



# Approaching Legal Multinomials from the Sociolinguistic Perspective – Insights into Authorship-Based Distinctions

Edyta Więclawska<sup>1</sup>

Accepted: 23 March 2023 / Published online: 30 April 2023  
© The Author(s) 2023

## Abstract

The paper explores the hypothesis that multinomials can act as authorship-based style distinguishing markers in legal communication. Specifically, the analysis focuses on identifying the quantitative distribution patterns of structural categories of multinomials as typical for two authorship categories and on their communicative function. The two authorship categories that are contrasted here are legal professionals/experts and lay people. The analysis is conducted in the corpus-based methodology with a custom-designed corpus of English, authentic texts found in the legal trade, in the domain of company registration proceedings. The findings confirm that multinomials that are conventionally considered to be a feature of professional legal communication are also cognitively salient in lay communication. Further, the texts drafted by the two categories of authors are profiled by structurally distinct multinomials. Functionally, it has been demonstrated that the structurally distinct types of multinomials that are found quantitatively salient in the two authorship categories are used predominantly for specific stylistic and/or pragmatic functions. Stylistically, multinomials contribute to conventional and ritual patterns which are used to meet the formality standards that have evolved in specific legal professions where authority is of particular importance. Pragmatic factors which account for quantitative salience of specific, structurally profiled categories of multinomials involve mainly reduplication of multinomials that embody norm-related concepts, which is required on the ground of intertextuality and ensures the materialisation of legal effect.

**Keywords** Multinomials · Authorship-based distinctions · Legal stylistics · Corpus-based study

---

✉ Edyta Więclawska  
edytawieclawska@poczta.fm; ewieclawska@ur.edu.pl

<sup>1</sup> Zakład Języków Specjalistycznych, Uniwersytet Rzeszowski (University of Rzeszow), Al. mjr. W. Kopisto 2 B 35-315, Rzeszów, Poland

## 1 Introduction

The field of phraseology has been attracting considerable attention of researchers in recent years, and these – apart from proposing increasingly more refined methodological frameworks for the well-established linguistic themes – bring to the fore the need for working on more homogeneous corpora, introducing digital processing of data, adopting inter-methodological research frameworks and addressing specific registers and interdisciplinary perspectives.

This research project focuses on multinomials as a phraseological category accommodating multi-word sequences profiled by coordination of two to five same-word class items, linked by *neither*, *nor*, *by*, *and*, *or* or comma, motivated by synonymy or a type of complementarity relation (antonymy, grammatical pairs, sequential patterns). In the literature of the subject the term multinomials is traditionally used in a more specific sense for the sequences fitting in the above specified structural profile but exceeding two-component structures of the said type, that is (trinomials/triplets, quadruplets or quints) [3, p. 108; 30; 31, pp. 2–12; 36, p. 2]. By way of generalisation the label ‘multinomials’ adopted in the foregoing covers both bi-component sequences and longer strings. The term ‘binomials’ is occasionally used in the discussion only when making part of the fixed collocations standing for specific sub-types, i.e. *binomials proper* (i.e. synonymously motivated) or *true binomials* (i.e. conjoined by *and*).

Specifically, here an authentic set of legal multinomials are studied in the sociolinguistic paradigm and the population examined, as covered by the analysis, is limited to company registration discourse. The aim is to study the use of multinomials in the said communicative environment for their authorship-based distinctions, which involves contrasting the styles of legal experts and lay people.

The research questions asked by the author are: (1) Does the use of multinomials by the two categories of authors show quantitatively distinct patterns? (2) How can the quantitative distinctions in point be accounted for in terms of the communicative functions pursued by the authors and what patterns emerge with regard to the construal of professional group identities?

It is hypothesised that the structure-based distribution model for the two authorship categories is significantly heterogeneous, and the salience of some structural categories can be accounted for in the author’s choices for stylistic and/or pragmatic uses.

The hypotheses may be said to be justified on the ground of the relevant findings made so far, and these point to significant dynamics in multinomials regarding their structure, sociological variation or communicative functions. Hence, it has been evidenced that the repertoire of multinomials varies by genre and over time [24, p. 206]. The variationist capacity was evidenced across individual Old English authors [32]. Scholars also emphasise the varied and diverse structure of multinomials, which in many respects escapes a single typology [14, 18, 28, 35]. Finally, the analysis of contemporary data on multinomials is expected to yield an updated account of the actual use of multinomials which – in view of the Plain Language Movement postulates to eliminate ritual expressions from the legal communication – may be assumed to have evolved [19, p. 61].

The methodological paradigm adopted for this study of multinomials draws from corpus sociolinguistic research [33, 39] and it relates to such concepts as identity issues [10], speech style variation, communicative divergence, group level identities or accommodation theory [15]. If we narrow this explanatory methodological framework to studies on legal language per se or somewhat more generally – to studies on legal language as one of the specialised languages – then we need to relate to variationist jurilinguistic studies [16, 17], to the concepts of *professional discourse* [7], (*professional*) *discourse community* [20, 21] and (*professional*) *group identity* [8, 11, 13, 42]. The first concept characterises the communicative environment of all the language materials analysed (texts making up the corpus). The label ‘discourse’ implies that we are dealing with specific use of a language in a natural, authentic, socioculturally coherent context. The label ‘professional’ denotes primarily that it is related to paid-work and concerns work-related topics but it also implies that it is to some point homogeneously situated, goal-oriented, conventionalised and engages socially ordered groups. Discourse community is the social structure, including human factor, which construes and maintains the conventionalised use of language of a (professional) discourse community, has a set of common public goals, mechanisms of communication among members, its own genres and participatory mechanisms to provide information and feedback as well as a degree of relevant content and discursual expertise [21, p. 500]. Any (professional) discourse community surfaces when it develops a (professional) group identity. In our case we are dealing with a community within company registration discourse, that is individuals, acting within the capacity of specific professional roles, related to their work position and scope of competences. They all work within the same theme, and set of procedural operations, thus making a discourse community. However, even within such a neatly delineated area, which is functionally and discursively homogeneous, systemic distinctions should be noted in the said discourse community in view of there being a naturally arisen sociocultural dichotomy: expert vs. lay factor [38]. The sociodiscursive approach adopted in the foregoing implies enabling us to examine how the company registration discourse manifests itself sociolinguistically with regard to the internal distinctions of the professional discourse community.

More narrowly, the research project also fits into the strand of authorship factor analyses in general language, in the context of authorship attribution studied from the highly technical perspective [e.g. 2, 9] or specifically legal language addressing the domain of phraseology [29] or other aspects of language [40].

## 2 Methodology

The methodology used involves the processing of a custom-designed, monolingual, thematically homogeneous corpus, compiled of English texts from the legal trade for the purpose of company registration in Poland (1,124,204 tokens, 932,839 words). The technical operations conducted to verify the theses posed involved systemic and supervised extraction of the candidate terms defined as all multinomials occurring in the corpus, composed of a maximum of five constituents, authored by the two specific categories of entities, hereinafter referred to as authorship macrocategories

to keep them distinct from the concept of individual categories of authors clustered. The extraction was based on exploiting the potential of a computer query language system for corpus search, as available in the Sketch Engine tool, where the custom-designed corpus was previously uploaded. The pre-processing stage, preceding the Sketch Engine operations, involved manual annotation of the texts making up the corpus, and this included authorship-related data. In order to make the extraction process meet the needs of supervised extraction, the author used the automatic system of part-of-speech tagging. After extraction of the data from Sketch Engine the candidate terms were manually coded for their structural properties (part-of-speech profile, number of conjuncts, type of conjunction and semantic motivation). In the foregoing this stage is referred to as *initial coding* in contrast with the *secondary coding* which involves systemic and quantity-conditioned clustering of initial values identified for the four structure-related variables. The corpus compilation involved exhaustive and systemic search of court files representative for the Polish nationwide context, that is English/Polish company registration discourse. This study exploited only the monolingual (English) text reserves of the corpus. The texts were retrieved from two divisions of the National Court Register, Register of Entrepreneurs, as found representative for the Polish context. ‘Systemic’ means that the search criterion and mechanism ensured retrieval of the relevant texts for all the branches of foreign companies as active in 2017 and operating in the capacity of the two representative court divisions (spatio-temporal context). Further, ‘exhaustive search’ means that all the texts on the court files that met the specific criterion were retrieved. Such choice of the sample (corpus texts) allowed us to identify actual frequencies and proportions as typical for the Polish context and diagnose the actual tendencies. Consequently, the exhaustive character of the court files search eliminated the danger of errors stemming from the choice of texts studied.<sup>1</sup>

The final statistical analysis covered secondary clusters corresponding to the distinctive structural properties of the candidate terms and clusters corresponding to the authorship macrocategories – legal experts vs. lay people. The two authorship macrocategories that were examined for the prevalence of specific structural types of multinomials were selectively sampled out of the total population making up the corpus and they are widely held to be promising ground for stylistic distinctions running through legal language. Specifically, the two authorship-related data sets are expected to display features shared on the ground of institutional affiliation and/or style imposed by a specific communicative environment and not an author’s individual conventions. Additionally, the criterion of quantitatively comparable samples was applied. Hence, legal experts (cluster 1) cover 520 candidate terms, and lay people (cluster 2) cover 578 candidate terms. Legal experts include the professional categories of legalisation officers and registration authorities at the executive and secondary level (legalisation officer, registration authority – head, registration authority – lower level). Legal genres produced by these authorship categories involve predominantly authentications, company extracts and verifications. The other authorship macro-

<sup>1</sup> For related analysis of the authorship-induced variation in company registration discourse with regard to binomials and multinomials, see [41]. The analysis there exploits a different scope of the sociocultural potential of the corpus.

category is referred to as lay people and it covers people involved in running the administration and management of a company (company officer, company founder and proxy holder). These authors produced mainly resolutions, reports, foundation acts, financial documentation and declarations of will.

The educational background in law (at the level enabling to pursue professional activities under one of the nationally recognised professional titles) was the main eligibility criterion for the macrocategory of expert legal drafters as opposed to lay people, and it operates as an explanatory term here although the communicative setting also comes into play. Having the specific educational background in law is related to pursuing legal professional activities under one of the professional titles, such as members of the Bar. This authorship macrocategory was assumed to represent expert communication in the domain of law and it was set against the in-company communication scheme accommodating lay people. The expert/lay dichotomy rests on a generalisation assuming that there is no professional factor, in the sense adopted here, in the in-company communication. By way of explanation, the design of the methodology in this respect rests on the assumption that the affiliation to the in-company communicative environment (communicative setting) complements the rationale for the said dichotomic categorisation of authors. Company officers having an educational background in law prove to communicate out of the rigour of language conventions featuring professional communication and they fit in the in-company style, which confirms the reliability of the methodology. This observation was confirmed by the manual analysis of the random text samples.

The correlation-oriented analyses involved deriving the frequency distribution schemes for the structural types of multinomials (part-of-speech profile, number of conjuncts, type of conjunctions and semantic motivation) in contrast for the two authorship macrocategories. Derivation of the frequency distribution models in the contrastive perspective was preceded by conducting calculations for the statistical significance of the correlations in question. Thus the quantitative data discussed below were tested for their statistical significance in the context of the aggregate model for the individual structural categories. Hence, chi square scores for the models covering part-of-speech profile, number of conjuncts profile, type of conjunctions profile and motivation profile are 36.88; 45.1; 48.96 and 18.91 respectively, and this yields a p value score that is in generally lower than 0.0001 (part of speech and type of conjunction, or equal to 0.001 (the final category)). The only exception here is the variable of the number of conjuncts, where the p value score is at the level of 0.08, and this renders the calculations for this structural category as statistically insignificant. As such, the paper does not include the relevant visual representation for the length of multinomials and this aspect will only be mentioned by reference.

Findings are inferred from analysing the quantitative data as visualised in the forest plots (odds ratio scores) together with the qualitative data. The qualitative analysis involved analysis of the immediate linguistic context, which enabled conceptual categorisation of multinomials and recognition of their pragmatic function at the sentence level and with regard to generic context. The distinctions in the frequency distribution schemes are interpreted in the light of the concept of professional group identity construal which involves systemic appropriation of specific structural categories of multinomials.

### 3 Results

Figure 1 presents the areas where the model of quantitative distribution of multinomials is not homogeneous in the two populations. The part-of-speech feature shows varied distribution across the two authorship macrocategories. In short, the higher the odds ratio score is the more probable it is for a given structural category to occur in cluster 2, that is in lay users. An odds ratio score that is close to '1' means equal probability of occurrence of a given category of multinomial in the two populations.

The first noteworthy tendency that emerges from the visualisation is that nominal and verbal multinomials do not discriminate between the two authorship macrocategories. The individual p values are at the level of 0.0727 and 0.0631 for nouns and verbs respectively. We may say that both categories of language users produce texts which fit into the stylistic pattern with regard to the participation of nominal and verbal multinomial structures. With regard to distinctive patterns here the comparative model shows that lay users are more prone to use adverbial and prepositional multinomials than expert users. In numbers, both these part-of-speech categories of multinomials occur 1.93 times more often in cluster 2 than in cluster 1. The population of adverbs for cluster 2 can be defined with regard to the following semantic fields: COMPOSITIONAL (*fully or partly*), MANNER (*directly or indirectly*), PARTICIPATIONAL (*jointly or severally*), SCOPAL (*generally or unconditionally*). The exhaustive sample covering the prepositional multinomials identified for lay people

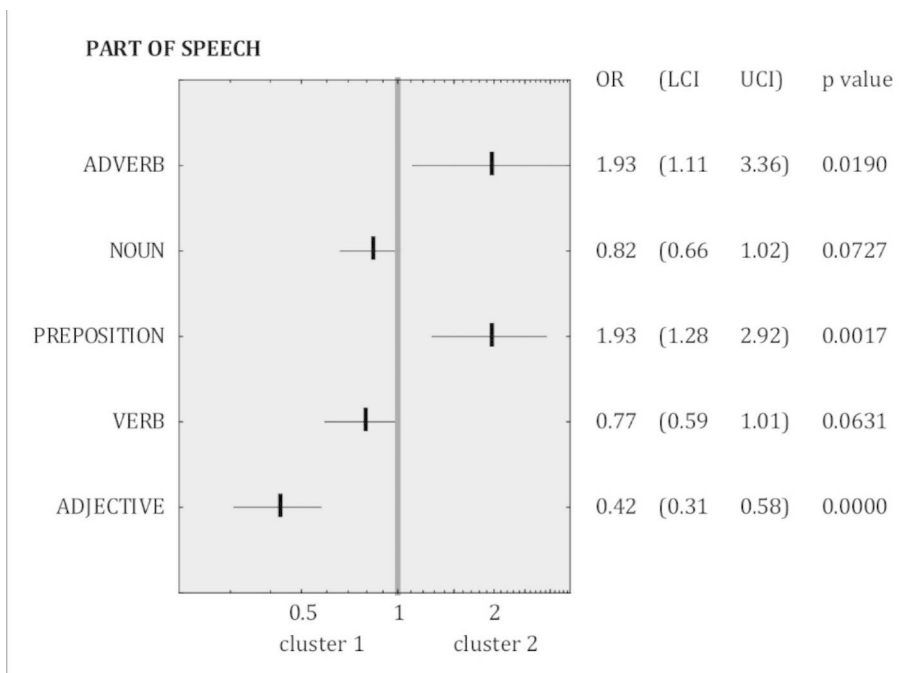


Fig. 1 Cluster-analytic results of part-of-speech distinctions

includes the following semantic categories: TEMPORAL (*at or before*), SPATIAL (*within or outside*), RELATIONAL (*from or in connection with, by or on behalf of*), CONSTITUTIONAL (*with or without*), COMPLIANT WITH LAW (*subject to and in accordance with*). The immediate linguistic context for the representatives of the quantitatively salient candidate terms here is found to be highly repetitive and the cases that are found representative for prepositional (Exhibit 1 and 2) and adverbial multinomials (Exhibit 3 and 4) respectively involve:

Exhibit 1: #93,668 • UK • 166 • 1 • 011 • 2 • 1 • M • 4 • 2004 • 3 • 4 • 2005 • 1 • 1885<sup>2</sup> [...] and accordingly shall in no way be limited or restricted (except where otherwise expressed in such paragraphs) **by reference to or inference from** the terms of any other paragraph or the name of the company [...].

Exhibit 2: #44,446 • UK • 83 • 1 • 004 • 5 • 1 • K • 6 • 1969 • 3 • 4 • 2004 • 1 • 15,267.

[...] and sanction the issue of Preference Shares which are, or at the option of the Company are to be, liable to be redeemed, **subject to and in accordance with** the provisions of the Statutes. The Special Resolution sanction [...].

Exhibit 3: • UK • 206 • 1 • 015 • 1 • 1 • K • 4 • 2005 • 3 • 4 • 2005 • 1 • 1695.

[...] as may from time to time be determined. (k) To pay for any property or rights.

acquired by the Company, either in cash or **fully or partly** paid-up shares, with or without preferred or deferred or special rights or restrictions [...].

Exhibit 4: #245,315 • Greece • 360 • 1 • 026 • 8 • 3 • M • 4 • 2008 • 10 • 3 • 2008 • 1 • 963 [...] of the Company Branch in Poland, as well as Mr. XXX, Manager of the Company Branch in Poland, acting either **jointly or severally**, in the name and on behalf of our Company, to sign any document, power [...].

The statistics show that the authorship-based distinctions concern less frequent part-of-speech categories (less prototypical) and – although in general the use of multinomials is commonly associated with expert communication in law that is assumed to be more ritual, conventional, stylistically traditional and sophisticated – here, with regard to the less prototypical candidate terms stronger representation is to be ascribed to lay users. In this category (cluster 2) the quantitatively salient categories are not embodiments of technical terms but – if we take prepositional binomials – they serve the purpose of constituting the syntactic frame of the sentences. Their function seems to be of pragmatic nature in that the tendency is noted that they are prefabs which are copied from the related legislative texts along the principle of intertextuality [22]. This shows that the lay communication is conducted with full awareness of scholarly principles of legal drafting, where attainment of legal effect is often conditioned by quoting and referencing to related documents.

Another distinction in the exploitation of multinomials by the two authorship-based macrocategories is noted for adjectives. This time the quantitative bias is

<sup>2</sup> The values falling within the framework of the initial coding system include in the order they are mentioned: token number, country, doc id, institutional name, krs, krs\_item, professional title, sex, source text word count, source text year, title, target text word count, target text year, type of translation, word count. Note the annotation mark referred to as ‘krs’ and ‘krs\_item’ refers to the files number corresponding to the specific text from which a candidate term was extracted.

negative for lay people. Namely, adjectival multinomials are used 0.42 times less frequently by lay people. The list of cognitively salient candidate terms can be categorised into semantic fields as follows: TEMPORAL (*annual or special; present or future; regular or special*), MISCELLANEOUS (*audio or video*), NORM-RELATED (*charitable or public; negotiable or transferable*), COMPOSITIONAL (*continuous and unbroken; real or personal; domestic or foreign*), MANNER (*direct or indirect*), HIERARCHICAL (*supreme, municipal or local*), CREDIBILITY (*true and correct; accurate or complete*).

The domination of adjectives in legal experts (cluster 1) here can be accounted for on the ground of more specific profession-based distinctions. Here we are dealing with highly formalised communication which is featured by institutionally imposed conventions. It was noted that there is a positive correlation between the length of a sentence and the use of adjectival multinomials. Further, the institutional dimension imposes specific stylistic conventions where multinomials are strong players. The examples below testify to the observation:

Exhibit 5: #34,060 • UK • 71 • 1 • 004 • 48 • 3 • K • 2 • 2006 • 1 • 2 • 2006 • 1 • 248 [...] to the documents on the file of the company in the custody of the Registrar of Companies, the company has been in **continuous and unbroken** existence since the date of its incorporation.</s><s>No action is currently [...].

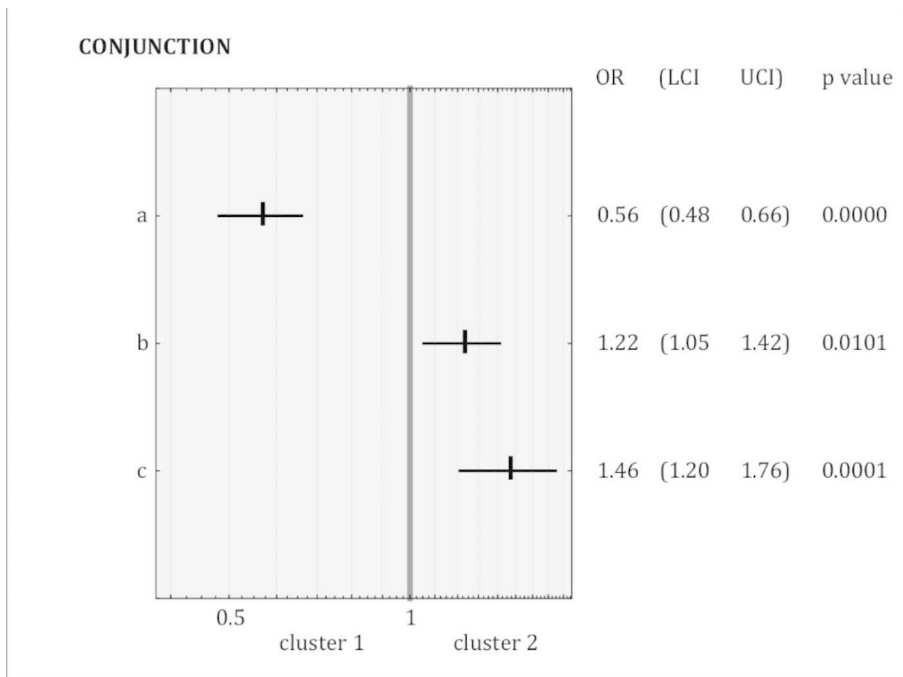
Exhibit 6: #56,075 • UK • 85 • 3 • 004 • 61 • 10 • K • 2 • 2009 • 1 • 2 • 2009 • 1 • 193 [...] faith without verification.</s><s>For this reason the Registrar cannot guarantee that the information on the register is **accurate or complete**.</s><s>English</s><s>Power of Attorney</s><s>Staines, 29 June 2012 [...].

The type of conjunction is another structural feature of multinomials which is subjected to extensive analyses. This covers the issue of conjunction omission [23, 35, 37], the use of predominant categories of conjunctions, with the distinction into the conventional conjunctions (*and* or *or*) or rarer conjunctions, including the correlative conjunctions (*neither... nor*), the use of comma [6, p. 68; 25, p. 3], and structure-based distinctions regarding conjunction variation capacity [6, p. 68]. Some studies are rather strict in defining the repertoire of conjunctions that can be admitted as structural components of multinomials. Others are more open, proposing conjunction typologies composed of *and*, *or* and ‘similar conjunctions’ [27, p. 82]. Yet, although the cases of the so-called ‘similar conjunctions’ are quantitatively marginal this aspect seems to be interesting to investigate in the context of contemporary patterns in points, as emerge these days.

After preliminary examination of the frequency of occurrence regarding conjunctions, the codes assigned initially were clustered to form more general categories, which enabled us to conduct statistically reliable analysis. Hence, in Fig. 2 below the codes on the vertical axis stand for: ‘a’ – conjunction *and*, ‘b’ – conjunction *or* and c – other conjunctions, including omission or comma. The frequency distribution model setting the two authorship macrocategories in contrast is as follows:

It emerges from the data set that all the conjunction-related values have discriminatory power in the sense of constituting a distinctive stylistic feature of the two authorship macrocategories by virtue of a p value score that is invariably lower than 0.05. The model shows a lower proportion of multinomials coordinated with the predominant category of conjunction *and* in lay people. The odds ratio score here is





**Fig. 2** Cluster-analytic results of conjunction-related distinctions

0.56 which is the ratio by which the conjunction *and* is rarer in cluster 2. As for the conjunction types coded with ‘c’ and ‘b’ these are prevalent in lay people with odds ratios at the levels of 1.46 and 1.22 respectively.

Closer qualitative analysis of the randomly selected excerpts shows that there is a positive interdependency between rare conjunctions and the length of a candidate term. Examples in point include almost exclusively verbal and nominal strings, since these appear as tri-, quadru- and quinquenomials. The authorship macrocategory comprising lay people (cluster 2) is definitely more prone to exploit the multinomials exceeding the 2 component word strings, often comprising less prototypical conjunctions. Such a rhetorical tool is found to have emerged from the habit of copying prefabricated strings of words that are exchanged in corporate communication and it is assumed to be a mechanism of facilitating in-company communication. Such a conclusion may run counter to common expectations, in that non-experts are found to recurrently use sophisticated, systemically and consistently coordinated components and it may be peculiar for the authorship macrocategories in question. As such the finding in question is better to be saved for the specific communicative environments covered by the analysis and left for validation in other communicative settings.

Another regularity, which is somehow related to what has been said, is that in lay people we note a higher percentage of multinomials proving to score lower with regard to the fossilisation parameter. The data show that the reversibility parameter is significant here. Also, the comma as a coordinating device proves to be prone to

show variation in that it proves to be interchangeably used with *and* and/or *or* in otherwise parallel multinomials. The same concepts expressed with multinomials prove to show variants with regard to the choice of conjuncts. Examples include *made, seconded or/and unanimously approved or person, firm and/or company*. The immediate context is presented below for better understanding:

Exhibit 7: #245,793 • Dania • 361 • 1 • 027 • 1 • 5 • K • 2 • 2011 • 10 • 2 • 2012 • 1 • 213 Any Other Business </s><s>None Adjournment</s><s>There being no further business to come before the meeting, upon motion duly **made, seconded and unanimously approved**, the meeting was adjourned [...].

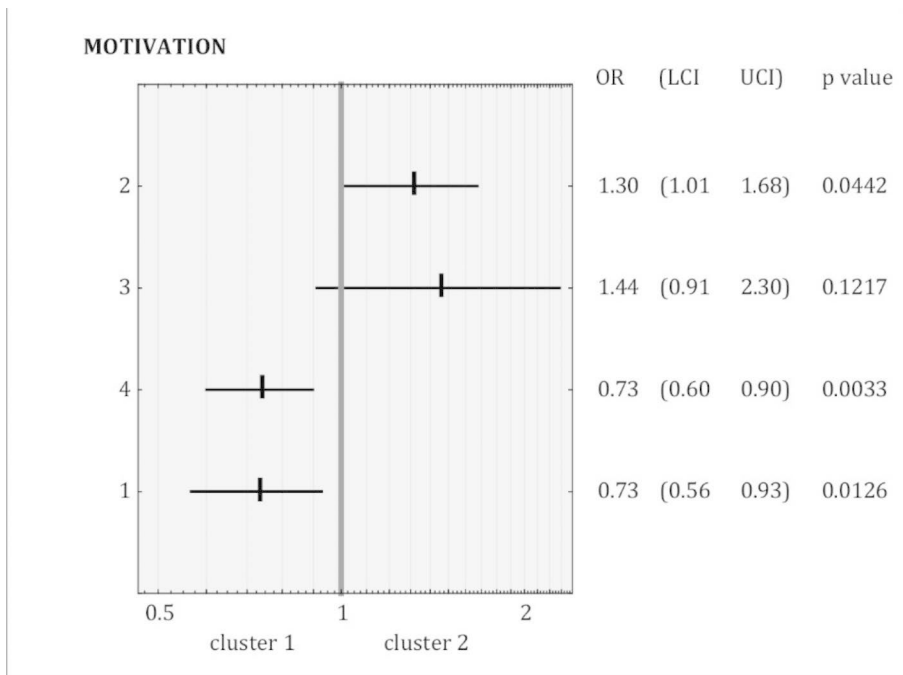
Exhibit 8: #123,104 • UK • 206 • 1 • 015 • 1 • 1 • K • 4 • 2005 • 3 • 4 • 2005 • 1 • 1695 or corporation.</s><s>(g) To lend and advance money or give credit on any terms and with or without security to any person, **firm or company** (including without prejudice to the generality of the foregoing any holding company [...].

Because the repertoire of candidate terms exemplifying the prevalence of less typical conjunctions here is extensive it is impossible to enumerate all the semantic fields they fit in to. The most common and abundantly represented cases belong to CORPORATE ENTITIES (*subsidiary, holding or fellow subsidiary*), LEGAL INSTRUMENTS (*debentures, mortgage debentures or debenture stock*) and NON-OPERATIONABILITY (*negligence, default, breach of duty*).

The third structural property of multinomials that has been exploited for the contrastive analysis here is motivation. The research here concerned the category of semantic motivation which, apart from the etymological and phonological paths of research, constitutes the main research domain in point. The overview of the typologies of semantic motivation as often referred to these days enables us to formulate some general conclusions. The binomials proper [1], that is synonymous word strings as a category, prove to recur through the majority of typologies [25, pp. 11–13; 36, pp. 70–71]. Antonymy shares this scenario [24, p. 188; 25, pp. 11–13; 36, pp. 70–71]. Complementarity is listed as a separate category [24, p. 188], occasionally encompassing hyponymy and/or contiguity and/or enumeration, which in other typologies make up their own categories [6, p. 52; 24, p. 188; 25, pp. 11–13] after Chapman [26, pp. 90–96].

As emerges from the short overview, there is no consistent scheme here and it is claimed explicitly that – unlike word class profile and conjunctions – ‘semantic relations are [...] difficult to systematise’ [25, p. 11]. In order to capture the specificity of the population of candidate terms covered by the analysis, the author assumes 4 basic categories of semantic motivation and these follow the system of initial coding which was retained. The numbers spread over the vertical axis stand for synonymy (‘1’) or complementarity which splits into: antonymy (‘2’), grammatical pairs (‘3’) and sequential patterns (‘4’). Figure 3 visualises the model in point.

The model that emerges here is heterogeneous. Three of the values show scores which are statistically significant, with the p value being lower than the prescribed threshold of 0.05. It is only the category of grammatical pairs that is not found to have discriminatory power, which shows that the style of each of the two authorship macrocategories is featured by the comparable number of representatives fitting in here. These include *person or persons, company or companies, purpose or purposes,*



**Fig. 3** Cluster-analytic results of semantic motivation distinctions

*time or times* or verbal patterns showing grammatical tense distinctions, such as *are or have been, serving or have served*.

The model is positively biased to the benefit of lay people (cluster 2) with the motivation category of antonymy (code '2') with an odds ratio score at the level of 1.30. The antonymously motivated multinomials identified in lay people that constitute quantitatively the most significant functional category are classified as NORM RELATED (*charitable or public, civil or criminal*) TEMPORAL RELATIONS (*deceased or former; present or future; at or before; on or before*), PROCEDURE RELATED (*registered or recognised; required or permitted to be*) OPERATIONS (*purchase or sale; act(s) or omission(s); issue or sale; redemption or purchase*), GRADATION (*acquire or propose to acquire; place or guarantee placing of*), AMENDMENTS (*excluded or varied; renewed, revoked or varied*).

The justification for the prevalence of antonymously motivated multinomials in lay people results from them following the principle of intertextuality in the sense of reduplicating the specific category of legislation-based, prefabricated formulae. This confirms the finding already formulated with regard to the word class distribution pattern, where we stated that lay communication in the domain of law makes use of the professional drafting techniques, like the rule of intertextuality. The antonymical pairs analysed qualitatively in lay people are well-established combinations and not low frequency cases coined ad hoc, and as such fall in the group that is prone to intensive reduplication. The data show that lay people prove to be more selective when it

comes to recurrent use of well-established multinomials and they more eagerly reduplicate multinomials which are not binomials proper (tautological), well-recognised for their purely ornamental function. The antonyms that show marked dominance in lay people (cluster 2) may be said to have pragmatic application. Some of the candidate terms are composed of normative terms (*civil or criminal; charitable or public*), and as such constitute an important intertextual link with other related prescriptive texts, while others are effective stylistic instruments to indicate fine distinctions (*at or before*) and/or complexity of cases, along the principle of precision in legal communication. Exhibit 9 provides an illustration in point:

Exhibit 9: #97,649 • UK • 168 • 2 • 011 • 4 • 1 • M • 4 • 2004 • 3 • 4 • 2005 • 1 • 3420.

[...] of Part II of Schedule 15 A of the Companies Act 1985, and in particular as to the provision.

of documents to each member **at or before** the time the resolution is supplied to him for signature, is not complied with.

The left side of Fig. 3 features the structural profile of multinomials which are identified less frequently in lay people's corporate communication. These are synonymous candidate terms (code '1') and sequential word strings (code '4'). In both cases the odds ratio is 0.73, which means that the said structural types of multinomials appear that many times rarer in lay people than in expert lawyers. The most representative examples of synonymous multinomials in the category of legal experts (cluster 1) that are quantitatively significant include by reference to the semantic fields they fit in: AUTHENTICATION (*accuracy or completeness, power or capability*) ACKNOWLEDGEMENT (*books and records*) PLEONASTIC (*business and affairs*) APPRAISAL (*force and effect*) RELATIONAL (*over or in respect of*).

Exhibit 10: #11,814 • US Cal • 17 • 3 • 002 • 1 • 1 • K • 5 • 2002 • 7 • 5 • 2002 • 1 • 8839 [...] attorney in fact.</s><s>A validly executed proxy that does not state that it is irrevocable shall continue in full **force and effect** unless (i) revoked by the person executing it, before the vote pursuant to that [...].

These pleonastic multi-word strings, also referred to as tautological binomials or binomials proper, are considered by many as the most prototypical category in this group, with some of them dating back to medieval times when the driving force for their construal often rested in conjoining the Anglo-Saxon and Norman equivalents. Professional writing (and thus that of legal experts) is more prone to feature such canonical candidate terms because institutional legal education goes in this direction. Professional means sophisticated, conventional and ritualistic. The functions that may be ascribed to synonymous multinomials may be said to be of stylistic nature in that such formulae serve the purpose of rendering the office of the professionals as authorities in the field of law and thus promote respect for the said office. Secondly, we may identify a sub-category here which comprises candidate terms that may be considered to be of pragmatic application in the sense of materialising an act in law enacted by a specific legal text. The multinomials making (part of) authentication formulae recurrently used by legalisation officers may serve as an example here (*true and correct copy*).

The quantitative salience of sequential patterns (code 4) in highly professional writing (cluster 1) may be viewed as a lexicogrammatical tool used to present a spe-

cific set of facts and it may be said to be employed for stylistic purposes. Cases in point include, for example, long strings of words, as illustrated in the example below:

EXHIBIT 11: #170,149 • UK • 285 • 1 • 022 • 8 • 1 • K • 4 • 2006 • 3 • 4 • 0 • 3 • 1939.

To subscribe for, take, purchase, or otherwise acquire, hold, sell, deal with and dispose of, place and underwrite **shares, stocks, debentures and bonds** issued [...].

The most cognitively salient candidate terms of the sequential type fit into the following semantic fields: AUTHENTICATION (*signature, seal or stamp*), OWNERSHIP FORM (*promissory note, bills of exchange, bill of lading, warrants, debentures*), IDENTIFICATION DATA (*number and class, time and place; date and place*), CORPORATE IDENTITY (*person, firm or company*), LEGAL INSTRUMENTS (*mortgage, charge or lien*), PROFESSIONAL APPELLATIONS (*receiver or manager*).

The prevalence of sequential chunks in the writing of legal experts may be accounted for from two perspectives. Firstly, such an enumerative stylistic pattern is perfectly comprehensible and favoured by educated text drafters, because they are competent in the reception and production of longer phraseological formulae which often rest on combining a series of terms embodying norm-based concepts. The syntagmatic arrangement in point falls under the regime of legal reasoning, whereby individual findings of the fact are related to specific legal norms and thus earn technical labels, with the norms having been previously construed by selection and arrangement of specific legal provisions (norm construction/derivation). The said combinations are fairly fixed, showing some distinctions in the conjunctions and/or order of conjuncts, but for the majority of them at least the framework is stable and shows to be consistently reduplicated. The second argument accounting for the quantity-related tendency in question here is text economy, which experts knowing the tricks of the trade can afford without risk of misunderstanding. To explain, sequential patterns often constitute an alternative to other stylistic patterns that would be far more descriptive, longer and thus redundant for expert users.

## 4 Conclusions

The conclusions that may be drawn from the analysis touch upon two issues: (i) the heterogeneous character of legal language in the context of authorship-based distinctions and (ii) the representation of multinomials in contemporary legal communication, as delineated by theme and communicative environment, taking account of pragmatic correspondences between their structural profile, the context and the users.

With regard to the first aspect, the mainstream direction of variation-oriented legilinguistic research has so far addressed mainly the domain of text typologies and genre-specific discursive practices with seminal works of key figures in the field [4, 12, 34]. Research into the interdiscursive stylistic conventions [5] includes successful attempts to capture the ground for the varied patterns of stylistic trends with justification in institutional settings, for example, which obviously exceeds the framework of primary typologies but, as the data above shows, constitutes only one of the alternatives in the variationist research. Further, variationist aspects of legal communication

are also identified in the context of translationese and global Englishes. The discussion here shows that the distinctions may also be authorship-biased, with author as a collective category exploiting stylistic conventions specific for a communicative environment, and the proof lies in the use of specific, structurally-distinct categories of multinomials. The quantitative data show that there are systemic and consistent variation patterns in the said respect. They confirm the heterogeneous character of legal language or legal discourse if we focus on a specific communicative environment and at the same time they show that the distinctions are not chaotic but can be categorised.

The second perspective (cf. ii) addressing the source and nature of the expert/lay dichotomy can be accounted for on the ground of the pragmatic and sociocultural background. The specific distribution patterns are to be held as a material used for building identity profile, both on the side of expert and lay entities. It has already been ascertained with regard to historical texts that multinomials may act as style markers of specific authors [32]. In our case, the discussion confirms the operation of authorship-based distinctions materialised as the frequency patterns in the use of specific structural types of multinomials, which may be assumed to be employed to serve different communicative functions of stylistic or pragmatic nature. Two conclusions come to the fore here. Firstly, we find that multinomials that are commonly held as markers of highly conventional style are not reserved as a stylistic phenomenon to expert communication. They are not relicts of the past, reserved only for highly conventional and traditional legal communication significantly based on scholarly academic rhetoric, thus featuring primarily the expert legal style. Specific structural categories of legal multinomials play a role in building the group identity of lay users, in our case communicating in the company environment. This confirms the positive operation of the intertextuality-based framework of legal communication across various authorship categories that, in principle, sets the requirement to stick to the terminological regime by recurrent use of specific terms. Secondly, it points to a significant impact of institutional traditions. The immersion in a corporate culture induces the dissemination of specific stylistic traditions involving also the use of sophisticated formulae even among non-expert drafters who, following the corporate work style, often work under time pressure and so the stylistic professionalism involving the use of sophisticated conventional formulae is not the priority for them. Thirdly, as supported by the qualitative data, the distinctions in the said distribution patterns are often pragmatically-laden but pragmatics is to be understood here as the potential to easily and effectively construe a text, possibly on the ground of multinomials being a part of a larger prefabricated formulae (i.e.) and not necessarily acting as carriers of important legal concepts. Contrary to the expectations, the tendency is that lay communication is characterised by increased use of purely ornamental multinomials (representatives of adverbial and prepositional multinomials) the employment of which is not to be assessed as critical but found useful as a material to construct repetitive structures, acting as syntactic frames for a sentence, generic schemes or as fillers that can make a sentence look more professional. It also remains of importance that such formulae, often being processed, do not pose a significant cognitive effort for the users and for the same reason they are successfully processed by electronic means of communication. On the other pole we have the most conventional nomi-

nal or adjectival multinomials, often carriers of norm-related concepts, salient in the population of legal experts featured by distinctive precision of wording. Further, the number of the most classic types, such as true binomials or binomials proper is higher in the expert population. The findings show that the expert/lay distinctions often do not run through the most obvious dichotomies. The somewhat more simplistic stylistic format of lay communication does not necessarily imply decreased use of rather ornamental linguistic formulae devoid of clear subject matter function. Further, lay users prove to be less rigorous when it comes to the reversibility parameter.

The added value of this research is not only associated with the quantitative and qualitative data and patterns emerging out of these, and with the very verification of the hypotheses posed, but it is to be assessed in the context of its contribution to the research in the domain of legal communication that is company registration discourse, which is rigidly delineated with regard to theme and communicative environment. The findings are believed to be of value in didactics and translation methodology in that they may constitute the ground for profiling the curricula and tools for teaching legal language and legal translation in order to facilitate the acquisition of legal English and legal translation competence.

## Declarations

1. The author has no relevant financial or non-financial interests to disclose.
2. The author has no competing interests to declare that are relevant to the content of this article.
3. The author certifies that she has no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.
4. The author has no financial or proprietary interests in any material discussed in this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Basaneže, Katja Dobrić. 2018. Binomials in EU competition law. In *Language and Law. The Role of Language and Translation in EU Competition Law*, eds. Silva Marino, Lucja Biel, Martina Bajčić, and Vilemini Sosoni, 225–249. Springer International Publishing.
2. Bhargava, Mudit, Pulkit Mehndiratta, and Krishna Asawa. 2013. Stylometric Analysis for Authorship Attribution on Twitter. In *Big Data Analytics. Second International Conference, BDA 2013 Mysore, India, December 2013 Proceedings*, eds. Mudit Bhatnagar, and Srinath Srinivasa, 37–48. Springer International Publishing.
3. Bhatia, Vijay Kumar. 1993. *Analysing Genre*. London: Longman.
4. Bhatia, Vijay Kumar. 2006. Legal genres. In *Encyclopedia of Language and Linguistics*, ed. Keith Brown., Amsterdam: Elsevier.

5. Bhatia, Vijay Kumar. 2017. *Critical genre analysis. Investigating interdiscursive performance in professional settings*. London/New York: Routledge.
6. Chapman, Don. 2017. Fixity and Flexibility in Wulfstan's binomials. In *Binomials in the history of English fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 41–62. Cambridge: Cambridge University Press.
7. Cheng, Winnie. 2014. Corpus analyses of professional discourse. In *The Routledge Handbook of Language and Professional Communication*. eds. Vijay Kumar Bhatia, and Stephen Bremner, 13–25. New York: Routledge & Francis Group.
8. Clark, Urszula, and Fikret Berkes. 2013. *Language and Identity in Englishes*. New York: Routledge & Francis Group.
9. Coyotl-Morales, Rosa, and María, Luis Villaseñor-Pineda, Manuel Montes-y-Gómez, and Paolo Rosso. 2006. Authorship Attribution Using Words Sequences. In *Progress in Pattern Recognition, Image Analysis and Applications*, eds. José Francisco Martínez-Trinidad, Jesús Ariel Carrasco-Ochoa, and Josef Kittler, 844–853. Springer International Publishing.
10. Dodswoth, Robin. 2014. Speech communities, social networks, and communities of practice. In *Research Methods in Sociolinguistics. A practical guide*, eds. Janeth Holmes, and Kirk Hazen, 262–275. Hoboken: Wiley Blackwell.
11. Edwards, John. 2009. *Language and Identity: an introduction*. Cambridge: Cambridge University Press.
12. Engberg, Jan. 1993. Prinzipien einer Typologisierung juristischer Texte. *Fachsprache 2*: 31–38.
13. Evans, David. 2016. Introduction. In *Language and Identity: discourse in the World*, ed. David Evans, 3014. London: Bloomsbury Publishing.
14. Fulk, Robert. 2017. Pragmatic and stylistic functions of Binomials in Old English. In *Binomials in the history of English fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 27–40. Cambridge: Cambridge University Press.
15. Giles, Howard. 2009. The process of communication accommodation. In *The New Sociolinguistics Reader*, eds. Nicolas Coupland, and Adam Jaworski, 276–287. Basingstoke: Palgrave Macmillan.
16. Goźdz-Roszkowski, Stanisław. 2011. *Patterns in linguistic variation in American Legal English*. Frankfurt am Main: Peter Lang.
17. Goźdz-Roszkowski, Stanisław. 2014. *Evaluative patterns in English and Italian judicial language*. In *Language and Law in Professional Discourse: Issues and Perspectives*. eds. Vijay Kumar Bhatia, Giuliana Garzone, Rita Salvi, Girolamo Tessuto, and Christopher Williams, 10–28. Cambridge: Cambridge Scholars Publishing.
18. Grañ, Johannes, and Martin Volk. 2021. Binomial adverbs in Romance and Germanic Languages – A corpus-based study. In *Corpora in Translation and Contrastive Research in the Digital Age: Recent advances and explorations*, eds. Julia Lavid-López, Carmen Maíz-Arévalo, and Juan Rafael Zamorano-Mansilla, 326–342. Amsterdam: Benjamins Translation Library.
19. Hansen-Schirra, Silvia, and Silke Gutermuth. 2015. Approaching comprehensibility in translation studies. In *Translation and Comprehensibility*, eds. Karin Maksymski, Silke Gutermuth, and Silvia Hansen-Schirra, 53–76. Berlin: Frank and Timme.
20. Holmes, Janet, and Marra Meredith. 2014. The complexities of communities in professional workplaces. In *The Routledge Handbook of Language and Professional Communication*. eds. Vijay Kumar Bhatia, and Stephen Bremner, 112–128. New York: Routledge & Francis Group.
21. John, Ann M. 1997. Discourse communities and communities of practice: membership, conflict, and diversity. In *Text, Role, and Context: developing academic literacies*, 51–70. Cambridge: Cambridge University Press.
22. Kjør, Lise Anne. 2007. Phrasemes in legal texts. In *Phraseologie. Ein Internationales Handbuch der Zeitgenössischen Forschung/Phraseology. An International Handbook of Contemporary Research*, eds. Harald Burger, Dmitrij Dobrovol'skij, Peter Kuen, and Neal R. Norrick, 506–515. Berlin/New York: Walter de Gruyter.
23. Van Koevering, Katherine, Austin R. Benson, and Jan Kleinberg, J. 2020. Frozen Binomials on the Web: Word Ordering and Language Conventions in Online Text. In *WWW '20: Proceedings of The Web Conference 2020*, 606–616. <https://doi.org/10.1145/3366423.3380143>.
24. Kopaczyk, Joanna. 2013. *The Legal Language of Scottish Burghs*. Oxford: Oxford University Press.
25. Kopaczyk, Joanna, and Hans Sauer. 2017. Defining and exploring binomials. In *Binomials in the history of English: fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 1–25. Cambridge: Cambridge University Press.



26. Koskeniemi, Inna. 1968. *Repetitive word pairs in Old and Early Middle English prose*. Turun Yliopisto.
27. Kotake, Tadashi. 2017. Binomials or not? Double glosses in Farman's gloss to the Rushworth Gospels 82. In *Binomials in the history of English fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 82–97. Cambridge: Cambridge University Press.
28. Kubaschewski, Elizabeth. 2017. Binomials in Caxton's Ovid (Book I). In *Binomials in the History of English Fixed and Flexible*, Joanna Kopaczyk, and Hans Sauer, 141–150. Cambridge: Cambridge University Press.
29. Larner, Samuel. 2017. Formulaic word n-grams as markers of forensic authorship attribution. In *Phraseology in Legal and Institutional Settings*, eds. Anne Wagner, Vijay Kumar Bhatia, Stanisław Goźdz-Roszkowski, and Gianluca Pontrandolfo, 258–276. London/New York: Routledge.
30. Lohmann, Arne. 2012. A processing view on order in reversible and irreversible binomials. *Vienna English Working Papers* 21: 25–50.
31. Motschenbacher, Heiko. 2013. Gentlemen before ladies? A Corpus-Based study of Conjunct Order in Personal Binomials. *Journal of English Linguistics* 41 (3): 212–242.
32. Ogura, Michiko. 2017. Binomials, word pairs and variation as a feature of style. In *Binomials in the history of English fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 63–82. Cambridge: Cambridge University Press.
33. Romaine, Suzanne. 2009. Corpus linguistics and sociolinguistics. In *Corpus Linguistics – An International Handbook*, eds. Anke Lüdeling, and Meria Kytö, 96–112. Berlin/New York: Walter de Gruyter.
34. Šarčević, Susan. 1997. *New Approach to Legal Translation*. Alphen aan den Rijn: Kluwer Law International.
35. Sauer, Hans. 2019. In defence of Lydgate: Lydgate's use of binomials in his TROY BOOK (PART 1). *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 136: 227–244.
36. Sauer, Hans, and Birgit Schwan. 2017. Heaven and earth, good and bad, answered and said: a survey of English binomials and multinomials (part 1). *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 134: 83–96.
37. Sprau, Melani. 2017. Binomials and Multinomials in Sir Thomas Elyot's the Boke named the Gouvernour 201. In *Binomials in the history of English fixed and flexible*, eds. Joanna Kopaczyk, and Hans Sauer, 201–221. Cambridge: Cambridge University Press.
38. Turnbull, Judith. 2014. Expert to Lay Communication: Legal Information and Advice on the Internet. In *Language and Law in Professional Discourse: Issues and Perspectives*. eds. Vijay Kumar Bhatia, Giuliana Garzone, Rita Salvi, Girolamo Tessuto, and Christopher Williams, 62–77. Cambridge: Cambridge Scholars Publishing.
39. Więclawska, Edyta. 2019. Sociolinguistic and grammatical aspects of English Company Registration Discourse. *Humanities and Social Sciences* 26 (4): 185–195.
40. Więclawska, Edyta. 2021. Quantitative distribution of verbal structures with reference to the authorship factor in legal stylistics. *Studies in Logic Grammar and Rhetoric* 66 (1): 147–165.
41. Więclawska, Edyta. 2023. *Binomials in English/Polish Company Registration Discourse. The Study of Linguistic Profile and translation Patterns*. Vandenhoeck and Ruprecht Verlage. ISBN 978-3-8471-1536-6.
42. Zhang, Zuo Cheng. 2015. Unpacking professional identities for Business English students. In *Language and Identity across Modes of Communication*. eds. Dwi Noverini Djenar, Ahmar Mahboob, and Ken Cruickshank, 291–312. Berlin/New York: Walter de Gruyter.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.