THE “WORD FACTORY”
A STUDY OF THE PROCESSES ENGAGED IN THE FORMATION OF LEGAL TERMS

by

Isabel Alice Walbaum Robinson

THE “WORD FACTORY”. A STUDY OF THE PROCESSES ENGAGED IN THE FORMATION OF LEGAL TERMS.¹

by
Isabel Alice Walbaum Robinson

Abstract

This paper explores the language of the law from the point of view of the factors that contribute to the building of its specialized lexicon and the options language offers to create, institutionalize and incorporate new words into the technical legal inventory. Sources of word-formation responsible for specialized lexicon accretion in the corpus are of two kinds: language-based and discipline-based. The former involves structure (syntax), meaning (semantics), and uses (pragmatics) of language and the processes engaged in word formation, such as compounding, trans-categorization, fusion, clipping, the creation of binomials and multi-word lexical units (henceforth, MLUs). The latter involves the creation of words engendered from within the discipline itself, such as landmark cases and professional contributions.

As a language user, the legal professional is faced with two types of cognitive processing. On one hand, ‘decoding’ text meaning when carrying out language-related activities such as reading documents or listening to speech. On the other, ‘encoding’ language for the purpose of writing (e.g. briefs, opinions, articles) or interacting with a colleague, taking part in a debate or discussing a point of law. Awareness of both lexical and syntactic features of a language for specific purposes (henceforth, LSP), as well as its word-forming processes, provides the legal professional a vantage for interpreting, comparing and using the ‘tools of the trade’, legal terms. Word-formation is an important sector of linguistics that reflects back on the nature and characteristics of language itself. It gives an ‘added dimension’ to lawyers, in particular English as a Foreign Language (EFL) comparative lawyers, who fulfil professional obligations that call for the technical, rather than intuitive use of language.

Keywords: language for specific purposes, LSP, law English, legal lexicography, lexicology, lexical inventory, lexicon enlargement, word formation, institutionalization, lexicalization.

¹ The material for this article was first collected in 2008 while involved in a two-year research project (2006-2008) at my university, financed jointly by the Italian Ministry of Education, University and Research (MIUR) and the Peretti Foundation. The research objectives regarded the adoption and implementation of foreign legal language in the university curricula for the training of the XXI century Italian and European lawyer.

♦ Università degli Studi di Roma Tre, Facoltà di Giurisprudenza, Dipartimento di Linguistica.
Introduction

Harris (1994: 156) claims that the legal profession is “a profession of words” and that “language penetrates the legal system, and ultimately and utterly depends on some form of linguistic negotiation.” While both the comparative lawyer and the linguist consider language an important ‘tool of the trade’, the approach taken by each with respect to language is quite different. The interface between these two disciplines means, for the linguist, to come in contact with the nature, principles, doctrine and philosophical foundations that underpin the legal sciences. Conversely, for the comparative lawyer, interdisciplinary engagement means developing a sensitivity towards the syntactic, semantic and pragmatic features of language as well as the rules favouring or sanctioning the creation or renewal of words in the legal inventory.

Those who read this paper need to be informed that the study is not focused on comparative law but on the lexical and structural features of the legal lexicon involved in the adoption of items into the legal inventory. Although the focus is on law English, the underlying assumption forwarded here is that knowing how words are formed in any language promotes a more efficient use of language, helping raise the level of analysis beyond intuitive guess-work often practiced by legal professionals (Solan 2005). Following is a synopsis of the four processes engaged in the ‘word factory’ discussed here: neologisms, borrowings, abbreviations and legal acquisitions.

1. Defining the terms

In lexicography, a distinction is made between three related terms, lexicon, dictionary, and vocabulary. By lexicon is meant “the sum total of the words in a language.” Dictionary and vocabulary, often used interchangeably to denote ‘the entire stock of lexical entries in a single volume,’ are not in fact synonymous. By dictionary is meant a reference source for words usually arranged alphabetically representing the lexicon of a language. By vocabulary is meant the language used by a particular sector (e.g. computer vocabulary), a particular author (e.g. Shakespeare’s vocabulary), or a particular body of knowledge or discipline (e.g. legal vocabulary) (De Mauro, 2005: 11). In this paper, however, reference will be made to ‘legal vocabulary’ and ‘legal dictionary’ without differentiation, given the fact that the authors of the legal volumes in the corpus have chosen to call them ‘dictionary’ rather than ‘vocabulary’.
Lemma, another word used here, is the dictionary equivalent of lexical unit. It is defined as “a meaningful linguistic unit that is an item in the vocabulary.”

The meaning of a word is complex, and needs to be analyzed from more than one perspective. Zgusta (1971) distinguishes between a word’s designation, connotation and range of application. The author defines designation as “all the relations existing between the single word and the single parts of the extra-linguistic world, as conceived by the speakers of the language” (ibid: 27). Designation includes three dimensions: (a) the form of the word expressed or capable of being communicated to a person by written or oral communication; (b) the designatum or the concept the word triggers; and (c) the denotatum or the corresponding segment of the extra-linguistic world the word is related to (ibid: 34). What happens when a new word is invented? Within the above framework, a new invention, i.e., a new word, would result in a new denotatum, which in turn would be the cause for a new lexical unit with a new form and a new designatum.

Connotation refers to those semantic properties, register and style, of a word that fall outside the core area of designation. For example, the legal term to defraud acquires a whole new connotation if expressed with colloquialisms or «informal register» such as trick, swindle (to cheat out of money or property), or jip (to induce a person to pay too much for something of low quality or value).

Range of application has to do with the restrictions of the use of a word based on its individual properties and the specific context of situation it is used in. It entails knowing a language well enough so as to be able to spot differences in applicability of a word’s use. For example, the words pedigree and genealogy share the same basic meaning «line of descent» and connotation «formal register», but not the same range. Pedigree applies to animals, genealogy to human beings. Jiang (2004: 119) gives an interesting comparative example of the range of application of English meeting and Chinese huiyi. Meeting and huiyi share the basic meaning «occasion where persons gather to discuss things» and connotation «formal register». The difference lies is the number of individuals gathered. Two individuals can have a meeting in English, whilst it takes a minimum of three for a huiyi. In Chinese, if only two persons gather, it is said they are having a tianhua (i.e. a talk).

2. Theoretical assumptions

Lexical units in a specialized vocabulary are motivated by specific functional and communicative objectives, with variations based on the activities and practices found in each disciplinary sub-field.

---

2 Lexical unit refers to the combination of a word’s form and its unique, identifiable meaning.
3 Merriam Webster Collegiate Dictionary, MWCD.
According to Sobrero (1993: 237-238), in a technical language variety, i.e., the language used to communicate arguments related to a particular discipline or professional activity (e.g. maths, biology, linguistics, law), inventory accretion is achieved by the creation of a number of sub-codes. Typical of a sub-code is that of having a specialized lexicon, configured as a true and proper nomenclature, i.e., a group of terms \(^5\) each of which has a conceptually explicit definition within a taxonomic hierarchy. There exist a number of shared lexical features, the author points out \((ibid: 244)\) that characterize LSPs: precision, economy, impersonality, conciseness, timelessness and conservatism. These features contribute to the principal remit of any specialized lexicon which consists in “univocally denoting concepts, objects, and activities that do not recur in common, day-to-day language,\(^6\)” without having to recur to the general-use lexicon \((ibid: 244-245)\).

In each sub-code, Sobrero (1993: 244-245) adds, denoting is accomplished by four linguistic processes of specialization: a) the use of terms already in existence in the general-purpose lexicon to which a specialized or ‘technical’ meaning is given; b) the creation of neologisms; c) borrowing of terms from other languages; and d) the creation and incorporation of abbreviations.

De Mauro (2005), a prominent linguist and academic, has written a book entitled \textit{La Fabbrica delle Parole} (En. \textit{The Word Factory}). The title captures the essential thinking of this paper, i.e., the question of lexicon enlargement and the means by which languages build specialization. According to the author, in order to enlarge and renew a word inventory, resorting to exogenous sources is often the most common way to proceed. Italian for example has from antiquity incorporated and adapted Latin and Greek terminology by the active incorporation of Latinisms and Greek terms. De Mauro adds that the true ‘factories of lexicon renewal’ occur through the semantic expansion of the meaning of a general-use word by means of linguistic transformations such as: a) trans-categorization\(^7\); b) the generation of complex lexemes (e.g. by the formation of multi-word lexical units); and c) neologisms via morphological, transformative processes such as derivation\(^8\), composition (also known as compounding)\(^9\), onomatopoeic terms, acronyms \((ibid: 146-147)\).

\(^{5}\) Term is understood here as a word or expression that has a precise meaning, peculiar to a science, profession or subject area (e.g. legal –) as defined in \textit{MWCD}. In this broad conception, it is equated to the form \textit{word}.

\(^{6}\) Translated by I.A. Walbaum Robinson.

\(^{7}\) It consists in the formation of a new word from an already existing one by means of category change. The new word is said to be “trans-categorized” in that it changes grammatical category (e.g., V > N transformations) such as nominalization discussed in sub-section 4.1.6.

\(^{8}\) The linguistic process which consists in the addition of an affix (prefix, infix or suffix) to an already existing word. For example, the addition of the suffix –itis in medicine to a word such as derma to form a word that defines an inflammatory disease of the skin, known as dermatitis. Likewise, the chemical elements with –ium suffixation to form actium, berkelium, chromium, germanium, or with –on to form argon, boron, carbon, silicon (Sobrero 1993: 244)

\(^{9}\) The combination of two free roots to form a new word (e.g., ‘eye’ + ‘witness’ > eyewitness).
Underlying lexical renewal and incorporation, the processes that contribute to the ‘shaping’ of the specialized vocabulary are: word-formation, institutionalization and the incorporation of previously non-institutionalized words, known as lexicalization.10

Whereas a substantial body of literature has attempted to characterize the language of the law by focusing on legal sociolinguistics (Bellucci, 2002), syntactic, semantic and pragmatic features of the language of the law and its complexities (Mortara Garavelli, 2001), diachronic processes of word-formation that lead to the incorporation of foreign words into the specialized lexicon (Baugh & Cable, 1951; Mellinkoff, 1963), less attention has been given to word-formation processes that have contributed to the creation of the legal sub-code(s). This study will argue that knowing about the syntactic, semantic and pragmatic features and underlying processes involved in specialized language word-formation allows language users and communicators to perfect their knowledge of the tools of the trade with which to efficiently interpret, compare, find equivalences, and interact.

3. Methodology

To explore the distinctive features of the language of the law, i.e., the syntactic and lexical features that characterize this subject-specific language, a corpus and a framework were developed. The characteristics and content of the corpus and its sources will be discussed in subsection 3.1. The framework (adapted from Sobrero 1993), designed to represent some of the core variables involved in LSP word-formation and the rationale behind it, are dealt with in subsection 3.2. Regrettably, space allows for only some broad generalizations of the language contained in the corpus. The research question and its running arguments are discussed in subsection 3.3. Together these three sub-sections constitute the methodological underpinnings of the study of the ‘word factory’.

---

10 Defined as the historical phenomenon resulting in the ‘institutionalized adoption’ of specialized terminology; a broad definition based on Quirk et al. 1985 and others to which I adhere.

11 Diachronic linguistics is the branch of linguistics that studies the historical transformation of language.
3.1 The corpus

The corpus includes the following dictionaries:

<table>
<thead>
<tr>
<th>Volumes</th>
<th>Contents</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merriam-Webster’s Dictionary of Law (henceforth, MWDL)</td>
<td>Number of words = 10,000+ from circa 15,000,000 citations in the Merriam-Webster offices supplemented by various on-line legal databases.</td>
<td>Designed to meet the needs of the professional and general user readership.</td>
</tr>
<tr>
<td>Black’s Law Dictionary (BLD)</td>
<td>Number of words, 10,000+ from a wide spectrum of original sources including a number of disciplinary areas from bankruptcy to securities law, from legal realism to critical legal studies.</td>
<td>Designed to produce an accurate, comprehensive dictionary that gives a balanced treatment of modern legal terms.</td>
</tr>
<tr>
<td>Merriam-Webster’s Collegiate Dictionary (MWCD)</td>
<td>Number of words = 165,000+ including: definitions = 225,000; new words and meanings = 10,000; idioms and phrases = 7,500; foreign words and phrases = 700</td>
<td>Collegiate dictionaries, initially targeted for the mature student, often consulted by professionals, workers and the public in general.</td>
</tr>
</tbody>
</table>

Other sources of consultation were the Online etymology dictionary and the Web.

The reason for the choice of dictionaries as corpus is consistent with the notion that lexicographic collections are representative sources of a large part of the total ‘terms of inventory’ contained in a given lexicon. As such, dictionaries are useful synchronic representations of what a technical inventory contains, i.e., a univocally defined collection of lemmas that has undergone three phases of change: coinage, conventionalization and incorporation.

3.2 The framework

The framework is sub-divided into four parts: a) types of integration; b) word-formation processes; c) syntactic features; and d) examples (see Table 1). Integration types include three of the four enlargement processes proposed by Sobrero (1993: 244-245), neologisms, foreign borrowings, and abbreviations and a fourth type, legal acquisitions. The latter was added for the purpose of filling a specific gap in the legal lexicon. It includes terms that relate to landmark cases and entries from professional contributions.

---

12 The first word formation process in Sobrero (1993: 245; cf. §2, a) above) that traces “the use of terms already in existence in the general-purpose lexicon to which a specialized or ‘technical’ meaning is given” used above all in the sciences, will not be discussed here. Quotation translated by I.A. Walbaum Robinson.
Table 1. Word-formation framework.

<table>
<thead>
<tr>
<th>Types of integration</th>
<th>Word-formation processes</th>
<th>Syntactic features</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 Neologisms</td>
<td>Open compound</td>
<td>Group entry: key term (superordinate); differentiation (co-subordinate)</td>
<td>consideration; adequate ~, fair ~; good ~</td>
</tr>
<tr>
<td></td>
<td>Closed compound - noun</td>
<td>Two root combinations, single meaning</td>
<td>eyewitness</td>
</tr>
<tr>
<td></td>
<td>Closed compound - adverb</td>
<td>Two or three element compounds:</td>
<td>herein, heretofore; therein, theretofore</td>
</tr>
<tr>
<td></td>
<td>Multi-word lexical unit</td>
<td>Two, three, four, five word NPs(^\text{13})</td>
<td>judicial precedent, acceptance of responsibility, charitable remainder of trust, administrator with the will annexed</td>
</tr>
<tr>
<td></td>
<td>Binomial</td>
<td>Same word class terms joined by a conjunction</td>
<td>will and testament</td>
</tr>
<tr>
<td></td>
<td>Nominalization</td>
<td>Trans-categorization or word category change [e.g. V &gt; N]</td>
<td>attornment, accretion, procedure, ordinance; bailor, bailee</td>
</tr>
<tr>
<td>Type 2 Borrowings</td>
<td>Integrated/non integrated</td>
<td>Single term or phrase from Latin, French or Anglo-French (with/without form change)</td>
<td>L: bona fides, certiorari; Fr: action; A-Fr: indict</td>
</tr>
<tr>
<td>Type 3 Abbreviations</td>
<td>Initialism</td>
<td>Phrase &gt; initials by clipping + fusion</td>
<td>CISG</td>
</tr>
<tr>
<td></td>
<td>Acronym</td>
<td>Phrase &gt; abbreviation by clipping + fusion</td>
<td>COBRA</td>
</tr>
<tr>
<td>Type 4 Legal acquisitions</td>
<td>Landmark case</td>
<td>N + N &lt; proper noun + legislative product</td>
<td>Miranda rights</td>
</tr>
<tr>
<td></td>
<td>Professional contribution</td>
<td>Simple terms or phrases &lt; expert input</td>
<td>right to privacy</td>
</tr>
</tbody>
</table>

Types 1-4 illustrate nine different processes, resulting in the formation of compounds, multiword lexical units, binomials, nominalization, direct or integrated borrowing, clipping, and fusion, legal acquisitions and professional contributions. The last column shows examples of each classification.

\(^{13}\) NP = noun phrase; V = verb; N = noun.
I. A. Walbaum Robinson, The “Word Factory”...

3.3 The research question

To formulate the proper research question, several running arguments about legal discourse and the specialized lexis are considered. First, Sacco (2005: 20) who claimed that “no word in any language corresponds to a word in another language.” Second, Solan’s (2005) alternative proposal to the ‘intuitive’ practice on behalf of professionals of relying on both the strength of their knowledge of a given language, or if need be, the use of legal dictionaries, to identify word meaning. He claims (2005) that to know the full semantics of any given word, particularly if two or more languages are involved, means engaging two disciplines, linguistics and law. In Solan’s words (2005: 126), “while there are many disputes over the meanings of statutory words” among professionals, particularly judges in U.S. courts, very few of these are centred on syntax. In fact, the author points out (ibid: 127) that “disputes over aspects of meaning that derive from syntax are virtually absent with respect to statutes” (e.g. United States v. Dixon). Accordingly, Solan (2005) suggests, “the interface between these two systems [linguistics and law] is an important area of research in linguistics today”.

Based on the above, and consistent with the principal research issues in this paper, the main question this study formulates is: In which way does knowledge of the processes engaged in word-formation contribute to casting a light on the nature and characteristics of legal language? In the following section (§ 4), answers to this question will be provided by illustrating several lexical and structural processes responsible for the enlargement of the legal lexicon in the ‘word factory.’

4 The data

Aside from a single instance in which a word is expressed, known as hapax legomenon, when a word is used for the first time, the innovation is known as nonce word or coinage. A new word

is formed by applying regular word formation rules. It serves an immediate communicative need of, or solves a problem for, the speaker, whether it be economizing, filling a conceptual/lexical gap, or creating a stylistic effect. (Bussman 1996, cited in Brinton & Traugott 2005: 45).

Brauer (1983: 45) writes that a “form ceases to be a nonce formation as soon as speakers using it are aware they are using a term which they have heard already; that is to say, virtually immediately.” This gives way to the next stage, known as institutionalization, which takes place when “the nonce formation starts to be accepted by other speakers as a known lexical item” (ibid: 48). In the case of a specialized term, it is the professional community that either sanctions or assimilates, i.e., ‘institutionalizes’ the

---

14 This is not the same as a nonce formation which is a word made up “for the occasion.”
nonce formation. For some scholars, institutionalization is succeeded by what is known in English word-formation as *lexicalization*. A word is said to be lexicalized when form and meaning become fixed and incorporated into the lexicon.

This is how Brauer (1983) and Brinton & Traugott (2005) illustrate the path from the first time a word is expressed to the time it becomes “official” and may be found in a specialized vocabulary:

Fig. 1. Word-formation stages.

<table>
<thead>
<tr>
<th>Nonce</th>
<th>Institutionalization</th>
<th>Lexicalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>First instance in which a word is expressed.</td>
<td>Word is shared by others in the speech community / specialized knowledge area (e.g. law).</td>
<td>Word is incorporated into the LSP inventory and learnt. Form and meaning are fixed.</td>
</tr>
</tbody>
</table>

In the following sections, several kinds of neologism are analyzed: open compounds (4.1.1); closed compounds - nouns (4.1.2); closed compounds – adverbs (4.1.3); MLUs (4.1.4); binomials (4.1.5); and nominalizations (4.1.6).

4.1 Neologisms

4.1.1 Open compounds

Lemmas in general-purpose dictionaries are for the most part single words. Polysemy is dealt with by differentiating the ‘senses’ of a lemma within a single entry, or by creating separate entries for each ‘sense’ (e.g. *rule*, and separate entries for *rule of the road*, *rule of thumb*). Alternatively, in the legal vocabulary, specialization is represented by what is known as the ‘group entry.’ The group entry consists in a key term (superordinate) under which a number of subordinate open compounds are listed, each denoting a variation of the superordinate term.

Here is an example of the term *consideration* and three of its co-subordinates:

15 MWCD.
Lemma differentiation in the legal vocabulary (Source: BLD). 

Superordinate lemma

Consideration n. Something (such as an act, a forbearance, or a return promise) bargained for and received by a promisor from a promisee; that which motivates a person to do something, esp. to engage in a legal act.

Consideration, or a substitute such as promissory estoppel, is necessary for an agreement to be enforceable.

Co-subordinate lemmas

adequate consideration. Consideration that is fair and reasonable under the circumstances of the agreement. 

Cf. sufficient consideration.

fair consideration. 1. Consideration that is roughly equal in value to the thing being exchanged; consideration given for property or for an obligation in either of the following circumstances: (1) when given in good faith as an exchange for the property or obligation, or (2) when the property or obligation is received in good faith to secure a present advance or prior debt in an amount not disproportionately small as compared with the value of the property or obligation obtained. [...]

good consideration. 1. Loosely, valuable consideration; consideration that is adequate to support the bargained-for exchange between the parts. 2. Consideration based on natural love or affection or on moral duty. [...] 

The group entry format in (1) represents a synopsis of how to approach terminology differentiation in a given knowledge area by way of open compounding which involves the semantic specialization of key words.16 A different kind of neologism, the closed compound, will be discussed in subsections 4.1.2 and 4.1.3.

4.1.2 Closed compounds - nouns

Most noun closed compounds are single lexical units (e.g. eyewitness) made up of two joined free roots, to which a single meaning is given. Roots may also be joined by hyphenation (e.g. cross-examination). Closed compound found in the legal vocabulary reported in Table 2 have been classified into five distinct syntactic patterns:17

---

16 A few examples among more than 150 key words found in the legal dictionaries include: processes (arbitration, mediation), conflict of interest (nuisance, breach, tort), crime (manslaughter, rape, murder, assassination), offence (battery, libel, smuggling), means of possessing property (mortgage, title), intent (malice in fact, malice in law), right to land or other freehold (fee, estate).

17 The pattern range is not exhaustive but representative.
Table 2. Two-root closed compounds (*MWDL*).

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. N+N&gt;N</td>
<td>eyewitness, pawnbroker, householder, factfinder, draftsman</td>
</tr>
<tr>
<td>b. V+N&gt;N</td>
<td>wedlock, leapfrog, cross-examination</td>
</tr>
<tr>
<td>c. N+V&gt;\textit{ing}</td>
<td>rulemaking, housebreaking, lawmaking, blockbusting, logrolling</td>
</tr>
<tr>
<td>d. A+N&gt;N</td>
<td>cyberlaw, plainclothes, mortgage (dead + pledge)</td>
</tr>
<tr>
<td>e. A+V&gt;N</td>
<td>jaywalk, fair-play</td>
</tr>
</tbody>
</table>

The meaning of a closed compound is either transparent or idiomatic. For example, the compound *leapfrog* (literal meaning=«jump made by the frog») is used both to refer to the children’s game where jumping is involved, and the legal term, i.e., «the practice by which cases in England and Wales skip a court hierarchy and are adjudicated by the one above it». Similarly, the term *logrolling* (literal meaning=«the American custom of neighbours assisting one another in the task of rolling a log for burnings») has been given a technical meaning to express «the practice of including unrelated provisions in a legislative bill to attract a wider base of support and ensure the passage of a bill as a whole».

Another Type 1 compound (4.1.3) consists in the joining of an adverb and a preposition or two adverbs and a preposition.

### 4.1.3 Closed compounds - adverbs

These adverbs are made up by a place adverb (*here, there*) and a preposition (*in, by, of, to*) to form words as those listed in (2). In some cases a time adverb (*after, forn*) is added to the compound, resulting in the formation of words such as *hereinafter* < *here* + *in* + *after*. Legal English has long been characterized by the use of these old-fashioned morphological constructions of the type *hereon, thereon*. Tiersma points out (1999: 93) that for Old English speakers rather than to say *under it or under that* they preferred to say *hereunder or thereunder*:\(^{19}\)

(2) *Here* compounds and periods of incorporation:\(^{20}\):

\[a \quad \text{berewith} \quad [\text{O.E.}] \quad \text{berwith} > \text{here} + \text{with} \]
\[b \quad \text{beron} \quad [\text{O.E.}] \quad \text{beron} > \text{here} + \text{on} \]

---

18 A = adjective.
19 Although this kind of language is considered for the most part obsolete, Tiersma comments that lawyers still “continue to sprinkle their prose with *hereunder* and *thereins*”. The author recognizes the importance of compound adverbs in terms of “economy of expression” (1999: 94).
20 *MWDL*. 

Two and three component closed compounds pack spatial and temporal notions which otherwise would require extra words to express. In keeping with the principle of economy, rather than to say «the conditions in this contract are listed later in the document» a speaker has the option of «the conditions in this contract are listed hereinunder».

4.1.4 Multiword lexical units

MLUs are what Zgusta (1971: 145, 147, 154) refers to as ‘set word combinations’ or combinations that carry lexical meaning as a whole (i.e. as a single lexical unit). They function as a single word and have a single designative function within the sentence. The difference between an MLU (open mortgage clause) and an idiomatic expression (piercing the corporate veil, blackmail) is that whereas the meaning of the former (MLU) is derived from its components (i.e., the expression is transparent), the former (idiomatic expression) lacks transparency (i.e., it has a figurative meaning). Corporations do not have veils to pierce and mail is not coloured black).

MLU’s are also distinguished from collocations. A collocation consists in a non-lexicalized combination of two or more words that co-occur not by chance but regularly. Collocations are not part of the lexicon, and thus, are not found in a dictionary or vocabulary except in collocations dictionaries.21

It suffices to peruse any law vocabulary to encounter a significant number of MLUs. Expressions of this sort are made up of varying degrees of componential complexity. MLUs identified in the legal corpus reported in Table 3 range from two-word combinations: [A+N] associate judge (A, a), and variations (where one of the noun phrase elements is itself formed by two words joined by a hyphen): [A-A+N] high-low agreement (A, b) or [N-N+N] asset-depreciation range (A, c). Also found in the corpus were a large number of three-word combinations: [N+P+N] identity of interest (B, a); four-word combinations: [A+N+P+N] effective assistance of counsel (B, a); and 5-word combinations: [N+P+Det+N+A] administrator with the will annexed (D, a).

21 Native speakers can easily spot regular word combinations; for this, collocations constitute an important group of combinations to learn, particularly for English as a foreign language (EFL) speakers. For example, to a native speaker it sounds natural to combine the adjective ‘strong’ with ‘tea’ in, I prefer strong tea for breakfast. If a synonym is used, such as ‘powerful’ in, I prefer powerful tea for breakfast, ‘powerful tea’ is considered odd. Likewise, if a speaker instead of saying, I need to make a decision, were heard saying, I need to perform a decision, the collocation with ‘perform’ is not viewed as acceptable.
Table 3. MLU typology.

<table>
<thead>
<tr>
<th>Type of MLU</th>
<th>Pattern</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Two-word combinations</td>
<td>A+N</td>
<td>a. associate judge, ameliorative waste</td>
</tr>
<tr>
<td></td>
<td>A-A+N</td>
<td>b. high-low agreement</td>
</tr>
<tr>
<td></td>
<td>N-N+N</td>
<td>c. asset-depreciation range</td>
</tr>
<tr>
<td>B. Three-word combinations</td>
<td>N+P+N</td>
<td>a. identity of interest, acceptance of responsibility</td>
</tr>
<tr>
<td></td>
<td>A+N+N</td>
<td>b. open mortgage clause; hot cargo agreement</td>
</tr>
<tr>
<td>C. Four-word combinations</td>
<td>A+N+P+N</td>
<td>a. effective assistance of counsel</td>
</tr>
<tr>
<td></td>
<td>A+N+N+N</td>
<td>b. charitable remainder annuity trust</td>
</tr>
<tr>
<td></td>
<td>N+P+Det+N</td>
<td>c. accessory before the fact / accessory after the fact</td>
</tr>
<tr>
<td>D. Five-word combinations</td>
<td>N+P+Det+N+A</td>
<td>a. administrator with the will annexed</td>
</tr>
<tr>
<td></td>
<td>N+A+P+N+A</td>
<td>b. fee simple on condition subsequent</td>
</tr>
</tbody>
</table>

4.1.5 Binomials

These are sequences of word combinations, generally two, which belong to the same word class (noun, verb, adjective) coordinated by an additive (and), adversative (but) or alternative (or) conjunction. The most frequent conjunction is by far and (see 3 below). The language of the law makes generous use of binomials and multinomials. Acts of Parliament, for instance, invariably begin with the following enacting formulation:

(3)  
Be it enacted by the Queen’s most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows:—

In this formulation there are two binomials (by and with; advice and consent) and a multinomial (Lords Spiritual and Temporal, and Commons).

Binomials are also prevalent in the stipulation of wills, where the purpose is that of strengthening “the primary performative act of bequeathing” (Danet & Bogoch, 1994: 113). Several examples by the

---

22 P = preposition; Det = determiner (article).
I. A. Walbaum Robinson, The “Word Factory”...

authors are included in Table 4, in which four different binomial patterns are illustrated. Often, these near-synonym pairs are formed by the combination of a Latin and a Germanic origin word.

Type 1 patterns consist in two nouns joined by and where either the first term is English and the second Latin, or both are Latin. For example in will and testament, will is from Old English willan and its variant wyllan meaning «to wish, desire, want»; testament is from Latin testamentum meaning «a will». In part and parcel, part is from L. partem (nom. pars, gen. partis) meaning «part, piece, side, share» and parcel is V.L.\(^24\) particella, from L. particula meaning small portion. The same with null and void. Null comes from L. nullus «not any, none» and void is from L. vocivus «unoccupied, vacant». Types 2 and 3 are formed by the joining of two verbs (with/without ‘to’+infinitive form). In aid and abet, aid is from L. adiuware (pp. adiutus) meaning «to give», whereas abet is from Germanic beter «to bait» possibly from Low Franconian betan «incite».\(^25\) In Type 4, due and owing is made up of two adjectives joined by and.

Table 4. Types of binomial expressions joined by and.

<table>
<thead>
<tr>
<th>Typology</th>
<th>Pattern</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td>N+Conj(^{26})+N</td>
<td>will and testament, part and parcel, null and void</td>
</tr>
<tr>
<td>Type 2</td>
<td>V(<em>{inf}) with ‘to’+Conj+V(</em>{inf}) with ‘to’</td>
<td>to have and to hold</td>
</tr>
<tr>
<td>Type 3</td>
<td>V(<em>{inf}/no ‘to’)+Conj+V(</em>{inf}/no ‘to’)</td>
<td>aid and abet</td>
</tr>
<tr>
<td>Type 4</td>
<td>A+Conj+A</td>
<td>due and owing</td>
</tr>
</tbody>
</table>

Following is a frequently used [V>N] transformation in law English, known as nominalization.

4.1.6 Nominalization

The timeless nature of a legal term is achieved linguistically by the use of the passive rather than the active form of the verb, where the ‘agent’, ‘referent’, or ‘person’ responsible for a thought or action is not indicated (Lo Cascio 2007: 208).

This may also be obtained by nominalization. An explanation for the frequent use of this kind of word formation process in the legal dictionary is the need to express timelessness and impersonality by means of language (Sobreño 1993: 234-244). This is achieved according to Simone (2006: 9), by transforming a verb to a noun (i.e., Event > Entity), to obtain both lack of emotional expressivity and the reduction of part of the temporal information. The transformation causes parameter, argument distribution, temporality and agency changes. Nominalization is a complex linguistic transformation that affects the verb and the noun in the following ways:

---

\(^{24}\) V.L. = Vulgar Latin.

\(^{25}\) Information on the etymology of each foreign origin component of binomials examined is from the Online etymology dictionary. Site: [http://www.etymonline.com](http://www.etymonline.com). Accessed on 31.5.10.

\(^{26}\) Conj. = conjunction.
a) at the conceptual-cognitive level, [nominalization] converts a State-of-Affairs or an Event into a compact Entity; b) at the linguistic level, it converts a clause into a noun, modifying the original argument distribution and suppressing some parameters (e.g. the temporal information is cancelled; the agent is removed); c) it allows for a Theme; and d) it also produces new lexical entries (Simone, 2006: 9).

Payne (1997: 226-227) makes a distinction between several kinds of nominalization. Of particular interest here is the distinction between action, agent and patient nominalization. In the following subsections a representative list of examples of action (4), agent (5) and patient (6) legal nominalizations are provided.

4.1.6.1. Action nominalization
According to Payne (ibid: 224) “an action nominalization refers to the action, usually in the abstract, expressed by the verb root.” It is formed by the morphological strategy of adding a suffix to the root of the verb (e.g. -ment, -ation, -ure, -ance). An important linguistic effect of action nominalizations is that of obscuring the Event, whilst the Entity is given greater prominence (cf Simone, 2006, sub§ 4.1.6 above). More productive nominalizations are those in (4, a-b). Less productive are those in (4, c-d). The list below is by no means exhaustive.

(4)

a With operator –ment Letter A examples in MWDL: abandonment, abatement, abutment, acknowledgement, assessment, assignment, attachment, attornment
b With operator –ation accretion, intervention, litigation, obligation, presumption
c With operator –ure procedure < verb to proceed; failure < verb to fail
d With operator –ance ordinance < verb to order; conveyance < verb to convey

4.1.6.2. Agent nominalization
This kind of nominalization is a strategy employed by speakers of a language that involves the transformation of a verb to a noun by adding a suffix. In legal English the most frequently used agent nominalization is realized by adding the operator –or to the verb.

---

27 Translated by I.A. Walbaum Robinson.
28 Productivity is referred to the use of affixation and other grammatical devices to form new words from already existing ones.
I. A. Walbaum Robinson, The “Word Factory”...

(5)

With operator -or
bailor, donor, feoffor, franchisor, grantor, indemnitor,
mortgagor, obligor, offeror, payor

4.1.6.3. Patient nominalization
This kind is formed by adding the suffix -ee to the root of the verb.

(6)

With operator -ee29
bailee, donee, feoffee, franchisee, grantee, indemnitee,
mortgagee, obligee, offeree, payee

Suffixation pairs -or and -ee denote complementary roles in a legal relationship. The [N+-or] member of the pair «bailor» indicates the Agent of the action coded by the verb, whereas the [N+-ee] member «bailee» indicates the Patient (Beneficiary) of the action. These complementary pairs of Law French origin are productive and the effects of their application uniform semantically (e.g. a donee is given a donation; a franchisee is given a right to conduct a franchise; a grantee is the receiver of a grant; an indemnitee is compensated with an indemnity; an offeree is forwarded an offer). It is possible with pairs of this kind to form a number of new words which in some cases may remain occasional, whilst in other cases may become stabilized (Zgusta (1971: 130). Sub-section 4.2 deals with another important source of lexicon enlargement in law English, borrowings from other languages.

4.2 Borrowings

Discipline-based inventories are also enlarged by the adoption of integrated or non-integrated foreign borrowings. According to Melinkoff (1963), at one point in the historical legacy of law English, foreign borrowings helped build a vocabulary that did not count with enough terms to define legal concepts, actions, activities, events. Mass foreign borrowing are, according to Melinkoff (ibid: 13), a characteristic ontological feature of the English language.

Many of the terms incorporated into the legal inventory, ubiquitous in law vocabularies, have resulted in a large stock of foreign borrowings in use even today. The most important source of technical terms is Latin, followed by a vast number of French and Anglo-French terms. Examples are from Mellinkoff (1963: 14-15), MWDL and BLD.

29 Sometimes referred to as ‘past participle’ since it originates from the French past participle (Payne, 1997: 227).
Baugh & Cable (1951: 166) write that “French was so long the language of the law courts in England that the greater part of the English legal vocabulary comes from the language of the conquerors. The fact that we speak of *justice* and *equity* instead of *gerichte*, *judgment* rather than *dom* (doom), *crime* in place of *synn* [...] shows how completely we have adopted the terminology of French law.”

The preservation of non-integrated foreign borrowings bestows a sense of “tradition”31 to the legal lexicon, as a language which has, amidst rapid change, resisted the passage of time, and hence deserves great respect. An expression such as *pacta sunt servanda* «agreements must be kept» is indicative of the propensity of legal lexical items to capture concisely and in a few memorable words an entire doctrine, principle, authoritative dictum, or maxim. This is the reason why the numerous attempts to modernize the law vocabulary have met with opposition from the legal profession. Tiersma (1999: 197) comments that somehow language that has survived the test of time, appears more established than the innovative phrasing typical of the language of journalism and computer sciences with numerous coined words falling into rapid disuse.32

---

30 Words in this list are mostly from Old French > from Latin or sometimes other languages such as Germanc *seisin* in *MWDL*.

31 It is enough to sit in a graduation commission for an entire day listening to law students presenting their dissertation theses to realize how ingrained the practice of using foreign words is even among young lawyers-to-be.

32 The ‘special’ meaning of the language of the law is not, Tiersma (1999: 51) claims, fortuitous but culturally and historically based. The author claims that “[f]or the most part legal language follows the rules that govern English in general. At the same time it diverges in many ways from ordinary speech, far more than technical languages in other professions.”
I. A. Walbaum Robinson, The “Word Factory”...

4.3 Abbreviations

This section deals with a type of word-formation that has become rather productive in recent years. It consists in new lemmas formed by conflating the initial sounds or letters in a phrase into a unified lexeme (Brinton, 2000: 43) to form what in both MWDL and BLD are defined as “abbreviations.”

There are two kinds of abbreviation: initialisms and acronyms. The first consists in the conflation of the initial letters of a phrase expressing a title, convention, resolution, treaty, company name. The initials of each main word together form a new lexeme whereby each initial is pronounced separately. A few examples from the legal vocabularies examined are reported in (7). Acronyms are abbreviations that result from the conflation of initial sounds of words contained in a phrase, constituting a new word with its own pronunciation (8).

(7) Initialisms:

\[
\begin{align*}
CISG & \quad <C(onvention\ for\ the)\ I(nternational)\ S(ale\ of)\ G(oods) \\
ADR & \quad <A(lternative)\ D(ispute)\ R(esolution) \\
LLC & \quad <Limited\ L(iability)\ C(ompany)
\end{align*}
\]

(8) Acronyms:

\[
\begin{align*}
COBRA & \quad <C(onsolidated)\ O(mnibus)\ B(udget)\ R(conciliation)\ A(ct\ of\ 1986) \\
QTIP trust & \quad <Q(ualified)\ T(erminal)\ I(interest)\ P(roperty)\ t\ rust \\
RICO & \quad <R(acketeer\ I(nfluenced\ and)\ C(orrupt)\ O(rganization)\ A(ct) \\
MUD & \quad <Municipal\ U(tility)\ D(istrict)
\end{align*}
\]

In the next sub-sections, a new kind of technical term is introduced, legal acquisitions, originating from sources within the discipline, such as landmark cases (4.4.1) and professional contributions (4.4.2).

4.4 Legal acquisitions

This section deals with a set of signature terms found only in legal dictionaries. As a form of accretion, legal acquisitions report on the history of the discipline, particularly case-law and the lexicalization of terms by way of professional contributions. Legal acquisitions are an important source of vocabulary accretion. They consist in the creation of new lemmas by taking the name of one of the parties in a landmark case (Miranda rights) or by the contribution of professionals that calls for the coinage of new terminology (right to privacy).

33 Sobrero (1993: 245) views initialisms (Lt. sigle) as semantically autonomous words that blend the initial letters of each or some of the words in a phrase into a series of initials. Acronyms are formed by the blending of initial words or sounds into one (Ibid: 245). My own version is similar to Sobrero’s but with a slight difference in meaning as to how initialisms and acronyms are formed.
4.4.1 Landmark cases

In legal lexicology, an entry based on a proper name such as in a landmark case, known as ‘eponymous’ lemma, is the source of a consistent number of lexical entries in both MWDL and BLD.34 It makes reference to a landmark case (i.e., a case heard in court and subsequently entered into the judicial precedent order) and it is the result of the combination of one or more proper names, indicating a specific person(s) involved in a preceding case, that later becomes a legal product (e.g. act, doctrine, error, material, standard, rule, rights).

Questions have been raised as to whether a proper name should be included in a dictionary at all in the same way the English words «house», «neighbourhood» or «town» are. The main point to keep in mind in relation to the inclusion of proper names in the lexicon is, according to Lyons, whether names belong to a language or not, and if so, whether they “have a symbolic, etymological or translational meaning” (1977: 223). In response to this, it appears to me that entries of this kind are indicative of the fact that law vocabularies have a double function: as “repositories” of the sum total of the stock of words in a particular subject-specific field, and as a “cultural” system of reference. Inclusion of names with a symbolic, historical or cultural meaning, whereby reference is made to a specific Actor, Action, Object in a particular case which later becomes a landmark case, corresponds to the systematic way of managing disciplinary activities that ultimately results in coinage. A few examples are given below (Table 6).

---

34 In these terms, both endogenous sources (activity of the forum) and exogenous sources (case law) are engaged in the word-forming and lexicon incorporation process, bringing “human reality” from the courtroom to the legal inventory.
<table>
<thead>
<tr>
<th>Lemma/Product</th>
<th>Definition/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brady act</td>
<td>A federal law establishing a national system for quickly checking the background of a prospective handgun purchaser <em>(BLD).</em> 18 USCA §§ 921-930.</td>
</tr>
<tr>
<td>D’Oench doctrine</td>
<td>A doctrine in banking law where a party (as a borrower or guarantor) cannot assert an unrecorded agreement with a failed bank against attempts by the federal insurer (as the Federal Deposit Insurance Corporation) or its assigns to collect on a promissory note (as a loan) <em>(WCDL).</em> Case: <em>D’Oench, Dubme &amp; Co., Inc. v. Federal Deposit Insurance Company</em>, 315 U.S. 447 (1942).</td>
</tr>
<tr>
<td>Bruton error</td>
<td>The violation of a criminal defendant’s constitutional right of confrontation by admitting into evidence a nontestifying codefendant’s confession that implicates both of them, where the statement is not admissible against the defendant under any exception to the hearsay rule <em>(BLD).</em> Case: <em>Bruton v. United States</em>, 391 U.S. 123, 88 S. Ct. 1620 (1968).</td>
</tr>
<tr>
<td>Jencks material</td>
<td>A prosecution witness’s written or recorded pretrial statement that a criminal defendant, upon filing a motion after the witness has testified, is entitled to have in preparing to cross-examine the witness <em>(BLD).</em> Case: <em>Jencks v. United States</em>, 353 U.S. 657, 77 S.Ct. 1007 (1957).</td>
</tr>
<tr>
<td>Miranda rights</td>
<td>The rights (as the right to remain silent, to have an attorney present, and to have an attorney appointed if indigent) of which an arresting officer must advise the person being arrested <em>(MWDL)</em> [omitted]. Case: <em>Miranda v. Arizona</em>, U.S. (1966).</td>
</tr>
<tr>
<td>McNabbs-Mallory rule</td>
<td>A doctrine in criminal procedure which establishes that an arrestee must be brought before a magistrate without unnecessary delay in order for a confession made during detention to be admissible <em>(WCDL).</em> Cases: 1) <em>McNabb v. United States</em>, 318 U.S. 332 (1943); 2) <em>Mallory v. United States</em>, 354 U.S. 449 (1957).</td>
</tr>
<tr>
<td>Jackson standard</td>
<td>In criminal law, the principle that the standard of review on appeal – when a criminal defendant claims that there is insufficient evidence to support the conviction – is to determine whether, after considering the evidence in the light most favourable to the prosecution, any rational trier of fact could have found the essential elements of the crime beyond a reasonable doubt <em>(BLD).</em> Case: <em>Jackson v. Virginia</em>, 443 U.S. 307, 99 S. Ct. 2781 (1979).</td>
</tr>
</tbody>
</table>

The next sub-section discusses lexicon accretion from contributions by experts in the legal sector.

### 4.4.2 Professional contributions

This type of term is also unique to the legal vocabulary. Rather than involving the work of the linguist, accretion here comprises the work of the legal expert or professional. An example of a nonce term that
has become all-important in recent years throughout Europe and around the world is the MLU ‘right to privacy’ (It. diritto di protezione dei dati personali), a phrase introduced by two lawyers. The background to the creation of ‘right to privacy’ goes back to the 1890s in Boston, where a noted society woman, married to Samuel Warren, a lawyer and former graduate from the Harvard School of Law, had become estranged from her spouse. She was at the centre of Boston gossip after a series of press articles written about the couple’s personal life was released. The situation deteriorated to the point of becoming unbearable for Mr. Samuel Warren, who decided to take action. Along with his best friend Brandeis, another Harvard graduate, the two published a landmark article entitled ‘The Right to Privacy’ in the Harvard Law Review.\(^3\) As a result, the concept of ‘right to privacy’, formerly understood in a narrower sense as the «right of a person to possess property» was expanded to mean «the right to protect one’s reputation and private life from defamation». An excerpt of the article sheds a light on the importance of the argument that led to the institutionalization of the phrase:

>[T]he intensity and complexity of life, attendant upon advancing civilization, have rendered necessary some retreat from the world, and man, under the refining influence of culture, has become more sensitive to publicity, so that solitude and privacy have become more essential to the individual; but modern enterprise and invention have, through invasions upon his privacy, subjected him to mental pain and distress, far greater than could be inflicted by mere bodily injury (Warren & Brandeis 1890).

The excerpt is indicative of how professionals as experts may at times become responsible for nonce formations by way of the re-interpretation of a point of law. Brauer (1983: 43) notes that “whether or not the new term gains currency will depend upon a number of factors. One of these is the status of the person who used the term.” She claims that acceptance of a new word also depends on the attitude towards the word triggered by the speech community that receives it, and by whether or not the word fills a gap in the lexicon. Resuming Zgusta (1970), ‘right to privacy’ is an example of a new term resulting in a new denotatum that has required the creation of a new MLU with its own designatum.

5 Conclusions

The creation and renewal of a lexical inventory such as the legal hinges as much on rules of grammar and principles of linguistics as it does on knowledge of the ‘rule of law’ and of world legal systems. When working with technical language, greater accuracy and transparency is obtained if interpretation,
comparison or the search for equivalences is not left solely to intuition. Following Zgusta (1970), it has become clearer that in working with language, analysis engages a wide spectrum of parameters. These include the identification of a word’s designation - its morphology, word class, semantic meaning, pragmatic uses - the concept(s) it is related to, the knowledge area and legal system in which the word is conceptualized, defined and used, what the word connotes and its range of application. This paper corroborates the importance of the disciplinary interface between linguistics and the law, stating that more than a cross-roads of ideas, inter-disciplinary convergence represents an ‘added dimension’ in dealing with words and the life words acquire once they are conventionalized and incorporated into the lexicon. Like morning dew reflecting an entire valley in a single drop of water, a word encloses vast areas of knowledge in a single symbolic representation. Learning about the lexicon enlargement processes involved in the ‘word factory’ is useful as a resource in dealing with the complexities of an LSP such as the language of the law.

Acknowledgements

I wish to thank Vincenzo Zeno Zencovich (Università degli Studi Roma Tre, Faculty of Law) for insightful suggestions on some legal issues discussed in this paper. I would also like to thank Edoardo Lombardi Vallauri (Università degli Studi Roma Tre, Department of Linguistics) for his relevant comments on an earlier version of the paper.
References


Dictionaries


