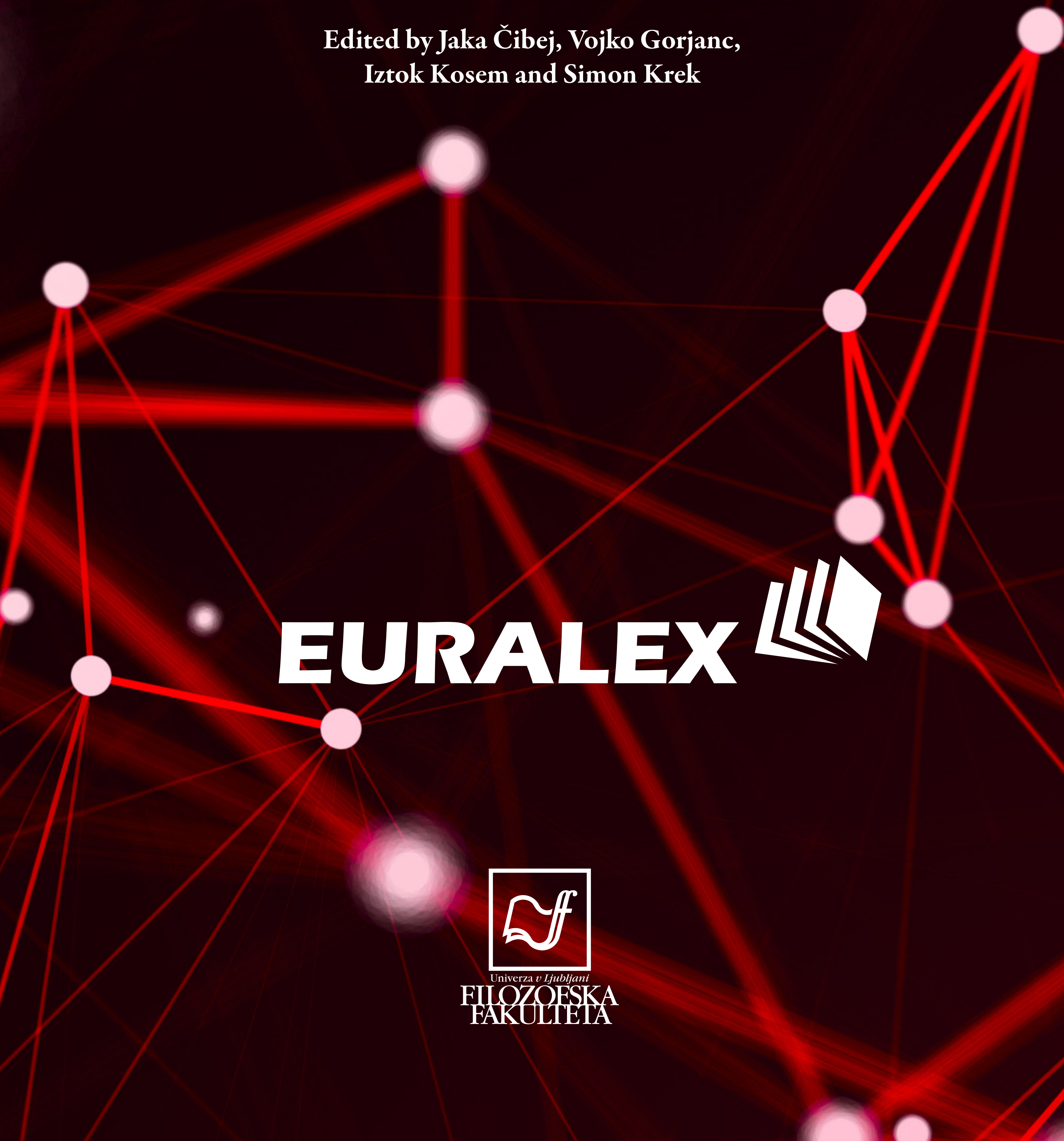
Proceedings of the XVIII EURALEX International Congress

Lexicography in Global Contexts

17-21 July 2018, Ljubljana

Edited by Jaka Čibej, Vojko Gorjanc,
Iztok Kosem and Simon Krek



EURALEX

Univerza *v Ljubljani*
FILOZOFSKA
FAKULTETA

Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts

Edited by: Jaka Čibej, Vojko Gorjanc, Iztok Kosem and Simon Krek
Reviewers: Andrea Abel, Zoe Gavriilidou, Robert Lew and Tinatin Margalitadze
English language proofreading: Paul Steed
Technical editor: Aleš Cimprič

Published by: Znanstvena založba Filozofske fakultete Univerze v Ljubljani / Ljubljana University Press, Faculty of Arts
Issued by: University of Ljubljana, Centre for language resources and technologies
For the publisher: Roman Kuhar, Dean of the Faculty of Arts, University of Ljubljana

Ljubljana, 2018
First edition, e-edition
Publication is free of charge.

The editors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P6-0215).

This text was written using the ZRCola input system (ZRCola.zrc-sazu.si),
developed at the Research Centre of the Slovenian Academy of Sciences and Arts in Ljubljana (www.zrc.sazu.si) by Peter Weiss.

DOI: 10.4312/9789610600961
E-book is available at: https://e-knjige.ff.uni-lj.si/

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani
COBISS.SI-ID=24619494
ISBN 978-961-06-0097-8 (epub)
ISBN 978-961-06-0096-1 (pdf)

Acknowledgements

We would like to thank all those who have made the XVIII EURALEX International Congress possible, by contributing to the reviewing, to the logistics and by financially supporting the event. In particular, we would like to thank our sponsoring partners and patrons:

A.S. Hornby Educational Trust

Ingenierie Diffusion Multimedia Inc.

Oxford University Press

ELEXIS – European Lexicographic Infrastructure

CLARIN.SI – Common Language Resources and Technology Infrastructure, Slovenia

Alpineon d.o.o.

DELIGHT d.o.o.

TshwaneDJe

Faculty of Arts, University of Ljubljana

Programme Committee

Andrea Abel (European Academy of Bozen/Bolzano, EURAC)

Polona Gantar (University of Ljubljana, Faculty of Arts)

Zoe Gavriilidou (Democritus University of Thrace, Xanthi)

Vojko Gorjanc (University of Ljubljana, Faculty of Arts)

Iztok Kosem (University of Ljubljana, Faculty of Arts / Trojina)

Simon Krek (Chair) (University of Ljubljana, Center for Language Resources and Technologies / Jožef Stefan Institute, Artificial Intelligence Laboratory)

Robert Lew (Adam Mickiewicz University in Poznań, Faculty of English)

Tinatin Margalitadze (Ivane Javakhishvili Tbilisi State University)

Reviewers

Adam Rambousek, Agáta Karčová, Agnes Tutin, Ales Horak, Alexander Geyken, Amália Mendes. Amanda Laugesen, Andrea Abel, Anne Dykstra, Annette Klosa, Antton Gurrutxaga, Arvi Tavast, Carla Marello, Carole Tiberius, Carolin Müller-Spitzer, Chris Mulhall, Christine Moehrs, Corina Forascu, Danie Prinsloo, Edward Finegan, Egon Stemle, Elena Volodina, Francesca Frontini, Francis Bond, Frieda Steurs, Geoffrey Williams, Henrik Lorentzen, Ilan Kernerman, Iztok Kosem, Janet Decesaris, Jelena Kallas, Jette Hedegaard Kristoffersen, John McCrae, Jorge Gracia, Julia Bosque-Gil, Julia Miller, Klaas Ruppel, Kristina Strkalj Despot, Kseniya Egorova, Lars Trap-Jensen, Lionel Nicolas, Lothar Lemnitzer, Lut Colman, Magali Paquot, Margit Langemets, Maria Khokhlova, Marie-Claude L'Homme, Michal Měchura, Michal Kren, Miloš Jakubiček, Monica Monachini, Nataša Logar Berginc, Nicoletta Calzolari, Nikola Ljubešić, Nora Aranberri, Oddrun Grønvik, Orin Hargraves, Orion Montoya, Patrick Hanks, Patrick Drouin, Paul Cook, Paz Battaner, Philipp Cimiano, Pilar León Araúz, Piotr Zmigrodzki, Pius ten Hacken, Polona Gantar, Radovan Garabík, Robert Lew, Roberto Navigli, Ruben Urizar, Rufus Gouws, Sass Bálint, Sara Može, Simon Krek, Stella Markantonatou, Svetla Koeva, Sylviane Granger, Špela Arhar Holdt, Tamás Váradi, Tanneke Schoonheim, Tatjana Gornostaja, Thierry Fontenelle, Tinatin Margalitadze, Tomaž Erjavec, Ulrich Heid, Valentina Apresjan, Vincent Ooi, Vojko Gorjanc, Xabier Artola Zubillaga, Xabier Saralegi, Yongwei Gao, Yukio Tono, Zoe Gavriilidou idm

Hornby

elexis

european lexicographic
infrastructure

alpineon

Oxford Dictionaries

delighT.
OFFICE SOLUTION
Ljubljana • Zagreb • Belgrade • Podgorica • Skopje

CLARIN.SI

University *of Ljubljana*
FACULTY OF ARTS

TshwaneDJe

Contents

POSTER PRESENTATIONS 787

Foreword

EURALEX, European Association for Lexicography was founded in 1983 and the year 2018 marks its thirty-fifth anniversary. From its second congress in 1986, the association organises a biannual congress series. Its 18th edition, EURALEX 2018 International Congress, was held between 17th-21st July 2018 in Ljubljana, Slovenia. It was organised jointly by the Centre for Language Resources and Technologies (CLRT) at the University of Ljubljana, and Trojina Institute for Applied Slovene Studies. Both institutions are dedicated to scientific research, and the development and maintenance of digital language resources and language technology applications for contemporary Slovene. Trojina Institute was founded in 2004 with the primary objective of promoting contemporary, goal-oriented research of the Slovene language, and the University of Ljubljana founded the Centre in 2015 to ensure a systematic long-term development of technologies, resources and tools for Slovene.

The motto of EURALEX 2018 was "Lexicography in global contexts", emphasising changes in the field of lexicography related to digital transformation, and the associated need to bring together lexicographic efforts on a global level. This has been done in recent years through the Globalex initiative, a constellation of lexicographic associations that includes representatives from all continental associations of lexicography: Afrilex, Asialex, Australex, Dictionary Society of North America, and Euralex. Similar development can be witnessed in the decision of European Commission in 2017 to fund a four-year project dedicated to the establishment of the European Lexicographic Infrastructure (ELEXIS), which was also presented at the congress.

This volume of proceedings includes congress papers submitted in three categories: papers, posters, and software demonstrations. During the review process each submitted contribution was evaluated by two independent blind referees. In case of doubt, a third independent opinion was involved. Similar to previous congresses, contributions were submitted on various topics of lexicography, including, but not limited to, the following fields:

- The Dictionary-Making Process
- Research on Dictionary Use
- Lexicography and Language Technologies
- Lexicography and Corpus Linguistics
- Bi- and Multilingual Lexicography
- Lexicography for Specialised Languages, Terminology and Terminography
- Lexicography of Lesser Used languages
- Phraseology and Collocation
- Historical Lexicography and Etymology
- Lexicological Issues of Lexicographical Relevance
- Reports on Lexicographical and Lexicological Projects

Four plenary lectures were given at the congress, with two plenary papers also included in this volume. In the Hornby lecture and paper, Sylviane Granger from the Centre for English Corpus Linguistics, Université catholique de Louvain, discusses the value of adding learner corpus data to the lexicographer's monolingual and bilingual corpus base. Plenary lecture and paper by Lars Trap-Jensen from Danish Society of Language and Literature, also former president of Euralex, discusses three major revolutions that lexicography has witnessed in the last hundred years. The remaining two plenary lectures were presented by Judy Pearsall, Dictionaries Director at Oxford University Press, titled "One model, many languages? An approach to developing global language content" and Edward Finegan, professor emeritus of linguistics and law at the University of Southern California, on "Legal Interpretation via Corpora: Are Judges Failing Lexicography 101?"

The organising committee would like to thank all plenary speakers for setting the tone of the congress, and to other contributors for submitting very interesting work. We would also like to thank all the colleagues who reviewed the papers and the colleagues who participated in the work of the EURALEX 2018 programme committee. As in past EURALEX editions, the Hornby Trust generously sponsored one of the plenary lectures in honour of A.S. Hornby, a pioneering figure in learner's dictionaries for non-native speakers. All patrons and sponsors who supported us for this edition are listed on a dedicated page within these proceedings.

As the chair of the congress, I would like to acknowledge precious work of the members of the organising committee who joined efforts with me to make EURALEX 2018 a successful event: Špela Arhar Holdt, Jaka Čibej, Kaja Dobrovoljc, Polona Gantar, Vojko Gorjanc, Iztok Kosem and Nataša Logar.

Simon Krek
Chair, XVIII EURALEX International Congress
July 5, 2018

Towards a Glossary of Rum Making and Rum Tasting

Cristiano Furiassi
University of Turin
E-mail: cristiano.furiassi@unito.it

Abstract

A lexicographic work exclusively dedicated to the making and tasting of rum has not been published to date. With the ambitious aim of filling this editorial gap in mind, this article focuses on the implementation stage of a specialized glossary of rum-related terms in the English language. Preceded by an overview of the historical, geographical and linguistic factors that made rum a renowned global product, the computer-assisted terminology acquisition procedures applied in order to extract rum-related terms from an *ad hoc* corpus are described. By merging computer-assisted term extraction with data collected from experts' knowledge, fieldwork and the existing specialized literature on rum, a list of candidate headwords was drafted. The replicability of the methodology applied makes this pilot study generalizable, thus fostering the compilation of specialized glossaries connected to other fields or disciplines.

Keywords: computer-assisted term extraction, glossary, rum, specialized lexicography

1 Introduction

A lexicographic product exclusively focusing on rum, namely a reference tool where both rum amateurs and connoisseurs can look up notions on the making and tasting of rum, is still missing on the market. Following an introductory section on the various historical and geographical aspects concerning rum, which also includes linguistic information about the word *rum* itself, the main aim of this article is to describe the implementation stage of a specialized glossary of rum-related terms, namely "[a] type of REFERENCE WORK which lists a selection of words or phrases, or the terms in a specialised field, usually in alphabetical order, together with minimal definitions or translation equivalents" (Hartmann & James 2002: 63).

More precisely, the article deals with the selection of headwords, the most salient macrostructural feature of any glossary, by showing how a list of candidate items may be obtained by exploiting a specialized corpus containing texts about rum written in English through the combination of (partly) automatic, namely "corpus-driven" (Krishnamurthy 2008: 231), and (mostly) semi-automatic, namely "corpus-based" Tognini Bonelli 2001: 65), techniques.[1]

The term-extraction procedures described are limited to specialized texts about rum written in English. However, the fact that rum production has spread on a large scale, also involving the French-speaking and the Spanish-speaking Caribbean, makes rum a global product *par excellence*. Therefore, by applying the same procedures, an additional step would lead to the compilation of a multilingual glossary of rum.

[1] A similar, though more sophisticated, approach to specialized lexicography regarding alcoholic beverages, namely wine, was adopted by Leroyer (2015; 2018) for his *Oenolex Wine Dictionary*.

2 Rum as a Global Product

In many rum-producing countries and especially in its Caribbean birthplace, Barbados, rum often represents a national symbol deeply rooted in the local culture.[2] In fact, in the 15th century it was Christopher Columbus who brought to the Caribbean a large amount of sugarcane from Spain, specifically the Canaries. At the beginning of its production, in the 17th century, rum was considered a drink of little value and unpleasant taste, lacking the prestige of more refined distillates made in Europe. However, by the 18th century, besides having become a precious export, the importance of rum grew both in the North American continent and Europe.

As opposed to other world-famous spirits, such as, for instance, cognac, gin, vodka and whisk(e)y, readily associable with France, England, Russia and Scotland or Ireland respectively, rum – an icon of the "cultural fragmentation" (Furiassi 2014: 91) typical of the Caribbean – spread throughout the world and eventually reached all continents to the point that nowadays the general public seems to ignore its exact origin and can hardly associate it with a particular territory. Within North America, a few distilleries may be found in the United States. In the Caribbean, most of the Greater Antilles and virtually all the Lesser Antilles are renowned for producing rum. Moreover, various mainland territories of Central America, including Belize, Costa Rica, Guatemala, Nicaragua and Panama, produce rum. In South America, rum is distilled in Argentina, Brazil, Colombia, French Guiana, Guyana, Paraguay, Peru, Suriname and Venezuela. In Asia, Japan, Nepal, the Philippines and Thailand are involved in the making of rum, while in Oceania, Australia, Fiji and New Zealand are rum-producing countries. In Africa, Madagascar, Mauritius, Réunion, South Africa and Saint Helena are also known for the production of this drink. In Europe, rum distillation is limited to Las Palmas, one of the Canary Islands. Finally, it is worth mentioning that even remote islands, such as Bermuda, are celebrated rum makers.

Lexicographically, the word *rum* has been defined alternatively as an alcoholic drink, liquor or spirit, as the following quotations show: "[a]n alcoholic spirit distilled from molasses and other sugar-cane products, prepared chiefly in the Caribbean and parts of Central and South America; a serving or variety of this" (*OED*); "an alcoholic liquor prepared by fermenting molasses, macerated sugarcane, or other saccharine cane product, distilling, coloring with caramel, and aging" (*Merriam-Webster*); "[a]n alcoholic drink industrially distilled from the juice of the sugar-cane, blended and cured in barrels" (*DCEU*).[3]

As far as its earliest attestation is concerned, the first written account of the word *rum* in English dates from 1654 (*OED*, *Merriam-Webster*). Nonetheless, it must be noted that the culture surrounding rum is also ingrained in the French- and Spanish-speaking Caribbean. In fact, its French and Spanish cognates, namely *rhum*, first recorded in 1688 (*TLFi*), and *ron*, dating from about 1770 (*DECH*), derive from English *rum*, as attested in authoritative lexicographic sources such as, for instance, the *FEW* and the *TLFi* for French and the *DECH* and the *DRAE* for Spanish.

3 Data Retrieval and Methodology: The *Caribbean Rum Corpus* (*CRC*)

Among others, Gamper and Stock (1998: 147) claim that "[t]he manual acquisition of terminological material from the domain-specific text material is a very time-consuming task. […] Computer-assisted

[2] As reported by Smith (2008: 13), "[…] evidence indicates that the British island of Barbados and the French island of Martinique were the cradles, if not the birthplaces, of Caribbean rum".

[3] Originally the shortening of *rumbullion*, an Early Modern English word perhaps originated in Devonshire and meaning 'a great tumult' or 'uproar', over the centuries rum was called by many different names, mostly referring to its close association with seafaring, buccaneering and the infernal regions: *Barbados water*, *devil's death*, *grog*, *(hot) hellish liquor*, *kill-devil*, *navy neaters*, *Nelson's blood*, *pirate's drink*, *rumbullion*, *rumbustion* and *taffia* or *tafia* – all included in the glossary headword list (see table 1).

term acquisition improves both the quantity and the quality of terminological work". Consequently, the first action taken to obtain a wordlist of rum-related terms was to design and compile a specialized – or "special" (Tognini Bonelli & Sinclair 2006: 210) – corpus, namely the *Caribbean Rum Corpus* (*CRC*).

Texts contained in the *CRC* include material from websites created by rum experts,[4] that is the official websites of 25 rum makers throughout the English-speaking Caribbean – all listed in the reference section, thus providing "adequate coverage of the field in question" (Bowker 2003: 162).[5] Here follows the list of rum makers grouped by territory: *Anguilla Rums* (Anguilla); *Antigua Distillery* (Antigua and Barbuda); *Bacardi*, *Todhunter-Mitchell Distillery* (Bahamas); *Cockspur*, *Foursquare Rum Distillery* (where *Doorly's*, *E.S.A. Field*, *Mahiki*, *Old Brigand*, *The Real McCoy* and *R.L. Seale* are produced), *Mount Gay Distillers*, *St. Nicholas Abbey* (Barbados); *Gosling* (Bermuda); Arund*el Estate Callwood Distillery*, *Pusser's* (British Virgin Islands); *Clarke's Court*, *Grand Havana Rum*, *Westerhall Estate* (Grenada); *Appleton Estate*, *Blackwell*, *Captain Morgan*, *Coruba*, *Myers's*, *Worthy Park Estate* (Jamaica); *Elements Eight Rum*, *St. Lucia Distillers* (St. Lucia); *10 Cane*, *Angostura*, *Caroni*, *Zaya* (Trinidad and Tobago).

Depending on the degree of usability of each website – how practical it was to extract plain text, most of the makers listed above were considered except for *Caroni*, whose website is non-existent since the distillery closed in 2002, and *Anguilla Rums*, *Mount Gay Distillers* and *St. Nicholas Abbey*, whose websites could not be exploited for technical reasons – the automatic extraction of texts was not allowed. At the end of the collection procedure, the *CRC* amounted to 33,625 tokens and 4,202 types: for the task at hand, the choice of texts and number of running words seem to meet both the representativeness and reliability requirements which a specialized corpus must satisfy in order to be useful for linguistic and lexicographic investigation (Biber 2008: 63-64; Bowker & Pearson 2002: 45).[6]

4 Computer-assisted Term Extraction

Once the *CRC* was collected, the data gathered were processed by means of *WordSmith Tools*.[7] The *CRC* wordlist (4,202 types), obtained via the *WordList* tool, was compared with two wordlists extracted from two general corpora of the English language, i.e. the *Freiburg-Lancaster-Oslo-Bergen Corpus of British English* (*FLOB*) and the *Freiburg-Brown Corpus of American English* (*FROWN*), the former containing texts typical of British English, the latter based on American English. Although compiled much earlier, namely in the early 1990s, the two reference corpora selected were considered functional for the lexicographic purpose at hand, as the size, granularity and text types included made term extraction viable. In addition, the *FLOB* and the *FROWN*, albeit somehow dated, were considered instead of the *British National Corpus* (*BNC*) and the *Corpus of Contemporary American English* (*COCA*), among others, because of usability criteria: in practice, the availability of all texts belonging to the *FLOB* and the *FROWN* allowed both corpora to be processed by *WordSmith Tools*.

Two separate non-lemmatized keyword lists were thus obtained using the *KeyWords* tool: one, containing 472 positive keywords, resulting from the comparison between the *CRC* and the *FLOB* wordlists, and another, containing 475 positive keywords, resulting from the comparison between the *CRC*

[4] See Bergenholtz (1995: 19-20), Bowker & Pearson (2002: 27-28) and Gotti (2011: 25-28) for a distinction of levels of expertise in the encoding/decoding of specialized texts.

[5] Atkins & Rundell (2008: 80) suggest that "a carefully designed web corpus can provide reliable language data".

[6] Although, at present, the corpus may look small, it must be noted that the domain under investigation is highly specialized. However, the *CRC* could be expanded in a future phase by extending the range of websites considered to those of rum producers in other English-speaking parts of the world, such as Australia, Fiji, New Zealand, the Philippines, South Africa and the United States.

[7] Despite the existence of various types of text analysis software such as, for instance, *AntConc* and *TextSTAT*, *WordSmith Tools* is among the few – to the author's knowledge – which allow the analyst to conveniently compare corpus wordlists in order to detect the keyness of certain items, and was specifically selected to help in such endeavors (see footnote 8).

and the *FROWN* wordlists.[8] After being merged, the positive keywords generated by *WordSmith Tools* (512 tokens) were of paramount importance for extracting single- and/or multi-word terminological units which will then become headword candidates of the rum glossary.

Keyword lists were expressly drawn to highlight "topic sensitivity" (Ringbom 1998: 48): it was essential to detect topic-sensitive items, that is "words that are closely linked to the topics dealt with" (Furiassi 2004: 194) in the *CRC*, namely rum. However, in order to keep only topic-sensitive content words, the resulting list, which still included some noise, i.e. undesired items, had to be further reduced (328 items) by manually eliminating function words and proper nouns related to rum brands/makers and toponyms.[9]

While most of these items can be intuitively associated with the specialized language of rum, e.g. *barrel*, *distillation*, *molasses*, others are also common in general English, e.g. *gold*, *scent*, *wood*, though obviously acquiring a specialized meaning in a rum-oriented context. Therefore, via the *Concord* tool, a concordance output was provided for each item in the wordlist thus obtained in order to establish whether it should qualify as a headword in the glossary.

In addition, the aim of the present study was not restricted to the selection of single words contained in the *CRC* keyword list (see footnote 9). Indeed, in order not to miss recurrent "collocations" (Sinclair 1991: 109-121), high-frequency word clusters were also obtained for each positive keyword: clusters range from a minimum of a two-word combination to a maximum of a four-word combination. Finally, from both the noise-free *CRC* keyword list and the manually-selected clusters, a wordlist of candidate items suitable for inclusion as headwords in a glossary of rum was gathered.[10]

5 The Selection of Headwords

Since LSP lexicography cannot rely entirely on corpus data, a final list of candidate headwords was drafted only after combining computer-assisted term extraction from the *Caribbean Rum Corpus* (*CRC*) – and the *Guyana Rum Corpus* (*GRC*), expressly collected at a later stage (see Section 5.2) – with data gathered from experts' knowledge, fieldwork and the existing specialized literature on rum published in English.

5.1 Headword Selection from the *CRC*

Alongside the initial corpus-based term-extraction procedures, carried out through the *KeyWords* tool provided by *WordSmith Tools*, the wordlist of candidate headwords selected from the *CRC* was mostly the result of a semi-automatic procedure since the following decisions were made:

[8] Chung (2003: 221) states that "[…] the corpus comparison approach using word types is a reasonably simple and practical way of identifying terms". More specifically, Furiassi (2004: 201), maintains that "[t]he comparison of two wordlists provides information about the keyness of each word in a corpus […]. Positive keywords are items that occur more often than would be expected by chance in comparison with the reference corpus".

[9] Unfortunately, the *KeyWords* tool provided by *WordSmith Tools* only extracts single-word units automatically. However, although function words, e.g. *of*, were discarded at this stage, they may still be included in the rum glossary as part of multi-word units gathered manually by selecting their typical clusters detected through the *Concord* tool, e.g. *gram of alcohol* (see Table 1).

[10] Groundbreaking NLP processing tools for term extraction, which work on lemmatized, POS-tagged wordlists extracted from corpora through statistical methods, were made available after the present research was conceived. Indeed, term extractors such as *OneClick Terms*, powered by *Sketch Engine*, and *TermoStat*, which exploit a hybrid method, i.e. statistical plus linguistic, to identify candidate terms, would undoubtedly contribute to the implementation stage of a glossary of rum. In particular, as far as terms to be considered as candidate headwords are concerned, it would then be mandatory to verify whether the same corpus data processed by fully-automatic term extractors produce similar outputs or, most certainly, the glossary is improved by including additional headwords. Moreover, once a final list of headwords is obtained, term extractors are also likely to enrich the lexical information for each headword in the glossary, e.g. word-class assignment.

- all words linked to rum making and rum tasting were included;
- abbreviations and acronyms were taken into account only if closely connected to the specialized language under scrutiny;
- headword status was also granted to multi-word lexical units resulting from cluster selection.

Following these criteria, a list, which contains 295 headwords and 81 sub-headwords, was drafted.

5.2 Headword Selection from the *GRC*

A glossary of rum would not be complete without considering distillers based in English-speaking Guyana, another famous rum-producing territory connected with the Caribbean.[11] Therefore, a smaller corpus including texts extracted from the websites of Guyanese rum producers, namely *Demerara Distillers* (maker of award-winning *El Dorado Rum*), and *XM*, was compiled. Despite the fact that the *XM* website could not be exploited since the automatic extraction of texts was not allowed, the *Guyana Rum Corpus* (*GRC*), consisting of 5,327 tokens and 1,323 types, underwent the same semi-automatic term-extraction procedures applied to the *CRC*, thus allowing the retrieval of 10 new headwords – including acronyms, i.e. *Authentic Caribbean Rum™ (ACR™)*, *butterscotch*, *Savalle still*, *signature rum*, *texture*, *toffee*, *uncrystallised sugar* and *West Indies Rum & Spirits Producers' Association (WIRSPA)*, and two new sub-headwords, i.e. *exotic fruit* and *flavo(u)rful*.

5.3 Headword Selection from Experts' Knowledge, Fieldwork and Specialized Literature

Corpus-based LSP lexicography, also known as computer-assisted term extraction or computer-assisted terminology acquisition, must be complemented by information retrieved from experts' knowledge, fieldwork and specialized literature.

Therefore, experts' knowledge gathered from detailed visits of Barbadian rum distilleries represented a valuable source of information, as it provided a number of headwords that could not be selected otherwise and that could only be extrapolated by interviewing tour guides, master blenders and master distillers, watching the documentaries shown as part of the guided tours and analyzing the various signs and posters on display inside distilleries. In addition, the specialized literature on the topic published to date was taken into account, namely Barty-King and Massel (1983), Arkell (1999), Plotkin (2001), Ruthström (2001), Broom (2003), Coulombe (2004), Williams (2006), Curtis (2007), Smith (2008), Miller et al. (2009), Liberman (2010), Laurie (2011), Foss (2012), Hopkins (2012), Maier (2013) and Smiley, Watson and Delevante (2014).

Consequently, with the aid of experts' knowledge and specialized literature, 19 additional headwords, namely *Coffey still*, *condensation*, *cooper machine*, *cut*, *de-ionized water*, *de-mineralized water*, *earthy*, *harmonious*, *head*, *heart*, *master distiller*, *pastry*, *pepper*, *peppermint*, *reduction*, *single cask*, *single distillation*, *subtle* and *tail*, and four additional sub-headwords, *oaky*, *ripe fruit*, *toasted wood* and *woody*, were collected.

6 Findings: The Rum Glossary Headwords

Eventually, it is worth mentioning that the "lexicographer's intuition" (Sinclair 2003: 167) was of paramount importance to decide whether a lexical item or collocation had to be included or excluded

[11] Even though Guyana, officially the Co-operative Republic of Guyana, is geographically in South America, politically, culturally and linguistically it is considered part of the Caribbean. Guyana is also among the founder members of the C*aribbean Community of Commonwealth States* (*CARICOM*); indeed, the headquarters of *CARICOM* are in Guyana, in the capital city of Georgetown, within the Demerara-Mahaica region.

from the final glossary headwords. Headwords in Table 1 are listed in alphabetical order (horizontally, from left to right) with grey shading signaling the first term for each letter of the alphabet. Each headword appears in bold; alternative spelling variants are shown in italics next to the headword. Some headwords required the insertion of sub-headwords: sub-headwords are shown in roman below the corresponding headword. Headwords (and sub-headwords) which are semantically linked to other headwords (and sub-headwords) included in Table 1 are cross-referenced: any cross reference is indicated by an arrow, i.e. →, followed by the respective headword (or sub-headword).

Table 1: Rum Glossary Headwords.

absolute alcohol	**ABV** → alcohol by volume → vol.	**aged rum**	**ageing (process)** *aging (process)*
alcohol	**alcohol by volume** → ABV → vol.	**alcohol recovery column**	**alcoholic fermentation**
alcoholic strength	**aldehyde**	**almond**	**amber**
apricot dried apricot	**aroma** → nose aroma profile	**aromatic**	**ACR™** → Authentic Caribbean Rum™
Authentic Caribbean Rum™ → ACR™	**balanced** balanced rum	**Barbados water**	**barrel** → cask
batch	**batch distillation** → pot still distillation	**batch number**	**batch rum**
black rum → dark rum	**blend**	**blending** blending information blending instruction blending process	**boiling pot**
Boston glass	**bottle**	**bottling** bottling strength	**bounty rum**
bouquet	**bourbon barrel** → bourbon cask	**bourbon cask** → bourbon barrel	**brand**
brand positioning → product range	**bronze**	**brown**	**brown sugar**
butterscotch	**buttery**	**by-product**	**cane sugar**
capacity → size	**caramel**	**carbon dioxide**	**carbon filtration**
cask → barrel	**champagne glass**	**character**	**charcoal filtration**
charred oak barrel → charred oak cask	**charred oak cask** → charred oak barrel	**chocolate**	**cinnamon**
citrus	**clean**	**clove**	**cocktail glass**
cocoa	**coconut**	**coconut rum**	**coffee**
Coffey still	**Collins glass**	**colour** *color*	**column distillation** → continuous still distillation
column still	**complexity** → sophisticated complex	**compound**	**concentrated alcohol**
condensation	**congener**	**connoisseur**	**content**

continuous still → pot still continuous still rum	**continuous still distillation** → column distillation	**cooper machine**	**copper alembic pot**
copper kettle	**copper pot still**	**coupette glass**	**cream**
crushed cane	**cut**	**dark rum** → black rum	**dash**
de-ionized water → de-mineralized water	**de-mineralized water** → de-ionized water	**devil’s death**	**distillation** distillation method distillation process
distilled drink	**distinct** *distinctive*	**double distillation** double distillate double distilled rum	**earthy**
estate	**ethyl**	**exact**	**expertise**
external water jacket	**extra old** → XO extra old rum	**factory**	**fermentation** fermentation process
fermented wash	**fertile soil**	**filtration** filtration process	**finish**
first press	**flavour** *flavor* → taste flavourful *flavorful*	**flavour compound** *flavor compound*	**flavouring agent** *flavoring agent*
fresh	**fruit** exotic fruit fresh fruit fruity honeyed fruit ripe fruit	**full** full-bodied	**gentle**
gentle filtration → light filtration	**ginger**	**glass**	**gold** gold rum
golden	**gram of alcohol**	**grog**	**hand blend**
hand-crafted	**harmonious**	**harshness**	**harvested** hand harvested
hazelnut	**head**	**heart**	**heavy** heavy bodied heavy rum
heavy pot still heavy pot still rum → light pot still	**hellish liquor** → hot hellish liquor	**highball glass**	**high proof rum**
high wine retort → low wine retort	**hint**	**honey**	**hot hellish liquor** → hellish liquor
hurricane glass	**infused**	**ingredient**	**instruction**
intense	**International Wine & Spirits Competition** → IWSC	**International Wine & Spirits Festival** → IWSF	**IWSC** → International Wine & Spirits Competition
IWSF → International Wine & Spirits Festival	**juice**	**kill-devil**	**label**
labour-intensive crop *labor-intensive crop*	**legacy**	**lemon** lemon peel lemon rind	**light** light bodied light rum

light filtration → gentle filtration	**light pot still** light pot still rum → heavy pot still	**lime** lime peel → lime rind lime rind → lime peel	**limited edition** → limited reserve
limited number → limited production	**limited production** → limited number	**limited reserve** → limited edition	**liqueur** *liquor*
long	**long drink**	**low wine retort** → high wine retort	**making process**
manufacturing process	**maple**	**margarita glass**	**market** European market local market mass market mid-market top market US market
marrying process	**mash**	**master blender**	**master distiller**
maturation process	**medal**	**medium** medium bodied medium rum	**mellow**
milled	**minimum aged rum**	**mixed** mixed drink	**mixing glass**
mixing rum	**molasses**	**naturally filtered**	**navy neaters**
neat	**Nelson’s blood**	**nose** → aroma	**note**
nut nutty toasted nut	**nutmeg**	**oak** oaky	**oak barrel** → oak cask
oak cask → oak barrel	**old** old rum	**old-fashioned glass** → rocks glass	**orange** orange peel → orange rind orange rind → orange peel
organic compound	**original**	**overproof rum**	**oxidation**
packaging packaging detail	**painkiller**	**palate**	**part**
passion fruit	**pastry**	**peach**	**pepper**
peppermint	**pirate’s drink**	**plant**	**plantation** plantation distillery plantation rum
platinum → PT platinum rum	**pot**	**pot still** → continuous still pot still rum	**pot still distillation** → batch distillation
premium premium rum	**primary water treatment system** → secondary water treatment system	**production capacity**	**product line** → production line
production line → product line	**product range** → brand positioning	**profile**	**PT** → platinum

quality rum	**raisin** honeyed raisin ripe raisin sweet raisin	**raw**	**recipe**
recovery column	**red**	**reduction**	**reserve**
residual impurity	**rich**	**rim**	**rocks glass** → old-fashioned glass
rounded rounded rum	**rum**	**rumbullion**	**rumbustion**
rum making rum making process	**rum steam**	**saccharomyces**	**Savalle still**
scent	**seal**	**secondary water treatment system** → primary water treatment system	**select** *selected*
sherry cask	**short glass**	**shot glass**	**signature drink**
signature rum	**single barrel** → single cask	**single cask** → single barrel	**single distillation** single distillate single distilled rum
single-label	**single rum**	**sipping rum** → tasting rum	**size** → capacity
smoky	**smoothness** smooth	**soft**	**soil**
sophisticated → complexity	**spice** → spiced	**spiced** *spicy* → spice spiced rum *spice rum*	**spirit** spirity
stalk	**steam** steam engine	**still** still maturation	**storage** storage container storage facility storage tank
straight straight rum	**strain**	**strength**	**subtle**
sugar sugar factory	**sugar cane** *sugarcane* sugar cane juice *sugarcane juice* sugar cane plantation *sugarcane plantation*	**sulphate**	**sultana**
superior superior rum	**super premium** super premium rum	**sweetness** sweet	**taffia** *tafia*
tail	**tall glass**	**taste** → flavour *flavor* taste profile	**tasting note**
tasting rum → sipping rum	**terroir**	**texture**	**toasted**
tobacco	**toffee**	**tot**	**triple distillation** triple distillate triple distilled rum
tropical ageing *tropical aging* tropically aged	**tropical fruit**	**uncrystallized sugar**	**vanilla**

velvety	**versatile** versatile rum	**vibrant**	**vintage** vintage blend vintage rum
vol. → ABV → alcohol by volume	**volatile sulphur compound**	**wash**	**water treatment system**
West Indian rum	**West Indies Rum & Spirits Producers' Association** → WIRSPA	**wheat bread**	**white** white rum
WIRSPA → West Indies Rum & Spirits Producers' Association	**wood** toasted wood woody	**wooden distillation**	**World Spirits Competition** → WSC
WSC → World Spirits Competition	**XO** → extra old	**yeast** yeast strain	**zest**

7 Conclusion

The initial steps of the "implementation" stage (Svensén 2003: 99) of a specialized glossary of rum – limited to the English language – were described. All combined, the compilation of the *CRC* and the *GRC*, the application of corpus-based term-extraction procedures, the exploitation of experts' knowledge through fieldwork, the analysis of specialized literature and the subsidy of the lexicographer's insight proved fruitful. Consequently, among thousands of candidate items, 324 headwords and 87 sub-headwords were eventually considered for inclusion, thus accomplishing the main goal of this piece of research. At a later stage, it will be possible to move towards the microstructure of the glossary, that is the editing of each entry: a more detailed treatment of headwords – not yet produced – implies that all entries will be provided with a definition, instances of usage in authentic texts and, where necessary, especially in cases that require the illustration of highly specialized appliances used in rum distillation, images will be added – at present, since this article is mostly a work-in-progress report on a glossary-making project, the plans to make the resource available and the strategy for distributing it through the appropriate channels are not yet underway.

The procedures implemented in this pilot study focusing on the macrostructure of a specialized glossary of rum-related terms, meant to be the starting point for the compilation of a specialised glossary of rum, seem generalizable. The same methodology, possibly complemented by the application of fully-automatic term extractors (see footnote 10), may indeed be replicated to enable the compilation of specialized glossaries connected to other subjects or domains, as already successfully attempted, among others, by Gamper and Stock (1998), Cabré Castellví (1999); Bourigault, Jacquemin and L'Homme (2001), Peñas, Verdejo and Gonzalo (2001), and Chung (2003).

8 Desiderata

The ubiquitous nature of rum, especially its popularity throughout the French- and Spanish-speaking Caribbean in addition to the English-speaking Caribbean, naturally calls for a further, more ambitious project, that is the compilation of a multilingual glossary of rum. Following the same criteria adopted for the selection of rum-related terminology in the English language, a French and

Spanish supplement should be considered in order to appeal to the worldwide audience of rum enthusiasts.

As for French, after compiling an analogous specialized corpus based on texts retrieved from the websites of rum distillers based in the French-speaking Caribbean, as well as Madagascar, Mauritius, Réunion and Seychelles, the wordlist provided may be compared to the wordlist produced from the *Corpus français* (*CF*) or the forthcoming *Corpus de référence du français contemporain* (*CRFC*), to be considered as reference corpora of general French (see Siepmann, Bürgel & Diwersy 2015: 64). As far as Spanish is concerned, the wordlist obtained from the specialized corpus gathered by extracting texts form the websites of rum makers in the Spanish-speaking Caribbean, as well as in Central and South America and the Canaries, may be set against the wordlist triggered by the *Corpus de referencia del español actual* (*CREA*), a general corpus of the Spanish language. After implementing the appropriate semi-automatic procedures, applying automatic term extractors and including the pertinent specialized literature written in French and Spanish respectively, the keywords obtained would lead to the drafting of a French and Spanish list of candidate headwords, thus collocating the content of the glossary in a multilingual perspective.

References

10 Cane. Accessed at: http://www.10cane.com [25/05/2018]

[*AIS*] *Associazione Italiana Sommelier* Accessed at: http://www.aisitalia.it [25/05/2018]

Angostura. Accessed at: http://www.angostura.com [25/05/2018]

Anguilla Rums. Accessed at: http://www.pyratrum.com [25/05/2018]

[*AntConc*] Anthony, L. (2018). *AntConc 3.5.7*. Tokyo: Waseda University. Accessed at: http://www.laurenceanthony.net/software/antconc/ [25/05/2018]

Antigua Distillery. Accessed at: http://antiguadistillery.com [25/05/2018]

Appleton Estate. Accessed at: http://www.appletonestate.com [25/05/2018]

Arkell, J. (1999). *Classic Rum*. London: Prion Books.

Arundel Estate Callwood Distillery. Accessed at: http://www.bareboatsbvi.com/cgb_callwood_distillery.html [25/05/2018]

Atkins, B. T. S., Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Bacardi. Accessed at: http://www.bacardi.com [25/05/2018]

Barty-King, H., Massel, A. (1983). *Rum: Yesterday and Today*. London: Heinemann.

Bergenholtz, H. (1995). Basic Issues in Specialized Lexicography. In H. Bergenholtz, S. Tarp (eds.) *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Amsterdam & Philadelphia: John Benjamins, pp. 14-47.

Biber, D. (2008). Representativeness in Corpus Design. In T. Fontenelle (ed.) *Practical Lexicography: A Reader*. Oxford: Oxford University Press, pp. 63-87.

Blackwell. Accessed at: http://www.blackwellrum.com [25/05/2018]

[*BNC*] Davies, M. (ed.). (2004-2018). *British National Corpus*. Provo: Brigham Young University. Accessed at: http://corpus.byu.edu/bnc [25/05/2018]

Bourigault, D., Jacquemin, C. & L'Homme, M.-C. (eds.). (2001). *Recent Advances in Computational Terminology*. Amsterdam & Philadelphia: John Benjamins.

Bowker, L. (2003). Specialized Lexicography and Specialized Dictionaries. In P. Van Sterkenburg (ed.) *A Practical Guide to Lexicography*. Amsterdam & Philadelphia: John Benjamins, pp. 154-164.

Bowker, L., Pearson, J. (2002). *Working with Specialized Language. A Practical Guide to Using Corpora*. London & New York: Routledge.

Broom, D. (2003). *Rum*. San Francisco: Wine Appreciation Guild.

Captain Morgan. Accessed at: http://www.captainmorgan.com [25/05/2018]

Cabré Castellví, M. T. (1999). *Terminology: Theory, Methods and Applications*. Amsterdam & Philadelphia: John Benjamins.

[*CF*] *Corpus français*. Leipzig: Universität Leipzig & Neuchâtel: Université de Neuchâtel. http://wortschatz.uni-leipzig.de/ws_fra [25/05/2018]

Chung, T. M. (2003). A Corpus Comparison Approach for Terminology Extraction. In *Terminology*, 9(2), pp. 221-246.

Clarke's Court. Accessed at: http://www.clarkescourtrum.com [25/05/2018]

[*COCA*] Davies, M. (ed.). (2008-2018). *Corpus of Contemporary American English*. Provo: Brigham Young University. Accessed at: http://corpus.byu.edu/coca [25/05/2018]

Cockspur. Accessed at: http://www.cockspurrum.com [25/05/2018]

Coruba. Accessed at: http://www.coruba.co.nz [25/05/2018]

Coulombe, C. A. (2004). *Rum: The Epic Story of the Drink that Conquered the World*. New York: Citadel Press.

[*CREA*] *Corpus de referencia del español actual*. Madrid: Real Academia Española. Accessed at: http://corpus.rae.es/creanet.html [25/05/2018]

Curtis, W. (2007). *And a Bottle of Rum: A History of the New World in Ten Cocktails*. New York: Three Rivers Press.

[*DCEU*] Allsopp, R. (ed.). (2003) [1996]. *Dictionary of Caribbean English Usage.* Mona: University of the West Indies Press.

[*DECH*] Corominas, J., Pascual, J. A. (1983). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.

Demerara Distillers. Accessed at: http://demeraradistillers.com [25/05/2018]

[*DRAE*] (2001). *Diccionario de la lengua española*, 22nd edn. Madrid: Real Academia Española. http://lema.rae.es/drae [25/05/2018]

El Dorado Rum. Accessed at: http://theeldoradorum.com [25/05/2018]

Elements Eight Rum. Accessed at: http://www.e8rum.com [25/05/2018]

[*FEW*] Von Wartburg, W. (2003). *Französisches Etymologisches Wörterbuch*. Accessed at: http://apps.atilf.fr/lecteurFEW [25/05/2018]

[*FLOB*] Hundt, M., Sand, A. & Siemund, R. (eds.). (1998). *Freiburg-Lancaster-Oslo-Bergen Corpus of British English*. Freiburg: Albert-Ludwigs-Universität Freiburg. Accessed at: http://icame.uib.no/flob [25/05/2018]

Foss, Richard. (2012). *Rum: A Global History*. London: Reaktion Books.

Foursquare Rum Distillery. Accessed at: http://foursquarerum.com [25/05/2018]

[*FROWN*] Hundt, M., Sand, A. & Skandera, P. (eds.). (1999). *Freiburg-Brown Corpus of American English*. Freiburg: Albert-Ludwigs-Universität Freiburg. Accessed at: http://icame.uib.no/frown [25/05/2018]

Furiassi, C. (2004). Spoken and Written Learner English: A Quantitative Analysis of ICLE-IT and LINDSEI-IT. In M. T. Prat Zagrebelsky (ed.) *Computer Learner Corpora. Theoretical Issues and Empirical Case Studies of Italian Advanced EFL Learner's Interlanguage*. Alessandria: Edizioni dell'Orso, pp. 193-208.

Furiassi, C. (2014). Caribbean English Vocabulary: Setting a Norm through Lexicographic Practice. In A. Molino, S. Zanotti (eds.) *Observing Norm, Observing Usage: Lexis in Dictionaries and in the Media*. Bern: Peter Lang, pp. 89-107.

Gamper, J., Stock, O. (1998). Corpus-based Terminology. In *Terminology*, 5(2), pp. 147-159.

Gosling. Accessed at: http://www.goslingsrum.com [25/05/2018]

Gotti, M. (2011). *Investigating Specialized Discourse*, 3rd edn. Bern: Peter Lang.

Grand Havana Rum. Accessed at: http://www.grandhavanarum.com [25/05/2018]

Hartmann, R. R. K., James, G. (2002). *Dictionary of Lexicography* London & New York: Routledge.

Hopkins, T. (2012) [2004]. Rum. In A. F. Smith (ed.) *The Oxford Encyclopedia of Food and Drink in America*, 2nd edn. Oxford: Oxford University Press, vol. II, pp. 158-160.

Krishnamurthy, R. (2008). Corpus-driven Lexicography. In *International Journal of Lexicography*, 21(3), pp. 231-242.

Laurie, P. (2011) [2001]. *The Barbadian Rum Shop: The Other Watering Hole*, 2nd edn. Oxford: Macmillan.

Leroyer, P. (2015). Turning the Corpus into a Functional Component of the Dictionary: The Case of the Oenolex Wine Dictionary. In *Procedia – Social and Behavioral Sciences*, 198, pp. 257-265.

Leroyer, P. (2018). The Oenolex Wine Dictionary. In P. A. Fuertes-Olivera (ed.) *The Routledge Handbook of Lexicography*. London & New York: Routledge, pp. 438-454.

Liberman, A. (2010). The Rum History of the Word "Rum". In *OUPblog*, 6th October 2010. Accessed at: https://blog.oup.com/2010/10/rum [25/05/2018]

Maier, E. (2013). *OED* Word Stories: 'rum'. Accessed at: http://public.oed.com/aspects-of-english/word-stories/rum [25/05/2018]

[*Merriam-Webster*] Gove, P. B. (ed.). (2002). *Webster's Third New International Dictionary Unabridged*. Springfield: Merriam-Webster. Accessed at: http://unabridged.merriam-webster.com [25/05/2018]

Miller, A., Brown, J., Broom, D. & Strangeway, N. (2009). *Cuba: The Legend of Rum*. London: Mixellany.

Mount Gay Distillers. Accessed at: http://www.mountgayrum.com [25/05/2018]

Myers's. Accessed at: https://www.diageo.com/en/our-brands/brand-explorer/#myers [25/05/2018]

[*OED*] Simpson, J., Weiner, E. (eds.). (1989-2018). *The Oxford English Dictionary*. Oxford: Oxford University Press. Accessed at: http://www.oed.com [25/05/2018]

OneClick Terms. (2016-2018). Brno & Brighton: Lexical Computing. Accessed at: https://terms.sketchengine.co.uk [25/05/2018]

Peñas, A., Verdejo, F. & Gonzalo, J. (2001). Corpus-based Terminology Extraction Applied to Information Access. In P. Rayson, A. Wilson, T. McEnery, A. Hardie & S. Khoja (eds.) *Proceedings of the Corpus Linguistics 2001 Conference*. Lancaster: UCREL, pp. 458-465.

Plotkin, R. A. (2001). *Caribe Rum. The Original Guide to Caribbean Rum and Drinks*. Tucson: BarMedia.

Pusser's. Accessed at: http://www.pussers.com [25/05/2018]

Ringbom, H. (1998). Vocabulary Frequencies in Advanced Learner English: A Cross Linguistic Approach. In S. Granger (ed.) *Learner English on Computer*. London & New York: Longman, pp. 41-52.

Ruthström, B. (2001). *Tafia*, *ratafia* and *rum* – liquor words of dizzy origin. In *Indogermanische Forschungen*, 106(1), pp. 262-275.

Siepmann, D., Bürgel, C. & Diwersy, S. (2015). The *Corpus de référence du français contemporain* (*CRFC*) as the first genre-diverse mega-corpus of French. In *International Journal of Lexicography*, 30(1), pp. 63-84.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. (2003). Corpora for Lexicography. In P. Van Sterkenburg (ed.) *A Practical Guide to Lexicography*. Amsterdam & Philadelphia: John Benjamins, pp. 167-178.

[*Sketch Engine*] Kilgarriff, A. (2003-2018). *Sketch Engine*. Brno & Brighton: Lexical Computing. Accessed at: http://www.sketchengine.co.uk [25/05/2018]

Smiley, I, Watson, E. & Delevante, M. (2014). *The Distiller's Guide to Rum*. Hayward: White Mule Press.

Smith, F. H. (2008) [2005]. *Caribbean Rum. A Social and Economic History*, 2nd edn. Gainesville: University Press of Florida.

St. Lucia Distillers. Accessed at: http://www.saintluciarums.com [25/05/2018]

St. Nicholas Abbey. Accessed at: http://www.stnicholasabbey.com [25/05/2018]

Svensén, B. (2003). Dictionary Projects. In R. R. K. Hartmann (ed.) *Lexicography: Critical Concepts*. London & New York: Routledge, vol. I, pp. 97-108.

[*TermoStat*] Drouin, P. (2010-2018). *TermoStat*. Montréal: Université de Montréal – Observatoire de linguistique Sens-Texte. Accessed at: http://termostat.ling.umontreal.ca [25/05/2018]

[*TextSTAT*] Hüning, M. (2015). *TextSTAT 3.0*. Berlin: Freie Universität Berlin. Accessed at: http://neon.niederlandistik.fu-berlin.de/en/textstat/ [25/05/2018]

[*TLFi*] (1994). *Le trésor de la langue française informatisé*. Paris: CNRS editions. Accessed at: http://atilf.atilf.fr [25/05/2018]

Todhunter-Mitchell Distillery. Accessed at: http://www.burnshouse.com [25/05/2018]

Tognini Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam & Philadelphia: John Benjamins.

Tognini Bonelli, E., Sinclair, J. (2006) [1993]. Corpora. In K. Brown (ed.) *Encyclopedia of Language & Linguistics*. Boston: Elsevier, vol. III, pp. 206-219.

Westerhall Estate. Accessed at: http://www.westerhallrums.com [25/05/2018]

Williams, I. (2006) [2005]. *Rum: A Social and Sociable History of the Real Spirit of 1776*, 2nd edn. New York: Nation Books.

[*WordSmith Tools*] Scott, M. (2018). *WordSmith Tools 7.0*. Liverpool: Lexical Analysis Software. Accessed at: http://www.lexically.net/wordsmith [25/05/2018]

Worthy Park Estate. Accessed at: http://www.worthyparkestate.com [25/05/2018]

XM. Accessed at: http://www.xmrumguyana.com [25/05/2018]

Zaya. Accessed at: http://www.infiniumspirits.com [25/05/2018]

Acknowledgements

The author would like to thank Jeannette Allsopp and Jason Siegel (*University of the West Indies* at Cave Hill, Barbados) for their careful revision of the manuscript and advice on key bibliographic material. Thanks are also due to Marek Łukasik (*Pomeranian University* in Słupsk, Poland) for his valuable comments and precious suggestions on the methodological implant. The author is grateful to Cristina Castielli, a student at the *University of Turin* (Italy), who contributed to the compilation of the *Caribbean Rum Corpus* (*CRC*) as part of her MA thesis. A final word of thanks goes to the personnel and tour guides at *Mount Gay Distillers* visitor center (St. Michael, Barbados) and distillery (St. Lucy, Barbados), *St. Nicholas Abbey* (St. Peter, Barbados) and *Foursquare Rum Distillery* (St. Philip, Barbados). Cristiano Furiassi's investigation was made possible by a three-month research stay – from September 2015 to December 2015 – at the *Richard and Jeannette Allsopp Centre for Caribbean Lexicography* of the *University of the West Indies* at Cave Hill, Barbados, jointly sponsored by the *University of Turin* and *Fondazione CRT* through the second edition of the *World Wide Style* (*WWS*) fellowship. In May 2016 Cristiano Furiassi was also awarded a certified sommelier diploma by the *Associazione Italiana Sommelier* (*AIS*), which provided him with more thorough knowledge of world spirits, including rum.