabulary has prioritized legal terminology at the expense of other vocabulary and has not focused on its use, Goźdź-Roszkowski provides a more comprehensive examination of the lexical variation within legal genres. Starting with an analysis of word types, Goźdź-Roszkowski finds that there is a range of variation, with legislation and contracts on the low end of word types used and academic journals and professional articles on the high end. Upon further investigation, Goźdź-Roszkowski attributes the diversity in word choice to the use of specialized vocabulary, where the two genres that function to create and modify

legal relations—legislation and contracts—utilize the fewest types of specialized vocabulary.

The remaining part of the chapter identifies keywords relevant to each legal genre in context (Appendix A lists the 100 keywords for each genre) and then separates the keywords within each genre into functional categories such as "citation keywords," "keywords expressing evaluation," "self-mention keyword 'we'," and "legal terms as keywords" (p.65). Goźdź-Roszkowski finds, in general, that the only functional category shared by all legal genres is that of legal terms and that the categories each possess a range of keywords from highly-specialized (e.g., *Miranda*, *penalty*, and *treaty* in academic articles) to general language (e.g., *patient*, *university*, and *values* in the same genre). The categories proposed in this chapter reveal the major trends in lexical composition within each genre, from legislation and briefs possessing several classes of keywords that serve clearly functional purposes to professional articles and textbooks that are marked by terminological density, supporting Goźdź-Roszkowski's position that "the legal lexicon should not be perceived as consisting primarily of terms in the sense defined by the traditional theory of terminology," but as one in which the lexicon serves diverse roles within and across legal genres (p.107).

Recognizing the important relationship between multi-word expressions and genres, Goźdź-Roszkowski next examines Multi-Word Patterns in Legal Genres in Chapter 4. He starts "from the premise that the analysis of how multi-word expressions are used in contexts provides a reliable indicator of variation between different text types, genres and registers" (p.109). In legal genres, in particular, Kjaer (1990) pointed out that the failure to use specific legally-prescribed combinations of words can alter or invalidate the entire meaning of a document; however, very little research has been done on this highly important topic in legal genres. This chapter fills that void by investigating the distribution and use of lexical bundles within legal genres.

In order to produce findings that are comparable with previous research on lexical bundles, Goźdź-Roszkowski adopts parameters similar to those in related studies (e.g., Biber, 2006; Cortes, 2004; Hyland, 2008); in particular, he focuses on four-word sequences and

sets the frequency cut-off point at 40 times per million words. Bundles meeting the frequency requirement have to occur in at least five different texts. His analysis identified 915 bundles, with contracts possessing the highest percentage of words in bundles (10.1%). By examining the distribution of bundles among genres, Goźdź-Roszkowski reveals, for example, that when comparing genres of similar sizes (e.g., contracts, legislation, and opinions), contracts and legislation have the greatest range of different bundles, the largest number of bundles, and the highest proportion of words contained in bundles, pointing to the high degree of formulaicity and repetitiveness in these two genres. Opinions, on the other hand, possess a relatively small range of different bundles, but rely on the greatest proportion of frequent bundles, evidence of the genre's operative function in communicating judicial decisions, which are highly procedural in nature. The variation in distribution and reliance on lexical bundles found among genres leads to a more detailed examination of the structural characteristics of bundles.

Based on Biber *et al.*'s (1999) framework and his own prior research on lexical bundles in a 500,000 word corpus of judgments from the House of Lords (2006), Goźdź-Roszkowski identifies and analyzes the ten major structural characteristics of lexical bundles found within the ALC, which include, for example, noun phrase with *of*-phrase fragment, prepositional phrase expressions, verb phrase with active verb, and (verb phrase +) *that*-clause fragments. He finds very little variation across genre types in that they all make frequent use of noun phrase and prepositional phrase bundles—supporting the notion that legal genres are nominally dense—and infrequent use of verb phrase bundles.

The remainder of the chapter analyzes the functions of bundles as they are categorized within three broad types previously proposed by Biber *et al.* (2004), Biber (2006), and Hyland (2008): "lexical bundles marking legal reference, stance and text-oriented bundles" (p.117). Referential bundles, found with the highest proportion in legislation, contracts, and professional articles, directly reference abstract or physical concepts within the domain of law; stance bundles, which express attitudes or assessments, were found to vary most among genres with academic journals and opinions containing the highest

percentage of bundles; and text-oriented bundles, which signal relationships between textual segments, were found with the highest percentage in academic journals and opinions wherein a third of all bundles were used in this manner.

Goźdź-Roszkowski concludes that his findings "show considerable variations in the frequency of forms, structures and functions of lexical bundles across different types of legal writing" (p.142). Of particular note is the fact that legislation and contracts contained more formulaic expressions than any other genre to date. Also revealing were differences in shared bundles across genres (e.g., contracts shared less than 20% of its bundles with other genres, meaning more than 80% of the bundles were unique to that genre, while academic journals shared more than 80% of its bundles with other genres); these findings, in conjunction with previous lexical distribution and functional analyses, further demonstrate that legal genres are, in fact, highly distinct genres within the broad category of legal language.

In the final two analytic chapters—Chapter 5: Multi-Dimensional Variation across Different Genres and Disciplines and Chapter 6: Multi-Dimensional Patterns of Variation across Legal Genres—Goźdź-Roszkowski approaches the study of legal genres from the perspective of co-occurrence patterns, or "dimensions" (Biber, 1988). Specifically, he uses MD analysis to examine the distribution of linguistic features across individual texts and genres and then identify groups of regularly co-occurring features in order to interpret their most commonly shared functions. In order to provide the most comprehensive description of legal genres, Goźdź-Roszkowski first analyzes the co-occurrence patterns in legal genres along the dimensions identified in Biber's (1988) study of non-legal genres (e.g., history textbooks, professional letters, and ecology research articles); and second, he performs a new MD analysis on legal genres in order to reveal co-occurrence patterns that may be new and relevant to the domain of law.

In Chapter 5, Goźdź-Roszkowski provides comparisons of the seven legal genres to four specialist and disciplinary non-legal genres (textbooks and journal articles in biology and history) and five non-specialist non-legal genres (conversation, general fiction, popular non-fiction, official documents, and academic prose). The dimensions used in this analysis are 1) Involved vs. Informational Production, 2)

Narrative vs. Non-narrative Discourse, 3) Explicit (Elaborated) vs. Situation-dependent Reference, 4) Overt Expression of Persuasion/Argumentation, and 5) Impersonal vs. Non-impersonal Style (Biber, 1988, 1995; Conrad, 2001). Goźdź-Roszkowski finds that while legal genres inherently vary across the dimensions, in general, all legal genres tend to be "clearly informational with little concern for affective or interactive features. They are all relatively non-narrative and they [are] marked by explicit reference and abstract or impersonal style" (p.181). Also of note is that their mean scores often far exceed those of non-legal genres, making them, for example, "*the most* informational and *the most* impersonal in terms of the MD analysis" (p.181, italics in original). What is especially interesting in the cross-disciplinary comparison of genres is that legal genres tend to be more spread out along the dimensions than many of the non-legal genres, further supporting the hypothesis that there is considerable variation within the domain of legal language.

In Chapter 6, Goźdź-Roszkowski performs a new MD analysis to identify patterns of variation within legal genres that may be of particular importance to that domain, resulting in a three-dimensional model of variation within legal discourse. Specifically, these dimensions include 1) Narrative, Stance-focused vs. Informational and Normative Discourse, 2) Instructive and Advisory Discourse, 3) Abstract, Elaborated and Operative vs. Content-focused Lexically Specific Discourse.

Dimension 1 contained 24 total features with 17 on the positive side (e.g., demonstrative pronouns, 3rd person pronouns, all stance adverbs, mental verbs, and past tense) and 7 on the negative side (e.g., prepositions, nominalizations, quantity nouns, and the modal *shall*) of the factor. Unsurprisingly, contracts and legislation had negative loadings for this dimension and were found to be the least narrative genres (i.e., the most informational); yet, the other five legal genres (textbooks, opinions, journal articles, briefs, and professional articles) all had positive loadings and were found to be the most narrative genres with textbooks and opinions leading the scale, even using features in a manner similar to history textbooks that narrate historical accounts.

Dimension 2, which exists more along a continuum since the two negative features have larger loadings on Dimension 3, contained ten

features. These included, for example, non-past tense forms, *be* as a main verb, necessity modals, and 2nd person pronouns. Similar to Dimension 1, textbooks had the highest positive score on Dimension 2 and legislation had the largest negative score with the remainder of the genres clustered within +/- 3 of the center.

Finally, Dimension 3 contained 15 features, with 10 on the positive side (e.g., prepositions, perfect aspect verbs, agentless passives, and communication verbs) and 5 on the negative side (e.g., word length, nouns, attributive adjectives, and *that* relatives). Interestingly, while textbooks still had the highest positive score for Dimension 3, legislation was joined by briefs and academic journals on the negative end, with academic journals actually having the most negative score.

Through this analysis, Goźdź-Roszkowski confirmed that legal genres vary remarkably along the new dimensions; however, one of the most striking results here is the demonstrated need to perform new MD analyses on different genres (as opposed to simply using existing dimensions). He revealed that while there was some similarity to Biber's (1988) original dimensions (most notably regarding non-narrative aspects of the discourse), Goźdź-Roszkowski found that his Dimensions 2 and 3 had no direct counterparts in Biber's dimensions. His new patterns revealed a dynamic split between legal genres in Dimension 1 with academic, expository, and persuasive genres falling on the narrative, stance-oriented side and informational, normative genres falling on the other. Additionally, in Dimension 2, he demonstrated that legal genres fell along a continuum that was linked to instructive and advisory discourse, with those genres written for non-expert audiences having the highest scores. Finally, Dimension 3 demonstrated the complexity of legal discourse by "cut[ting] across the so far fundamental distinction between operative and expository/persuasive genres" (p.225). These fine-grained nuances would not have been revealed without preforming a new MD analysis on the legal genres.

Finally, in Chapter 7, Synthesis and Final Conclusions, Goźdź-Roszkowski concludes that "despite the widely-held perception of legal language as constituting a relatively homogenous and uniform linguistic phenomenon, the situation appears to be much more complex" as exemplified by the "highly systematic patterns of use

occurring across legal genres" (p.227). Goźdź-Roszkowski leaves the reader with a succinct overview of these highly systematic patterns of variation for each of the legal genres.

Overall, Goźdź-Roszkowski presents a very well organized, much needed description of the patterns of variation in American legal genres. His methodologies are well attested within corpus research and his execution of such methods is clearly described and concise (even to the inclusion of his keyword lists in Appendix A and MD statistical results in Appendix B). However, there is one aspect of his study that could have used further exploration. Specifically, while the decision to focus on 4-word bundles was sound, especially for cross-study comparison purposes, an opportunity was missed to explore the extent of formulaicity that occurs within legal genres by not examining longer bundles. Within his own study, Goźdź-Roszkowski builds an argument for examining longer sequences of words, stating that "the highly formulaic and repetitive nature of many legal genres … makes such texts particularly amenable to a methodology which focuses on very frequent uninterrupted sequences of word forms" (p.110) and that, other than legislation and contracts, "no other genre or text type has been shown to contain such a large proportion or formulaic expressions" (p.142). When compared to Biber *et al.*'s (1999) findings, 3-4 word bundles were the most pervasive within the registers of conversation and academic prose, but the number of bundles decreased sharply in the 5-6 word range. Given Goźdź-Roszkowski's own confirmation of the formulaic, repetitive nature of language in legal genres, especially contracts and legislation, it would have been highly informative to investigate the upper end of bundle length given the scarcity of studies in this vein (i.e., bundle length in legal genres or bundle length within specialized vs. non-specialized registers). Perhaps future opportunities will arise to investigate this research trajectory as it would make an additional, valuable contribution to studies on legal genres and lexical bundles in English.

In sum, Goźdź-Roszkowski's *Patterns of Linguistic Variation in American Legal English* is a much needed and highly praised addition to studies of language variation and studies of legal language, or, what we may now more appropriately call, studies of 'legal genres'.

References

Bhatia, V. K. (2004). *Worlds of Written Discourse. A Genre-based View*. New York: Continuum.

Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D. (1995). *Dimensions of Register Variation: A Cross-linguistic Comparison*. Cambridge: Cambridge University Press.

Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Philadelphia: John Benjamins Publishing.

Biber, D. and Conrad, S. (2003). Register variation: A corpus approach. In D. Schiffrin, D. Tannen, and H. Hamilton (eds.), *The Handbook of Discourse Analysis*. Malden, MA: Blackwell Publishing, 175-196.

Biber, D. and Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., and Cortes, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. In P. Wilson, P. Rayson, and T. McEnery (eds.), *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*, New York: Peter Lang, 71-92.

Biber, D., Conrad, S., and Cortes, V. (2004). *If you look at…*: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25: 371-405.

Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. Harlow: Longman.

Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad and D. Biber (eds.), *Multi-Dimensional Studies of Register Variation in English.* Harlow: Longman, 94-107.

Conrad, S. and Biber, D. (2001). *Variation in English: Multi-dimensional studies*. Harlow: Longman.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23: 397-423.

Goźdź-Roszkowski, S. (2006). Recurrent word combinations in judicial argumentation: A corpus-based study. In D. Bartol, A. Duszak, H. Izdebski, and J-M Pierrel (eds.), *Langue, Droit, Société (Language, Law, Society)*. Presses Universitaires de Nancy, 139-152.

Halliday, M. A. K. (1988). On the language of physical science. In M. Ghadessy (ed.), *Registers of Written English*. London: Pinter, 162-178.

Hyland, K. (2008). *As can be seen*: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27: 4-21.

Kjaer, A. (1990). Phraseology research-state of the art: Methods of describing word combinations in language for specific purposes. *Terminology Science and Research*, 1(1-2): 3-32.

Tammy Gales received her Ph.D. in Linguistics from the University of California, Davis and is currently an Assistant Professor of Linguistics at Hofstra University. Her current research interests integrate corpus analysis and various forms of discourse analysis, especially Appraisal, to address topics of language and the law as they pertain to diversity in the U.S., to investigate forensic linguistic questions concerning the nature of threatening communications in contemporary American society, and to reveal language ideologies about the discourses within each genre. Email: Tammy.A.Gales@Hofstra.edu.

Patterns of linguistic variation in American legal English: A corpus-based study. Frankfurt am Main, New York: Peter Lang. Hamann, H. (2014). Â Computer assisted legal linguistics: Corpora and empirical methods as a new instrument in the legal toolbox. Law & Social Inquiry. Journal of the American Bar Foundation (ABF). Vogel, F., PÃ¶tters, S., & Christensen, R. (2015). Richterrecht der Arbeit - empirisch untersucht. You will learn about Patterns of Linguistic Variation in American Legal English: A Corpus-Based Study by Stanislaw GoÅºdÅº-Roszkowski. Ebook Free Download.Â Patterns of plant species richness within families and genera in lowland neotropical forests: Are similarities related to ecological factors or to chance? By Stevenson, Pablo R.; Castellanos, MarÃa Clara; Agudelo T., Juliana. A review of "Patterns of Linguistics Variation in American Legal English. A Corpus Based Study" by StanisÅ‚aw GOÅ¹DÅ¹-ROSZKOWSKI. Article (PDF Available) Â· November 2015 with 78 Reads. DOI: 10.14746/cl.2014.19.07.Â and is devoted to the analysis of linguistic variation patterns in American legal English. The book is composed of seven chapters with the first being the introduction and the. last the conclusions. What should be emphasised here is the fact that it is one of the few. works devoted to research in legal language, in which corpus-based methodology has. been applied. In the first chapter the author reveals why he has decided to investigate legal. language used in seven different genres of legal texts instead of investigating legal. language as such.