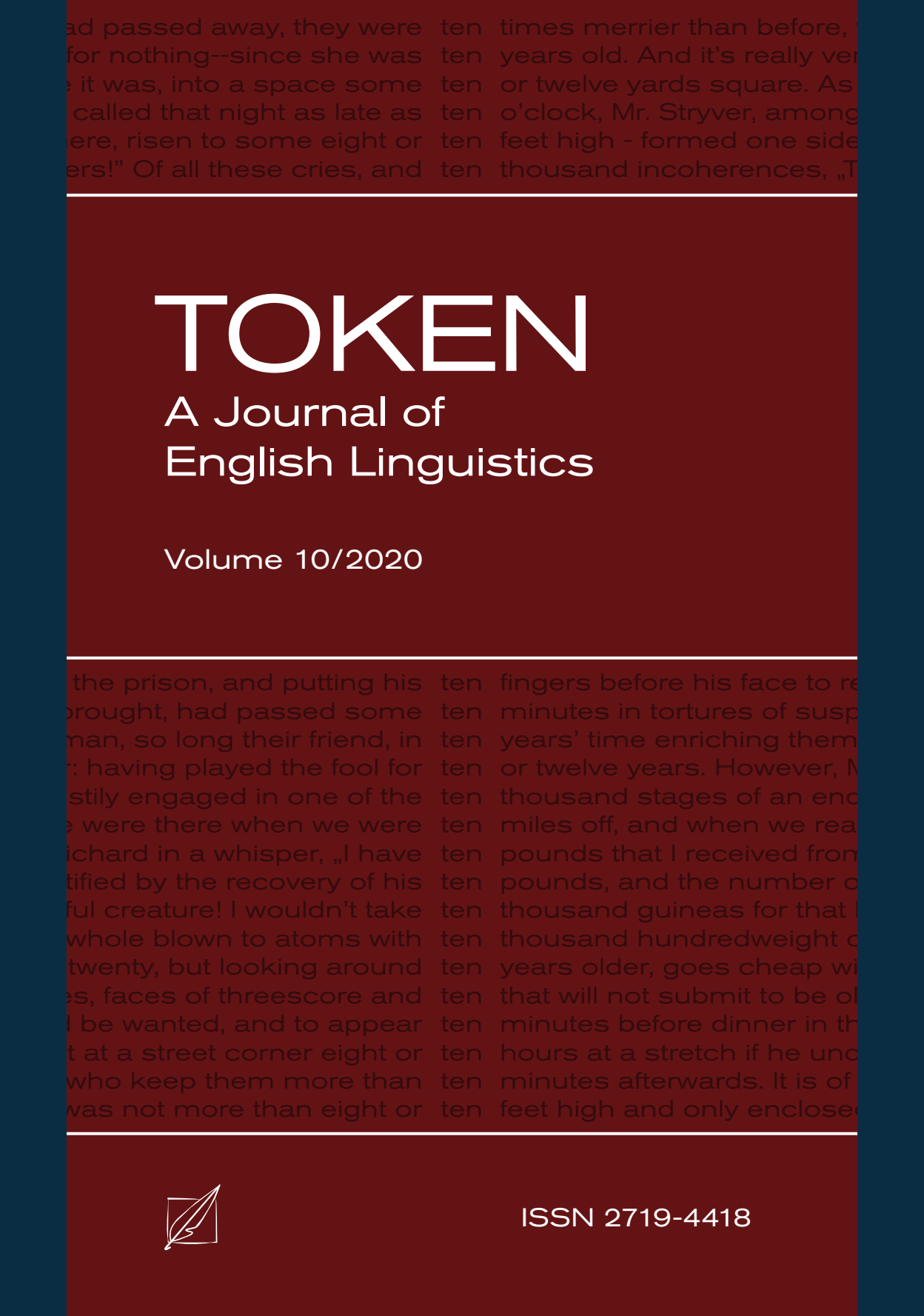

ISSN 2719-4418

Token: A Journal of English Linguistics
Volume 10

JAN KOCHANOWSKI UNIVERSITY OF KIELCE

Token: A Journal of English Linguistics

Volume 10

Edited by
John G. Newman
Marina Dossena
Sylwester Łodej

Jan Kochanowski University Press

Kielce 2020

Cover design Jakub Patryk Łodej, Anna Domańska

Formatting Marzena Buksińska

Editorial Correspondence
Jan Kochanowski University
Institute of Linguistics and Literary Studies
ul. Uniwersytecka 17
25-406 Kielce
Poland
https://token.ujk.edu.pl

Preface

The publication of Volume 10 marks a jubilee issue of *Token: A Journal of English Linguistics*. On this occasion, the editorial team would like to extend sincere thanks to those who have been instrumental in their support of *Token* from its inception, the members of the Advisory Board, the several guest editors, the conference organizers who offered collections of papers for special issues, the reviewers of submissions, the authors of articles published, the administrators of the sponsoring institutions Jan Kochanowski University, the University of Bergamo, and the University of Texas Rio Grande Valley, and the greater academic readership of the journal. You, esteemed colleagues, have been fair winds and following seas to our endeavor.

English linguistics broadly conceived has from the inaugural issue defined the scope of *Token*, and a brief glance back at some of its articles reveals that sweep, but it also evinces disciplinary expansion. Among *Token*'s volumes to date can be found: consequential syntactic analyses, like those of John Anderson (Volume 1), Daisuke Suzuki (Volume 3), and Nuria Calvo Cortés (Volume 10); penetrating morphological studies, like those of Chris Palmer (Volume 2) and Julia Schultz (Volume 4); insightful phonological reviews like that of Łukasz Stolarski (Volume 1); and elucidating lexical investigations like those of Joanna Esquibel and Anna Wojtyś (Volume 1), Elisabetta Lonati (Volume 7), Ewa Ciszek-Kiliszewka (Volume 2), and Paulina Ołownia (Volume 5); as well as enterprising surveys of English translation semantics, such as those of Judith Turnbull (Volume 7) and Polina Shvanyukova (Volume 7), English medical discourse, such as those of Giovanni Iamartino (Volume 8) and Cecilia Lazzaretti and Franca Poppi (Volume 9), linguistic evidence for social identities, such as that of Minna Nevala and Matylda Włodarczyk (Volume 6), pragmatics of content words, such as that of Gordana Dimkovic-Telebakovic (Volume 2), Internet English discourse, such as that of Isabel Ermida (Volume 6), and cognitive semantics, such as that of Wendy Anderson (Volume 10). The full list of the papers published in Volumes 1-10 follows these introductory lines.

Today, thanks in some measure to hard-won data base indexing in SCOPUS, EBSCO, ERIH PLUS, MLA, and elsewhere as well as to its golden open access policy, *Token* can boast an unquestionably global reach and a truly international identity. Guaranteeing the journal an intercontinental dissemination from the beginning, the initial editorial partnership between Sylwester Łodej and John Newman itself represented both Poland and the US. When Marina Dossena joined the editorial team at Volume 4, Italy was represented there also. Original Advisory Board membership added the United Kingdom, Italy, Germany, Finland, Spain, Austria, Ukraine, and Japan representation, and authors of Volume 1 papers hailed from some of these countries but also France and New Zealand. Since then, and including scholars who have kindly refereed papers submitted to *Token*, national representation has widened further to include Canada, The Netherlands, Norway, Sweden, Belgium, Switzerland, Hungary, Malta, Australia, and China as well. The increasing breadth of scholarly representation is perhaps one of the most significant markers of *Token's* international development.

Not to be remiss at this moment, we must put down that, sadly, three members of the Advisory Board, the eminent Jacek Fisiak, the renowned Dieter Kastovsky, and the decorated Farzad Sherifian passed on during the period. Also, with deep sorrow we note the passing of Matti Rissanen, Professor Emeritus of the University of Helsinki, and Jim Miller, Professor Emeritus of the University of Edinburgh – dear friends and colleagues who assisted our editorial work with their expertise and advice. We mourn them as good souls, and as shining men of letters. In their places, our forum knows hollow but echoing spaces.

Raising our chins to the horizon, we celebrate the day, and we look forward to future submissions with interest and curiosity for new trends in the discipline: ten volumes on, we are still as energetic and enthusiastic in our venture to circulate significant new work in the field of English linguistics. With the invaluable support of our colleagues, whether as authors, readers, book reviewers, or peer reviewers, *Token* will continue to be one of the best enterprises we have ever undertaken.

John G. Newman
Sylwester Łodej
Marina Dossena

Yoko Iyeiri
Sheila Dooley
Polina Shvanyukova
Chris C. Palmer

Aneta Iwan (UJK Press)
Marzena Buksińska (Typesetter)
Grzegorz Krukowski (Webmaster)

Papers published in *Token*: *Journal of English Linguistics*, Volumes 1-10

VOLUME 1

John M. Anderson *University of Edinburgh* (Emeritus) "Parasitic passives of intransitives in English".

Marina Dossena *University of Bergamo* "Sense and sensibility: verbal morpho-syntax in nineteenth-century Scottish emigrants' letters and the intersection of standard and vernacular usage".

Stefania Maci *University of Bergamo* "The language of tourism regulations in the European Union".

Andreea S. Calude & Gerald Delahunty *University of Reading* and *Colorado State University* "Inferentials: fixed or not?".

Patrice Larroque *Paul Valery University, Montpellier* "The prepositions at and to: opposite points of view".

Joanna Esquibel & Anna Wojtyś *University of Warsaw* "Devil aka Satan: an enemy or fiend? On the rivalry between the familiar and the foreign in early English".

Magdalena Murawska *Adam Mickiewicz University, Poznań* "Patient imaging in medical case reports".

Daisuke Suzuki *Kyoto University* "A corpus-based study of modal adverbs in English: from the viewpoint of grammaticalization".

Elena Salakhyan Eberhard Karls *University of Tübingen* "Tübingen Corpus of Eastern European English (TCEEE): from a small-scale corpus study to newly emerging non-native English variety".

Oskar Gawlik *Jan Kochanowski University, Kielce* "On the complementation of start, begin and continue in spoken academic English".

Łukasz Stolarski *Jan Kochanowski University, Kielce* "Size-sound symbolism in car names".

VOLUME 2

Chris C. Palmer *Kennesaw State University* "Historical sociolinguistic approaches to derivational morphology: A study of speaker gender and nominal suffixes in Early Modern English".

Brian Lowrey *University of Picardie Jules Verne* "The Old English causative verb hatan and its demise".

Stefania Maci *University of Bergamo* "Popularising scientific discourse for an academic audience: The case of Nobel lectures".

Ewa Ciszek-Kiliszewska *Adam Mickiewicz University*, *Poznań* "Middle English preposition and adverb twix".

Laura Esteban-Segura *University of Murcia* "A diachronic study of the prepositions among and amongst".

Patrice Larroque Paul Valery *University of Montpellier* "The Representation of Modality in Non-Standard English".

Magdalena Murawska *Adam Mickiewicz University*, Poznań "A multitude of voices and worlds: Towards a new model of the medical case report".

Masatoshi Konno *Tohoku University* "Inverted reporting verbs in the sentence-initial position in English: Focus on phrasal verbs".

Takashi Hamada *Keio University* "Evaluative Meaning of Negative Expression: Asymmetry between readable and unreadable".

Gordana Dimković-Telebaković *University of Belgrade* "The semantics and pragmatics of motion verbs in air traffic English and general English."

Alexandra Uzoaku Esimaje *Benson Idahosa University* "Patterns of lexical collocation in sermonic texts".

Mayuko Yoneda *Kyoto University* "Analogically-driven strong and weak verb transformations in the history of English: The role of synonyms".

Łukasz Stolarski *Jan Kochanowski University, Kielce* "Style-shifting as a function of multiple factors: A corpus based study".

Shala Barczewska *Jan Kochanowski University, Kielce* "The 1925 Scopes Trial as a Discursive Event: Does reference to the 1925 trial affect our view of teachers in the contemporary debate over evolution?"

VOLUME 3

Joan Beal & Marco Condorelli *University of Sheffield* "Cut from the same CLOTH? Variation and change in the CLOTH lexical set".

Massimo Sturiale *University of Catania* "The social construction of Standard (Spoken) English: Eighteenth-century orthoepists as a 'discourse community'".

Marina Dossena *University of Bergamo* "'Dispensers of knowledge'. An initiatory investigation into nineteenth-century popular(ized) science".

Polina Shvanyukova *University of Milan* "'A cargo of coffee, sugar, and indigo': Transatlantic business correspondence in nineteenth-century business letter-writing manuals".

Isabel Moskowich & Begoña Crespo *University of La Coruña* "Stance is present in scientific writing, indeed. Evidence from the Coruña Corpus of English Scientific Writing".

Sofia Zea *University of La Coruña* "Attributive adjectives in eighteenth-century scientific texts from the Coruña Corpus".

Kevin McCafferty *University of Bergen* "'I was away in another field […] got'. A diachronic study of the be-perfect in Irish English".

Nataša Stojakovič *University of Sarajevo* "Diachrony and idiosyncrasy: The subjunctive in the first half of the nineteenth century".

Daisuke Suzuki *University of Kyoto* "A historical study of English modal adverbs: Evidence from a combination of diachronic corpora".

Ayumi Nonomiya *University of Sheffield* "THOU and YOU in eighteenth-century English plays".

VOLUME 4

Dania Jovanna Bonness *University of Bergen* "'How is her eyes [?] are they still closed [?]' Subject-verb agreement in nineteenth-century Irish English".

Anna Franca Plastina *University of Calabria* "The discourse of comforting: The case of online health support groups".

Anna Ciechanowska *PWSTE Jarosław* & Grzegorz Andrzej Kleparski *University of Rzeszów* "On the semantic features of prison slang".

Annalisa Bonomo *"Kore" University of Enna* "The English language and the 'Multilingual Turn'".

Julia Schultz *University of Heidelberg* "Word-formations of recent borrowings from French to English: An analysis based on data from the Oxford English Dictionary".

Bogumiła Pacyna *Adam Mickiewicz University, Poznań* "Parallel and comparable corpora in investigating modal verbs in legal and literary discourse".

Sylwia Pielecha *University of Warsaw* "Reflexivity in Ælfric's Lives of Saints: A Corpus study".

Aleksandra Zofia Kowalczyk *University of Rzeszów* "On the projection of FOODSTUFFS on the macrocategory BODY PARTS".

VOLUME 5

Elisabetta Lonati *Università degli Studi di Milano* "The language of medicine in the Philosophical Transactions: Observations on style".

Patrice Larroque *Toulouse-Jean Jaurès University*, Toulouse "The dialogic nature of nonstandard English".

Annalisa Bonomo *"Kore" University of Enna* "English in multilingual education programs: some cases and perspectives".

Mayowa Akinlotan *Vrije Universiteit Brussels* "Genitive alternation in New Englishes: The case of Nigerian English".

Julia Schultz *University of Heidelberg* "The impact of Arabic on the English lexicon since 1801".

Daisuke Suzuki *Setsunan University* "Maybe: Development and Topic Marking".

Paulina Ołownia *Jan Kochanowski University, Kielce* "The lexical field WINE in American and British English: A corpus-based study".

VOLUME 6

Minna Nevala & Matylda Włodarczyk *University of Tampere* and *Adam Mickiewicz University, Poznań* "Exploring social identities in public texts".

Hanna Limatius *University of Tampere* "'There really is nothing like pouring your heart out to a fellow fat chick': Constructing a body positive blogger identity in plus-size fashion blogs".

Tony McEnery & Helen Baker *Lancaster University* "The poor in seventeenth-century England – a corpus based analysis".

Anni Sairio *University of Helsinki* "A learned woman's warfare: Social identity in eighteenth-century anonymous satire".

Daniela Cesiri *Ca' Foscari University of Venice* "'Haha, what a twit I am'. The construction of a social identity in the comments sections of UK food blogs".

Matylda Włodarczyk *Adam Mickiewicz University, Poznań* "Social identities in an institutional network: Colonial Office correspondence on the Cape Colony in the early nineteenth century".

Minna Palander-Collin & Ina Liukkonen *University of Helsinki* "Constructing the defendant role in the trial proceedings of the Old Bailey: Guilty or not guilty".

Isabel Ermida *University of Minho* "'Get the snip – and a job!' Disagreement, impoliteness and conflicting identities on the Internet".

VOLUME 7

Marina Dossena *University of Bergamo* and Polina Shvanyukova *University of Florence* "Introduction".

Nicholas Brownlees *University of Florence* "The Gazette de Londres: Disseminating news and exercising news management through translation".

Massimo Sturiale *University of Catania* "Late Modern newspapers as a mirror of linguistic (in)stability and change".

Marina Dossena *University of Bergamo* "'Sassenach', eh? Late Modern Scottish English on the borders of time and space".

Elisabetta Lonati *University of Milan* "Words of religious dissent in eighteenth-century Italian translations of Chambers's Cyclopaedia".

Polina Shvanyukova *University of Florence* "Teaching Business English in nineteenth-century Italy".

John Douthwaite *University of Genoa* "The method and practice of translational stylistics".

Judith Turnbull *Sapienza University of Rome* "Museum communication: The role of translation in disseminating culture".

Larissa D'Angelo *University of Bergamo* "Translating cultural references in the Italian dubbing and re-dubbing of E.T. The Extra-Terrestrial".

VOLUME 8

Giovanni Iamartino *University of Milan* and Irma Taavitsainen *University of Helsinki* "Introduction".

Giovanni Iamartino *University of Milan* and Giulia Rovelli *University of Insubria* "A Physical Dictionary of 1655: When translating medical science is not enough".

Lucia Berti *University of Milan* "Italy and the Royal Society: Medical papers in the early Philosophical Transactions".

Alicia Rodríguez-Álvarez *Universidad de Las Palmas de Gran Canaria* "The medical entries in John Kersey's abridged Dictionarium Anglo-Britannicum (1708) or how to retain highly demanded lexical material in a short dictionary".

M. Victoria Domínguez-Rodríguez *Universidad de Las Palmas de Gran Canaria* "An overview of medical terminology in Nathan Bailey's An Universal Etymological English Dictionary (1721)".

Elisabetta Lonati *University of Milan* "Stabilising scientific lexicon in eighteenth-century British encyclopædias and specialised dictionaries: Focus on medical terminology".

Magdalena Zabielska *Adam Mickiewicz, Poznań University* "'I resolved to cut. But there was before my eyes the fear of haemorrhage' – subjective, emotional and author-centred discourse of the late nineteenth-century case reports from the British Medical Journal".

Anna Franca Plastina *University of Calabria* "Case reporting: A historical discourse analysis of the functional uses of if-conditionals in medical-officer-of-health reports".

Kim Grego *University of Milan* "From 'dying well' to 'inducing a good death': euthanasia in the British press (1864-1949)".

Book reviews:
K. Aaron Smith and Susan M. Kim This Language, A River: A History of English, Peterborough, Ontario: Broadview, 2018, 360 pp. Reviewed by Felicia Jean Steele, *The College of New Jersey*, USA.

Julia Schultz The Influence of Spanish on the English Language since 1801: A Lexical Investigation, Newcastle upon Tyne: Cambridge Scholars, 2018, xi + 314 pp. Reviewed by José A. Sánchez Fajardo, *University of Alicante*, Spain.

VOLUME 9

Marina Bondi & Franca Poppi *University of Modena and Reggio Emilia* "Introduction".

Giuliana Diani *University of Modena and Reggio Emilia* "Metadiscourse in web-mediated health communication".

Cecilia Lazzeretti & Franca Poppi *University of Modena and Reggio Emilia* "Children with autism or autistic children? Indexicality in the websites for parents of children with neurological conditions".

Annalisa Sezzi & Marina Bondi *University of Modena and Reggio Emilia* "'I am going on a ketogenic diet'. Communicating dietary requirements for pediatric patients".

Silvia Cavalieri, Maddalena Marchiò, Marina Bondi, Giuseppe Biagini *Sapienza University of Rome University of Modena and Reggio Emilia* "Assessing caregiver informative materials on the Ketogenic Diet in Italy: a textual ethnographic approach".

Stefania M. Maci *University of Bergamo* "Parents' narrative about congenital heart diseases: Acquiring knowledge and sharing empathy".

Magdalena Zabielska *Adam Mickiewicz University, Poznań* and Magda Żelazowska-Sobczyk *University of Warsaw* "Hearing aids and cochlear implants – A study of the discourse of patients' English ENT online support forums".

Gillian Mansfield *University of Parma* "The dynamics of popularised discourse for diabetics: Online forums for information and emotional support".

Elisa Mattiello *University of Pisa* "Figuration and obesity: Warning bells from TheGuardian.com".

Judith Turnbull *University of Modena and Reggio Emilia* "Online health information for the elderly".

Winnie Cheng *Hong Kong Polytechnic University* "A study of closings in nurse-elderly resident consultations at a mobile integrative health centre".

VOLUME 10

Wendy Anderson *University of Glasgow* Metaphor in the digital age: Opening the flood-gates.

James Stratton *Purdue University* Fiction as a Source of Linguistic Data: Evidence from Television Drama.

Christina Samson *University of Florence* From private to public: Letters contextualising the 1857-58 mutiny in the British press.

Marina Dossena *University of Bergamo* Peaceful coexistence? Ideology in the Representation of Languages and Varieties in Late Modern Literature.

Paula Schintu *University of Salamanca* "Theers gud stuff amung uz Darbysher foaks": Dialect Enregisterment in 19th-century Derbyshire

Annalisa Scatà *Independent scholar* Ideologies of linguistic representation in Late Modern English: The case of James Fenimore Cooper.

Elisabetta Cecconi *University of Florence* Overlexicalization and semantic variation in the Early Modern English naming of Native Americans.

Paloma Tejada Caller *Universidad Complutense of Madrid* A Cultural Linguistics approach to the "discovery of childhood" in sixteenth and seventeenth century Britain.

Nuria Calvo Cortés Complutense *Universidad Complutense of Madrid* Maria Edgeworth's choice of auxiliary verb in perfect tenses.

Sonja Kuosmanen *University of Helsinki* Terms of Reference and Discursive Representations: A Case Study with Saddam Hussein in the 1991 Gulf War

Valeria Franceschi *University of Verona* Making sure everybody is on the same page: interactional communication strategies in BELF encounters.

Patrizia Anesa *University of Bergamo* Legal English as a Lingua Franca in Academia: The Strategic Use of Repetitions in Lectures.

Book review:
Stefan Dollinger Creating Canadian English: The Professor, the Mountaineer, and a National Variety of English Cambridge University Press, Cambridge. 2019, xviii + 283 pp. Reviewed by Carol Percy, *University of Toronto*, Canada.

Metaphor in the digital age: Opening the flood-gates

Wendy Anderson
University of Glasgow

ABSTRACT

The digital age and recent developments in historical lexicography bring exciting new possibilities for the study of metaphor, in relation to both the scale of analysis and the identification of linguistic and conceptual patterning that was previously hidden from view. This article illustrates this by using the "Metaphor Map of English" to investigate metaphor in the semantic field of moving water. The Metaphor Map offers an unparalleled overview of the place of metaphor in the recorded language history of English and across semantic space, as it takes as its source of data the entire *Historical Thesaurus of English*, itself largely based on the *Oxford English Dictionary.* This semantic field (labelled 'Tides, waves and flooding' in the Metaphor Map) is a productive source of metaphor, and has been since the Old English period, and is also, though less commonly, the target of metaphor. It reveals a clear picture of the complex interrelationships between semantic categories and the ripples of metaphor development over time.

Keywords: metaphor, *Historical Thesaurus of English*, Metaphor Map of English, lexicography, semantics, digital humanities.

1. Introduction

The digital age is an exciting time for linguistics. Early advances in computer technology in the second half of the twentieth century made the growth of corpus linguistics possible. This opened up new perspectives on language in text and a better understanding of, for example, lexicogrammatical patterning, collocation and semantic prosody, phenomena which could only be glimpsed hazily before corpus analysis tools allowed us to manipulate

DOI: 10.25951/4356.

text in large quantities and move away from a purely linear reading of texts. As corpus linguistics has expanded and arguably blended with the newer, broader endeavour of Digital Humanities, new possibilities for the analysis of language and text have come to light.

The digital analysis of the data contained in large-scale reference works such as dictionaries and thesauruses is part of this. The project that forms the focus of this article is of exactly this type: 'Mapping Metaphor with the *Historical Thesaurus*'[1] (henceforth 'Mapping Metaphor') exemplifies what we can do when we have a complete historical thesaurus of a language, in the digital age.[2] Mapping Metaphor draws on the entire contents of the *Historical Thesaurus of English*, and indirectly on the material contained in the *Oxford English Dictionary* (*OED*, 2nd edn.), to map out metaphor in English across semantic space and over time. In short, it aims to establish the place and role of metaphorical word senses in the language system of English.

The Mapping Metaphor project identified nearly 12,000 metaphorical connections between the semantic categories of the *Historical Thesaurus*. While this article will give a brief overview of the picture of metaphor as a whole, clearly we cannot do justice to the richness of the resource here. Therefore, we will narrow in on a single semantic category and focus on metaphor in that category as it relates to time and semantic space. This category, named 'Tides, waves and flooding', pulls in the lexical material from the *Historical Thesaurus* relating to a number of types of moving water (not including rivers and streams which sit in an adjacent category and are metaphorically rich in their own right). As we will see, some metaphors in 'Tides, waves and flooding' are long-standing and traceable back to Old English; others have emerged at different points in the history of English. The category as a whole is neither one of the most densely packed with metaphor, nor one in which the extent of metaphor is slight. Rather it

[1] 'Mapping Metaphor with the *Historical Thesaurus*' was funded by the UK Arts and Humanities Research Council between 2012 and 2015 and carried out at the University of Glasgow, UK (AHRC grant number: AH/I02266X/1). The project team was large, including the present author as Principal Investigator, Christian Kay, Carole Hough and Marc Alexander as Co-Investigators, Ellen Bramwell as Postdoctoral Research Assistant, Brian Aitken as Digital Humanities Resource Officer, Rachael Hamilton as PhD student, Flora Edmonds as Technician, and many postgraduate and undergraduate students who contributed to the huge effort of manually identifying metaphor in the digitally manipulated lexical data.

[2] This research was presented at the 'English Historical Lexicography in the Digital Age' event at the Università degli Studi di Bergamo, 11-13 April 2019. I would like to thank Professor Marina Dossena and the organising team for the invitation to speak at that event and for encouragement with this article.

represents a semantic area with a moderate amount of metaphor that has permeated and characterised the English language over a long time.

2. Metaphor and the *Historical Thesaurus of English*

To use a metaphor appropriate to our focus here, the *Historical Thesaurus of English* represents a high-water mark in historical lexicography. It was created, under the direction of Professor Michael Samuels and then Professor Christian Kay at the University of Glasgow over a period of 45 years, leading to print publication of the complete *Historical Thesaurus* in 2009 (Kay et al. 2009), and work continues, with subsequent versions made available online (*Historical Thesaurus of English* online). In essence, the *Historical Thesaurus* presents the entire contents of the *Oxford English Dictionary* (2nd edn.), supplemented for the earlier period by *A Thesaurus of Old English* (Roberts – Kay 1995), organised in a complex hierarchical semantic system.

The possibilities that the *Historical Thesaurus* opens up for research into lexis and meaning, synchronically and diachronically, are endless. What we saw in it, however, was a way of viewing the extent and complexity of metaphor in the language system of English across the history of the language. This became the aim of the Mapping Metaphor project.

The *Historical Thesaurus* has a number of qualities that made it an ideal starting point for the pursuit of metaphor. Most important – indeed the *sine qua non* – is its semantic organisation. The lexical items in the *Historical Thesaurus* are arranged into a complex hierarchy of 225,131 nested semantic categories.[3] At the highest level, there are three primary divisions, The External World, The Mental World and The Social World. These are broken down into, or made up from, depending on one's perspective, 37 second-level divisions (major semantic groupings such as Animals, Emotion and Morality), 377 third-level divisions (broadly 'basic' semantic categories such as Fish, Anger and Virtue) and so on down to a maximum of twelve levels, depending on the category in question. In fact, this hierarchy proved to be too complex for the purposes of Mapping Metaphor, and instead we flattened the structure, incorporating only the highest two divisions exactly, and collapsing all the lower divisions into a slightly adapted third level.

[3] Figures for the *Historical Thesaurus* relate to version 4.0, which was the version that initially informed the Mapping Metaphor project. Full details of version changes can be found at https://ht.ac.uk/versions-and-changes/ (accessed July 2019).

The other qualities of the *Historical Thesaurus* that make it suitable for investigating metaphor are its sheer size and its long time-span. It encapsulates the entirety of semantic space – all concepts that have been lexicalised in English and recorded in the *OED*, from the first of the third-level divisions, 'Region of the earth', to the last, 'Dancing', and everything in between. In the version of the *Historical Thesaurus* used for Mapping Metaphor, this includes 793,742 word senses. In terms of time-span, the lexemes in its scope extend from the eighth century to the present day, roughly 1300 years of English. The early materials include items attested in the standard dictionaries of Old English on which the *Historical Thesaurus* drew for the period before c1150, especially Bosworth – Toller (1898) and Clark Hall – Merritt (1960). This very broad historical sweep allows us to see metaphor across an extended period of time.

2.1 Identifying metaphor

Using this dataset, the project team's aim was to identify all the systematic metaphorical connections that speakers of English have made since the earliest times of the language. The guiding principle was that patterns of metaphor could be identified from considering the lexical items that appear in two or more semantic categories of the *Historical Thesaurus*. That is, the project exploited the polysemy created by metaphor. For example, the polysemous word forms *rain*, *tempest*, *cloudy* and *storm* appear in several semantic categories of the *Historical Thesaurus*. Significantly, they all appear in semantic categories related to Weather (the source domain) and Emotional suffering (the target domain). This suggests a robust metaphorical connection between the concepts of weather and emotion. The individual words also appear in other places, however: *cloudy* appears in semantic categories relating to Intelligibility (where it is metaphorical and has the sense 'abstruse', 'obscure') and *rain* appears in semantic categories related to Abundance (where it is also metaphorical and relates to a profusion of something).

The process involved breaking the *Historical Thesaurus* data down into 415 semantic categories of similar levels of generality, e.g. Weather, Emotional suffering, Sight, Taste, Music, Excitement, and so on, and then automatically comparing all the lexical entries in every category with all the lexical entries in every other category, thus creating sets of shared lexis for every pair of categories. The following stage was to work through all of these data files to identify where this shared lexis was due to metaphor. There are various

reasons why word forms appear in more than one semantic category, and metaphor is only one of these, alongside homonymy and forms of polysemy other than that motivated by metaphor. So for example, *rain* appears not only in categories of Weather, Emotional suffering and Abundance, but also in Farming where it denotes a narrow strip of land. This last sense is however homonymous with the others and comes, as the *OED* notes, from a borrowing from early Scandinavian (*OED rain* n.2) whereas the sense related to water vapour is an unrelated word inherited from Germanic (*OED rain* n.1). There is no semantic connection, let alone a metaphorical connection, merely an accidental identity of word forms. The metaphor identification stage of the process was largely manual and very time-intensive, as while computers are invaluable in the pattern-matching task of identifying repeated lexis, they are not good at identifying metaphor.

Of course, humans also struggle with identifying metaphor, or at least agreeing on what is metaphorical and what is not. Caballero and Ibarretxe-Antuñano have expressed this succinctly as follows: "metaphoricity may be seen as a matter of degree: not all metaphorical language is regarded as such by all people, underlining the role of context and social convention in metaphor awareness and identification" (Caballero – Ibarretxe-Antuñano 2013: 274).

It is not surprising, of course, that metaphor should be a matter of degree; if we take a prototype approach to language, we expect "fuzziness and gradualness" (Taylor 1995: 121). To the social and contextual differences that Caballero and Ibarretxe-Antuñano mention, we can also add the further difficulties that arise when handling historical language data. While language users may have a good (if not always conscious) feel for metaphorical usage in the language of their own time, there are more significant limitations when it comes to historical language.

As a project, therefore, we took quite an inclusive approach to metaphor, established a team of coders with various types of expertise and specialism, and also built a lot of cross-coding and cross-checking into our procedures from the outset. In the first main stage of analysis, all category pairs were coded twice (once from the start-point of one category, and once from that of the other), normally by different coders. All category pairs were then also checked by an independent checker. Then, in a final stage, we reviewed all of the metaphorical connections that we had identified in the earlier stages, and, again basing our deliberations on the evidence of the *Historical Thesaurus* and *OED*, selected lexical examples to substantiate each metaphorical connection, identified the source and target category in each

case (that is, established the directionality of the metaphor), and identified the date of the earliest evidence for each metaphor.

While the project team's understanding of metaphor is of course heavily influenced by the notion of a conceptual metaphor and by conceptual metaphor theory, which has been the dominant approach to metaphor in recent decades, the metaphors identified in our data are not all metaphors in the sense most familiar in conceptual metaphor theory, namely Lakoff and Johnson's TARGET IS SOURCE formulation (1980), of which LIFE IS A JOURNEY is probably the most famous example. Rather, they are more appropriately described as 'metaphorically related category pairs'. For example, the metaphorical connection between the semantic category 'Relative position' and 'Prosperity and success' cannot be summed up by a single TARGET IS SOURCE conceptual metaphor. The lexical items shared by those semantic categories suggest the presence of at least two conceptual metaphors within this metaphorically related category pair: a dominant metaphor PROSPERITY IS UP (instantiated by items such as *height*, *up*, *upward* and *top*) and a less strongly evidenced conceptual metaphor that we might call, perhaps somewhat awkwardly, PIONEERING IS BEING AT THE FRONT OF A MOVING OBJECT (instantiated by *leading edge*, which has the metaphorical sense of being in the vanguard of technological development). This lack of one-to-one correspondence between metaphorical connections that hold between semantic categories in a thesaurus and Lakoffian conceptual metaphors is a result of the data-driven and bottom-up nature of the Mapping Metaphor methodology, and the heavy reliance on the semantic categories developed, also through a bottom-up process, in the *Historical Thesaurus*.

As metaphor is a matter of degree, a key question here is how we knew when what we were seeing in our data was indeed metaphor. As a project team, we found ourselves constantly considering the nature and boundaries of metaphor. While there is significant recent work on criteria for identifying metaphor, much of this stems from research into metaphor in text or discourse and does not transfer well to research on metaphor in the language system (see in particular the Metaphor Identification Procedure (MIP) developed by the Pragglejaz Group 2007, and its refinement MIPVU, developed at the Vrije Universiteit, Amsterdam, see Steen et al. 2010). Nevertheless, we had various types of information to draw on to evidence metaphorical connections. No type was sufficient in itself, but together they gave us confidence in the metaphorical nature of connections.

One of the main types of support for a metaphorical connection came from the *Historical Thesaurus* and *OED* attestation dates for word senses. That is, we expected to see historical priority for literal senses and later dates

for extended, metaphorical word senses. As no reference work draws on absolutely comprehensive evidence of language in use, however, there is always the possibility that expected ante-datings are missing and we cannot always trust apparent historical priority. Similarly, we could draw with some confidence on the typical concrete-abstract directionality of metaphorical extensions as a clue to a metaphorical connection between categories but had to combine this information with human judgement. Concreteness and abstraction are relative notions, are frequently argued over by metaphor scholars (for a discussion see Dancygier – Sweetser 2014: 64-67) and do not neatly apply to all types of transfer, including those between concrete concepts (what the *OED* calls 'transferred' as opposed to figurative senses).

A major clue to the presence of metaphor lay in the patterns that emerged from looking at whole sets of lexical data together. There was stronger evidence for a systematic metaphorical connection between categories if it could be seen to be instantiated in several words. This can be illustrated with a small sample of data.

Table 1. Selection of data from 'Tides, waves and flooding', showing lexical items that also appear in 'Prosperity and success'

Lexical item	Part of speech[1]	Attestation dates	Sense
Rise	n	1721-	Advancement/progress
tide-wave	n	1861	..significant
tidal-wave	n	1884-	..significant
rising	n	1595-	.rise in prosperity/power/rank
rising	aj	1631-	.rising in prosperity/power/rank
silver	aj	1659/60	..of times/places
rise	n	1632-	.rise in prosperity/power/rank
increase	vi	1388-1722	.rise in prosperity/power/rank
rise	vi	1303-	.rise in prosperity/power/rank
buoy	vi	1709 + 1742 fig.	.rise to the top
peak	n	1902-	..highest point
high-water mark	n	1814-	..highest point

[1] In the Part of speech field, n = noun; aj = adjective; vi = intransitive verb. In the Attestation dates field, a dash after a date indicates that a sense remains current, and 'fig.' indicates a figurative sense. In the Sense field, dots before a sense gloss indicate hierarchical level.

Table 1 shows a small snippet of the data on which the Mapping Metaphor team drew when establishing a metaphorical connection between the categories 'Tides, waves and flooding', and 'Prosperity and success'. On the left, it shows a set of lexical items that appear in both categories. The other three columns all relate to each lexical item in the second category, 'Prosperity and success', and show its part of speech, the attestation dates for the particular sense of the lexical item in the second edition of the *OED*, and an indication of the sense in that category, drawn from the most immediate sense heading in the *Historical Thesaurus* hierarchy (fuller sense information could be readily accessed when needed). It is clear from this that, just as one would expect, there is a metaphorical connection between the two categories in question, with a number of words with literal senses in 'Tides, waves and flooding' being used with extended, metaphorical senses in 'Prosperity and success': *tide-wave*, *tidal-wave* and *high-water mark* stand out particularly. This connection is perhaps also supported by *buoy*, although one might argue that the source category for *buoy* is 'Navigation', found within the later Social World division, rather than 'Tides, waves and flooding'. The connection between 'Tides, waves and flooding' and 'Prosperity and success' is particularly clear-cut; in other cases, the process of identifying metaphorical connections was less straightforward, either because there were very few lexical items instantiating the connection or because the list of lexical items shared by the two categories was so extensive that identifying the metaphor was challenging.

2.2 The Metaphor Map of English

The identification and analysis of metaphor outlined above allowed us to create the project's online resources. These are the Metaphor Map of English (henceforth MME, based on the *Historical Thesaurus* data drawn originally from the *OED*) and the Metaphor Map of Old English (MMOE, based on the various data sources that come together as *A Thesaurus of Old English*). It was decided to create separate Maps for these time periods both because the data come from two quite different types of source and because the two sets of data are quite different in scale, bringing the risk that the very substantially smaller set of OE data could be swamped by that from the later period. Fig. 1 shows the interface of MME.

In Fig. 1, all of the 37 second-level categories are shown around the outside of the circle: working clockwise from the top, categories 1A to 1Q

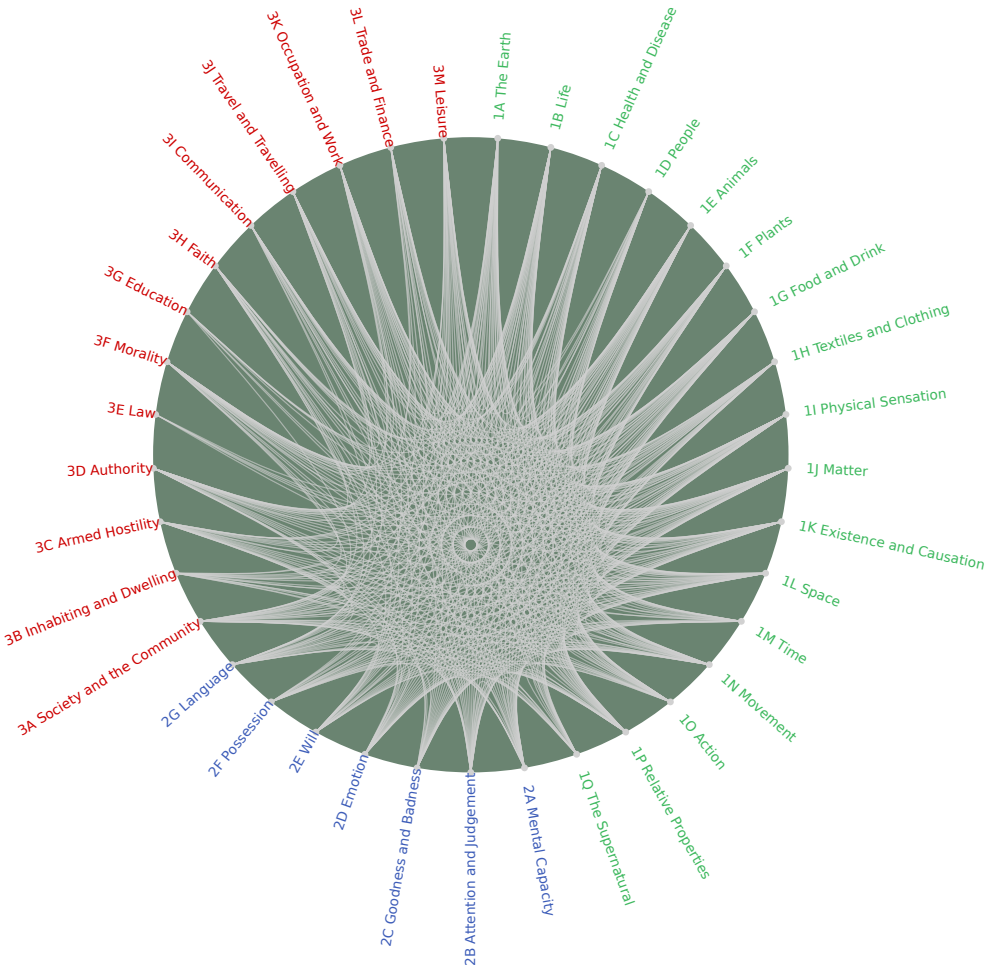

Figure 1. Metaphor Map of English, top-level view

represent the External World, 2A to 2G represent the Mental World, and 3A to 3M represent the Social World. Each category can be expanded to show its component third-level categories, which number 415 in total. The Metaphor Maps allow the user to see the detail of all of the metaphorical category connections that the project identified. Fig. 2 shows an expanded version, with third-level category 'Tides, waves and flooding' (which has the identifying code 1A13) opened up, and the high-level metaphorical links with second-level categories highlighted in yellow. Users can click on the yellow links to see further detail of the third-level categories with which 'Tides, waves and flooding' has specific metaphorical connections.

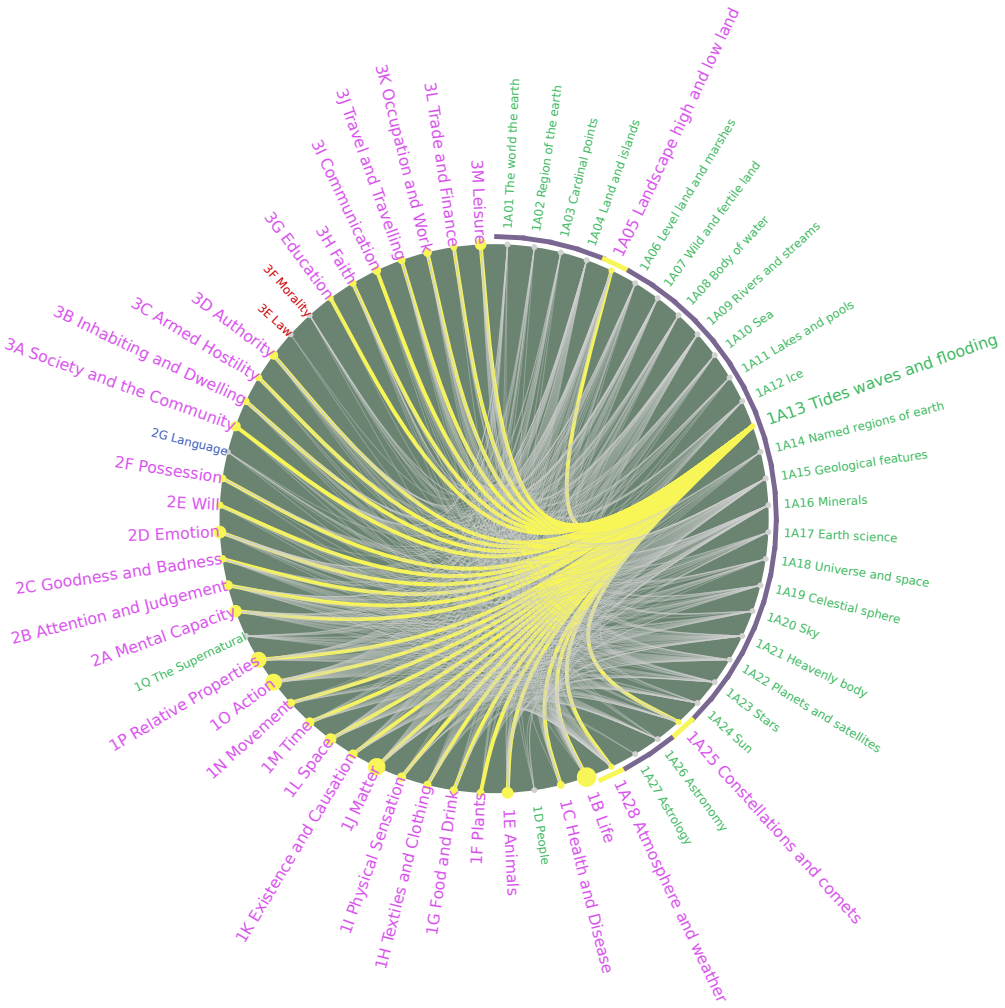

Figure 2. Metaphor Map of English, showing links to and from 'Tides, waves and flooding'

3. Metaphor across space and time

Section 4 below will focus on metaphor in one particular semantic category, 'Tides, waves and flooding', which was used to illustrate the Metaphor Maps above. First, however, it is useful for context to gain a sense of the scale of the Metaphor Maps by briefly surveying the picture that they show of metaphor across semantic space and over time.

Considering metaphor across semantic space, MME records around 12,000 metaphorical connections between semantic categories, and MMOE

records around 2,500. These are not evenly distributed across semantic space, however. Table 2 shows the ten categories in MME and MMOE, respectively, that enter into the largest numbers of metaphorical category connections. Category codes are given, to allow a better understanding of how each category fits into the overall Mapping Metaphor structure. The first letter of the code indicates the primary division (1 = The External World; 2 = The Mental World; 3 = The Social World), the following letter indicates the second-level category (e.g. 1L codes second-level category Space, which lies in 1 The External World), and the final pair of numbers indicates the third-level category.

Table 2. Categories with the highest numbers of metaphorical category connections

Metaphor Map of English (MME)		Metaphor Map of Old English (MMOE)	
Category code and name	Number of metaphorical category connections	Category code and name	Number of metaphorical category connections
1L04 Shape	245	1L06 Relative position	132
1G01 Food and eating	242	1L04 Shape	97
1L06 Relative position	233	1N06 Movement in a specific direction	95
1F01 Plants	217	1L03 Size and spatial extent	80
1L03 Size and spatial extent	201	3H01 Faith	62
1B11 Body parts	190	1F01 Plants	60
1N06 Movement in a specific direction	186	1B11 Body parts	58
1E09 Birds	179	1C02 Ill-health	54
1A28 Atmosphere and weather	177	1G01 Food and eating	52
1J03 Weight, heat and cold	171	1L05 Place and position	51

As can be seen here, there are four categories in MME which are part of a metaphorical connection, as either the source or the target of metaphor, with more than half of the other 414 categories: 'Shape', 'Food and eating', 'Relative position', and 'Plants'. In MMOE, the numbers of metaphorical

connections are much lower, owing to the smaller data set, and only 'Relative position' is part of more than 100 metaphorical connections, though 'Shape' and 'Movement in a specific direction' come close. Notably, seven of the ten most metaphorical categories in MME also appear in the equivalent list for MMOE, suggesting a continued strong place for metaphor in these semantic areas across the lifetime of the English language. As the categories are not of equal size with respect to the number of lexical items that they contain, it is also useful to look at normalised figures to see where in semantic space metaphor is most dense. Table 3 shows the ten categories in each of the two Metaphor Maps with the highest number of metaphorical connections per 100 lexical tokens.

Table 3. Categories with the highest numbers of metaphorical category connections per 100 lexical tokens

Metaphor Map of English (MME)		Metaphor Map of Old English (MMOE)	
Category code and name	Metaphorical connections per 100 lexical items	Category code and name	Metaphorical connections per 100 lexical items
1J08 Strength	45.39	1J08 Strength	141.67
1J10 Hardness	33.00	1A18 Universe and space	111.11
1I09 Touch	32.19	1J11 Softness	100
1B05 Absence of life	30.56	1J10 Hardness	83.33
1J31 Fireworks	29.23	3A01 Society	83.33
1J09 Weakness	28.02	1B23 Vascular system	71.43
1J27 Illumination	25.50	1J05 Lack of density	66.67
1B09 The human body	24.69	1J32 Transparency and opacity	66.67
1A18 Universe and space	24.62	1D06 Adult and middle-aged person	64.71
1B28 Manner of death	24.58	1J09 Weakness	64.71

The categories in Table 3 are quite different from those in Table 2, showing that the largest categories are not generally the most metaphorically dense. Again, however, there is considerable overlap between MME and MMOE, with four of the top ten in common. Particularly notable are 'Strength' and

'Hardness', in which metaphor appears as highly prevalent consistently from Old English onwards.[4]

Viewing the data over time, around two-thirds of the metaphorical category connections found in Old English are maintained into the later period, and many new metaphorical connections emerge continuously over the centuries that follow. Table 4 gives an overview of new metaphorical category pairs for each fifty-year period in MME.

Table 4. New metaphorical connections across time in the Metaphor Map of English

Period	Number of		
	new connections	of which strong / (%)	of which weak / (%)
Old English	779	610 (78.31%)	169 (21.69%)
1150-1199	41	30 (73.17%)	11 (26.83%)
1200-1249	335	234 (69.85%)	101 (30.15%)
1250-1299	144	97 (67.36%)	47 (32.64%)
1300-1349	514	342 (66.54%)	172 (33.46%)
1350-1399	778	447 (57.46%)	331 (42.54%)
1400-1449	529	266 (50.28%)	263 (49.72%)
1450-1499	304	137 (45.07%)	167 (54.93%)
1500-1549	788	374 (47.46%)	414 (52.54%)
1550-1599	1623	593 (36.54%)	1030 (63.46%)
1600-1649	1283	307 (23.93%)	976 (76.07%)
1650-1699	695	134 (19.28%)	561 (80.72%)
1700-1749	419	75 (17.90%)	344 (82.10%)
1750-1799	381	55 (14.44%)	326 (85.56%)
1800-1849	856	116 (13.55%)	740 (86.45%)
1850-1899	1078	79 (7.33%)	999 (92.67%)
1900-1949	882	61 (6.92%)	821 (93.08%)
1950-1999	453	17 (3.75%)	436 (96.25%)
Total	**11882**	**3974**	**7908**

[4] Note, however, that factors such as text scarcity for Old English may obscure other categories that may have been similarly prevalent.

It further breaks these connections down into 'strong' and 'weak' connections. The distinction between strong and weak metaphorical connections is not a precise measure, but an assessment of relative strength by the coding team, taking account of the extent of the lexical evidence for a metaphorical connection and its presence in a range of genres, as opposed to isolated occurrences in a specialised genre such as poetry.

4. Tides, waves and flooding

We can look at both of these axes – semantic space and time – from the perspective of an individual category. 'Tides, waves and flooding' is one of 415 categories in the Metaphor Maps. Its code, 1A13 indicates that it sits in the first primary division, the External World, and the first second-level category The Earth. Here it sits alongside other third-level categories such as 'Lakes and pools', 'Ice', 'Minerals' and 'Named regions of earth'. Semantically, the category contains the component of the *Historical Thesaurus* that concerns water as it moves in relatively natural and uncontrolled ways, specifically the subcategories of the *Historical Thesaurus* listed in Fig. 3.

> Flowing; Flow/flowing; Flow; Cause to flow; Current; Tide; Tidally; Type of tide; State of sea; Have/be in specific kind of motion; Direction of sea; Wave; Move in waves; Pour in waves; Foam/surf; Form surf; Movement of waves; Move restlessly about; Whirlpool; Turn into whirlpool; Sudden rush of water; Rush; Flood/flooding; In flood; Flood/overflow; Structure protecting from water/flooding; Flowing water

Figure 3. The semantic scope of the category 'Tides, waves and flooding'

'Tides, waves and flooding' enters into a metaphorical connection with 108 other categories in the Metaphor Map of English (that is, about a quarter of the others) and 26 of the other categories in the Metaphor Map of Old English. Clearly, therefore, there are a number of very long-standing metaphorical connections that can be traced back to the earliest stages of the language; there are also many connections which appear as new on the basis of the evidence we have and much new lexical evidence for existing connections emerging in every century up to the present. We can get a sense from this of waves of metaphor development over time.

Some categories in the Metaphor Map can be characterised as predominantly providing the source of metaphor (for example, 'Food', 'Birds', 'Plants', 'Body parts' – notably concrete/tangible concepts), and others as being the target of metaphor (for example, 'Emotional suffering', 'Behaviour and conduct' – notably abstract concepts). 'Tides, waves and flooding', like many other categories, has a mixed profile, though in this case with a strong tendency towards being the source of metaphor. Let us look at each of these briefly, and then at the more complex cases where 'Tides, waves and flooding' apparently enters into bidirectional metaphorical connections with other categories.

4.1 'Tides, waves and flooding' as source

'Tides, waves and flooding' is more typically the source than the target of metaphor, as one would expect from a category in the second-level category The Earth, which predominantly contains concepts of material things. In MME, lexical items from 93 other categories have been identified as instantiating a metaphorical connection in which 'Tides, waves and flooding' is the source. In MMOE, there are 17 such categories. It should not perhaps come as a surprise that metaphors with a source in 'Tides, waves and flooding' should be so prevalent in British English, given the importance of water in a country with a significant coastline. These metaphors are also very long-standing: for example, Potter (1988) has studied the use of *wylm* ('that which wells up or surges') and *weallan* ('to well') in the Old English epic poem *Beowulf*, where it provides the source of a pervasive emotion metaphor:

> It is, in fact, establishment of a correspondence between the elemental matter of water or fire and the throbs of the human heart, in waxing and in waning, which makes the *wylm-weallan* word-complex not only generally poetic and imagistically creative, but precisely metaphoric, a sophisticated poet's handle on an immemorial knowledge. (Potter 1988: 192)

In cases where 'Tides, waves and flooding' is the source, the respective target categories can be seen to fall into groups, drawing especially on (1) light (in the period covered by MME), (2) movement, and (3) emotion, and we can also pick out a further loose grouping (4) of a large number of connections with miscellaneous abstract categories. The examples below

demonstrate this, giving a sample of target categories in MME, MMOE or both, and instances of the metaphorically-related lexical items that connect the categories, together, where necessary, with a brief explanation of the sense as it relates to each target category.

(1) *Light*

1J25 Light	(MME) *streaming*; *streamy*; *flood* [movement of light is like movement of water]
1J27 Illumination	(MME) *stream* [as above]
1J30 Artificial light	(MME) *flood* [as above]

(2) *Movement and position in space*

1L06 Relative position	(MMOE) *flowan*; (MME) *conflux*; *gulf* [movement of people]
1N02 Types of movement	(MME) *maelstrom* [irregular movement like intensive movement of water]
1N03 Progressive movement	(MME) *flow*; *stream*; *surging*; *surge* [forward movement like water]
1N06 Movement in a specific direction	(MMOE) *weallan*; (MME) *flow*<*flowan*; *flood-gate* [copious emission of things like movement of water]
3J01 Travel and journeys	(MMOE) *flowan*; (MME) *flow*; *flowing*; *flux*; *flood*; *conflux*; *stream*; *afflux*; *float* [movement of vehicles like movement of water]
3J02 Transport	(MME) *surge*; *tidal* [movement of traffic like movement of water]

(3) *Emotion*

2D01 Emotion	(MME) *flow*; *pour*; *wave*; *tidal wave* [exhibit emotion; rising of emotion like movement of water]

2D02 Strong emotion/lack of emotion	(MMOE) *wielm*; *weallende* [violent emotion like rolling movement of waves][5]
2D03 Excitement	(MMOE) *afysed* [excited, from agitated water] (MME) *set afloat*; *ground-swell*; *wave*; *splash* [nervous excitement from movement of water][6]
2D06 Emotional suffering	(MME) *flood-gate* [source of tears]
2D07 Anger	(MMOE) *wielm* [strong emotion from movement of waves]

(4) *Miscellaneous abstract concepts*

1B26 Death	MME) *billow* [death as overwhelming flood]
(1K01Existence and its attributes	(MMOE) *upspring* [come into existence]; (MME) *insurge* [come into existence]; *current*; *stream* [in line with prevailing tendency]
1K04 Causation	(MME) *backwash* [unintended consequences]
1O07 Completion	(MME) *seventh wave* [culminating act]
1O11 Difficulty	(MME) *flood-gate* [that which can prevent]; *vortex* [something from which it is difficult to be extracted]; *cross-current* [an opposing force]

[5] Kövecses (2010: 108) notes that EMOTION IS A NATURAL FORCE is a typical conceptual metaphor that characterises emotion, and that HAPPINESS IS A NATURAL FORCE is a more specific related metaphor. For more on the close metaphorical links between intense emotion and water, see also Omori (2008) who finds "the most plausible source concept for characterizing the uncontrollability of emotion is A HUGE MASS OF MOVING WATER IN THE NATURAL WORLD" (2008: 137). Omori also suggests that the concept is significant across cultures and confirms that Japanese speakers also "have a wealth of words for expressing emotion in terms of A HUGE MASS OF MOVING WATER" (2008: 139).

[6] For a more detailed discussion of metaphors of Excitement, see Anderson (2016).

1O16 Prosperity and success	(MME) *overflow* [go beyond bounds]; *high-water mark* [highest point of something]; *tide-wave*; *tidal-wave* [significant advancement or progress is like movement of water *en masse*]
1O18 Adversity	(MMOE) (*ge*)*drefan* [to afflict, from movement of agitated sea]; (MME) *ebbing*; *ebb* [fall from prosperous condition]; *low-water-mark* [lowest point of something]
1P29 Sufficiency and abundance	(MMOE) *walling* < *weallende*; *wall with* < *weallan*; *flow* < *flowan*; (MME) *flood*; *affluent* [abundance of water transferred to general abundance]
3D05 Authority, rebellion and freedom	(MME) *insurge* [rise in revolt]; *with the stream* [submissive]; *cataclysm* [political unrest, from the Great Flood or deluge]; *undam* [free from restraint]

In almost all cases, it is the characteristic movement of water, with more or less force, that is at the heart of the metaphorical connection. This is true also of the metaphors that have their source in the category 'Rivers and streams'. Metaphors with their source in other water categories in the Metaphor Map, such as 'Body of water' and 'Sea', however, highlight other features of water, such as the vast size of a body of water, its depth or shallowness, or its freshness or stagnancy. Taken together, the evidence of the water categories support Mittlefehldt's statement that "it is this multifaceted quality of water that makes it so appropriate as a metaphor for our creative selves" (2003: 139).

4.2 'Tides, waves and flooding' as target

'Tides, waves and flooding' is much less commonly the target of metaphor than the source. Nevertheless, there are various ways in which the concepts contained in this category are expressed in metaphorical terms. In MME, lexical items from 23 other categories have been found to instantiate

a metaphorical connection in which 'Tides, waves and flooding' is the target, and in MMOE, the corresponding figure is 10.

As with the target categories in the section above, so too here the source categories can be seen to fall into clear groupings of related concepts. These draw especially on (5) living beings and their physical characteristics, (6) abstract qualities of humans, and (7) physical characteristics of material things. The examples below show a sample of source categories, again accompanied by one or two of the metaphorically related lexical items connecting the categories, and a brief explanation of the sense as it relates to moving water.

(5) *Physical characteristics of living beings*

1B01 Life	(MMOE) *quick* < *cwic*; (MME) *living*; *lively* [moving water is a living thing]
1B11 Body parts	(MME) *head*; *shoulder* [parts of a wave]
1B12 Skin	(MME) *wrinkled*; *dimple* [the 'texture' of water]
1E04 Animal bodies	(MME) *crest* [break with foam]
1E15 Horses and elephants	(MME) *horse* [to carry away be flooding]; *sea-horse* [a white-crested wave]
1I02 Sleep	(MME) *lull* [make water calm]

(6) *Abstract qualities of humans*

2D07 Anger	(MMOE) *wood* < *wod*; (MME) *rage* [moving water as human emotion]
3A10 Social discord and harmony	(MME) *troublesome*; *brawling* [moving water behaves like an uncontrolled group of humans]

(7) *Physical characteristics of material things*

1A05 Landscape, high and low land	(MMOE) *ridge* < *hrycg*; (MME) *valley* [crests and troughs in the sea]
1H01 Textiles	(MME) *frizado* [a woollen fabric, applied here to water flowing with agitated movement, producing appearance of coarse woollen cloth]

1H02 Clothing	(MME) *wimpling*; *wimple* [a stream meanders like moving fabric of a veil]; *hood* [cap of foam]
1J11 Softness	(MME) *softness*; *soft* [moving water has the appearance of a smooth texture]
1K03 Weight, heat and cold	(MME) *scalding*; *boiling* [an expanse of moving water with the 'texture' of boiling water in a pan]
3K07 Materials and fuel	(MME) *glassy*; *hyaline* [moving water with the texture of glass]

The characteristic movement of water is again key to the understanding of the metaphorical connections between the categories in question here. Water as it moves in the form of tides and waves is expressed in human terms, as conveying anger and discord. But texture and shape are also very important here, as we make sense of water in mass as bearing resemblance to other – perhaps more everyday – objects and materials.

4.3 Apparently bidirectional metaphors

Looking separately at 'Tides, waves and flooding' as source and target as we have done above, however, obscures part of the picture, as can be glimpsed from some of the examples, including 'Anger'. In some of the metaphorical category connections in the Metaphor Maps, the category of 'Tides, waves and flooding' is *both* source and target. That is, some of the category connections are bidirectional. It must be remembered, however, that we are talking here about metaphorical connections between semantic categories in the Metaphor Maps (and ultimately in the *Historical Thesaurus*), each of which contains many individual concepts, and *not* necessarily conceptual metaphors in the Lakoffian sense.

In MME, there are eight category connections in which 'Tides, waves and flooding' is both source and target, and in MMOE there is only one. The former can be illustrated by the connection between 'Tides, waves and flooding' and 'Landscape, high and low land'. Here, lexical items like *wave*, *wavy* and *surging* have their source in 'Tides, waves and flooding' and have transferred senses likening elements of the landscape to moving water, and *valley*, with the transferred sense of the trough of a wave, has its source in 'Landscape, high and low land'. In MMOE, a similar situation is found with

'Strong emotion and lack of emotion', in which *wood* (from OE *wod* 'mad') transfers to the conceptualisation of the roughness of waves, and *wielm* and *weallende* show transfer from the rolling movement of waves to violent emotion. This last example is perhaps closest to what we would think of as a truly bidirectional metaphor, in which the two sources are conceptually very close to each other and so are the two targets.

We also see category connections in which 'Tides, waves and flooding' is the source in one of the Metaphor Maps but the target in the other, that is, the directionality changes from the Old English period to the subsequent period. The category of 'Disadvantage and harm' illustrates this. In MMOE, 'Tides, waves and flooding' is the target of metaphor, instantiated by *wiþersæc* extending its meaning from a sense of unfavourableness to the collision or flowing together of waters. In MME, on the other hand, there is evidence of 'Tides, waves and flooding' as the source, illustrated by *sea-mark* (a marker of a high tide, transferred to a warning sign more generally) and *undam* (to release water, transferring to a general sense of depriving something of protection). The opposite happens with 'Social discord and harmony': in MMOE, 'Tides, waves and flooding' is the source, with the concept of contention or strife borrowing from the surging, heaving movement of waves (*gewealc*), and in MME, *troublesome* and *brawling* show transfer from social discord to the agitated flow of water. Illustrating the same shift from MMOE to MME, we can consider also 'Clothing'. In MMOE, 'Tides, waves and flooding' is the source, with the transfer being from a rolling wave (*wealca*) to a billowing flimsy fabric. In the long period represented by MME, 'Tides, waves and flooding' shows up as the target, with two clothing words transferred at different times to the rippling, meandering of water (*wimple*, *wimpling*), and to a cap of sea-foam (*hood*).

5. Conclusion

'Tides, waves and flooding' is just one of over four hundred categories in the Metaphor Map, but it gives insights into the complex interrelationships between semantic categories, and the ripples of metaphor development over time. In particular, it illustrates that while some metaphorical connections are very long-standing, traceable back to the Old English period, new connections emerge up to the present, and will no doubt continue to emerge in the future. These metaphorical connections are instantiated by lexical evidence which itself varies over time, with new words emerging to express existing connections. Considering 'Tides, waves and flooding' as

the source of metaphor, a core of words in the semantic category can be seen to reliably carry metaphorical meaning across many semantic categories and across time. This core includes words like *flood*, *flow* (and related forms such as *flowing*, *overflow*, *flux*, etc.), *stream*, *surge*, *tide* and *wave.* This is supported by a wider periphery of lexemes used metaphorically but appearing as connections between fewer pairs of semantic categories. These lexemes include *flood-gate*, *high-water mark* and *vortex*, all of which occur several times as instantiating a metaphorical connection between a pair of categories in the Metaphor Maps, and a long tail of words that have been picked out only once by the Mapping Metaphor coders as instantiating a metaphorical category connection, including *bow-wave*, *land-flood*, *sea-froth*, *tenth wave*, *tide-mark*, *voraginous*, *water-wall*, and *whirlpool*, relating to more specific concepts. These last examples relate to much more specific concepts than *flow*, *stream*, *wave* and their like: these may perhaps make for more vivid and effective metaphors, but further research would be needed to explore this.

Looking beyond 'Tides, waves and flooding', there is scope for much more detailed analysis than the Mapping Metaphor team has been able to carry out so far, both of further individual categories or set of related categories, and of how metaphor has entered and survives in the language system of English as a whole. More generally, as demonstrated by the Mapping Metaphor with the *Historical Thesaurus* project, the digital age and developments in historical lexicography bring exciting new possibilities for the study of metaphor. They are enabling us to explore metaphor at a much larger scale than before, and they allow us to look at metaphor in new ways, for example where digital resources are interconnected like the *Historical Thesaurus of English* and the Metaphor Map of English. This may be the key to future responses to some of the thorniest questions in metaphor theory, beyond the scope of this article, such as to what extent the notion of dead metaphor is meaningful, whether it is meaningful to speak of bidirectional metaphor, or whether we can more precisely define what we mean by a semantic domain by exploiting the categorisation work of a major reference thesaurus. The digital age appears to have opened the flood-gates for research into metaphor.

REFERENCES

Sources

Bosworth, J. – T.N. Toller (eds.)
1898 *An Anglo-Saxon Dictionary.* London: Oxford University Press.

Clark Hall, J.R. – H.D. Merritt
1960 *A Concise Anglo-Saxon Dictionary* (4th edn.). London: Cambridge University Press.
Historical Thesaurus of English
http://historicalthesaurus.arts.gla.ac.uk, accessed September 2019
Kay, C. et al. (eds.)
2009 *Historical Thesaurus of the Oxford English Dictionary*. Oxford: Oxford University Press.
Mapping Metaphor with the *Historical Thesaurus*
http://mappingmetaphor.arts.gla.ac.uk, accessed September 2019
Oxford English Dictionary (*OED* online)
http://www.oed.com, accessed September 2019
Roberts, J. – C. Kay – L. Grundy
2000 *A Thesaurus of Old English* (2nd end.) Amsterdam: Rodopi.

Special studies

Anderson, W.
2016 "Waves of excitement, waves of metaphor". In: W. Anderson – E. Bramwell – C. Hough (eds.) *Mapping English Metaphor through Time*. Oxford: Oxford University Press, 115-136.
Caballero, R. – I. Ibarretxe-Antuñano
2013 "Ways of perceiving, moving, and thinking: Revindicating culture in conceptual metaphor research", *Journal of Cognitive Semiotics* 5 (1-2): Special Issue *Conceptual Metaphor Theory: Thirty Years After*. Edited by R. Fusaroli – S. Morgagni, 268-290.
Dancygier, B. – E. Sweetser
2014 *Figurative Language*. Cambridge: Cambridge University Press.
Kövecses, Z.
2010 *Metaphor: A Practical Introduction* (2nd edn.). Oxford: Oxford University Press.
Lakoff, G. – M. Johnson
1980 *Metaphors We Live By*. Chicago: University of Chicago Press.
Mittlefehldt, P.J.
2003 "Writing the waves, sounding the depths", *Interdisciplinary Studies in Literature and Environment* 10 (1), 137-142.
Omori, A.
2008 "Emotion as a huge mass of moving water", *Metaphor and Symbol* 23, 130-146.
Potter, J.
1988 "'Wylm' and 'weallan' in *Beowulf*: A tidal metaphor", *Medieval Perspectives* 3, 191-199.
Pragglejaz Group
2007 "MIP: a method for identifying metaphorically used words in discourse", *Metaphor and Symbol* 22 (1), 1-39.

Steen, G.J. et al.
2010 *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. Amsterdam: John Benjamins.
Taylor, J.R.
1995 *Linguistic Categorization* (2nd edn.). Oxford: Oxford University Press.

Address: WENDY ANDERSON, English Language & Linguistics, University of Glasgow, 12 University Gardens, Glasgow, G12 8DW, Scotland, UK.
ORCID code: orcid.org/0000-0003-3188-4876.

Fiction as a source of linguistic data: Evidence from television drama

James M. Stratton
Purdue University

ABSTRACT

Growing evidence suggests that the language used in fictional television can be a fair representation of contemporary language use and changes within the linguistic system. To explore this relationship further, the present study uses variationist quantitative methods to examine the composite system of intensifiers, as well as adjectives of strangeness, in the British fictional TV series *Misfits*. Results indicate that the distribution and constraints of the two variable systems are similar in both scripted and non-scripted language. With respect to intensifiers, amplifiers were more frequent than downtoners, younger speakers used intensifiers more frequently than older speakers, and women had higher intensification rates than men. Predicative adjectives were also intensified more frequently than attributive adjectives, and the top three boosters were *so*, *really*, and *very*. As for adjectives of strangeness, consistent with findings from vernacular speech, the adjective *weird* made up over 70 percent of the semantic field, and was favored predominantly by younger speakers. The present study therefore provides empirical support for the use of naturalistic fictional language as a proxy for studying language variation and change.

Keywords: scripted language, variationist sociolinguistics, British English, intensifiers, adjective variation.

1. Introduction

Linguistic variability is ubiquitous and can be observed across different domains of grammar. To describe a proposition as 'great', speakers of English have a variety of functionally equivalent adjectives at their disposal: *great*, *fantastic*, *cool*, *brilliant*, *class*, *ace*, *topnotch*, *epic*, *sick*, *boss*, *wicked* and *sound*.

DOI: 10.25951/4357.

Recent studies have shown that the decision to use one adjective over a synonymously equivalent counterpart is systematically conditioned and constrained by various linguistic and social factors (Tagliamonte – Pabst 2020; Stratton forthcoming[a]). These same adjectives can also be intensified by a range of functionally equivalent intensifier variants: *very*, *really*, *so*, *dead*, *right*, *proper*, *bloody* and *well*.[1] Like with adjectives, intensifiers are also conditioned by linguistic and social constraints (Peters 1994; Ito – Tagliamonte 2003; Tagliamonte – Roberts 2005; Tagliamonte 2008; Méndez-Naya – Pahta 2010; D'Arcy 2015; Stratton 2018, 2020a, 2020b, 2020c). Weinreich et al. (1968: 100) famously referred to this structured systematicity as "orderly heterogeneity", which forms the major theoretical assumption of variationist sociolinguistics.

After examining the constraints on the intensifier system in York (Ito – Tagliamonte, 2003), Tagliamonte – Roberts (2005) turned to the TV series *Friends* to examine this system in North America. Among several findings, *so* was found to be an incoming variant, a change which, in line with the general principles of language change (Labov 1990: 210-215), was spearheaded by women. However, a broader implication of the study was the emerging view that the analysis of linguistic phenomena through a fictional lens is a viable and fruitful research methodology for tapping into language variation and change. Since then, several authors have followed in their footsteps and have examined different variable domains using datasets from scripted television (Heyd 2010; Reichelt – Durham 2017; Stange 2017; Stratton 2018).

While variationist methods have been used to examine the intensifier system in fictional language (Tagliamonte – Roberts 2005; Reichelt – Durham 2017; Stratton 2018), to the best of my knowledge, no studies have run regression models on fictional data, which have included both fixed and mixed/random effects.[2] Since the failure to include random intercepts can have a significant effect on the outcome of an analysis, it is important to re-evaluate these previous findings with accountable statistical rigor. Moreover, beyond the analysis of intensifiers, to date, no studies have examined whether semantic fields, such as adjectives of strangeness

[1] Following variationist terminology, 'variants' refers to two or more ways of saying the same thing (Labov 1972: 188).

[2] The term 'mixed effects' refers to the combination of both 'fixed effects' and 'random effects'. In a model in which 'speaker' is run as a random/mixed effect, each individual is treated as a source of random variability. Failure to include 'speaker' as a mixed effect can be problematic when more than one observation (i.e., data point) is collected from each speaker.

(Tagliamonte – Brooke 2014), have similar distributions and constraints in fictional language as unscripted vernacular speech.

Therefore, to test the assumption that "media language actually does reflect what is going on in language" (Tagliamonte – Roberts 2005: 296), the present study uses variationist quantitative methods to examine the intensifier system and adjectives of strangeness in the British fictional TV series *Misfits*. Two research questions were formulated. First, what is the distribution of the intensifier system in the *Misfits Corpus* in terms of frequency and function, and is intensifier use constrained by similar conditioning factors as in real-life vernacular speech? Second, what is the distribution of adjectives of strangeness in the *Misfits Corpus*, and are these constrained by similar conditioning factors as in vernacular speech? By comparing the distribution and constraints on these two linguistic variables in fictional language and real-life vernacular speech, the present study tests the claim that language observable in fictional television is a fair representation of contemporary language use and language change.

2. Previous literature

2.1 Intensifiers

Intensifiers are devices which scale a quality upward or downward from an assumed norm (Bolinger 1972: 17; Quirk et al. 1985: 589-590). Since they are generally thought to be "a vehicle for impressing, praising, persuading, insulting, and generally influencing the listener's reception of the message" (Partington 1993: 178), they are subject to constant renewal, recycling, and replacement (Tagliamonte 2008; Stratton 2020a). Quirk et al. (1985: 590) divide intensifiers into "amplifiers" and "downtoners". Amplifiers scale upward from an assumed norm, as in *that's very good*, whereas downtoners scale downward from an assumed norm, as in *that's kinda good*. Amplifiers can also be further sub-divided into "boosters" and "maximizers" according to the degree of amplification. Boosters "denote a high degree on a scale" whereas maximizers "denote the upper extreme point" on a scale. In terms of distribution, previous studies on English have found that amplifiers are more frequent than downtoners, and more specifically, boosters are more frequent than maximizers (e.g., Peters 1994; D'Arcy 2015). Crosslinguistic evidence also seems to suggest that this type of distribution appears to be consistent across languages (e.g., Stratton 2020d).

A cornerstone of a variationist sociolinguistic analysis is the consideration of both internal and external factors of variation. Previous studies on the intensification of adjectives have found that highly developed and frequently used intensifiers appear predominantly in predicative position, as in, *it was really interesting*, as opposed to attributive position, as in *a really interesting book* (e.g., Ito – Tagliamonte 2003; Tagliamonte – Roberts 2005; Tagliamonte 2008). Moreover, based on Dixon's classification of adjectives (1977: 31), intensifier frequency has also been found to correlate with an increased number of intensified semantic classes (e.g., Partington 1993: 183; Ito – Tagliamonte 2003: 268; Stratton forthcoming[b]). In contrast, outgoing and less frequently used variants collocate with a fewer number of semantic classes and occur with a limited number of unique heads (e.g., Stratton 2020a: 219-221).[3]

As for the external conditioning factors, even outside of the field of variationist sociolinguistics, it is easy to recognize that language use can correlate with the external factor geography. For instance, in England, to intensify an adjective, an adolescent speaker might use *proper*, as in *that film was proper boring* (Stratton 2020b), whereas in parts of Scotland, an adolescent speaker might prefer *pure*, as in *that film was pure boring* (Macaulay 2006). In contrast, using variants such as *hella* and *totes*, as in *that movie was hella/totes boring* indexes North American speech (Bucholtz et al. 2007). However, in addition to geography, other external factors have been found to correlate with intensifier use, such as socioeconomic status (Macaulay 1995), sex (Tagliamonte – Roberts 2005; Tagliamonte 2008; Fuchs 2017; Stratton 2020d), and age (Tagliamonte 2008; Palacios-Martínez – Núñez-Pertejo, 2012; Núñez-Pertejo – Palacios-Martínez 2018).

In general, women have been found to use intensifiers more frequently than men (Tagliamonte 2008; D'Arcy 2015; Fuchs 2017), which also seems true crosslinguistically (Stratton 2020d). While correlation should not be confused with causation, explanations usually fall into one of two schools of thought. On the one hand, it is thought that women might use intensifiers more frequently than men to make up for potential

[3] The original semantic classes, as defined by Dixon (1977: 31) were: value (*good*, *bad*), dimension (e.g., *big*, *small*), physical property (e.g., *hard*, *soft*), speed (e.g., *fast*, *slow*), human propensity (e.g., *happy*, *kind*), age (e.g., *young*, *old*) and color (e.g., *red*, *yellow*). However, Dixon has nuanced his seven-category distinction over time. In Dixon (2005: 484-485), eleven categories are delineated, where similarity (e.g., *similar*, *different*), volition (e.g., *deliberate*, *accidental*), difficulty (e.g., *easy*, *challenging*) and qualification (e.g., *appropriate*, *rational*) are added.

suppression in society (Lakoff 1975; Erikson et al. 1978), but on the other hand, it is also hypothesized that their highly frequent use may be a product of high sociability and expressivity (Carli 1990). Age has also been found to be a strong predictor of intensifier use, with younger speakers having higher intensification rates than older speakers (Ito – Tagliamonte 2003: 265; Barnfield – Buchstaller 2010: 261-262; Stratton 2020d: 207). Recent diachronic evidence also suggests that the propensity to intensify may have increased gradually over time (Stratton forthcoming[b]). Studies have also shown that adolescent speakers use intensifiers differently to adults (Tagliamonte 2008; Núñez-Pertejo – Palacios-Martínez 2018) partly because "teenagers are cliquish to the nth degree" (Tagliamonte 2016a: 3), so depending on the British variety, adolescents have their own in-group language.

2.2 Adjectives of strangeness

To describe a proposition as "strange", speakers of English have several variants to choose from: *strange*, *weird*, *unusual*, *eerie*, *peculiar*, *creepy*, *bizarre*, *odd*. In their analysis of adjectives of strangeness in Toronto English, Tagliamonte – Brooke (2014) found that *weird* made up 70 percent of the system. The distribution of these adjectives in apparent time indicated that *strange* is moving out of favor and is being replaced with *weird*. Similar distributions were also found in British English where *weird* appears to be increasing in use whereas variants such as *peculiar* and *strange* appear to be going out of favor.

In similar research, adjectives of positive evaluation have been found to correlate with social factors, such as sex and age (Tagliamonte – Pabst 2020; Stratton forthcoming[a]). In their study on Canadian English, Tagliamonte – Pabst (2020) found that adjective choices correlated with the age of the speaker, with *cool* favored predominantly by younger speakers, whereas other variants, such as *terrific*, were used predominantly by older speakers. In a recent study on German, similar results were found, with age but also sex correlating with adjective choices (Stratton forthcoming[a]). These studies therefore reveal that the choice to use one adjective over a synonymously equivalent one is subject to the same linguistic and social constraints as other variable phenomena. One of the goals of the present study is to determine whether the distribution of strangeness adjectives is similar to that observed in unscripted language (Tagliamonte – Brooke 2014), but also to examine whether variants such as *weird* are favored mostly by younger speakers.

2.3 Fictional language

While language in fictional television is scripted, "it is the scriptwriters' aim to give a recognisable, and at the same time, fair representation" of the variety being portrayed (Stratton 2018: 795). "Writers use linguistic features that are typical of naturally-occurring conversation to achieve realistic dialogues" (Baños 2013: 526) and if the language does not appear authentic, "viewer identification with the show characters can be negatively impacted, thus, potentially, affecting the success of the show" (Quaglio 2009: 13). In addition to the study by Tagliamonte – Roberts (2005), there is a large body of growing empirical evidence to support the claim that language in fictional television can often reflect the linguistic changes taking place in naturally-occurring speech at least with respect to form, frequency, and distribution (Quaglio 2009; Bednarek 2010; Reichelt – Durham 2017; Stratton 2018). Beyond synchronic studies, fictional dialogue has long been used as a proxy for studying diachronic change (Jucker 1995; Culpepper – Kytö 2010), and given that television has become an integral part of modern culture, its influence on language and society is inevitably profound (Marshall – Werndly 2002: 2). If fictional language is a fair representation of real-life language, based on previous literature, one should expect to find the following in scripted speech:

- Amplifiers are more frequent than downtoners
- Boosters are more frequent than maximizers
- Younger speakers have higher intensification rates than older speakers
- Women use intensifiers more frequently than men
- The three most frequently used intensifiers are *so*, *really*, and *very*
- The intensifier *very* is favored predominantly by older speakers
- Intensifiers collocate more widely with predicative adjectives than attributive adjectives
- Highly frequent and developed intensifiers intensify a higher number of semantic classes
- The adjective *weird* makes up over 2/3 of the system of strangeness adjectives
- Younger speakers prefer *weird* over other functionally equivalent adjectives

To test these hypotheses, the British TV series *Misfits* was used. *Misfits* is a contemporary fictional drama which centers around delinquent

adolescents who, for various reasons, are subject to mandatory community service in London. Over five seasons, a series of unfortunate events leads to the killing of four probation workers. Although the majority of the dialog comes from the young delinquents, there is also supporting dialog from some adult or elderly speakers.[4] The protagonists represent an "ethically and geographically mixed group" which corresponds "to a 'mixed' Great Britain" (Zotevska 2013: 6). For instance, Alisha is from London, Finn is from Liverpool, Kelly, Rudy and Alex are from northern England, Nathan is from Ireland, Simon speaks with an Estuary accent, and Curtis represents MLE 'multicultural London English'. While it is not entirely clear where the speakers Abbey and Jess are from, they also have southern British accents. These ten characters, in addition to the supporting cast ($n = 5$), make up the *Misfits Corpus*, which consists of ca. 110,000 words. The TV series aired in November 2009 and concluded its fifth and final series in December 2013. However, due to the limited time depth, time is not treated as a variable in the present study.

The language in *Misfits* represents layered variability in terms of the intensifiers used. Some examples from the corpus are reported in (1) and (2). In (1a-e), the utterances are the same other than for the chosen intensifier variant. Because the intensifiers are functionally equivalent in meaning, that is, they are all boosters, they are apt for a variationist sociolinguistic analysis. The examples show both inter- and intra-speaker variability. The intensifiers are also used to intensify a variety of adjectives, which can appear predicatively (2a-c), attributively (2d-f), or function as their own discourse unit (2g-h). In the present study, the latter type is referred to as bare adjective intensification.

(1) (a) It's **dead** weird (Kelly, s1-ep1 [12:36])
(b) It's **so** weird (Cancer Patient, s5-ep7 [15:58])
(c) It's **really** weird (Abbey, s5-ep4 [38:57])
(d It's **fucking** weird (Curtis, s3-ep3 [19:18]])
(e) It's **too** weird (Abbey's GF, s5-ep4 [22:20]])
(f) I feel **really** weird (Kelly, s1-ep1 [6:55])
(g) …did any of you feel like **dead** weird? (Kelly, s1-ep1 [20:35])

(2) (a) I'm going to get some chocolate because I'm **very** upset (Rudy, s4-ep-1 [13:38])

[4] In the present study, 'adolescent' refers to people aged 16-24.

(b) If you're **so** happy, then why were you at the support group? (Finn, s5-ep5 [33:22])
(c) I just met this girl who's **really** nice (Rudy, s4-ep7, [26:24])
(d) You're a **very** attractive young man (LPW, s4-ep5 [36:00])
(e) That were a **bloody** brilliant holiday! (Rudy, s5-ep4 [29:31])
(f) I have **very** important probation worker business to attend to (Rudy, s4-ep1 [6:39])
(g) …**so** predictable… (Nathan, s1-ep4 [17:28])
(h) …**very** bossy… (Rudy, s3-ep4 [13:37])

3. Methodology

A personalized *Misfits Corpus* was created by downloading the transcripts from *www.springfieldspringfield.co.uk*, a website which provides a compendium of popular TV and movie scripts. The show was then watched closely with the transcripts to check for missing data, errors, and to code each utterance by speaker. While intensifiers can intensify several parts of speech, following previous research (e.g., Ito – Tagliamonte 2003; Tagliamonte – Roberts 2005), the present study focused specifically on the intensification of adjectives. However, not all adjectives are part of the variable context since not all adjectives can be intensified. Therefore, circumscription of the variable context was necessary, which, following previous literature, meant the exclusion of negative, comparative, and superlative tokens because these are not functionally equivalent in meaning (e.g., Ito – Tagliamonte 2003).

What remained was a list of intensifiable adjectives, some of which were intensified, some of which were not. Looking to the left of each adjective for the presence of an intensifier revealed whether it had been intensified or not. The intensification, or lack thereof, was coded respectively (*yes* [1] – *no* [0]). Adjectives of strangeness were also coded respectively. Each intensifiable adjective was manually coded for the appropriate sociolinguistic metadata (speaker, sex and age) and linguistic factors (such as syntactic position). Each intensifier was also coded according to the taxonomy of Quirk et al. (1985: 590) so that comparisons among functionally equivalent variants could be made. Following previous variationist work (e.g., D'Arcy 2015), a logistic regression was run in *Rbrul* (Johnson 2009), with occurrence versus absence of intensification run as the application value. Since age was not a statistically significant factor in Tagliamonte – Brooke (2014), the model was run on intensification only.

4. Results

4.1 Intensifiers

There were 1596 intensifiable adjectives in the corpus. However, of these, 89 were removed because they came from cameo characters who contributed too few words to be included in any meaningful representative quantitative analysis. What remained were 1506 adjectives produced by 15 speakers, of which 556 were intensified (Table 1). The overall intensification rate of adjectives was therefore 37%, which, while on the high end, is in the range of what has been observed previously in English (Ito – Tagliamonte 2003; Tagliamonte – Roberts 2005; D'Arcy 2015; Tagliamonte 2016b; Stratton 2018). The intensification rate in apparent time is reported in Figure 1, which, like previous research (e.g., Ito – Tagliamonte 2003: 265; Barnfield – Buchstaller 2010: 261-262), shows a higher preference for intensification among younger speakers than older speakers.

Table 1. The overall distribution of intensification in Misfits

Total $N = 1507$			
Intensified		Not Intensified	
%	N	%	N
37	556	63	950

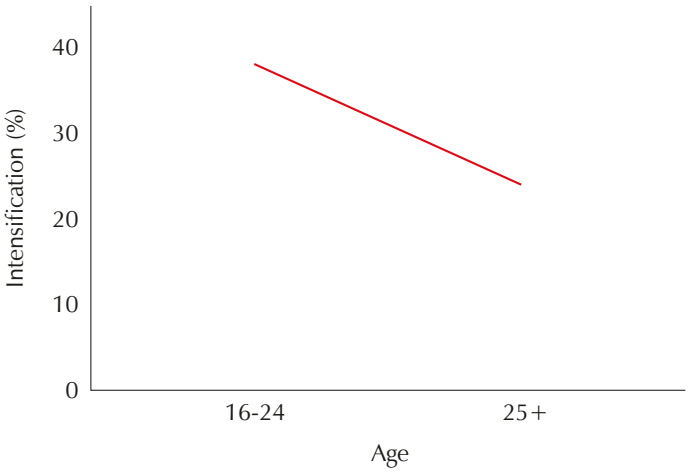

Figure 1. Intensification rate by age

As for differences in intensification among men and women, like in previous research (e.g., Tagliamonte 2008; Fuchs 2017), women intensified more adjectives than men. Women had an intensification rate of 52% (n = 216/418) whereas men had an intensification rate of 31% (n = 340/1088). The breakdown of the intensification rate by character, proportionally to the number of intensifiable adjectives each speaker produced, is reported in Table 2.

The 556 intensifiers are ranked by frequency in Table 3. When classified according to their scalar function, it becomes clear that, as found in previous studies (Peters 1994: 271; D'Arcy 2015: 460), amplifiers were more frequent than downtoners (Figure 2). Furthermore, within the subset of amplifiers, boosters (n = 434/459) were more frequent than maximizers (n = 25/459). As for the specific variants, as has been observed in studies on naturally occurring speech (e.g., Ito – Tagliamonte 2003; Tagliamonte 2008), the three most frequently used boosters were *so*, *really*, and *very*. The fourth most frequently used booster was *fucking*, which has also been reported as the most frequent intensifier (after *very*, *really*, and *so*) (Aijmer 2018: 75). Like in Aijmer (2018: 75), *fucking* was used almost more frequently by young women than young men.

Table 2. Intensification rate by character

Name	Sex	Total Adj	Intensified	Rate
Abbey	F	51	23	45%
Alisha	F	94	55	59%
Jess	F	88	43	49%
Kelly	F	83	44	53%
Sally	F	25	10	40%
Girl I	F	67	37	55%
Abbey's GF	F	10	4	40%
Alex	M	38	9	24%
Curtis	M	90	31	35%
Finn	M	146	47	32%
Nathan	M	234	76	33%
Last P-W4	M	29	5	17%
Seth	M	22	4	18%
Rudy	M	485	157	32%
Simon	M	44	11	25%
Total		1506	556	

In terms of the ranking by age, the most frequently used variant among young speakers was *really*, followed by *so*, followed by *very*. While these were also the three most frequently used variants among older speakers, older speakers preferred *very* over *really* and *so*. This ranking order is in line with previous work, where the use of *so* is led predominantly by younger speakers, whereas *very* is used more frequently by older speakers (e.g., Ito – Tagliamonte 2003; Tagliamonte – Roberts 2005; Tagliamonte 2008). The ranking of variants by sex also lined up with previous research. The variant *very* made up 31% of the male booster system but only 10% of the female booster system. In contrast, *so* made up 20% of the male system, and 26% of the female system. Therefore, consistent with previous work on intensifiers (e.g., Tagliamonte – Roberts 2005) the data show that women are leading the change toward using *so*, whereas men use variants which have been around for much longer (e.g., *very*). The distribution of *very* in apparent time also suggests that *very* is an outgoing variant whereas *so* is the favored variant. As for low frequency variants, descriptively speaking, *well* and *bloody* were used more frequently by men, whereas women preferred the booster *dead*.

Table 3. The frequency of adjective intensifiers in Misfits

Intensifier	N	%
so	93	17
really	92	17
very	91	16
fucking	43	8
a bit	33	6
pretty	31	6
too	26	5
all	26	5
bloody	14	3
quite	14	3
dead	11	2
totally	10	2
absolutely	8	1.5
Double intensifiers[1]	19	3
Other	80	5.5
Total	556	100

[1] 'Double intensifiers' are consecutive intensifiers. Examples from the dataset include: *that is so fucking romantic, a really fucking serious falling out,* and *I feel really fucking weird*.

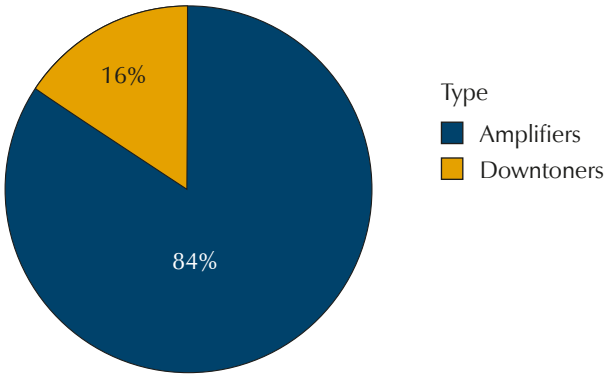

Figure 2. Frequency: Amplification vs. downtoning

(3) (a) Sorry, that was **really** unprofessional (Sally, s1-ep5 [4:12])
(b) Why do you think I was always **so** mean to him? (Nathan, s2-ep3 [18:00])
(c) They're being **all** coy and shy (Finn, season 5, ep. 5 [15:28])
(d) I was **fucking** ecstatic (Last Probation Worker, s4-ep5 [14:40])
(e) Well I can't drink lager, it's **too** gassy (Seth, s3-ep3 [22:04])
(f) Yeah, that's **dead** smart (Kelly, season 1, ep. 3 [26: 28])
(g) My aunt and uncle are **well** religious (Curtis, s1-ep5 [5:48])[5]
(h) I gave her a **right** good seeing to (Nathan, s1-ep2 [36:04])
(i) He's a rat-faced glue sniffer. Yeah, he's **proper** ratty (Kelly, s1-ep3 [6:20])

As for the syntactic distribution, predicative intensification was more common than attributive intensification, a finding consistent with previous work (e.g., Stratton 2020d; Stratton forthcoming[b]). As for the collocational width, of the semantic classes delineated by Dixon (2005: 484-485), human propensity, value, and physical property were intensified most frequently, which is also consistent with findings on naturally occurring speech (e.g., Ito – Tagliamonte 2003: 268-270; Méndez-Naya 2008: 44).

To examine whether the aforementioned linguistic (syntactic position, semantic classification) and social factors (sex, age) had a significant effect

[5] Other interesting examples of the intensifier *well* include: *I was well stressed* (s1-ep3 [38:08]), *I thought you two were well loved up* (s3-ep3 [18:45]) and *you're well butch* (s1-ep3 [31:29]). From the stress, it is clear *well* was functioning as an intensifier given that in its current intensifying use in British English *well* is always stressed (*OED*, *well*, adv. and *n*. IV16c.)

on the use of intensifiers, a mixed effects logistic regression model was run in *Rbrul* (Johnson 2009). Syntactic position had two levels [attributive, predicative] and semantic classification had five levels [human propensity, value, physical property, difficulty, dimension]. Bare intensification was not included as a level in the factor of syntactic position because there were only a few instances in the corpus. For semantic classification, only five semantic classes were included for the same reason. Sex had two levels [male, female] and age had two levels [16-24, 25+]. Intensification was run as the application value. The output of the model is reported in Table 4.

Table 4. Logistic regression of the factors conditioning intensification

Input .136 Total *N* 700	N	%	FW
Linguistic			
SYNTACTIC POSITION			
predicative	397	37.5	.74
attributive	303	.07	.26
Range			48
SEMANTIC CLASS			
human propensity	206	29.6	.63
value	264	25.0	.57
physical property	154	21.4	.52
difficulty	13	15.4	.42
qualification	63	12.7	.37
Range			26
Social			
SEX (.012)			
female	116	39.7	.59
male	584	21.2	.40
Range			19
AGE (.008)			
16-24	651	25.2	.65
25+	49	.12	.36
Range			29
Random Effect *Speaker*	*SD* = .3005 *n* = 15		

A significant effect for all four factors was found. As the factor weights (FW) indicate, women used intensifiers at a significantly higher frequency than men, younger speakers intensified significantly more adjectives than older speakers, predicative adjectives had a significantly higher intensification rate than attributive adjectives, and adjectives of human propensity, value, and physical property were intensified most frequently. The range for the factor syntactic position (48) indicates that syntactic position had the most significant effect, of the linguistic factors, on the use of intensifiers. The range for age (29) indicates that, of the social factors, age had the strongest effect on the use of intensifiers.

4.2 Adjectives of strangeness

Of the 1506 adjectives, 113 were adjectives of strangeness. Like in work on naturally occurring speech (Tagliamonte – Brooke 2014), this system was dominated predominantly by one variant, *weird*, which made up 77% of the system. Moreover, like in Tagliamonte – Brooke (2014), *strange* was the second most frequently used variant, where it made up 10% in *Misfits* and 14% in Toronto English. The distribution of variants from *Misfits* is reported in Table 5.

Table 5. The distribution of strangeness adjectives

Variants	Tokens	%
weird	81	77
strange	10	10
wrong	5	5
funny[1]	3	3
creepy	2	2
unusual	2	2
spooky	1	1
rapey	1	1
Total	105	100%

[1] Tagliamonte – Brooke (2014) removed *funny* from their analysis due its polysemy. However, the three tokens included here were unambiguous. An advantage of using fictional television as a source of linguistic data is that its meaning can often be inferred from the context, paralinguistic information (i.e. the character's facial feature), prosody/intonation, and general information about the storyline.

As for the distribution by age, adults used *strange* 50% of the time, versus a use of only 8% by younger speakers. In addition to using *weird* and *strange*, young speakers also made use of some low frequency variants such as *funny* and *creepy*. Like in Toronto English, men and women used *weird* at almost the same rate. Of the 81 intensifiable tokens of *weird*, 38 (47%) were intensified. Examples of use appear in (4).

(4) (a) I liked you which is really **weird** because you can be such a total dick (Jess)
(b) Wouldn't that be a bit **weird**? (Finn)
(c) I feel well **weird** (Kelly)
(d) You're like, proper **strange** (Girl I)
(e) People are **strange** mate (Rudy)
(f) What's with the **strange** lingering silence? (Nathan)
(e) Have you seen anything **unusual**? (Sally)
(f) The **creepy** coach guy…. (Alisha)
(g) but not in a bloody **creepy** way! (Rudy)

5. Discussion and concluding remarks

The present study set out to examine the validity of using fictional television as a reliable source of linguistic data for examining aspects of language variation and change. In comparing both the intensifier system and the system of strangeness adjectives in the *Misfits Corpus* with naturally occurring speech, few quantitative differences were found. All ten hypotheses outlined in Section 2.3 were supported. Amplifiers were more frequent than downtoners, boosters were more frequent than maximizers, younger speakers had significantly higher intensification rates than older speakers, and women used intensifiers at a significantly higher frequency than men. The three most frequently used boosters in British English are *so*, *really*, and *very*, which was also the case in the *Misfits Corpus*. The variant *fucking* has also reportedly become the fourth most frequently used variant among British adolescents (Aijmer 2018: 75), which was also true in the *Misfits*. Moreover, older speakers used *very* more frequently than *really* and *so*, predicative adjectives were intensified more frequently than attributive adjectives, and the most frequently used intensifiers collocated the highest number of semantic classes. The distribution of strangeness adjectives in the *Misfits*

Corpus also mirrored distributions in naturally occurring speech, with *weird* making up over 70% of the system.

Therefore, at least insofar as the *Misfits* is concerned, there is little reason to believe that the language used in fictional television is significantly different to real-life vernacular speech. Previous claims about the reliability of using fictional language as a proxy for examining language variation and change (e.g., Tagliamonte – Roberts 2005) are therefore supported. Although the present study used data from only one TV show, and different genres may have different reliability rates, the quantitative similarities between the *Misfits* and real-life speech are clear. It is unlikely that scriptwriters are consciously aware of current linguistic research regarding the correlation between intensifier use and, for instance, women, and age, yet these correlations bleed through into the scriptwriting. Thus, on the one hand, there is potentially empirical evidence to suggest that scriptwriting can manifest some intuitively correct and perceptible social assumptions about language use in characterization, even if these core intuitions are tacit and below the level of consciousness. However, on the other hand, one possible reason for the similarity between the language in contemporary fictional drama and naturally occurring language is the actors and scriptwriters are typically native speakers of the language. Therefore, even though the language is scripted, their portrayal of the language often brings about a reasonable representation of the language; unconscious features of the language are still transmitted and thus observable in the data.

There is, however, one obvious caveat. If, by way of contrast, linguistic features present in a contemporary re-enactment of a historical period (i.e., a historical drama like *Downton Abbey*) were compared with authentic data on the historical period, one might expect there to be a much lower reliability rate. The language of a naturalistic contemporary drama reflects, to a large extent, naturally occurring speech because the scriptwriters and actors are contemporary speakers of the language. However, this is clearly not the case for 20th and 21st century historical dramas. Some fictional sources have no reference point (e.g., *Game of Thrones*), while others might be based on a real historical period (e.g., *Downton Abbey*) but are portrayed by actors displaced in time, and often space. These genres aside, the present study provides empirical support for the use of naturalistic contemporary fictional dramas as a proxy for examining language variation and change; sources which can be particularly useful in the absence of corpora on real-life vernacular speech.

REFERENCES

Sources

Springfield! Springfield!
www.springfieldspringfield.co.uk, accessed December 2019

Special studies

Aijmer, K.
2018 "That's well bad: Some new intensifiers in spoken British English". In: V. Brezina – R. Love – K. Aijmer (eds.) *Corpus Approaches to Contemporary British Speech: Sociolinguistics Studies of the Spoken BNC2014*. London: Routledge, 60-95.

Baños, R.
2003 "That is so cool': Investigating the translation of adverbial intensifiers in English Spanish dubbing through a parallel corpus of sitcoms", *Perspectives: Studies in Translation Theory and Practice* 21, 526-542.

Barnfield, K. – I. Buchstaller
2010 "Intensifiers on Tyneside: Longitudinal developments and new trends", *English World Wide* 31 (3), 252-287.

Bednarek, M.
2010 *The Language of Fictional Television: Drama and Identity*. London: Bloomsbury.

Bolinger, D.
1972 *Degree Words*. The Hague, Paris: Mouton.

Bucholtz, M. et al.
2007 "Hella nor call or totally so call? The perceptual dialectology of California", *Journal of English Linguistics* 35 (4), 325-352.

Carli, L.
1990 "Gender, language, and influence", *Journal of Personality and Social Psychology* 59, 941-951.

Culpeper, J. – M. Kytö
2010 *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.

D'Arcy, A.
2015 "Stability, stasis and change", *Diachronica* 32 (4), 449-493.

Dixon, R.M.W.
1977 "Where have all the adjectives gone?", *Studies in Language* 1, 19-80.
2005 *A Semantic Approach to English Grammar* (2nd edn). Oxford: Oxford University Press.

Fuchs, R.
2017 "Do women (still) use more intensifiers than men? Recent changes in the sociolinguistics of intensifiers in British English", *International Journal of Corpus Linguistics* 22 (3), 345-374.

Heyd, T.
2010 "How you guys doin'? Staged orality and emerging plural address in the television series *Friends*", *American Speech* 85 (1), 33-66.
Ito, R. – S.A. Tagliamonte
2003 "Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers", *Language in Society* 32 (2), 257-279.
Johnson, D.E.
2009 "Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis", *Language and Linguistics Compass* 3 (1), 359-383.
Jucker, A.H. (ed.)
1995 *Historical pragmatics: Pragmatic Developments in the History of English*. Amsterdam, Philadelphia: John Benjamins.
Labov, W.
1972 *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
1990 "The intersection of sex and social class in the course of linguistic change", *Language Variation and Change* 2 (2), 205-254.
Lakoff, R.
1975 *Language and Women's Place*. New York: Harper and Row.
Macaulay, R.
1995 "The adverbs of authority", *English World-Wide* 16, 37-60.
2006 "Pure grammaticalization: The development of a teenage intensifier", *Language Variation and Change* 18 (3), 267-283.
Marshall, J. – A. Werndly
2002 *The Language of Television*. London: Routledge.
Méndez-Naya, B.
2008 „'The which is most and right harde to answere': Intensifying *right* and *most* in earlier English". In: R. Dury – M. Gotti – M. Dossena (eds.) *English Historical Linguistics 2006. Vol. II: Lexical and semantic change*. Amsterdam, Philadelphia: John Benjamins, 31-51.
Méndez-Naya, B. – P. Pahta
2010 "Intensifiers in competition: The picture from early English medical writing". In: I. Taavitsainen – P. Pahta (eds.) *Early Modern English Medical Texts: Corpus Description and Studies*. Amsterdam: John Benjamins, 191-214.
Núñez-Pertejo, P. – I. Palacios-Martínez
2018 "Intensifiers in MLE: New trends and developments", *Nordic Journal of English Studies* 17 (2), 116-155.
Palacios-Martínez, I. – P. Núñez-Pertejo
2012 "He's absolutely massive. It's a super day. Madonna, she is a wicked singer. Youth language and intensification: A corpus-based study", *Text and Talk* 32 (6), 773-796.

Partington, A.
1993 "Corpus evidence of language change: The case of intensifiers". In: M. Baker – G. Francis – E. Bonelli (eds.) *Text and Technology. In Honour of John Sinclair*. Amsterdam: John Benjamins, 177-192.
Peters, H.
1994 "Degree adverbs in early modern English". In: D. Kastovsky (ed.) *Studies in Early Modern English*. Berlin: Walter de Gruyter, 269-288.
Quaglio, P.
2009 *Television Dialogue: The Sitcom Friends vs. Natural Conversation*. Amsterdam: John Benjamins.
Quirk, R. et al.
1985 *A Comprehensive Grammar of the English Language*. London and New York: Longman.
Reichelt, S. – M. Durham
2017 "Adjective intensification as a means of characterization: Portraying in-group membership and Britishness in *Buffy the Vampire Slayer*", *Journal of English Linguistics* 45 (1), 60-87.
Stange, U.
2017 "'You're so not going to believe this': The use of GenX *so* in constructions with future *going to* in American English", *American Speech: A Quarterly ofLinguistic Usage* 92 (4), 487-524.
Stratton, J.M.
2018 "The use of the adjective intensifier *well* in British English: A case study of *The Inbetweeners*", *English Studies* 99 (8), 793-816.
2020a "A diachronic analysis of the adjective intensifier *well* from Early Modern English to Present Day English", *Canadian Journal of Linguistics* 65 (2), 216-245.
2020b "'That's proper cool': The emerging intensifier 'proper' in British English", *English Today,* 1-8.
2020c "Corpora and diachronic analysis of English". In: E. Friginal – J.A. Hardy (eds.) *The Routledge Handbook of Corpus Approaches to Discourse Analysis.* Abingdon, New York: Routledge, 202-218.
2020d "Adjective intensifiers in German", *Journal of Germanic Linguistics* 32 (2), 183-215.
(forthcoming[a]) "Tapping into German adjective variability: A variationist sociolinguistic approach", *Journal of Germanic Linguistics*.
(forthcoming[b]) "Old English intensifiers: The beginnings of the English intensifier system", *Journal of Historical Linguistics*.
Tagliamonte, S.A.
2008 "So different and pretty cool! Recycling intensifiers in Canadian English", Special issue of *English Language and Linguistics: Intensifiers* 12, 361-394.

2016a *Teen Talk: The Language of Adolescents*. Cambridge: Cambridge University Press.

2016b "So sick or so cool? The language of youth on the internet", *Language in Society* 45 (1), 1-32.

Tagliamonte, S.A. – J. Brooke

2014 "A weird (language) tale: Variation and change in the adjectives of strangeness", *American speech* 89 (1), 4-41.

Tagliamonte, S.A. – C. Roberts

2005 "So weird; so cool; so innovative: The use of intensifiers in the television series *Friends*", *American Speech* 80 (3), 280-300.

Tagliamonte, S.A. – K. Pabst

2020 "A cool comparison: Adjectives of positive evaluation in Toronto, Canada and York, England", *Journal of English Linguistics* 48 (1), 3-30.

Weinreich U. et al.

1968 *Empirical Foundations for a Theory of Language Change*. Austin: University of Texas Press.

Zotevska, E.

2013 *Representation of British Teenage Slang in the TV-series* Misfits*: A Communicative Approach to Teenage Slang*. Unpublished B.A. thesis, Gothenburg: University of Gothenburg.

Address: JAMES M. STRATTON, Department of Linguistics School of Languages and Cultures Purdue University 640 Oval Drive, West Lafayette IN, 47906, USA Purdue, e-mail: jstratt@purdue.edu
ORCID code: orcid.org/0000-0002-7255-5545.

From private to public: Letters contextualising the 1857-58 mutiny in the British press

Christina Samson

University of Florence

ABSTRACT

This study focuses on private letters which turned public by their publication in British newspapers and contextualised the 1857-1858 mutiny in India. In doing so, they provided first-hand information of the dramatic events being experienced, and shaped the news presented as well as the readers' opinions, thus foregrounding how letters can be considered a situated activity, in that they are written for a specific recipient and purpose. By adopting a corpus linguistic approach integrated with a qualitative interpretation, the study analyses the use of repeated keywords and their clusters' discursive environment in a small specialised corpus of letters written by men and women during the mutiny. The emerging data suggest the language used attempts to generate emotive and critical reactions in the readership while creating a sense of borderless community across the empire by connecting personal concerns to a wider sense of public engagement while legitimising and or delegitimizing the 1857-58 mutiny.

Keywords: letters, British press, mutiny, corpus linguistics, place names, sepoys, weapons.

1. Introduction

For centuries conflicts have been fought by military means on battlefields as well as over public opinion through the press (Taylor 1997; Thussu – Freedman 2003) where all forms of information are drawn in, either voluntarily or involuntarily, to drive public acceptance and support a country's actions. Conflicts have always been highly newsworthy, as they build an allure by

DOI: 10.25951/4358.

tempting readers to side with the good heroes against evil within appealing narratives that persuade the public to take a stand in the conflict (Nohrstedt 2009). As a result, the public's sympathies or antipathies depend on how it identifies with the different parties in conflict, and which strategy best gains attention and emotional engagement, as in the case of private letters turning public in the press during the 1857-58 mutiny[1].

The distinction between private and public communication was a blurred dichotomy until the 19th century, when a cultural shift took place and private goods and interests were no longer to be confused with public institutions (Del Lungo Camiciotti 2010). This was determined, on the one hand, by an increasing intervention of the state and its legal apparatus in social spaces typified by local communities; on the other hand, by the spread of silent reading which encouraged solitary meditation and individual piety (Del Lungo Camiciotti 2010).

Within the private/public dichotomy, letters represent a continuum. They are considered private when two parties are involved and reciprocal communication is not to be divulged, but in the case of private letters contextualising the 1857-58 mutiny in the press, the private communicative context undergoes a re-contextualisation. The events are discursively represented within social practices that regulate collective interaction in what can be termed a multi-layered context. The concept, drawing on Pahta and Taavitsainen (2010), involves both textual contexts and socio-historical conditions of text production with its societal, situational, historical, ideological and material sides including the writers' and readers' language attitudes and their social and situational context.

Within this perspective, this study analyses a corpus of private letters written during the 1857-1858 mutinies in India and published at the time in the press. By adopting a corpus driven methodology integrated with a qualitative interpretation, the repeated keywords and their recurring clusters' discursive environments are investigated as they contribute to contextualising the representation of the dramatic events in the newspapers.

The remainder of this paper is organised as follows: section two provides the historical background to the letters forming the corpus analysed in this study. Section three focuses on letter writing and letters in the press, section four describes the corpus and the methods adopted, whereas section five analyses the data. Section six concludes the paper.

1 The 1857-58 mutiny was actually characterised by several uprisings.

2. The context

The outbreaks of unrest among the Indian troops marked the beginning of a crisis which in imperial terms came to be known as the Indian or sepoy 'mutiny', or as the first 'national-popular imperialist war fought by Britain in its Empire (Dawson 1995), or, in nationalist terms, as the 'First War of Independence' (Blunt 2000). The causes of the uprising were and are still contested. For instance, Bhargava (1992) claims that imperial histories have tended to focus on the rumour that cartridges for new Enfield rifles had been greased with beef and pork fat. Having to bite into such cartridges before using them meant both Hindu and Muslim infantry soldiers known as sepoys were forced to break their religious faith. By contrast, most contemporary debates (Nielsen 2020; Dutta – Rao 2015; Major – Bates 2013, to mention a few) about the causes of the 'mutiny' focus on the organization of the Bengal army which was characterised by a widening distance between British officers and sepoys and the annexation of the province of Oudh in 1856. The year was characterised by intense growing disaffection among Indian infantry soldiers against the British East India Company which, till then, ruled on behalf of the British Crown.

In the following year, 1857, detachments of the Bengal army mutinied in the garrison town Meerut, 40 miles northeast of Delhi, killing several British officers and setting fire to the cantonment, before marching to Delhi and declaring the Mughal king, Bahadur Shah II, the reinstated ruler of Hindustan. Such actions have been considered consequential to the British deposing several noble Indians from their thrones without attracting significant support from the Indian population. By 1858, the revolts spread throughout central and northern India, taking place in Bengal by stretching across Delhi, Lucknow, and Cawnpore, as the rebels captured large tracts of the North-Western Provinces and Awadh (Oudh), where the 'mutiny' was also characterized by widespread agrarian unrest (Blunt 2000).

The key episode was the Cawnpore mutiny in which the East India Company forces and civilians were caught unprepared for an extended siege and were forced to surrender to the rebel forces under Nana Sahib, an aristocrat, in return for a safe passage to Allahabad. However, their evacuation from Cawnpore turned into a massacre on 27 June 1857 along the Ganges river as the 120 British women and children captured by the sepoys were killed in the Bibighar massacre. Their remains were thrown down a nearby well in an attempt to hide evidence, as the East India Company rescue force from Allahabad approached Cawnpore which was retaken in mid-July 1858.

In order to re-establish the British law, the Company forces engaged in widespread retaliation against the captured sepoys and local civilians who had supported the revolt (Blunt 2000).

During the several mutinies in 1857-58, letters written home by women and men were published in the press, delivering the public with first-hand information and personal perspectives on the dramatic events taking place.

3. Letter writing and private letters in the press

Letter writing is a very old practice and one of the most pervasive literate activities in society, in that it crosses formal and informal contexts, as shown by a vast variety of letters found in most domains of life. The earliest letters were closely linked to the bureaucratic needs of expanding empires (Ogborn 2008), but personal letters or familiar letters were not unknown even in ancient times. For long, they provided a means of expression for a vast range of social classes ranging from the elite to those outside the mainstream society while facilitating the development of states and empires, but also, it is suggested, helping to destroy them (Earle 2016; Goodman 2005).

Letter writing is anything but a static process. It can be viewed as activity rather than product (Nevalainen – Tanskanen 2007), a form of highly context-sensitive, personal and social interaction; in addition, the shift of focus onto letter writing as an activity shows the extent to which writers are the agents responsible for the outcome of the process (Nevalainen – Tanskanen 2007)[2]. Letters can be considered concrete, unmediated when private, historical artefacts which are strongly rooted in particular contexts and form the hidden underpinnings of much historical research. They can also be viewed as a social practice displaying the signs of a distinct environment in which they are embedded when they appear in the press.

The publication of letters in various text types goes back to the early 1620s in England, a time when the country was already playing a leading role in the development of the press in Europe, from the early newsletters to the first newsbooks (Ambrosi – Tessardo 1991). From the mid-nineteenth century, though, letters in the newspapers were clearly demarcated from reporting, as the rise of professional journalists provided news that could be distinguished from amateurs' contributions based on issues of personal concern. Letters to the editor are, thus, something distinct from the news,

[2] Cf. Samson 2020 for further details on letter writing.

given the ways in which they are produced, how those involved understand them, and how they present themselves to the readers (Nielsen 2010). Letters have, however, been directly connected to a wider sense of public engagement and have frequently been published when, for example, news took long to reach England from the colonies, acquiring therefore a referential-expressive function.

This has led us to view letters which appeared as personal communication in the press as actually being carefully crafted and curated products of editorial processes, or according to Wahl-Jorgensen (2002), construed through the co-creation of news workers and letter writers. As such, they turn from unmediated into highly mediated texts through journalistic routines, including those of editorial selectivity (Wahl-Jorgensen 2002; Gregory – Hutchins 2004). The latter privileged letters that related to content already on the news agenda – in this case the Indian Mutiny – and were therefore "moulded" to fit journalistic criteria, given their public significance (Brownlees et al. 2010) and their importance as gauges of public opinion and as the conduit for many editorially orchestrated campaigns (Conboy 2017).

Letters, in fact, can have more than a purpose. On the one hand, as in the case of the mutiny, they morally orient readers with the newspaper's editorial position, while serving as a space for a variety of opinions and revealing glimpses of how ordinary people made sense of major events and crises unfolding around them. Jones (1996) views letters as an essential means of feeling the national pulse and as a form of sensitive and complex political representation. This positioned readers' letters as a manifestation of pre-existing conversations already occurring elsewhere, a view shared by mid-Victorian editors. As Hampton (2004) argues, an ideal of politics by public discussion on the questions of the day permeated mid-Victorian elite society and the model of the rationally debating citizen was central in the creation of discourses of journalism in the early nineteenth century (Chalaby 1998; Wahl-Jorgensen 2007), leading to journalism of representation (Hampton 2004) with a strong social and political impact.

On the other hand, epistolary narratives helped to contextualise the mutinous events through reciprocal productions of place (Tuan 1991; Caquard 2011; Entrikin 1991; Herman 2001). The orientation elements connected writers and readers (Herman 2001), while helping to express human experience and lived space, since narrative itself can be a "spatially symbolic act" (Tally 2011) and (re)orientation can be managed and achieved in or through narratives.

Research on letters published in British newspapers in the 19th century has mainly focussed on the identification of letters of opinion and letters asking for advice in the Readers Letters page (Baczynski 1987), on the relationship between letters and the development of popular press (Bromley – Stephenson 1998), on the function of correspondence sections in the provincial press and their reader value (Jackson 1971), on the strong element of rituality in the topics approached in letters and their political influence (Tunstall 1977), or on the editorial choices of letters (McNair 2000), to mention a few. By contrast, there is a paucity of corpus linguistic analyses of letters published in the press specifically during the 1857-58 mutinies in India. This study attempts to fill the existing gap.

4. Corpus and methods

In order to analyse the linguistic features of letters written during the 1857-58 Indian mutinies, I developed a small specialised corpus – INMULE – of approximately 42,000 tokens. All the letters were downloaded from the British Newspaper Archive (www.britishnewspaperarchive.co.uk/) and were saved in txt format. Table 1 lists the British newspapers[3] from which the letters were downloaded.

[3] *The Morning Advertiser* was first published in 1794 by the London Society of Licensed Victuallers. It was devoted to trade interests, rather than to the support of a political party. Its circulation, however, fostered by the society, was, in mid-19th century second only to that of *The Times*. Founded in 1794 as *The Publican's Morning Advertiser*, it is the UK's oldest continuously produced paper. In 1858 the paper became the first newspaper to subscribe to Reuters' news service.

The Morning Post was a conservative daily newspaper published in London from 1772 to 1937. Initially a Whig paper, it was purchased by Daniel Stuart in 1795, who made it into a moderate Tory organ. A number of well-known writers contributed, including Samuel Taylor Coleridge, Charles Lamb, James Mackintosh, Robert Southey and William Wordsworth. In the seven years of Stuart's proprietorship, the paper's circulation rose from 350 to over 4,000. During the 1850s, the *Post* was very closely associated with the Palmerston ministry. The paper was noted for its attentions to the activities of the powerful and wealthy, its interest in foreign affairs, and in literary and artistic events.

The London Evening Standard was founded in 1827 as *The Standard* and it gained eminence for its detailed foreign news, contributing to a rise in circulation.

The Evening Mail was launched in 1823, it proved to be the longest lasting evening paper in Ireland. The paper was an instant success seeing its readership hit 2,500 in a month, making it, when few could read and the only people who bought papers were the gentry and aristocracy, the city's top seller. Its conservative readership ebbed and flowed during the century.

Table 1. 1857-58 British Newspapers

The Morning Advertiser
The Morning Post
The London Evening Standard
The Evening Mail
The Morning Chronicle
The Globe
The London Daily News

The letters appear to have been private, as they were addressed to relatives in Britain and were prevalently written by East India Company army officials and their wives who followed their husbands across British India in the various cantonments, that is, military stations wherein they lived. Other letters were by missionaries and other unspecified civilians.

The methodological approach I adopted in this study is a mixed one. It started with a corpus-driven analysis which commits the researcher to the integrity of the data as a whole, and the descriptions of language emerge from the corpus itself (Tognini-Bonelli 2001; Sinclair 2004). In this way, the centrality of the texts forming the corpus is pivotal, as findings are directly derived from the corpus and not filtered through existing concepts. Furthermore, I used Wordsmith Tools 7 (Scott 2016), a commercial software suite, to generate a word list which, to attain a key word list, I successively compared with a Corpus of Late Modern British English (CLMETEV) of 15 million tokens/words including various text genres such as personal letters, literary fiction, and scientific writing by men/women belonging to different social classes of 18th-19th century British society, ranging between 1710-1920. Key words derive from the comparison of the frequency of each word in the INMULE word-list with the frequency of the same word/s in the reference word-list. A word is considered key in a keyword list if it is

The Morning Chronicle was founded in 1769 by William Woodfall as publisher, editor, and reporter whose journalism slanted toward the Whig party in the House of Commons.

The Globe was a British newspaper that ran from 1803 to 1921. During the 1820s it supported radical politics, and was regarded as closely associated with Jeremy Bentham. By the 1840s it received briefings from within the Whig administration.

The London Daily News was founded in 1846 by Charles Dickens, who also served as the newspaper's first editor. It was conceived as a radical rival to the right-wing *Morning Chronicle*.

unusually frequent in comparison with what one would expect on the basis of the larger word-list of the reference corpus (Scott 2016).

I further investigated the recurring key words in their collocational patterns, that is, the tendency of words, or group of words, to occur more frequently in some environments than others (Hunston 2011). These phraseological arrangements or clusters are based on the assumption that words are not to be seen as elements in isolation that can be slotted into syntactic frameworks, but as forming larger units of meaning (Sinclair 1996; Römer 2010). Since the meaning of words lies in their use and use cannot exist in isolation, use can only be recognised and analysed contextually and functionally. I therefore see language in this study as the vector of continuous repetitions forming clusters which, in turn, form extended patterns of meaning. These mirror the specific situational context of the mutinies in 1857-58 India that make the language unique to the particular environment of Bengal. I then integrated the quantitative analysis with a qualitative interpretation of the recurring data to foreground the letter writers' representations of and personal perspectives on the events and those involved in them.

5. Data analysis

Wordsmith Tools 7 (Scott 2016) detected 236 keywords which correspond to the relevance of the different aspects of the topic during the mutinies. The comparatively most frequent key words are related to place and space as indicated in Table 2. In the Table, the first column shows the key word; the second shows its frequency in the source text(s) – INMULE; the third, the percentage of the frequency; the fourth indicates the number of texts in which it was present in INMULE; the fifth, its frequency in the reference corpus (the CLMETEV) and in the last column the *p* value, that is, the keyness value of the item under consideration.

Table 2. INMULE Key words – *Cities*

Key word	Freq.	%	Texts	RC. Freq.	P
Delhi	87	0,21	22	28	0,0000000000
Cawnpore	56	0,13	14	4	0,0000000000
India	84	0,20	26	1.287	0,0000000000

The relative high frequency of place names shows spatial and locational awareness in the accounts of the mutinies and are not surprising since, as Knopf (2014) claims, geographic understanding, awareness and communication are key factors in military activities because they create spaces, places, environments and landscapes with references to a distinct moral order (Woodward 2005). This emerges in the recurring use of place names in the corpus which foregrounds the need of the letter addressers to name places in order to label, identify and contextualise the Indian dramatic events. In addition, geographical names usually mean something that goes beyond a place's topography which tends to abstract and reduce the complexity of a topographic place to a single or a few fundamental traits representing irreplaceable cultural values of vital significance to people's sense of being (Andersson 1994; Helleland 2012). However, place names also function at an emotive, ideological community-creating level and an analysis of the key clusters in Table 2 significantly highlights the various meanings geographical territories can acquire in representing the mutinies in colonial India.

5.1 Delhi

I firstly applied the Concord programme of WST 7 to the key place name *Delhi* to access information about its collocations, an extract of which is shown in Table 3.

Table 3. Concordances – *Delhi*

N	Concordance
1	One officer at *Delhi* escaped with three shots through his hat.
2	my affectionate nephew, *Delhi* has been recaptured by our men.
3	the lodgings been filled up. *Delhi* not taken.
4	out of five regiments before *Delhi*, only 2,000 Europeans can be mustered.
5	infantry, who took the route to *Delhi.* They were attacked

Delhi frequently collocates with grammatical words (*of*, *to*, *at*, *from*, *in*, *toward*), stative verbs (*be*, *have*), and a few private verbs related to intellectual not observable states (*believe*, *suppose*, *know*, *fear*), but most importantly with action verbs (*take*, *march*, *force*, *assault*, *return*, *arrive*, *accompany*, *kill*, *move*, *fill*), as well as with common nouns referring to the military context (*shots*, *men*,

lodgings, *regiments*, *infantry*, *officer*, *route*, *king*, *news*, *letters*, *soldiers*, *telegraph*, *corps*, *weeks*, *accounts*, *troops*, *mutiny*, *insurrection*, *infantry*) and adjectives (*bad*, *few*, *long*, *immense*, *European*). The collocates suggest that among the main concerns of the letters in the press there was the exigency to communicate the succession of the military events but also to help the writers to reorient themselves in the disorienting experience of conflict while simultaneously helping readers to better comprehend actions in it.

Given that the meaning of concordance strings can emerge only if examined within a wider context, I considered them within their source text and analysed the recurring cluster patterns. The latter allow us to highlight the connotations which give sense to the phraseology of place names and common nouns in different situational contexts and to underline their uniqueness in INMULE.

For example, Delhi's recurring comparatively most frequent cluster *there is now* co-occurs repeatedly with + n/s as, for instance, in excerpt (1):

(1) *There is now* a large force, about 15,000 men, in Delhi, and all long to hear of its fall. Lord Lake took it originally with much less than half that number. Everything seems to be concentrated with that view.

The existential *there* clause has more than one function. It states the existence or occurrence of a large force of 15,000 men, which is the notional subject of the clause as well as the pointing of the writer to a new element linked to the cluster (*a large force of 15,000 men*) that refers to the rest of the discourse. The information in the existential clause also includes a time adverbial *now* and spatial reference *in Delhi* which enhances the connection of the information to the context of Delhi, thus anchoring every dialogue in INMULE which otherwise would appear to be a loose collection of disconnected utterances. Furthermore, the use of *There is now* recalls Herman (2001) who uses the concept of spatialisation to indicate "spatial reference" applicable to narrative and letter communication wherein the writer prompts the reader to relocate from the site of the communicative interaction, the letter, to the space-time location of the narration.

These features, however, verbally encode not only space but also place – Delhi – which Gieryn (2000) defines as having geographic location and imbued value. Letter communication, in particular, uses deliberate forms of deixis to refer to the writer, intended reader(s), time, and space, invoking the "two worlds" of sender and receiver and managing the spatio-temporal distance between them (Barton – Hall 2000). In excerpt (1) deixis has therefore

two functions: drawing attention to the detailed high number of forces previously employed to make Delhi fall – *much less than half that number* – and indicating the distance from the current time as well as from current reality or facts by frequently using the present tense. This foregrounds the immediacy of the first-hand detailed information, that is, the way military forces are deployed while wishing the enemy's defeat. Furthermore, the immediacy of the information provided is also characterised by the writer's personal perspective, which emerges in the basic evaluative sequence (evaluation + entity/process evaluated): *a large force*, *all long to hear of its fall, took it originally with much less than half that number*, *seems to be concentrated*.

Another equally frequent cluster of the key word Delhi is *King of Delhi* which collocates with action verbs (*storm*, *overtake*, *hang*, *kill*) as in excerpt (2) and features multiple deictic shifts leading the reader through a series of events – *Delhi has been stormed*, *the king of Delhi is hung*, *mutinous troops were killed*:

(2) Delhi has been stormed by the European and other troops, and fearful retribution had overtaken the mutineers. The *king of Delhi* is hung; and about eight or ten thousand of the mutinous troops were killed.

In excerpt (2), the narration, i.e. action recording sentences in sequence, is typified by temporal relativity (Werlich 1983), that is, the use of past and present constantly relative to the discursive present of letter communication which characterises the mix of written and oral discourse constantly used in such texts. These are loaded with a highly positive connotation referring to the English troops successfully winning back Delhi and supporting the rule of one collectivity over another, wherein the 'other'– *the mutineers*, *king of Delhi*, *mutinous troops* – is ruled over and demarcated according to the process of differentiation (Samson 2020). Within this perspective, adjectives (*fearful*, *mutinous)*, nouns (*retribution*) and verbs repeatedly encoding brutal actions also acquire a positive connotation for the public.

5.2 Cawnpore

The most frequent concordance lines of *Cawnpore*, the second comparatively most frequent key word of INMULE, are shown in Table 4. The relative high frequency of the place name indicates not only a place on the map of British India but also its linkage with a particular personal experiential and subjective meaning for the letter writers.

Table 4. Concordances – *Cawnpore*

N	Concordance
1	From *Cawnpore* nearly all the residents made their way
2	So far, the *Cawnpore* mutiny had none of those outrageous features
3	the rebels are in three engagements, occupying *Cawnpore,* and capturing guns.
4	Sir Hugh Wheeler has been killed at *Cawnpore.*
5	things have remained in town quo, except that *Cawnpore* has been reoccupied.

Cawnpore collocates with grammatical words (*of*, *to*, *at*, *from*) stative verbs (*have*, *be*,), private verbs (*believe*, *read*, *fear*), fictive motion verbs (*march*, *send*, *reoccupy*, *come*, *make*), common nouns (*Europeans*, *massacre*, *time*, *road*, *tragedy*, *troops*, *guns*, *place*, *mutiny*) and some adjectives (*all*, *far*). The emerging cluster *of the Cawnpore* co-occurs recurrently with nouns (*of the Cawnpore* + n).

In excerpt (3), the cluster anaphorically co-occurs with person markers (*I*, *we*, *you*) underlining the pervasiveness of subjectivity markers (Herring et al. 2004) encoding personal negative opinions being recurrently expressed (*within an ace of losing the empire*, *I do not yet see my way through the crisis*) on the dramatic situation created by the mutineers as well as by the inadequate preparation, reaction and number of the English forces at the time in India:

(3) We have been within an ace of losing the empire, and I do not yet see my way certainly through the crisis. Have you read *of the Cawnpore* tragedy?

While subjectivity contributes to the impressionistic views of the letter writers by pointing at interpersonal bonds which structure meaning, the repeated participation of *you* reinforces a continued relationship between the letter addressor and recipient, as the question *have you read of the Cawnpore tragedy?* in excerpt (3) indicates. Speaking to the recipient as if s/he were physically present is typical of language gap-closing (Fitzmaurice 2002) in letters wherein features of both orality and written language co-exist making them appear as a form of utterance which is not yet an unmediated conversation on paper.

This implies a shift of deictic centres, whereby the addressor prompts his/her interlocutor to relocate from the here and now of the act of narration to other space-time coordinates, namely those defining the perspective

from which the events are narrated. However, since the INMULE letters were published in the press, they represent a case of what Ambrosi and Tessardo (1991) term as a secondary orality. They argue that, like primary orality (conversation) letters, they generate a strong sense of belonging to a community, although the latter is much larger and less clearly defined than in primary orality and, specifically, *we have been within an ace of losing the empire* suggests the audience should be seen as part of the 'in-group' classification characterising colonialism (Samson forthcoming) wherein certain traits are invested with social significance and attributed to or claimed by those whose group identity is thereby constituted. In this way, a set of logical distinctions become homologous to a hierarchy of social distinctions which is not only a system of signification, but also a structure of domination, according to Kress and Hodge (1979). The letters thus embody socially shared assumptions and practices that allow a high number of readers to construct their ways of being in society.

The second most frequent cluster *Cawnpore has been reoccupied* underscores motion, which, as mentioned, is a significant component of military manoeuvres that typify the nature of military operations and conflict letters wherein narration is one of the principal means of building and communicating projective or viewer-relative locations (Ochs et al. 1992). The cluster repeatedly co-occurs also with cognitive verbs, for example *suppose* (*Cawnpore has been reoccupied* + v), which seem to function as a framework anticipating and encapsulating the evaluation taking place toward the end of the sequence, as in excerpt (4).

(4) *Cawnpore has been reoccupied* and I suppose this time Lucknow relieved by General Havelock.

Moreover, the framework – most typically, subjectivity marker + verb phrase (Bondi – Diani 2010) – performs the primary function of unequivocally signalling the source of the evaluation, that is, the writer who, in this case, takes the responsibility for the implicitly positive evaluation of General Havelock developed by the subsequent element of the sequence in *this time Lucknow relieved*. In this sense, the 'framework' meaning element may be regarded as a form of self-attribution of the opinion expressed on an action. In addition, the evaluation implicitly refers to the expression of identity which is connected with the towns – *Cawnpore*, *Lucknow* – that entail perceived differences between 'us' (English) and 'them' (mutineers) occupying the towns.

5.3 India

From the analysis of the third key word, *India*, a change in the perspective of the facts related to the mutinies emerges as from the concordance lines in Table 5:

Table 5. Concordances – *India*

N	Concordance
1	and other parts of North *India*. The presidency of Fort
2	the Hindoos in this part of *India*, and was also the scene of
3	enacting through North *India* have been appalling, and
4	the safest point in North *India*, and scores of European
5	supposed to know anything about *India* who were ignorant.

The collocates of India show again a high frequency of active verbs that, however, are unrelated to military actions (*send*, *put*, *remain*, *write*, *discuss*, *follow*), nouns that often refer to daily life (*letter/s*, *government*, *company*, *part*, *news*, *saving*, *regiment*, *time*, *troops*, *East*, *task*, *things*, *soldier*, *sister*, *brother*, *North*, *convent*) and adjectives (*special*, *European*) that are as generic as the adverbs (*very*, *recently*). There is a frequent use of person markers (*we*, *you*, *us*) signalling strong interaction between the addressor and his/her recipients. Furthermore, the top cluster *the East India* constantly co-occurs, not surprisingly, with a noun forming the proper name *the East India Company* as shown in excerpt (5) which foregrounds the prominent role held by the Company before and during the mutinies.

(5) *The East India Company* had treated him with insult and injustice.

In example (5) *the East India Company* is closely linked to its deplorable actions towards the Indians, specifically the aristocrats, Nana Sahib for instance, that eventually led to the beginning of the mutinies against the English. Moreover, the highly negative connotation of the Company is underscored by the use of nouns (*insult*, *injustice*) which indirectly evaluate negatively the Company men who had no close contact with the local Indian culture and did not know how to interact with them appropriately.

A further glimpse of the Bengal context is provided by excerpt (6) in which, once again, a negative evaluation of *the East India Company* can be inferred, by linking anaphorically the cluster to the decision of the English

government to exclude gentlemen connected with it from any service to be provided. As a matter of fact, the Company lost all its administrative powers following the Government of India Act of 1858, and its Indian possessions and armed forces were taken over by the Crown. Rule of India shifted therefore from the directors of the Company to a Secretary of State for India advised by a council, whose members were appointed by the Crown. The latter also directly appointed the governor-general, or viceroy, and provincial governors in India. The East India Company itself was formally dissolved by an Act of Parliament in 1874 and the British Raj, the direct imperial rule of India by the British state, began.

(6) The government adopted measures which should give them the services not only of their own military and civil officers, but of independent English gentlemen not connected with *the East India Company.*

Apart from place names, nouns are also linked to crucial features emerging in the INMULE keywords detected by Wordsmith Tools 7 (Scott 2016), as shown in Table 6.

Table 6. INMULE Key words – *Mutineers* – *Arms*

Key word	Freq.	%	Texts	RC. Freq.	P
Sepoys	48	0,11	16	11	0,0000000000
Guns	73	0,17	21	1.135	0,0000000000

The comparatively most frequent nouns *sepoys* and *guns* are strictly connected to the verbal representations and meaning of places in which the mutinies took place, as well as to insights into the human stories within the battles, the movements, and combatant's reflections on what occurred, what was done and why, thus creating personal geographies.

5.4 Sepoys

The comparatively most frequent concordance lines of *Sepoys* are listed in Table 7.

Apart from the grammatical words (*of*, *to*, *at*), stative verbs (*be*, *have*), action verbs (*come*, *guard*, *disarm*, *make*, *join*, *burn*, *leave*), cognitive verbs (*think*, *hope*, *pray*) and nouns (*regiment*, *army*, *night*, *morning*), *Sepoys* co-occurs also

with determiners (*some*, *several*), pronouns (*they*), adjectives (*faithful*, *native*, *their*, *our*) and adverbs (*after*, *when*, *now*) which differ from the previous key words.

Table 7. Concordances – *Sepoys*

N	Concordance
1	from which the bulk of our *Sepoys* come, and it is now
2	ort William by treacherous *Sepoys* guard at its gates.
3	there are either disarmed *Sepoys* or none at all.
4	mutinous army of some 30,000 *Sepoys*. Four entire regiments
5	artillery and ammunition, the *Sepoys* have made no stand against

This can be seen in excerpt (7) wherein the most frequent cluster *of the sepoys* collocates with prepositions (*of the sepoys* + p) which refer anaphorically to a cognition evaluation that is developed through a personal (*I*) mental process, as in example (4). The person marker anticipates and encapsulates the evaluation (*no treachery to fear*) that cataphorically refers to safety of the English in the area:

(7) I think we have no treachery to fear, either on the part *of the sepoys* or of the citizens of this place.

In excerpt (8), instead the cluster sequence *of the sepoys* + p becomes the entity evaluated (*were faithful*) by the addresser in his/her narration and it acquires a positive connotation by referring cataphorically to the adjective *faithful* and the active verbs *cut a hole*, *took her out*. In this way, a partially positive, realistic view of the mutiny is provided, as not all the sepoys were against the English:

(8) Some *of the Sepoys* in her husband's regiment were faithful and cut a hole in the wall and took her out.

5.5 Guns

The incompetence of the English officers involved in dealing with the mutineers is further foregrounded by the use of the last key word, *guns*. Its first most frequent concordance lines are provided in Table 8.

Table 8. Concordances – *Guns*

N	Concordance
1	Cawnpore, and capturing *guns.* The rencontres were very
2	cavalry, and two 6-pounder *guns* and a howitzer; with this
3	Contingent, who have lots of *guns* with them. It would, therefore
4	to make free of our heavy *guns.* I have no doubt the place
5	company European and three *guns* but the fellows will

The most repeated collocates of *gun* are verbs related to the handling of ammunitions (*take*, *place*, *load*, *fall*, *drive*), the stative verb (*be*), nouns linked to weapons and the forces using them (*howitzer*, *cavalry*, *company*, *force*, *ammunition*, *loss*, *rounds*, *entrentchment*, *road*), person markers (*they*, *we*) and adjectives (*European*, *big*, *great*, *immense*, *our*, *their*) which typify projective locations that are frequently found in descriptions or reference to weaponry used and linked to the location of the writer.

The top cluster *guns and the* co-occurs with noun/s (*guns and the* + n) and the use of the person marker *we* indicating his belonging to the English fighting the 'other', that is the mutineers. In addition, the addressor expresses an implicitly positive evaluation of the fact that not only the guns but also two remaining magazines were taken over from the rebels, a common feature in accounts of projective locations which overlap with elements of fieldcraft in military geography.

(9) During the day we took *guns and the* remaining two magazines.

In excerpt (10), though, a negative picture of the context is provided by the narrative sequence of action verbs (*killed*, *wounded*, *fell*). These co-occur anaphorically and cataphorically with the cluster sequence, encoding the disastrous outcome of the battling characterised by the loss of men as well as of light and heavy ammunition. However, all the military movements are not only related to the notions of territory that must be conquered but are also explicit in the potential for and reality of death (Knopf 2014):

(10) One other officer was killed, and several wounded. Three *guns and the* howitzer fell into the enemy's hands.

6. Concluding remarks

The analysed data confirm the context-sensitive, personal and public interaction characterising the INMULE letters in the press which offer a detailed first-hand succession of dramatic events the English were involved in during the 1857-58 Indian mutiny. The letters are not only typified by the immediacy of the information provided from personal perspectives but also by the constant explicit or implicit writers' evaluations of the incidences occurring in specific places which contribute to shape the readers' views and how the world around them could or should be perceived. This foregrounds how letters can be considered a situated activity, in that they are written for a specific recipient and purpose.

As to this point, the recurring use of place names in the texts highlights their different functions. They are used to engender emotive and critical reactions towards the mutinies as well as to push the reader beyond their actual geographical location, by creating a sense of borderless community across the empire.

This sense of community is further enhanced by the evaluations of the military operations and the usage of weapons which recurrently distinguish the English from the others, thus generating in-group convergence and out-group divergence, a social identification process in which an individual recognises him/herself as a member of a social group or a larger collectivity, this being a crucial feature during any conflict calling for political stance. These letters can, therefore, be considered an effective way to connect personal concerns to a wider sense of public engagement while legitimising and/or delegitimizing an issue as problematic as the 1857-58 mutiny.

REFERENCES

Sources

British Newspaper Archive
https://www.britishnewspaperarchive.co.uk/, accessed August 2020

Corpus of Late Modern English Texts (CLMETEV) Compiled by H. De Smet.
https://perswww.kuleuven.be/~u0044428/, accessed August 2020

UK Government
https://www.parliament.uk/about/living-heritage/evolutionofparliament/legislativescrutiny/parliament-and-empire/

parliament-and-the-american-colonies-before-1765/east-india-company-and-raj-1785-1858/), accessed August 2020

Special studies

Ambrosi, E. – M. Tessardo
1991 *Dalla parte del lettore: Ricerca sulla posta dei quotidiani italiani*. Roma: Edizioni lavoro.

Andersson, T.
1994 "The existential conditions of place-names [Ortnamnens existentiella villkor]". In: G. Ulfsparre (ed.) *Ortnamn värda att vårda. Föredrag från Riksantikvarieämbetets symposium Ortnamnskultur. Stockholm 5-7 maj 1993 [Place-Names Worthy of Caring for. Symposium Place-Name Culture organised by the National Heritage Board. Stockholm 5-7 May 1993]*. Stockholm: Riksantikvarieämbetet, 7-21.

Baczynski, R.
1987 *Discontent and Liberal Opinion: Non Partisan Readers' Letters to British Local Newspapers since the Late 1960's*. London: Metaballon Books.

Barton, D. – N. Hall
2000 "Introduction". In: D. Barton – N. Hall (eds.) *Letter Writing as a Social Practice*. Philadelphia: John Benjamins Publishings, 1-14.

Bhargava, M.L.
1992 *Saga of 1857: Success and Failures*. New Delhi: Reliance Publishing House.

Blunt, A.
2000 "Embodying war: British women and domestic defilement in the Indian mutiny 1857-1858", *Journal of Historical Geography* 26 (3), 403-428.

Bondi, M. – G. Diani
2015 "I am wild about cabbage: Evaluative 'semantic sequences' and cross-linguistic (dis)similarities", *Nordic Journal of English Studies* 14 (1), 116-151.

Bromley, M. – H. Stephenson (eds.)
1998 *Sex, Lies and Democracy: The Press and the Public*. London: Addison-Wesley Longman.

Brownlees, N. – G. Del Lungo Camiciotti – J. Denton (eds.)
2010 *The Language of Public and Private Communication in a Historical Perspective*. Newcastle u. T.: Cambridge Scholars Publishing.

Caquard, S.
2011 "Cartography I: Mapping narrative cartography", *Progress in Human Geography* 37 (1), 135-144.

Chalaby, J.
1998 *The Invention of Journalism*. London: Palgrave Macmillan.

Conboy, M.
2017 "How the war made the mirror", *Media History* 23 (3-4), 451-468.
Dawson, G.
1995 "The Imperial adventure hero and British masculinity: The imagining of Sir Henry Havelock". In: T. Foley et al. (eds.) *Gender and Colonialism*. Galway: Galway University Press, 46-59.
Del Lungo Camiciotti, G.
2010 "Introduction". In: N. Brownlees – G. Del Lungo Camiciotti – J. Denton (eds.) *The Language of Private and Public Communication in a Historical Perspective*. Newcastle u. T.: Cambridge Scholars Publishing, 1-23.
Dutta, S. – H. Rao
2015 "Infectious diseases, contamination rumors and ethnic violence: Regimental mutinies in the Bengal Native Army in 1857 India", *Organizational Behaviour and Human Decision Processes* 129, 36-47.
Earle R. (ed.)
2016 *Epistolary Selves: Letters and Letter-Writers, 1600-1945*. London: Routledge.
Entrikin, J.N.
1991 *The Betweenness of Place: Towards a Geography of Modernity*. Baltimore: The Johns Hopkins University Press.
Fitzmaurice, S.M.
2002 *The Familiar Letter in Early Modern English: A Pragmatic Approach*. Amsterdam: John Benjamins Publishing.
Gieryn, T.F.
2000 "A space for place in sociology", *Annual Review of Sociology* 26, 463-496.
Goodman, D.
2005 "Letter writing and the emergence of gendered subjectivity in eighteenth- century France", *Journal of Women's History* 17 (2), 9-37.
Gregory, L. – B. Hutchins
2004 "Everyday editorial practices and the public sphere: Analysing the letters to the editor page of a regional newspaper", *Media International Australia Incorporating Culture and Policy* 112 (1), 186-200.
Hampton, M.
2004 *Visions of the Press in Britain 1850-1950*. Urbana and Chicago: University of Illinois Press.
Helleland, B.
2012 "Place names and identities", *Oslo Studies in Language* 4 (2), 95-116.
Herman, D.
2001 "Spatial reference in narrative domains", *Text* 21 (4), 515-541.
Herring, S.C. et al.
2004 "Bridging the gap: A genre analysis of weblogs". In: *Proceedings of the 37th Hawaii International Conference on System Sciences* (*HICSS-37*). Los Alamitos, CA: IEEE, http://csdl.computer.org/comp/proceedings/hicss/2004/2056/04/205640101b.pdf

Hunston, S.
2011 "Doing analysis in discourse and corpus: The case of evaluative language". Plenary talk given at *CL 2011 – The Corpus Linguistics 2011 Conference*. Birmingham, 20-22 July 2011.
Jackson, I.
1971 *The Provincial Press and the Community*. Manchester: Manchester University Press.
Jones, A.
1996 *Powers of the Press: Newspapers, Power and the Public in Nineteenth Century England.* Aldershot: Ashgate.
Knopf, C.M.
2014 "Sense-making and map-making: War letters as personal geographies", *NANO: New American Notes Online 6*, https://nanocrit.com/issues/issue6/sense-making-map-making-war-letters-personal-geographies
Kress, G. – R. Hodge
1979 *Language as Ideology*. London: Routledge and Kegan Paul.
Major, A. – C. Bates (eds.)
2013 *Mutiny at the Margins: New Perspectives on the Indian Uprising of 1857: Britain and the Indian Uprising*. London: Sage Publications.
McNair, B.
2000 "Journalism and democracy: A millennial audit", *Journalism Studies* 1 (2), 197-211.
Nevalainen T. – S.K. Tanskanen (eds.)
2007 *Letter Writing*. Amsterdam: John Benjamins Publishing.
Nielsen, D.
2020 "Remembering the 1857 Indian Uprising in civic celebrations". In: K. Haldane Grenier – A.R. Mushal (eds.) *Cultures of Memory in the Nineteenth Century. Consuming Commemoration.* London: Palgrave Macmillan, 183-201.
Nohrstedt, S.A.
2009 "New war journalism", *Nordicom Review* 30 (1), 95-112.
Ochs, E. et al.
1992 "Storytelling as theory-building activity", *Discourse Processes* 15, 37-72.
Ogborn, M.
2008 *Indian Ink: Script and Print in the Making of the English East India Company*. Chicago: University of Chicago Press.
Pahta, P. – I. Taavitsainen
2010 "Scientific discourse". In: A.H. Jucker – I. Taavitsainen (eds.) *Historical Pragmatics*. Berlin: Mouton de Gruyter, 549-586.
Römer, U.
2010 "Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews", *English Text Construction* 3 (1), 95-119.

Samson, C.
2020 "'The whole of Bengal is in revolt' A corpus based analysis of letters from the 1857-58 mutinies in India", *Lingue e Linguaggi* 36, 283-296.
Scott, M.
2016 *WordSmith Tools*, version 7. Stroud: Lexical Analysis Software.
Sinclair, J.M.
1996 "The empty lexicon", *International Journal of Corpus Linguistics* 1 (1), 99-119.
2004 *Trust the Text. Language, Corpus, and Discourse*. London: Routledge.
Tally, R.T. Jr.
2011 "On literary cartography: Narrative as a spatially symbolic act", *NANO: New American Notes Online* 1, https://nanocrit.com/issues/issue1/literary-cartography-narrative-spatially-symbolic-act
Taylor, P.
1997 *Global Communications, International Affairs and the Media since 1945*. New York: Routledge.
Thussu, D.K. – D. Freedman (eds.)
2003 *War and the Media: Reporting Conflict 24/7*. London: Sage Publications.
Tognini-Bonelli, E.
2001 *Corpus Linguistics at Work*. Amsterdam: John Benjamins Publishing.
Tuan, Y.F.
1991 "Language and the making of place: A narrative-descriptive approach", *Annals of the Association of American Geographers* 81 (4), 684-696.
Tunstall, J.
1977 "Editorial sovereignty in the British press: Its past and present". In: *Studies on the Press*, Royal Commission on the Press 1974-7 Working Paper 3. London: HMSO.
Wahl-Jorgensen, K.
2002 "Understanding the conditions for public discourse: Four rules for selecting letters-to-the-editor", *Journalism Studies* 3 (1), 69-81.
2007 *Journalists and the Public: Newsroom Culture, Letters to the Editor and Democracy.* Cresskill, NJ: Hampton Press.
Werlich, E.
1983 *A Text Grammar of English.* Heidelberg: Quelle und Meyer.
Woodward, R.
2005 "From military geography to militarism's geographies: Disciplinary engagements with the geographies of militarism and military activities", *Progress in Human Geography* 29 (6), 718-740.

Address: CHRISTINA SAMSON, Department of Education, Languages, Intercultures, Literatures and Psychology, Via S. Reparata 93-95, Università degli Studi di Firenze, 50129 Firenze, Italy.
ORCID code: orcid.org/0000-0001-9216-7809.

Peaceful coexistence? Ideology in the representation of Scots and North American languages in Late Modern literature*

Marina Dossena

University of Bergamo

ABSTRACT

This study discusses some ideological traits in Late Modern English literary discourse concerning contact with other languages or socially- and geographically-marked varieties across the North Atlantic. Beyond 'dialect literature' and occurrences of 'literary dialect' (Shorrocks 1996), other very popular works greatly contributed to the definition of how readers perceived different languages and varieties in terms of relative prestige. In addition, popular culture also helped to disseminate evaluations of linguistic features. Representations in such texts were often ambivalent, ranging from humorous (to the point of caricature) to nostalgic, elegiac tones, but they always drew attention to the exotic, distant quality of the forms under discussion. Indeed, glossaries often present these features as witnesses of a quickly vanishing past, in need of preservation like ancient relics. In my analysis I will consider both different languages and varieties of the same language, in an attempt to show how their more or less explicit evaluation contributed to the creation of their (often persistent) image among readers on both sides of the Atlantic.

Keywords: Late Modern English, Scotland, America, literary discourse, ideology.

* A preliminary version of this text was presented as a plenary lecture at the international conference on "Linguistic Approaches to Dialects in English Literature" (LADEL) held in Salamanca (26-28 October 2017). I am grateful to the organizers, led by prof. María Fuencisla García-Bermejo Giner, the participants, and the anonymous reviewers for their insightful comments; the usual disclaimers apply.

DOI: 10.25951/4359.

1. Introduction

My title borrows a phrase ('peaceful coexistence') dating from Cold War times as a starting point for the study of some ideological traits in Late Modern English literary discourse pertaining to Scotland and North America. The main research question aims to assess to what extent contact with other languages or socially- and geographically-marked varieties in that register was indeed peaceful, i.e. descriptive of linguistic landscapes, or was more or less explicitly biased, and therefore a tool of historical and political argumentation. It is true that literary scholars have studied similar phenomena, especially in relation to post-colonial contexts – see for instance Sorensen (2000) and Brown (2018); however, their approach has typically addressed ideology in a broader sense, going beyond linguistic issues to address more overarching themes.

In addition, if we exclude Derrick McClure, whose work has consistently placed Scotland within the framework of European literary studies (e.g. in McClure 1995a and 2000), the scholars who have investigated Scottish literary discourse in Late Modern times – such as Broadhead (2013) and Sorensen (2017) – have normally done so from an Anglo-centric or at least a Continental perspective. In fact, an additional level of analysis may be offered: one in which Scotland and North America may be seen to be part of the same cultural framework, and indeed the ways in which the languages of one are represented may find an echo in the representation of the other. In this respect, the linguistic study of literary discourse may prove of great relevance.

Unlike in studies of earlier stages of the language, where the analysis of literary texts has normally been part of scholarly investigations, not least on account of the relative paucity of materials at hand or of their popularity and accessibility, when Late Modern English is concerned historical linguists have tended to focus instead on grammarians, orthoepists and lexicographers – i.e., on prototypical codifiers (e.g. Tieken-Boon van Ostade 2008). Of course there are notable exceptions, but present-day studies of the language of literary figures have normally taken into consideration non-literary works (e.g. Tieken-Boon van Ostade 2014) or analyses have mostly concerned stylistics. A valuable approach has recently been taken by the project on *Dialect in British Fiction 1800-1836* (Hodson – Broadhead – Millward 2014; see also Hodson – Broadhead 2013). However, its rather limited time span (expanded in Hodson 2017) only offers preliminary insights into the role played by significant literary figures in the endorsement or stigmatization of linguistic features at a time when also literary critics and book reviewers could

have a say in the assessment of language variation (Percy 2010). Moreover, labelling this kind of fiction as 'British' corresponds to a Late Modern view of presupposed unity and uniformity across the British Isles, but for this very reason it is ideological in itself, as it irons out any distinctions between English, Irish and Scottish literature, regardless of the literary tradition that different texts may continue or challenge.

Beyond so-called 'dialect literature' and uses of 'literary dialect' (Shorrocks 1996), the success of works authored by figures like James Fenimore Cooper and Robert Louis Stevenson greatly contributed to the definition of how readers perceived the relatively greater or lesser prestige of different languages and language varieties.[1] While these authors were not unique in their treatment of language variation, their attention to language as a special poetic and narrative element is well-documented (e.g., see Warner 1969, Blakemore 1984, Rosenwald 1998, Dossena 2005: 131-133, Shields 2009, Schachterle 2011, and Dossena 2012 and 2013) and may thus be taken as emblematic examples of how their interpretation helped shape attitudes to languages and varieties. Besides, the role of popular culture in the dissemination of specific views on language variation is hardly negligible: use of socially- and geographically-marked features in songs, ballads, dime novels and penny dreadfuls is a valuable object of investigation for the study of how language representation could be more or less ideologically charged.

Both in literary works and in popular culture, the coexistence of languages and varieties was often represented as ambiguous and problematic. In addition to stereotypical uses for humorous purposes, in which differences were emphasized to the point of caricature, as in the famous case of Sam Weller's Cockney speech in Charles Dickens's *Pickwick Papers* (1836-37), the supposed distance between varieties could be stressed to evoke exotic scenarios, such as in the representation of the so-called 'Mountain Men', fur trappers and traders in nineteenth-century North America, where social class distinctiveness was represented and indeed emphasized by the geographical distance of the context (see Hubbard 1968).

On the other hand, a certain fascination with different varieties underpins the compilation of numerous Late Modern English glossaries,

[1] Obviously, such figures cannot be mentioned without reference to other authors, whose impact on literary discourse and on perceptions of cultural links would be very significant both on the Continent and across the Atlantic. It is the case, for instance, of the extensive Ossianic tradition (see Moore 2017), but also of Walter Scott, as we will see below, and of Robert Burns, whose popularity both in the US and in Canada would contribute to the creation of a specific (and often idealized) identity in the Scottish diaspora (see Dossena 2012).

such as those collected in the *Salamanca Corpus of English Dialect Texts* (2011-), in which variation represents both social and diachronic distance. In the Positivistic agenda of the times, language standardization was expected to improve as progress advanced, but this improvement entailed the disappearance of varieties which therefore needed to be preserved like archeological artefacts, bearing witness to a distant (and often idealized) past – see Dossena (2005: 83-115). Within this framework, political ideologies could also be at work, whether it was to highlight distinctiveness or to emphasize linguistic contiguity as a metaphor of national unity.

In my contribution I intend to focus both on different languages and on varieties of the same language, in an attempt to show how their more or less explicit evaluation contributed to the construal of their image among readers on both sides of the Atlantic – an image which may have persisted through time thanks to the popularity of the texts in which it was framed.

After an overview of how language commentary was informed by both political bias and sociological ideologies as to what models should be followed (Section 2), in Sections 3 and 4 I will discuss some examples of the ways in which literary discourse has contributed to the stability or the demise of specific languages and to the representation of their speakers: Section 3 will focus on Scots, while Section 4 will make a brief comparison with the case of Native American languages.

In this context the issue of authenticity will appear to be relatively unimportant, as authors either attempted to imitate credible, recognizable usage, or exaggerated it for humorous purposes, but in any case it should be remembered that the texts were authored by fully-literate people: they were not monolingual speakers of the languages or varieties under discussion, whose access to the printed medium was minimal on a receptive level and non-existent on a productive level. The choice of a socially- and/or geographically-marked code was thus deliberate and therefore grounded in ideology, which would also make the notion of enregisterment (Agha 2003) an ideological construal. Finally, some concluding remarks will be offered in Section 5.

2. Ideology in language commentary

Twenty-first-century (non-expert) readers may be justified in assuming that academic texts are neutral, scientific representations of their objects of study, as great emphasis has often been placed on the value of detached discussions

in which any instance of personal involvement should be avoided. However, numerous analyses have shown that even in academic texts the authors' views and attitudes do feature in more or less explicit ways: see for instance Hunston – Thompson (2001), Martin – White (2005), Hunston (2011) and White (2007, 2015). Whether it is to criticize or support the views expressed by other scholars, or indeed to convey one's own ideas in more convincing tones, stance emerges at both the lexical and the syntactic level in all types of academic texts (see for instance Hyland– Bondi 2006, Hyland – Diani 2009, and Bondi 2015 and 2017). Unsurprisingly, the expression of stance is definitely more obvious in texts where the argumentative function is given prominence, but even supposedly more neutral descriptions, such as those of language phenomena, can be forcefully evaluative. One striking example is Otto Jespersen's (1905: 39) explanation of why few Celtic items were adopted into English (cited by Filppula – Klemola 2014: 35):

> There was nothing to induce the ruling classes to learn the language of the inferior natives; it could never be fashionable for them to show an acquaintance with that despised tongue by using now and then a Celtic word. On the other hand the Celt would have to learn the language of his masters, and learn it well; he could not think of addressing his superiors in his own unintelligible gibberish, and if the first generation did not learn good English, the second or third would, while the influence they themselves exercised on English would be infinitesimal. (Jespersen 1905: 39, in Filppula – Klemola 2014: 35)

The sociolinguistic features of the historical context under discussion are presented in terms that are certainly unacceptable by twenty-first-century standards, but they use the strongly evaluative vocabulary of Late Modern linguistic commentary, where the social and often political bias of individual authors could be expressed quite directly. Similar remarks had been made for centuries in descriptions of Scots and Scots Gaelic, both of which could be described as barbarous and indeed as damaging for society in general. Only a few years after the Union of the Crowns, in 1609, the Statutes of Iona imposed schooling in the Lowlands for at least the eldest child, whether male or female, so as to eradicate "ignorance and incivility":[2]

[2] In fact, Lowlanders had associated Gaelic with "incivilitie" already in the sixteenth century, as shown in often hyperbolical terms in the *Flyting of Dunbar and Kennedy* (1508).

(1) it being undirstand that the ignorance and incivilitie of the saidis Iles hes daylie incressit be the negligence of guid educatioun and instructioun of the youth in the knowledge of God and good letters […] it is inactit that every gentilman or yeaman within the said Ilandis, or any of thame, haveing childerine maill or femell, and being in goodis worth thriescore ky, sall put at the leist their eldest sone, or haveing no children maill thair eldest dochter, to the scullis in the Lowland, and interneny and bring thame up thair quhill that may be found able sufficientlie to speik, reid and wryte Inglische.

(Register of the Privy Council 1609 Vol. IX, 28-29, in Innes 1993)

A few years later, in 1616, an Act of the Privy Council established parish schools both to promote religious education and to eradicate use of Gaelic:

(2) Forsameikle as the Kingis Majestie having a speciall care and regaird that the trew religioun be advancit and establisheit in all the pairtis of this kingdome and that all his Majesties subjectis especiallie the youth, be exercised and trayned up in civilitie, godliness, knawledge, and learning, that the vulgar Inglishe toung be universallie plantit, and the Irische language, whilk is one of the cheif and principall causes of the continewance of barbarite and incivilitie amongis the inhabitantis of the Ilis and Heylandis, may be abolishit and removeit; and quhair as thair is no measure more powerfull to further his Majesties princlie regaird and purpois that the establisheing of Scooles in the particular parroches of this Kingdom whair the youthe may be taught at least to write and reid, and be catechised and instructed in the groundis of religioun.

(Register of the Privy Council 1616 Vol. X, 671-672, in Innes 1993)

It was only at the onset of Late Modern times that a different image of Scotland and its languages began to emerge, in which the antiquity of such languages was emphasized, thus granting them the prestige that a rampantly anglicizing trend had denied them. Also in such cases, however, ideological bias was often the driving force behind what elements were foregrounded: already in 1677 the religious pamphlet *Ravillac Redivivus* (1678: 77)[3] included an annotation on the proximity between Scots, Northern English and older Saxon forms. The text, probably best known for being the one in which the first occurrence of *Scotticism* is found, referred to William Lisle's *Saxon*

[3] Published anonymously, it is attributed to George Hickes (1642-1715).

Monuments, published forty years before, but the lineage of English and Scots would be a matter of debate throughout the Late Modern period.

The antiquity of Scots was also emphasised in James Adams's *Vindication of the Scottish Dialect* (1799), a fiercely anti-French text which actually recommended a list of Scots lexical items for adoption into English (Dossena 2005: 85-90) when codifiers all over Britain had been strenuously attempting to eradicate Scotticisms. For Adams, Scots preserved "the Saxon original in spite of the attempts of the Norman invaders and tyrants who endeavoured totally to extirpate its antient form" (1799: 148). Indeed, George Chalmers made the same point citing Lisle, whose work had now been proving influential for almost two centuries, in his Introduction to the *Poetical Works of Sir David Lyndsay*:

(3) Lisle, the Saxon scholar, says, in the Preface to his *Ancient Monuments of the Saxon Tongue*, that he improved more, in the knowledge of Saxon, by the perusal of Gawain Douglas's Virgil, than by that of all the Old English he could find, poetry, or prose; because it was nearer the Saxon, and further from the Norman. (Chalmers 1806: 146fn)

The lively debate on the authenticity of Macpherson's Ossian poems influenced the controversy on the Celtic or Germanic roots of Scots. However, Scottish Teutonism was also tinted with political overtones: for instance, in 1722 Edmund Gibson dedicated the revised edition of his translation of William Camden's *Britannia* to George I in emphatically loyalist terms. As only a few years had passed since the Jacobite rebellion of 1715, "the terrible Storm" to which Gibson refers at the beginning of his text, he stressed the mutual Saxon origins of the House of Hanover and of Britain's language, law, customs, names and place names. In practice, he was indirectly stating that the exclusion of the Stuart line from succession was not only acceptable, but could also be justified on historical grounds:

(4) When we consider the terrible Storm which threaten'd these Protestant Kingdoms a few years ago, […]; We cannot but adore the Wisdom and Goodness of God, in laying such a Train of Providences, for our Deliverance in that Hour of Extremity. […].

It is this Alliance which has made *Us* happy in your Majesty and your Royal Family, and which entitles *You* to the Love of every Subject, as a Prince of our own Blood; […]. But the ensuing Work points out a Relation between your Majesty and these Kingdoms, of

> a far more Ancient Date. Not only our Histories, but our Language, our Laws, our Customs, our Names of Persons and Names of Places, do all abundantly testify, that the greatest part of your Majesty's Subjects here, are of Saxon Original. And if we enquire from whence our Saxon Ancestors came, we shall find, that it was from your Majesty's Dominions in *Germany*, where their Brethren who staid [sic] behind, spread themselves through a noble and spacious Country, which still retains their Name. So that the main Body of your People in both Nations, are really descended from one and the same common Stock; and now, after a Disunion of so many Ages, they live again under the Protection and Influence of the same common Parent.
>
> (Gibson 1722: To the King)

Also in the case of Samuel Johnson Teutonism was linked to anti-French bias, but it did not contradict interest in both Gaelic and Older Scots; not only did Johnson subscribe to William Shaw's *Galic and English and English and Galic Dictionary* (1780), but he also encouraged James Boswell to compile a dictionary of Scots, despite Boswell's ambivalent attitude to that language. Boswell did start this project in the 1760s, but he never completed it,[4] like he never went beyond his 'Proposals for a periodical paper in the Scots dialect', to be called *The Sutiman* (see Pottle – Abbott – Pottle 1993: 106). However, it is important to see how two figures that are often associated with the heyday of linguistic prescriptivism also paid attention to a more varied linguistic landscape on account of their antiquarian interest.

3. "For the sake of auld lang syne": Languages and varieties in literary discourse

The search for 'pure Saxon' persisted throughout the nineteenth century (Dury 1992): in 1888, when Charles Mackay published *A Dictionary of Lowland Scotch with an introductory chapter on the poetry, humour, and literary history of the Scottish language and an appendix of Scottish Proverbs*, the closer connection of Scots with older 'Anglo-Teutonic' vocabulary was emphasized, stressing that such vocabulary was obsolete in English, but still fully comprehensible in Scotland (1888: xii).

[4] The manuscript, believed to be lost, was rediscovered in the Bodleian Library in 2011 (see Rennie 2011 and 2012).

This kind of approach thus draws attention to an interesting ambivalence in attitudes to Scots that would underpin both literary and critical writing. Far from being the 'vulgar, barbarous jargon' that prescriptivists sought to eradicate, Scots could be an icon of patriotism and/or sentimentality, in what McClure (1995b: 57) has called the 'Pinkerton syndrome', referring to a comment that John Pinkerton made in 1786 in the preface to a selection of poetry from the Maitland MSS and where he wrote:

(5) none can more sincerely wish a total extinction of the Scotish *colloquial* dialect than I do, for there are few *modern* Scoticisms which are not barbarisms… Yet, I believe, no man of either kingdom would wish an extinction of the Scotish dialect in poetry. (Pinkerton 1786: I, xvii)

Pinkerton stressed the separation between ordinary and literary discourse that had been the object of metalinguistic comments since Early Modern times, when George Puttenham had given very clear indications as to what models poets should follow, not only in terms of geographical variety, but also in terms of sociological specificity:

(6) [Use of language] in our maker or Poet must be heedyly looked vnto, that it be naturall, pure, and the most vsuall of all his countrey: and for the same purpose rather that which is spoken in the kings Court, or in the good townes and Cities within the land, then in the marches and frontiers, or in port townes, where straungers haunt for traffike sake, or yet in Vniuersities where Scholers vse much peeuish affectation of words out of the primatiue languages, or finally, in any vplandish village or corner of a Realme, where is no resort but of poore rusticall or vnciuill people: neither shall he follow the speach of a craftes man or carter, or other of the inferiour sort, though he be inhabitant or bred in the best town and Citie in this Realme, for such persons do abuse good speaches by strange accents or illshapen soundes, and false ortographie. But he shall follow generally the better brought vp sort, […].
(Puttenham 1589: Book 3, ch. 2)

Of course it would be an anachronism to call Puttenham an 'Ur-prescriptivist': his book was explicitly addressed to courtiers, hence the kind of guidance that he provided ought to be seen in the framework of numerous other texts which aimed to educate users in the best manners, not least in relation

to language use. Among these, books aiming to form Renaissance princes paid great attention to the role of rhetoric, eloquence, and knowledge of literary texts written both in classical and in vernacular languages – see for instance the works of Stefano Guazzo and Baldassarre Castiglione, in which the concepts of *politesse* first begin to be outlined (see Culpeper 2017 and Paternoster – Saltamacchia 2017).

The ideological role of antiquity in literary representation changed again when the Pre-romantic movement took a different approach from what had been customary in previous decades. Classical models in literary expression and in architecture began to be seen as sublime ruins in both domains – worthy of preservation, no doubt, but also irretrievably distant, possibly on the verge of erasure, and therefore to be collected before it was too late. Dictionaries and glossaries became the linguistic counterpart of *Wunderkammern*; this was especially true in cases where folk lore and traditions had to be saved for posterity before progress and the resulting increase in linguistic and cultural uniformity could wipe them out – see the examples below, recorded in the *Salamanca Corpus* among numerous other instances:

(7) Harland, John.
1865. *Ballads and Songs of Lancashire: Chiefly Older than the 19th Century*. London: Whittaker & Co. 1865. EDD.
1867. *Lancashire Folk-lore: Illustrative of the Superstitious Beliefs and Practises, Local Customs and Usages of the People of the County Palatine*. London: Warne. EDD.
1875. *Ballads and Songs of Lancashire, ancient and modern, collected, Compiled and edited with notes*. Second edition corrected, revised and enlarged by T.T. Wilkinson, F. R. A. S. London: George Routledge and Sons and L.C. Gent. SC. EDD.

Even John Jamieson, in his *Etymological Dictionary of the Scottish Language*, aimed to illustrate "national rites, customs and institutions in their analogy to those of other nations" (1808: titlepage), thus making his dictionary a valuable resource for both linguists and ethnologists. In addition to glossaries of local traditions, folk lore, farming, flora and fauna, numerous collections of proverbs were also published throughout Late Modern times, in the presupposition that they offered insights into both local culture and the antiquity of certain forms (see Dossena 2000).

In other cases, small, inexpensive booklets were published with an openly entertaining purpose, offering dialogues between 'rustic' characters in order to amuse readers who did not necessarily hail from completely different parts of the country, but whose linguistic competence could be flattered through a representation of 'less-than-educated' usage – see for instance the following texts, again from the *Salamanca Corpus*, in which phonetic spelling is meant to evoke the kind of usage that readers should expect:

(8) Robison, Joseph Barlow (c. 1820-1883)
Owd Sammy Twitcher's Visit tu't Gret Exibishun e Derby Roat, Kompoazed an Hillustrated by a Darbysher Mon. (1870)
Owd Sammy Twitcher's Crismas Bowk for the year 1870. Full a fun, tales, etc. (1870)

Audiences could be entertained with the thought they could 'see [themselves] as others [saw them]', while feeling outside (and above) the circle of those very same speakers. William Donaldson's studies of popular literature in Victorian Scotland (Donaldson 1986 and 1989) have shown the relevance of such texts which featured very prominently in local newspapers and magazines, providing both humorous, often satirical, commentary and helping to preserve uses which could be recognizable, though probably somewhat inauthentic, like all constructed instances of spoken usage.

On the other hand, the literary production of Robert Fergusson first and of Robert Burns later is exemplary of how Scots could be employed to convey meanings that were not just pastoral (like in Allan Ramsay's *Gentle Shepherd*) or humorous. Their use of Scots in poems like *Auld Reekie* and *A Man's a Man* is perfectly within a literary tradition that could have continued along independent lines if it had not been for events that changed Scotland's political, cultural and linguistic horizon: the Union of the Crowns in 1603 and of Parliaments in 1707.

Literary critics shared this ambivalent attitude: while they could praise Scottish literary texts, they could have reservations about the language in which they were written. Robert Burns, for example, gained fame as a supposedly simple ploughman with an extraordinary literary talent, but whose language was claimed to be hardly accessible to English readers:

(9) One bar, indeed, his birth and education have opposed to his fame – the language in which most of his poems are written. Even in Scotland, the provincial dialect which Ramsay and he have used is now read

> with a difficulty which greatly damps the pleasure of the reader: in England it cannot be read at all, without such a constant reference to a glossary, as nearly to destroy that pleasure.
>
> (Mackenzie 1786 [1897]: 278)

As a matter of fact, several articles emphasized the specificity and difficulty of Burns's language; however, the critics' comments reflected their own ideological bias, as Burns is known to have monitored his linguistic choices quite closely, using Scots, English or Scottish English in different contexts and trying to avoid "the peculiarities of Scottish phraseology" (Currie 1846: li). Indeed, Burns appears to have been aware of the dangers of the 'over-scotticization' of Scottish poetry, something that might result in a stereotypical and oversimplified shortbread-tin image of Scotland.

Unfortunately, that is exactly what happened in the nineteenth century, when the defeat of the last Jacobite rebellion was gradually construed as a heroic-but-doomed effort, and a series of often mythical dichotomies were established. Although 'the Forty-five' had actually been a rather poorly orchestrated enterprise which ended with Charles Edward Stuart's escape back to France, it was romanticized in countless songs, poems, ballads, novels and even paintings – and linguistic considerations also played a part in the construction of that image. The myth of an opposition between Jacobite, Gaelic- or Scots-speaking Highlanders on one side, and Hanoverian, English-speaking Lowlanders and Englishmen on the other, was a convenient but hardly accurate representation, since Hanoverians were both in the Highlands and in the Lowlands, and English had been taught in the Highlands since at least the seventeenth century, albeit with varying levels of success.

While Jacobite poems and songs did get written in Gaelic,[5] especially in the nineteenth century novels began to represent Jacobite characters as speakers of Scots, thus giving an extraordinary boost to a certain representation of Scotland that is still with us today. Nor were such representations exempt from anachronistic features – in Stevenson's *Kidnapped*, for instance, one of the protagonists whistles a few bars of a well-known Jacobite tune, 'Charlie's my darling', despite the fact that the song was written many years after the events narrated in the book, set in 1752.[6]

---

5 Useful studies of Jacobite songs are provided by Donaldson (1988) and Pittock (2009).

6 Indeed, in 1817 James Hogg compiled a collection of Jacobite songs on commission from the Highland Society of London; in 1819 William Blackwood published it in Edinburgh as *The Jacobite Relics of Scotland*.

Scots thus functions as an identity-marker, placing the text in a well-defined social and geographical context – a strategy that appears to have been initiated in the novels authored by Sir Walter Scott, the third most frequently quoted source in the *Oxford English Dictionary* (henceforth *OED*) after *The Times* and William Shakespeare. Undoubtedly, Scott's novels did contribute to the popularization of numerous lexical items normally associated with Scotland. On the basis of *OED* data, among the items that Scott introduced into daily usage we find the following:

(10) **Gael, n. Etymology:** <Scottish Gaelic *Gaidheal* a member of the Gaelic race = Old Irish *Gaidel*, *Goidel*. … A Scottish Highlander or Celt; also, an Irish Celt.

[1596 J. Dalrymple tr. J. Leslie *Hist. Scotl.* (1888) I. 73 Calling thame al Scottis… albeit is plane and euident that mony hundir ȝeiris eftir, thay war called Gathelis fra Gathel.]

1810 Scott *Lady of Lake* v. 192 The Gael around him threw His graceful plaid of varied hue.

slainte, int. Etymology: < Gaelic *sláinte*, lit. 'health'. A Gaelic toast: good health!

1824 Scott *Redgauntlet* II. vii. 159 He then took up the tankard, and saying aloud in Gaelic, '*slaint an Rey*', just tasted the liquor.

sporran, n. Etymology: < Scottish Gaelic *sporan*, Irish *sparán* purse. A pouch or large purse made of skin, usually with the hair left on and with ornamental tassels, etc., worn in front of the kilt by Scottish Highlanders.

1817 Scott *Rob Roy* III. vii. 209 I advise no man to attempt opening this sporran till he has my secret.

As for new semantic values, Scott's contribution is found in the following entries, among many others:

(11) **beyond, adv., prep., and n.** C. *n.*

2. **the back of beyond**: a humorous phrase for ever so far off, some very out of the way place.

1816 Scott *Antiquary* I. ii. 36 You…whirl'd them to the back o'beyont to look at the auld Roman camp.

forty-five, n. a. the Forty-five: the year 1745, and the Jacobite rebellion of that year.
1832 Scott *Redgauntlet* II. xi. 247 Ye have heard of a year they call the *forty-five*.

Although Scott's Scots-speaking characters have often been seen as prototypical instances of dialect users in literature, his ideological stance in relation to Scottish culture went beyond literary discourse and actually contributed to the achievement of political aims. When *Waverley* was published anonymously in 1814, the success was extraordinary and the name of the author, officially announced only in 1827, was in fact an easy guess. In 1815 Scott obtained permission to search for the long-lost Honours of Scotland, the crown and regalia which had last been used for the coronation of Charles II in 1651, and he actually found them hidden in Edinburgh Castle, where they can still be seen today. This led to Scott organizing George IV's visit to Edinburgh in 1822, the first time a reigning monarch had been to Scotland since Charles II in 1650 (see Prebble 1988). It was also the occasion on which tartan was given new value, although the 1746 Act of Proscription had been repealed in 1782. The 1746 Act stated:

(12) from and after the first day of August, one thousand seven hundred and forty seven, no man or boy, within that part of Great Briton [sic] called Scotland, other than shall be employed as officers and soldiers in his Majesty's forces, shall on any pretence whatsoever, wear or put on the clothes commonly called Highland Clothes (that is to say) the plaid, philibeg, or little kilt, trowse, shoulder belts, or any part whatsoever of what peculiarly belongs to the highland garb; and that no tartan, or partly-coloured plaid or stuff shall be used for great coats, or for upper coats; and if any such person shall presume, after the said first day of August, to wear or put on the aforesaid garments or any part of them, every such person so offending, being convicted thereof […], shall suffer imprisonment, without bail, during the space of six months, and no longer; and being convicted for a second offence before a court of justiciary or at the circuits, shall be liable to be transported to any of his Majesty's plantations beyond the seas, there to remain for a space of seven years. (Act of Proscription, 19 Geo. II, ch. 39, sec. 17, 1746)

When George IV visited Edinburgh, according to the *Caledonian Mercury* "His Majesty was superbly dressed in the Highland costume, with trews of the Stuart tartan. […] the manly and graceful figure of his Majesty was

finely displayed in this martial dress". London's *Morning Post* for 22 August added that "his Majesty was dressed in a full Highland uniform, and wore the broad sword, pistols, and philebeg" (see Goff 2014). In practice, between Culloden and the repeal of the Act, the Highland garb had been allowed only in the Hanoverian army, i.e. it could be worn only on condition that loyalty be sworn to the ruling monarch; but when George IV wore it in Edinburgh, with this gesture the king merged identities and – paradoxically – transformed what used to be the most obvious marks of an opposing group into the new official outfit of 'North Britain'.

Sir Walter Scott is thus seen to have contributed in very substantial ways to the creation of a distinctly romanticized image of Scotland, or indeed of 'Scott-land', as Kelly (2010) has called it. The novelist's literary use of Scots should therefore be considered in the framework of a loyalist agenda which also chose to ignore the harsh reality of the Highland Clearances, although it was in those same years that forced evictions were taking place in many areas and depleted the country of those same speakers whose language was so appreciated, at least in theory.

The case is somewhat different for Robert Louis Stevenson, whose works were published at another high-water mark in the romanticization of Scotland: what some critics have actually labelled 'Balmoralization', following Queen Victoria's fascination with the area.[7] Although many of the Scots lexical items employed by Stevenson had indeed occurred in other literary works, his ideological approach differs from that of his predecessors, as he was neither a purist nor a staunch Unionist,[8] but paid close attention to the historical sources on which he relied for his novels and stories.

Stevenson knew very well that printers felt at liberty to 'standardize' whatever was submitted, ironing out spelling or morphological discrepancies in manuscripts, for the sake of uniformity and homogeneity. Also, dialect forms could be deliberately diluted for editorial reasons, so as to make the text accessible to a wider audience (Donaldson 1986: 146-147). His works, however, were very accurate in this respect, as he paid close attention to these issues. In the introduction to the Scots poems in *Underwoods* (1887) Stevenson problematized the inconsistency of Scots spelling and candidly

---

7 See for instance comments on this phenomenon in Hodler (2015).

8 We know that Stevenson did not approve of the label 'North Britain', as in a letter dated 1888 he wrote:

> Don't put 'N.B.'on your paper: put Scotland, and be done with it. Alas, that I should be thus stabbed in the home of my friends! The name of my native land is not North Britain, whatever may be the name of yours. (RLS to S.R. Crockett, c. 10 April 1888, in Booth and Mehew 1995: 156, original emphasis).

admitted that, as far as lexical choices were concerned, his usage was somewhat eclectic:

(13) if I wish the diphthong *Ou* to have its proper value, I may write *Oor* instead of *Our*; many have done so and lived, and the pillars of the universe remained unshaken. But if I did so, and came presently to *Doun*, which is the classical Scots spelling of the English *Down*, I should begin to feel uneasy; and if I went on a little farther, and came to a classical Scots word, like *Stour* or *Dour* or *Clour*, I should know precisely where I was – that is to say, that I was out of sight of land on those high seas of spelling reform in which so many strong swimmers have toiled vainly. To some the situation is exhilarating; as for me, I give one bubbling cry and sink. […]. As I have stuck for the most part to the proper spelling, I append a table of some common vowel sounds which no one need consult; […].

I simply wrote my Scots as well as I was able, not caring if it hailed from Lauderdale or Angus, from the Mearns or Galloway; if I had ever heard a good word, I used it without shame; and when Scots was lacking, or the rhyme jibbed, I was glad (like my betters) to fall back on English. (Stevenson 1887: x-xi)

It would thus be futile to look for 'dialect authenticity' in his works, as Stevenson challenged the ideology of 'purity'. At the same time, his texts are both effective and evocative; pieces like *Thrawn Janet* (published in *Cornhill Magazine* in 1881) and *The Tale of Tod Lapraik* (in *Catriona*, published in 1893) are actual tours-de-force in the art of native story-telling, and present stories within stories, with a mise-en-abyme effect that enables readers (or indeed listeners) to imagine themselves in the distant, spooky milieu of the protagonists. In *Thrawn Janet*, the shift from the introduction (in Scottish English) and the actual story, is not even marked explicitly, as if the narrator had always been present and had only been waiting for his turn to speak. As for *The Tale of Tod Lapraik*, it is a Gothic interlude in the main novel, in many ways like *Wandering Willie's Tale* in Scott's *Redgauntlet*, first published in 1824. In both cases the first-person narrative is presented as an accurate representation of life experiences – let's compare the two incipits:

(14) Scott:

But this that I am going to tell you was a thing that befell in our ain house in my father's time – that is, my father was then a hafflins

> callant; and I tell it to you, that it may be a lesson to you that are but a young thoughtless chap, wha ye draw up wi' on a lonely road; for muckle was the dool and care that came o' 't to my gudesire. […] Ye maun have heard of Sir Robert Redgauntlet of that ilk, who lived in these parts before the dear years. The country will lang mind him; and our fathers used to draw breath thick if ever they heard him named. He was out wi' the Hielandmen in Montrose's time; and again he was in the hills wi' Glencairn in the saxteen hundred and fifty-twa; and sae when King Charles the Second came in, wha was in sic favor as the laird of Redgauntlet? He was knighted at Lonon Court, wi' the king's ain sword; and being a red-hot prelatist, he came down here, rampauging like a lion, with commissions of lieutenancy (and of lunacy, for what I ken), to put down a' the Whigs and Covenanters in the country. (Scott 1855: 73)

Stevenson:

> My faither, Tam Dale, peace to his banes, was a wild, sploring lad in his young days, wi' little wisdom and little grace. He was fond of a lass and fond of a glass, and fond of a ran-dan; but I could never hear tell that he was muckle use for honest employment. Frae ae thing to anither, he listed at last for a sodger, and was in the garrison of this fort, which was the first way that ony of the Dales cam to setfoot upon the Bass. Sorrow upon that service! The governor brewed his ain ale; it seems it was the warst conceivable. Therock was proveesioned frae the shores with vivers, the thing was ill-guided, and there were whiles when they büt to fish and shoot solans for their diet. To crown a', thir was the Days of the Persecution. The perishin'cauld chalmers were a' occupeed wi' sants and martyrs, the saut of the yearth, of which it wasna worthy. (Stevenson 1893: 164-165)

In such cases, the use of Scots does not merely serve to characterize speakers, but actually to profile a genre, the Gothic story, which is consistent with the attention paid to folk lore, legends, superstitions and other traditions by collectors of lexical items for 'provincial' glossaries – see Dossena (forthcoming). A similar approach is seen in *The Merry Men*, published in *Cornhill Magazine* in 1882, where Scots forms frame the development of another tale of the supernatural and which thus may bring Stevenson also close to the Robert Burns of *Halloween* (1785) and *Tam O'Shanter* (1790).

4. Another wilderness across the ocean

The power of ideology in the representation of other languages and/or other varieties is seen to be at work also in American literature, and equally in often ambivalent ways.[9] Indeed, beyond literature, ideology underpins other artistic representations, educational materials, and even lexicography: Noah Webster's patriotic views are well-known and they are clearly illustrated in the presentation of what sources his dictionary meant to employ:

(15) I do not indeed expect to add celebrity to the names of *Franklin, Washington*, *Adams*, *Jay*, *Madison*, *Marshall*, *Ramsay*, *Dwight*, *Smith*, *Trumbull*, *Hamilton*, *Belknap*, *Ames*, *Mason*, *Kent*, *Hare*, *Silliman*, *Cleaveland*, *Walsh*, *Irving*, and many other Americans distinguished by their writings or by their science; but it is with pride and satisfaction, that I can place them, as authorities, on the same page with those of *Boyle*, *Hooker*, *Milton*, *Dryden*, *Addison*, *Ray*, *Milner*, *Cowper*, *Davy*, *Thomson and Jameson*. […] Franklin and Washington, whose language is their hereditary mother tongue, unsophisticated by modern grammar, present as pure models of genuine English, as Addison or Swift. But I may go farther, and affirm, with truth, that our country has produced some of the best models of composition. The style of President Smith; of the authors of the Federalist; of Mr. Ames; of Dr. Mason; of Mr. Harper; of Chancellor Kent; [the prose] of Mr. Barlow; of the legal decisions of the Supreme Court of the United States; of the reports of legal decisions in some of the particular states; and many other writings; in purity, in elegance and in technical precision, is equaled only by that of the best British authors, and surpassed by that of no English compositions of a similar kind.

(Webster 1828: Preface)

A patriotic approach in which the specificity of American English was stressed is equally found in books aimed at self-education, such as in *Beadle's Dime Speakers*, 25 inexpensive booklets published between 1859 and 1886 by a company best-known for its dime novels, which comprised titles like *The American Speaker* (1859), *The National Speaker* (1860), *The Patriotic Speaker* (1862), *The Spread Eagle Speaker* (1869) and *The Hail Columbia Speaker* (1876).

[9] On the similar approaches to Scottish and Native American cultures see Calloway (2008) and Lauzon (2008).

In the historical context of those same years, both before and after the Civil war, the US was also defining its identity through a long sequence of Indian wars which were reflected in an often contradictory representation of protagonists and events both in literature and in other forms of art. In paintings, sculptures and chromolithographs, for instance, Native subjects could be depicted either as fierce warriors or as 'vanishing Indians' (see Cartosio 2016); similarly, alongside troubling captivity narratives readers found conversion narratives that stressed the 'civilizing' value of religion (see Wyss 1999 and Campbell 2015). Emigrants' guides and travelogues provided what were supposed to be objective descriptions, but it was literature and popular narratives that established images of Native Americans (and of their languages) that would be just as powerful as those offered by the visual arts (see Dossena 2015). To give just two examples, we can compare the annotations we find in a travelogue (Campbell 1876) and in a historical account (McIntosh 1853). In the former, John Campbell described the phonological rendition of a native conversation in the Puget Sound area in the following terms:

(16) An old woman, clicking as men click when they talk in the Caucasus and at the Cape of Good Hope, with strange grunts and gutturals for language, chattered. The men grinned. They were the ugliest set of mortals that ever I saw. (Campbell 1876: 111-112)

In the latter, instead, John McIntosh called Algonquin "the Italian of the western continent" on account of its vowel system, and stressed traits that were seen to be characteristic of 'proper' languages: elegance, harmony, "richness of expression, […] variety of turns, […] propriety of terms, and […] regularity":

(17) The *Algonquin* language has not so much force as the *Huron*, but it has more sweetness and elegance, and may with great propriety be denominated the Italian of the western continent; for it abounds with vowels, which renders it soft, musical, and harmonious. Both the *Algonquin* and the *Huron* have a richness of expression, a variety of turns, a propriety of terms, and a regularity which seldom prevails in some of the more cultivated languages of Europe.
(McIntosh 1853: 93-94)

Within this framework, the works of James Fenimore Cooper are justly considered to be exemplary of this ambivalent (and ideologically-construed)

attitude to the languages of Native Americans (see Scatà, this volume). In *The Last of the Mohicans* (1826) the ways in which different characters, their languages and their cultural values are seen to interact places at least four different languages along a cline ranging from near-perfection (the language of the Mohican) to depravation (French) – English is better than French, but still inadequate to represent Mohican qualities, while the language of the Huron, who are siding with the French, is stigmatized on account of their having abandoned their ancestral culture, in order to take on that of an invading people who have no qualms about renaming their territory and imposing foreign place names. As for the character of Natty Bumppo, European by birth but adopted by the Mohican, he often relies on silent gestures like his Delaware companions, almost as if his ability to understand two different cultures placed him at the centre of a complex communicative network.

In stories concerning 'Mountain Men', i.e. fur trappers and traders, instead, linguistic representation again presents a different kind of exotic usage. Instances of this can be found in George Frederick Ruxton's narrative of his "Adventures in Mexico and the Rocky Mountains", first published in Blackwood's *Edinburgh Magazine* in 1848, and indeed the text in which the term 'mountain men' is first found.[10] Here the protagonists are shown to use a kind of sociolect not unlike what is found in urban narratives. Rather than geographical distance, what is showcased is a class distinction that does not envisage much schooling and in which both lexis and pronunciation define the in-group. In the excerpt below, eye-dialect, slang, and typical colloquialisms intertwine and offer readers a glimpse of this distant, but still comprehensible, world:

(18) "Whar's them mules from? They look like Californy." "Mexican country-away down south." "H-! Whar's yourself from?" "There away too." "What's beaver worth in Taos?" "Dollar." "In St. Louiy?" "Same." "H-! Any call for buckskin?" "A heap! The soldiers in Santa Fe are half froze for leather, and mocassins fetch two dollars easy." "Wagh! How's trade on Arkansa, and what's doing to the fort?" "Shians at Big Timber, and Bent's people trading smart. On North Fork, Waters

[10] See the entry in the *OED* for *mountain man*, n. […] b. *U.S.* A backwoodsman, a trapper; a pioneer. Also *fig.* 1847 G.F.A. Ruxton *Adventures Mexico & Rocky Mts.* xxv. 221 The depreciation in the value of beaver-skins has thrown the great body of trappers out of employment, and there is a general tendency among the mountain-men to settle in the fruitful valleys of the Rocky Mountains.

> got a hundred pack right off, and Sioux making more." "Whar's Bill Williams?" "Gone under they say: the Diggers took his hair
>
> (Ruxton in Hubbard 1968: 220)

In spite of the apparent heterogeneity of the various texts from which examples have been presented here, what they all have in common is the attempt to show linguistic distinctiveness as a token of authenticity, a feature to which a strong ideological value is attributed, whether it is to stress patriotic views or to emphasize distance from other cultures.

5. Concluding remarks

In nineteenth-century literature speakers of different languages and/or language varieties are often seen to embody the ambivalent attitudes that authors, critics and the reading public shared towards the cultures that such languages and varieties expressed. Literary representations reflected and indeed helped to perpetuate the kind of bias which idealized supposedly ancient forms, while at various levels educational policies stigmatized those same forms and sought to eradicate them. On both sides of the Atlantic grammarians and lexicographers strove to promote linguistic assimilation and uniformity; at the same time, compilers of glossaries and novelists made sure that lexical and phraseological items could be preserved as relics of ancient eloquence.

Both approaches could lead to the fabrication of myths – on one hand, of standard regularity; on the other, of antiquity, purity and (implicitly) of non-viability in the modern world. It was only in very few cases that socially- and/or geographically-marked forms could be employed as strategies for social critique and empowerment; in general, the ideology underpinning the representation of such forms in literary discourse reflected a kind of 'imperialist nostalgia' even in linguistic matters.

While observations along these lines have often been put forward in the study of post-colonial literature, in this study I have attempted to outline how representations of Scots fit into this framework and indeed how closely related they can be to the linguistic representation of Native Americans, offering a transatlantic perspective on topics that have not often been considered through a more encompassing approach. At the same time, it may be stressed that, in order to study these phenomena, the tools of quantitative investigation can only provide a starting point for further

interdisciplinary studies; to that end, historical linguists must acquire new tools and question assumptions from different perspectives; i.e., as aptly suggested by Filppula and Klemola (2014), they must redefine themselves, first and foremost, as language historians.

REFERENCES

Sources

Act of Proscription, 19 Geo. II, ch. 39, sec. 17, 1746, www.electricscotland.com/history/other/proscription_1747.htm, accessed December 2020

Adams, J.
1799 *The Pronunciation of the English Language*. [Repr. 1974]. In: R.C. Alston (ed.) *English Linguistics* 1500-1800. London: The Scolar Press.

Campbell, J.F.
1876 *My Circular Notes.* London: Macmillan & Co.

Chalmers, G.
1806 *The Poetical Works of Sir David Lyndsay […]*. London: Longman.

Currie, J. (ed.)
1846 *The Complete Works of Robert Burns, with an Account of his Life and a Criticism on his Writings, […].* Halifax: William Milner.

Gibson, E.
1722 *Britannia* (2nd edn.). London: Printed by Mary Matthews, for Awnsham Churchill, and sold by William Taylor.

Jamieson, J.
1808 *An Etymological Dictionary of the Scottish Language […] to which is prefixed A Dissertation on the Origin of the Scottish Language*. Edinburgh: Printed at the University Press for W. Creech, A. Constable & Co. […].

Mackay, C.
1888 *A Dictionary of Lowland Scotch with […] an Appendix of Scottish Proverbs*. Edinburgh: Ballantyne Press.

Mackenzie, H.
1786 "Extraordinary account of Robert Burns, The Ayrshire ploughman". [Repr. 1897]. In: R. Cochrane (ed.) *The English Essayists*. Edinburgh: W.P. Nimmo, Hay and Mitchell, 278-280.

McIntosh, J.
1853 [1843] *The Origins of the North American Indians: With a Faithful Description of their Manners and Customs, both Civil and Military, their Religions, Languages, Dress and Ornaments.* New edition. New York: Nafis and Cornish.

Oxford English Dictionary Online (OED-3)
www.oed.com, accessed December 2020

Pinkerton, J.
1786 *Ancient Scotish Poems, Never Before in Print […]. And […] an Account of the Contents of the Maitland and Bannatyne MSS*. London: Printed for Charles Dilly and William Creech at Edinburgh.

Puttenham, G.
1589 *The Arte of English Poesie*. London.

Ruxton, G.F.
1848 "Life in the Far West", *Blackwood's Edinburgh Magazine* 63-64, passim.

Salamanca Corpus of English Dialect Texts
2011- M.F. García-Bermejo Giner – P. Sánchez-García – J. Ruano-García – The DING Group, University of Salamanca (compilers), www.thesalamancacorpus.com/index.html, accessed December 2020

[Scott, Walter]
1824 [1855] *Redgauntlet. A Tale of the Eighteenth Century*. By the Author of "Waverley". In: Three Volumes. Vol. I (II-III). Edinburgh: Printed for Archibald Constable and Co.

Stevenson, R.L.
1887 *Underwoods*. London: Chatto and Windus.
1893 *Catriona*. London: Cassell & Co.

Webster, N.
1828 *An American Dictionary of the English Language*. New York: Published by S. Converse.

Special studies

Agha, A.
2003 "The social life of cultural value", *Language and Communication* 23, 231-273.

Blakemore, S.
1984 "Strange tongues: Cooper's fiction of language in *The Last of the Mohicans*", *Early American Literature* 19, 21-41.

Bondi, M.
2015 "Probably most important of all. Importance markers in academic and popular history articles". In: N. Groom – M. Charles – J. Suganthi (eds.) *Corpora, Grammar and Discourse. In Honour of Susan Hunston*. Amsterdam: John Benjamins, 161-182.
2017 "What came to be called: Evaluative what and authorial voice in the discourse of history", *Text and Talk* 37, 25-46.

Booth, B.A. – E. Mehew (eds.)
1994/1995 *The Letters of Robert Louis Stevenson*. New Haven: Yale University Press.

Broadhead, A.
2013 *The Language of Robert Burns: Style, Ideology, and Identity*. Lewisburgh, PA: Bucknell University Press.
Brown, D.W.
2018 *English and Empire: Literary History, Dialect, and the Digital Archive*. Cambridge: Cambridge University Press.
Calloway, C.G.
2008 *White People, Indians, and Highlanders. Tribal Peoples and Colonial Encounters in Scotland and America*. Oxford: Oxford University Press.
Campbell, D.
2015 "Captivity narratives", https://public.wsu.edu/~campbelld/amlit/captive.htm, accessed December 2020
Cartosio, B.
2016 "American artists look West". In: M. Dossena – S. Rosso (eds.) *Knowledge Dissemination in the Long Nineteenth Century: European and Transatlantic Perspectives*. Newcastle u. T.: Cambridge Scholars, 9-26.
Culpeper, J.
2017 "The influence of Italian manners on politeness in England, 1550-1620", *Journal of Historical Pragmatics* 18 (2), 195-213.
Donaldson, W.
1986 *Popular Literature in Victorian Scotland: Language, Fiction and the Press*. Aberdeen: Aberdeen University Press.
1988 *The Jacobite Song. Political Myth and National Identity*. Aberdeen: Aberdeen University Press
1989 *The Language of the People: Scots Prose from the Victorian Revival*. Aberdeen: Aberdeen University Press.
Dossena, M.
2000 "Sense, shortness and salt: Ideas of improvement in eighteenth- and nineteenth-century collections of Scottish proverbs", *Review of Scottish Culture* 12, 93-106.
2005 *Scotticisms in Grammar and Vocabulary*. Edinburgh: John Donald (Birlinn).
2012 "'A highly poetical language'? Scots, Burns, patriotism and evaluative language in nineteenth-century literary reviews and articles". In: C. Percy – M.C. Davidson (eds.) *The Languages of Nation: Attitudes and Norms*. Bristol: Multilingual Matters, 99-119.
2013 "'*Stour* or *dour* or *clour*': An overview of Scots usage in Stevenson's works and correspondence". In: J. M. Kirk – I. Macleod (eds.) *Scots: Studies in its Literature and Language*. Amsterdam: Rodopi, 87-101.
2015 "(Re)constructed eloquence. Rhetorical and pragmatic strategies in the speeches of Native Americans as reported by nineteenth-century commentators", *Brno Studies in English* 41 (1), 5-28.

forthcoming "Scots as the language of the uncanny: The case of nineteenth-century Gothic narratives". In: D. Montini – I. Ranzato (eds.) *The Dialects of British English in Fictional Texts*. London: Routledge.

Dury, R.
1992 "Saxonism and the preference for 'native' vocabulary". In: N. Pantaleo (ed.) *Aspects of English Diachronic Linguistics*. Fasano: Schena, 133-146.

Filppula, M. – J. Klemola
2014 "Celtic influences in English: A re-evaluation", *Neuphilologische Mitteilungen* 115 (1), 33-53.

Goff, M.
2014 "George IV in Highland dress", http://blogs.bl.uk/untoldlives/2014/01/george-iv-in-highland-dress.html, accessed December 2020

Hodler, M.
2015 "Performing Scottishness in the US: Highland games, diaspora, and nation", https://ussporthistory.com/2015/07/09/performing-scottishness-in-the-us-highland-games-diaspora-and-nation/, accessed December 2020

Hodson, J. (ed.)
2017 *Dialect and Literature in the Long Nineteenth Century*. Abingdon: Routledge.

Hodson, J. – A. Broadhead
2013 "Developments in literary dialect representation in British fiction 1800-1836", *Language and Literature* 22 (4), 315-332.

Hodson, J. – A. Broadhead – J. Millward (compilers)
2014 *Dialect in British Fiction 1800-1836*. Database development by K. Rogers – M. Pidd, University of Sheffield Humanities Research Institute, www.dialectfiction.org, accessed December 2020

Hubbard, C.
1968 "The language of Ruxton's Mountain Men", *American Speech* 43 (3), 216-221.

Hunston, S.
2011 *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London: Routledge.

Hunston, S. – G. Thompson (eds.)
2001 *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.

Hyland, K. – M. Bondi (eds.)
2006 *Academic Discourse Across Disciplines*. Bern: Peter Lang.

Hyland, K. – G. Diani (eds.)
2009 *Academic Evaluation: Review Genres in University Setting*. London: Palgrave Macmillan.

Innes, E.J.
1993 "The social, economic and political reasons for the decline of Gaelic in Scotland", www.scottishhistory.com/articles/highlands/gaelic/gaelic_page1.html, accessed December 2020
Jespersen, O.
1905 *Growth and Structure of the English Language*. Leipzig: B.G. Teubner.
Kelly, S.
2010 *Scott-land: The Man Who Invented a Nation*. Edinburgh: Polygon.
Lauzon, M.
2008 "Welsh Indians and savage Scots: History, antiquarianism, and Indian languages in 18th-century Britain", *History of European Ideas* 34 (3), 250-269.
Martin, J.R. – P.R.R. White
2005 *The Language of Evaluation. Appraisal in English*. London: Palgrave Macmillan.
McClure, J.D.
1995a *Scots and its Literature*. Amsterdam: John Benjamins.
1995b "The Pinkerton syndrome". In: J. Derrick McClure (ed.) *Scots and its Literature*. Amsterdam: John Benjamins, 57-67. Originally published in *Chapman* 41 (1985), 2-8.
2000 *Language, Poetry and Nationhood – Scots as a Poetic Language from 1878 to the Present*. East Linton: Tuckwell Press.
Moore, D.
2017 *The International Companion to James Macpherson and The Poems of Ossian*. Glasgow: Association for Scottish Literary Studies.
Paternoster, A. – F. Saltamacchia
2017 "(Im)politeness formulae and (im)politeness rules: Metadiscourse and conventionalisation in 19th-century Italian conduct books". In: E.M. Pandolfi et al. (eds.). *Studies on Language Norms in Context*. Frankfurt a. M.: Peter Lang, 263-301.
Percy, C.
2010 "How eighteenth-century book reviewers became language guardians". In: P. Pahta et al. (eds.) *Social Roles and Language Practices in Late Modern English*. Amsterdam: John Benjamins, 55-85.
Pittock, M.G.H.
2009 *The Myth of the Jacobite Clans: The Jacobite Army in 1745*. Edinburgh: Edinburgh University Press.
Pottle, M.S. – C.C. Abbott – F.A. Pottle (eds.)
1993 *Catalogue of the Papers of James Boswell at Yale University*. Edinburgh: Edinburgh University Press.
Prebble, J.
1988 *The King's Jaunt: George IV in Scotland, August 1822: 'One and twenty daft days'*. London: Fontana.

Rennie, S.
2011 "Boswell's Scottish Dictionary rediscovered", *Dictionaries: Journal of the Dictionary Society of North America* 32, 94-110.
2012 "Boswell's Dictionary update", *Dictionaries: Journal of the Dictionary Society of North America* 33, 205-207.
Rosenwald, L.
1998 "*The Last of the Mohicans* and the languages of America", *College English* 60 (1), 9-30.
Scatà, A.
2020 "Ideologies of linguistic representation in Late Modern English: The case of James Fenimore Cooper", this issue.
Schachterle, L.
2011 "James Fenimore Cooper on the languages of the Americans: A note on the author's footnotes", *Nineteenth-Century Literature* 66 (1), 37-68.
Shields, J.
2009 "Savage and Scott-ish masculinity in *The Last of the Mohicans* and *The Prairie*: James Fenimore Cooper and the diasporic origins of American identity", *Nineteenth-Century Literature* 64 (2), 137-162.
Shorrocks, G.
1996 "Non-standard dialect literature and popular culture". In: J. Klemola – M. Kytö – M. Rissanen (eds.) *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt a. M.: Peter Lang, 385-411.
Sorensen, J.
2000 *The Grammar of Empire in Eighteenth-Century British Writing*. Cambridge: Cambridge University Press.
2017 *Strange Vernaculars: How Eighteenth-Century Slang, Cant, Provincial Languages, and Nautical Jargon Became English*. Princeton: Princeton University Press.
Tieken-Boon van Ostade, I. (ed.)
2008 *Grammars, Grammarians and Grammar-Writing in Eighteenth-Century England*. Berlin: Mouton de Gruyter.
2014 *In Search of Jane Austen: The Language of the Letters*. Oxford: Oxford University Press.
Warner, F.B., Jr.
1969 "Stevenson's first Scottish story", *Nineteenth-Century Fiction* 24 (3), 335-344.
White, P.R.R.
2007 "Appraisal". In: J.-O. Östman – J. Verschueren (eds.) *Handbook of Pragmatics Online*, s. v. Amsterdam: John Benjamins, www.benjamins.com/online/hop, accessed December 2020
2015 *The Appraisal Website*, www.grammatics.com/appraisal/, accessed December 2020

Wyss, H.E.
1999 "Captivity and conversion: William Apess, Mary Jemison, and narratives of racial identity", *American Indian Quarterly* 23 (3/4), 63-82.

Address: MARINA DOSSENA, Dipartimento di Lingue, Letterature e Culture Straniere Università degli Studi di Bergamo, Piazza Rosate 2, 24129 Bergamo, Italy.
ORCID code: orcid.org/0000-0002-8025-6086.

"Theers gud stuff amung uz Darbysher foaks": Dialect enregisterment in 19th-century Derbyshire

Paula Schintu
University of Salamanca

ABSTRACT

The textual material included in the *Salamanca Corpus* bears witness to dialectal awareness in 19th-century Derbyshire, with an important number of literary texts that reflect the local people's habits of speech. Despite the fact that this variety ought to be of particular interest since it was used in an area which marks the transition between the North of England and the West Midlands, and the East and West Midlands, literary representations of the Derbyshire dialect remain largely unexplored (García-Bermejo Giner 1991, 1993 is the most remarkable exception). According to research in the field, the analysis of this type of representation is crucial to investigate the processes of enregisterment of dialect varieties, as Johnstone et al. (2006) and Johnstone (2009, 2013) have shown. They examine the enregisterment of Pittsburghese by looking at non-standard discourse in a range of modern sources. Less attention, however, has been paid to the study of this process in historical contexts, the works by Beal (2009, 2017, 2019), Ruano-García (2012, 2020, forthcoming), Clark (2013), Cooper (2013, 2016, 2020), and Beal – Cooper (2015) being among the exceptions. This study takes a preliminary approach to the enregisterment of 19th-century Derbyshire dialect by examining a selection of instances of dialect writing, most of which are included in the *Salamanca Corpus*. I aim at identifying the main linguistic forms associated with this variety in terms of spelling, morphology and lexis, as well as determining the extent to which 19th-century instances of dialect writing contribute to the enregisterment and dissemination of such linguistic forms and the values they index.

Keywords: Derbyshire dialect, enregisterment, nineteenth century, dialect literature, literary dialect.

1. Introduction: Indexicality, enregisterment, or the making of a dialect

Over the past decade, *enregisterment* – the "processes through which a linguistic repertoire becomes differentiable within a language as a socially

DOI: 10.25951/4360.

recognized register of forms" (Agha 2003: 231) – has become an increasingly important approach to the study of the mechanisms underlying the development and legitimation of the different linguistic varieties within the English language. As research has shown (Johnstone 2009, 2013; Beal 2009, 2019, among others), linguistic varieties are very often imbued with socio-cultural meaning that make them different from others and shape speakers' perceptions and attitudes towards different habits of speech. They endow varieties with legitimacy since speakers acknowledge and internalise the link between language and certain ideological connotations that make them perceive linguistic repertoires as stable. This, in turn, leads to their maintenance "across time and region via metapragmatic practices that reiterate [their value] and its link to social status and correctness" (Johnstone et al. 2006: 80).

Michael Silverstein (1976) describes various levels in the process of value assignment whereby certain socio-cultural notions are indexed or associated with linguistic varieties; he refers to three stages that he calls orders of indexicality. Johnstone (2009: 164) and Beal – Cooper (2015: 35), amongst others, have explained Silverstein's taxonomy, in which the first order of indexicality refers to the earliest step of the process, that in which a speech community is unaware of the correlation that an outsider would perceive between the set of linguistic forms they use and a certain social category. At the second order, there is awareness of this link thanks to factors such as language contact, and speakers start to rationalise, modify and accommodate their habits of speech taking into account criteria such as correctness, style, etc. When reaching the third order of indexicality, notions such as locality and social class are indexed to linguistic varieties, creating linguistic and socio-cultural stereotypes about specific speech communities. It is when this level of awareness has been reached that *enregisterment* may arise.

As highlighted by Johnstone (2009: 160), once a variety shows third order indexicality, its enregisterment is determined by discursive practices – or, as she puts it, "talk about talk"– in the form of oral or textual artefacts such as literature, dictionaries and other types of discourse that represent and exemplify it. The dissemination of these artefacts, Agha explains (2003: 243), is key in the process of enregisterment since it helps to spread, share and typify the linkages between language and social features, this way, in his own words, "making possible the large-scale replication of register stereotypes across social populations".

As shown by studies dealing with dialect enregisterment in historical contexts, the role of dialect writing is crucial in this process since literature

bears witness to the way language was used and perceived in the past. This is clearly acknowledged by Ruano-García, who highlights the role of dialect writing as a "clear conduit by which the correlation between language and sociocultural values, as well as the ideas derived from it, are foregrounded, circulated and consumed" (2012: 377). Similarly, Beal – Cooper stress that "a key element in observing processes of enregisterment is the production of dialect literature, literary dialect, and dialect 'commentary'", since "through these media we are afforded a glimpse into the social value of language features in historical periods" (2015: 52). In this sense, although literary representations of dialect do not provide detailed transcripts of the language, they provide valuable insight into how writers, as well as society, viewed and understood the dialect represented. They are rich sources of information about how regional varieties were perceived and the attitudes speakers had towards them.

In spite of this, and although literary representations of dialect have often been explored in order to improve our knowledge about certain historical traits (see, for instance, García-Bermejo Giner 1991, 1993 and 1994), relatively little attention has been paid to the role of dialect writing in the process of enregisterment; exceptions are Beal (2009, 2017, 2019), Ruano-García (2012, 2020, forthcoming), Clark (2013), Cooper (2013, 2016, 2020) and Beal – Cooper (2015). Accordingly, this paper aims to examine literary representations of 19th-century Derbyshire dialect from the point of view of enregisterment so as to gain insight into this variety's most salient features and the socio-cultural perceptions underlying its representation. My purpose is twofold. First, to identify which were the main linguistic and sociolinguistic characteristics of the dialect as represented in literature. Second, I seek to determine how the Derbyshire dialect was enregistered in the period analysed, whilst trying to ascertain the role of dialect writing in the process.

2. Literary representations of dialect in the 19th century: Dialect literature and literary dialect

The 19th century was a period of remarkable linguistic awareness due to the many social changes of the period. Beal – Cooper (2015: 42-43) explain that, during the first half of the century, industrialization pushed people to move to urban sectors, which made moving around the country more important than ever. Railway companies were quick enough to notice this new demand and take action. Railway systems were improved and expanded,

providing the English population with a quicker and more convenient way of travelling. As a result, the 1800s saw an unparalleled population flow circulating the country, thereby bringing the different varieties of English into contact.

Geographical mobility and language contact have been identified as key factors when it comes to enregisterment. They trigger the appearance of second-order indexical links by disrupting the otherwise closed social and linguistic networks and lead "people to link dialect and social identity more explicitly" (Johnstone et al. 2006: 94). Together with the advances in education that took place in 19th-century England, these factors led to the acknowledgement of linguistic difference and gave way to a growing dialect awareness among the population. Speakers became conscious of the linguistic diversity and the distinctiveness of their local varieties and started to be concerned about the loss of their particular linguistic identities due to the social and geographical upheaval that threatened to standardise the English language. Derbyshire author Joseph Barlow Robinson (c. 1820-1883) bears witness to this state of unease in the preface to his *Owd Sammy Twitcher's Visit tu 't Gret Exibishun e Darby* (1870b), in which he justifies the necessity of preserving local varieties by means of the written word (1):

(1) (…) the time is fast approaching when, by the spread of education, railways, and other means, all peculiarities will be lost, and merge into one general and universal manner of speech throughout the kingdom. A work of this character will then serve to give future generations some idea of those who lived before them, and prevent their many peculiarities from being totally lost (1-2).

Mirroring these concerns, the 19th century saw the production and publication of a great number of dialect works, mainly in the form of dialect literature (DL) and literary dialect (LD). As is well known, DL comprises those "works composed wholly (sometimes partly) in a non-standard dialect, and aimed essentially, though not exclusively, at a non-standard-dialect-speaking readership", whilst LD refers to "the representation of non-standard speech in literature that is otherwise written in standard English (…) and aimed at a general readership" (Shorrocks 1996: 386).

There are some key differences between these types of representation in terms of authorship, purpose and audience that are quite relevant and worth summarizing for the purpose of this paper. Concerning authorship and purpose, DL, on the one hand, is mostly produced by dialect natives,

which, to a certain extent, validates its linguistic realism and accuracy. As shown in (1) above, DL writers aim at reflecting manners of speech, instances of regional writing acting, thus, as relics that attempt to preserve and maintain a range of oral features. LD, on the other hand, is normally written by outsiders of the variety; authors with some knowledge of the dialect that typically use it in the dialogues of their works with characterization purposes in order to portray and identify dialect speakers so that they can be recognised by users and non-users of the variety alike. Hence, the traits employed must be salient enough so as to be understood and associated with a specific type of character embodying a particular set of extralinguistic features.

Instances of both LD and DL can be found in 19th-century Derbyshire. Although little can be traced about its earlier literary tradition, the county was, of course, no exception to the outburst of dialect writing in the 1800s. The first literary record of the Derbyshire dialect was made available by Thomas Tapping (1817-1886) in *The Rhymed Chronicle of Edward Manlove* (1851). The book adds a preface and a glossary of Derbyshire mining terms to a poem written by Edward Manlove (1615-1671). Although Tapping's edition was published in 1851, Manlove's poem dates back to 1653, being the earliest available literary representation of the dialect. During the first half of the 19th century, Richard Furness's (1791-1857) poem *Medicus Magus or the Astrologer. A Poem with a Glossary* (1836) was published, followed by *The Cat and the Vicar* in 1858. Both texts include instances of the Derbyshire dialect, some of which are recorded in the glossary appended to the first poem. Six years later, *Derbyshire Men* (1864) would be published in the journal *The Reliquary*; it is a short poem written by Walter Kirkland (1828-1899) that illustrates both the Derbyshire dialect and character.

However, it is in 19th-century prose that a higher awareness of the Derbyshire dialect as a distinct variety can be observed. Mary Howitt's (1799-1888) *My Uncle, the Clockmaker* (1844) seems to be the earliest novel including dialogues written in the dialect. The second half of the century presents the largest availability of dialect writing in the county. Joseph Sheridan Le Fanu (1814-1873) set one of his most famous novels, *Uncle Silas* (1864), in Derbyshire, and thus some of its dialogues include passages representing the county's speech. By the same token, *The Reliquary* published in its 1870-1871 edition *A Village Sketch, at Ashford-in-the-water, in Illustration of the Derbyshire Dialect*, a short story by Thomas Brushfield (1828-1910), whose main aim was to give evidence of the dialect spoken in the county. Similarly, Derbyshire author Frances Parthenope Verney (1819-1890) would also use

dialect in her novels *Stone Edge* (1868) and *The Greypool and other Stories* (1891); Mrs. Humphrey Ward (1851-1920), in turn, would represent it in her three-volume *David Grieve* (1892). Robert Murray Gilchrist (1867-1917) was a very prolific author with an interest in Derbyshire and the Peak District. He wrote a number of novels set here during the late 1800s and at the turn of the 20th century, some of them contain representations of the dialect. We may refer to *A Peakland Faggot. Tales of Milton Folk* (1897), *The Courtesy Dame* (1900), *Natives of Milton* (1902), and *Good-bye to Market* (1908).

DL material is scarce in Derbyshire. Indeed, only one author provides us with this type of representation in the 19th century. Between the years 1870 and 1881, Joseph Barlow Robinson (c. 1820-1883) wrote a series of five stories whose protagonist and narrator, Sammy Twitcher, takes us to different events and exhibitions in the county. Three of these texts – *Owd Sammy Twitcher's Visit tu't Gret Exibushun e Darby* (1870b), *Owd Sammy Twitcher's Second Visit tu't Gret Exibishun e Darby, wi' Jim* (1870) and *Owd Sammy Twitcher's Visit tu't Watter Cure Establishment at Matlock Bonk* (1871) – include glossaries explaining some of the dialect words used, the lists added to the second and third stories being revisions of the original compilation.

By exemplifying and circulating the dialect, these literary works contributed to its circulation not only within 19th-century Derbyshire but also beyond its borders. This way, literary representations of the variety brought it into contact with non-natives of the dialect, contributing to its legitimation and shaping linguistic and socio-cultural ideas about the county, its language and its speakers.

3. The enregisterment of 19th-century Derbyshire dialect: A preliminary approach

3.1 Primary data

In line with other studies that have explored the historical enregisterment of northern English (e.g. Beal 2017, 2019; Ruano-García 2012, 2020, forthcoming), this paper makes a quantitative and qualitative analysis of literary representations of the Derbyshire dialect published during the second half of the 19th century. In order to undertake this study, I have considered four literary renditions of the Derbyshire dialect, three of which are included in *The Salamanca Corpus*, taking into account the following criteria:

- Publication in the period 1850-1900.
- Genre: Texts written in prose fiction.
- Type of representation: DL and LD.

Table 1. Corpus material

Text type	N texts	N words
DL	2	16,299
LD	2	98,688
Total	4	114,987

As Table 1 shows, my primary data for this paper consist of four texts which amount to 114,987 words.[1]

3.2 Quantitative analysis

3.2.1 Spelling

A careful survey of the DL data makes it clear that the representation of the Derbyshire dialect was based upon a particular set of features that includes a consistently occurring set of spellings signalling dialect sounds. Table 2 summarises the repertoire of the most frequent spellings found in the corpus which I have classified according to their RP pronunciation and standard spelling.

As shown in Table 2, the spelling <u> for RP /ʌ/ is the second most frequent trait in the sample, suggesting an [ʊ] realization of the sound in words like *sum* or *tutch*. This suggests the lack of centralization and loss of lip rounding of ME /u/ in 1640 in the Derbyshire dialect, pointing to the lack of FOOT-STRUT split that, as is well known, is one of the main features that distinguishes northern from southern dialects (see Upton – Widdowson 2006: Map 7). Indeed, Ihalainen (1994: 213) notes that this is a typical characteristic of the linguistic north, which, as Clark explains, "includes the Midlands, incorporating the Birmingham-Wolverhampton conurbation, i.e., the West Midlands" (2008: 139). Furthermore, we can find instances of another feature which has traditionally been associated with the West Midlands in terms of phonology: the rounding before nasals (see Upton – Widdowson 2006: Map 2). The repeated rounding of /a/ into /ɒ/ in the texts analysed suggests an [ɒ] pronunciation which is uncommon in other varieties outside the West

[1] A *Peakland Faggot. Tales of Milton Folk* (Murray Gilchrist 1897): 32,147 words; *Owd Sammy Twitcher's Crismas Bowk for the Year 1870* (Robinson 1870a): 8,535 words; *Owd Sammy Twitcher's Visit tu 't Gret Exibishun e Darby* (Robinson 1870b): 7,764 words; *Stone Edge* (Verney 1868): 66,541 words. See the *Salamanca Corpus of English Dialect Texts* (2011–) for more information about these texts.

Midlands (Wakelin 1977: 96). Ihalainen highlights that this is "an exclusively west midland feature towards the end of the nineteenth century, when it disappeared from south-western English" (1994: 217). Its use in Derbyshire is documented by Pegge (1896), who records forms like *conno* 'cannot', *ony* 'any' and *mon* 'man' in the county (viii).

Table 2. Top spelling traits (DL) (> 150 tokens): raw data

Traits	Standard spellings	Types	Tokens	Some examples
<aa> for RP /aʊ/	<ou>, <ow>	51	324	*aat* 'out', *taan* 'town'
<u> for RP /ʌ/	<oCe>, <o>, <ou>, <oe>	37	268	*sum* 'some', *tutch* 'touch'
<oi> for RP/aɪ/	<iCe>, <i>	58	236	*woife* 'wife', *moind* 'mind'
<ow> for /u:/	<oo>, <wo>, <o>, <ou>	15	214	*rowf* 'roof', *dow* 'do'
<ee> for RP /eə/	<e + r + e>	12	206	*wheer* 'where', *theer* 'there'
<o + n> for RP /æ + n/	<a + n>	18	172	*mon* 'man', *con* 'can'
<ow> for RP /əʊ/	<o + l>	14	167	*owd* 'old', *howd* 'hold'

There is one feature that is particularly frequent in the corpus, as Table 2 shows: the use of <aa> to represent words otherwise pronounced RP /aʊ/ (e.g. *aat*, *taan*), which points to the [a:] realization of the sound in this county. The *English Dialect Grammar* (henceforth *EDG* [Wright 1905]) records "ā" among the different pronunciations of this sound in Derbyshire (Wright 1905: 146), and so does Ellis (1889: 425, 427), who testifies to its use in several parts of the county in words such as *daats*. This feature notably outnumbers the other forms found in the sample, with the exception of <oi> for words pronounced /aɪ/ (e.g. *woife*, *moind*). As Upton and Widdowson show (2006: Map 10), [ɔi] is the most common pronunciation of RP /aɪ/ in the dialect, and it is also recorded in both the *EDG* (Wright 1905: 128) and Ellis (1889: 425).

This is in line with the data found in the LD sample.

Table 3 shows that <aa> for <ou>/<ow> nearly doubles the next feature as the most frequent form in the LD data, and that it is three times more common in terms of types. Likewise, the evidence reflects both the lack of FOOT-STRUT split and the rounding before nasals as two of the top non-standard traits in this type of dialect representation, which is in line

with the data found in the DL texts. This goes some way to suggesting that, together with the other features (e.g. <ow> for RP /əʊ/ or <ee> for RP /eə/), spelling <aa> for RP /aʊ/, the lack of FOOT/STRUT split and the rounding before nasals were commonly understood as characteristic of 19th-century Derbyshire dialect not only by a dialect-speaking readership, but also by a wider audience.

Table 3. Top spelling traits (LD) (> 90 tokens): raw data

Traits	Standard spellings	Types	Tokens	Some examples
<aa> for RP /aʊ/	<ou>, <ow>	32	259	*maase* 'mouse', *haa* 'how'
<u> for RP /ʌ/	<oCe>	8	149	*summat* 'something', *un* 'one'
<oo> for RP/ʌ/	<oCe>	6	128	*loove* 'love', *coom* 'come'
<o + n> for RP /æ + n/	<a + n>	11	112	*hond* 'hand', *lond* 'land'
<ow> for RP /əʊ/	<o + l>	11	107	*towd* 'told', *gowd* 'gold'
<ee> for RP /eə/	<e + r + e>	6	99	*theer* 'there', *wheer* 'where'

Table 4 summarises the most recurrent phonological traits in the texts analysed.

Table 4. Top spelling traits (DL/LD) (> 250 tokens): raw data and NF / 1,000

Traits	DL		LD		Total	
	Tokens	NF	Tokens	NF	Tokens	NF
<aa> for RP /aʊ/ (*raand* 'round', *braan* 'brown)	324	19.87	259	2.62	583	5.07
<u> for RP /ʌ/ (*cuntry* 'country', *luv* 'love')	268	16.44	149	1.50	417	3.62
<ee> for RP /eə/ (*weer* 'where', *theer* 'there')	206	12.63	99	1.00	305	2.65
<o + n> for RP /æ + n/ (*grond* 'grand', *stond* 'stand')	172	10.55	112	1.13	284	2.46
<ow> for RP /əʊ/ (*cowd* 'cold', *sowd* 'sold')	167	10.24	107	1.08	274	2.38

When it comes to the comparison of both types of representation, <aa> spelling of words containing the diphthong /aʊ/ is clearly identified as the most salient feature of the dialect, followed by unsplit [ʊ]. Slight differences

can be observed with regard to spelling <ee> in words like *theer* or *weer*, which suggests an [i:] pronunciation of terms otherwise pronounced with the centring diphthong /eə/. This is recorded in the *EDG*, which gives "*ī*" as one of the possible realizations of the sound in Derbyshire (Wright 1905: 108). Despite the fact that this is the third most common trait in the whole dataset, DL texts use it more frequently than LD material, which employs the rounding before nasals and <ow> for RP /əʊ/ slightly more regularly.

3.2.2 Morphology and lexis

Dialectal morphology and lexis also receive attention in the works analysed. As regards morphological traits, the analysis of the data has revealed a repertoire of features which are consistently used in both the DL and LD material; they are summarised in Table 5.

Table 5. Top morphological traits (DL/LD) (> 500 tokens): raw data and NF / 1,000

Traits and examples	DL		LD		Total	
	Tokens	NF	Tokens	NF	Tokens	NF
2nd p. pron. *ye/yo*: "Wunna *ye* come an tak' the wapses' nest?"	34	2.08	774	7.84	808	7.02
past tense BE: "Hee *wor* a gud lowkin owd feller"	344	21.10	432	4.37	776	6.74
-na negation: *isna* 'is not', *shouldna* 'should not'	99	6.07	467	4.73	566	4.92

As shown in Table 5, the use of *ye* and *yo* for the 2nd person subject pronoun is the most recurrent feature in the sample. Interestingly, its frequency in the LD material notably outnumbers that of the DL data, which suggests that this trait was particularly salient and widely recognised as part of the Derbyshire dialect by non-natives of the variety. This is in line with contemporary non-literary accounts of the dialect. The *English Dialect Dictionary* (henceforth *EDD* [Wright 1896-1905]) acknowledges the use of these two pronouns in the county, whilst both Halliwell (1881: xiv) and Pegge (1896: 85) record *yo* as a characteristic form of Derbyshire speech. The following example may illustrate the use of these forms in the corpus texts (2):

(2) "I thowt as *ye*'d summat *ye* wanted sore to speak on to Nathan," burst out Bessie suddenly, remembering Roland's urgent messages, and wishing kindly to forward the business.

"'Twas my father wanting to know whether *yo* kep' them two sheep as is in the croft to joist," said Roland (Verney 1868: 35. My emphasis).

Non-standard past tense BE forms seem to be strongly associated with the dialect too, especially in DL, where the alternation between *was* and *were* is overwhelmingly more frequent than the rest of the features, both in the DL and the compared data. This morphological trait, thus, seems to be perceived as markedly Derbyshire not only by outsiders, but, most remarkably, by natives of the variety. However, there are differences in the way this feature is reflected in both text types: whilst only *-r* forms are found in the DL sample (e.g. "I thowt it *wor* a gud chonce"; "t'yung men an wimmin *wor* theer" [Robinson (1870b: 6), (1870b: 4), my emphasis]), the LD data point to a more variable system in which *was* and *were* are used in singular and plural contexts alike: "et *was* her doin'"; "et *were* a strange thing"; "them beech trees *wes* hard to do"; "mother and feyther *were* laid theer" (Murray Gilchrist [1897: 30], [1897: 6], [1897: 22], my emphasis). Nevertheless, there seems to be a tendency towards *-r* forms in the LD data, which, overall, reveals a "strong preference for generalized singular *were* forms" in Derbyshire (Pietsch 2005: 150). Finally, instances of *-na* negation, a characteristic north-west Midland feature (Ihalainen 1994: 218; Britain 2007: 84), are also considerably frequent in the corpus, which goes in line with contemporary and modern non-literary evidence of the dialect. We find examples such as: "ah *hadna* mitch wok e hond" (Robinson [1870a: 9], my emphasis). Halliwell (1881: xiv) recorded the use of *-na* forms such as *conner* 'cannot', *shanner* 'shall not' and *wooner* 'will not' in Derbyshire, as did Pegge (1896: 46), who referred to *munna* 'must not' in his list of Derbicisms. Some years later, the *Survey of English Dialects* (*SED*) would likewise testify to the use of *-na* negation in the county (Ihalainen 1994: 218).

Concerning lexis, Table 6 shows the most recurrent traits documented in the corpus. Terms like *mun*, *nowt*, *sin* and *summat* are consistently employed in all the works analysed, suggesting that they were generally accepted as part of the dialect both in and out of the community in which they were used. However, the use of the feminine subject pronoun *hoo* stands out as the most salient lexical trait in the texts considered, especially as regards the DL material, in which its frequency is four times higher than in the LD data. As is well known, this feature is one of the most widely recognised West Midland

traits (see, for example, Ihalainen 1994: 218-219; Upton – Widdowson 2006: Map 34). The *EDG* (Wright 1905: 273) testifies to the use of *hoo* in Derbyshire, while Pegge (1896) recorded several examples of this pronoun, as in "*hoo'l ne'er o'er't*, she will never get over it" (xii). It is worth noting that all these words are documented in non-literary accounts of the dialect. *Mun*, *nowt/nought* and *summat/summut* are recorded in both the glossary appended to *Owd Sammy Twitcher's Visit tu't Gret Exibushun e Darby* (Robinson 1870b) and in Pegge's work, which also glossed *sin*. Furthermore, all the terms are recorded and quoted from Derbyshire in the *EDD* (Wright 1896-1905), which, remarkably, cites from two of the works analysed to illustrate the use of *nowt* and *sin*: *Owd Sammy Twitcher's Visit tu't Gret Exibushun e Darby* (Robinson 1870b) and *A Peakland Faggot. Tales of Milton Folk* (Murray Gilchrist 1897), respectively.

Table 6. Top lexical traits (DL/LD) (> 50 tokens): raw data and NF / 1,000

Traits and examples	DL		LD		Total	
	Tokens	NF	Tokens	NF	Tokens	NF
hoo/how 'she': "but *hoo* were Johanna's dowter"	92	5.64	161	1.63	253	2.20
mun 'must': "Handyfist *mun* be a cliver feller"	11	0.67	140	1.41	151	1.31
nowt 'nothing': "ye heerd *nowt* but good on him"	14	0.85	49	0.49	63	0.54
sin 'since': "Et 's thretty year *sin'* he died"	12	0.73	45	0.45	57	0.49
summat/summut/su'mmut 'something': "I'll tell yo' *summat*"	19	1.16	34	0.34	53	0.46

All things considered, the evidence suggests that the literary representation of the Derbyshire dialect drew on a specific set of phonological, morphological and lexical features. These include spelling patterns such as <aa>, <u>, <ee>, <o + n> and <ow> for RP /aʊ/, /ʌ/, /eə/, /æ + n/ and /əʊ/, respectively, and a repertoire of morphological and lexical features which comprises the use of 2nd person pronoun *ye/yo*, non-standard past tense BE forms, *-na* negation, the feminine pronoun *hoo*, and terms like *mun*, *nowt*, *sin* and *summat*. Interestingly, these traits are the most salient features in both text types, which seems to confirm that they were identified as characteristic of the dialect during the period analysed both by insiders and outsiders of the variety.

3.3 Qualitative analysis

As has been shown, the analysis of both DL and LD representing the dialect spoken in 19th-century Derbyshire can give us some insight into its most salient linguistic features. Needless to say, the use of dialect in writing is often associated with a certain type of character, allowing readers to see how and to what purpose the language is used. All the texts selected for this study link the Derbyshire dialect to a very specific type of persona: farmers, peasants and country folk. In fact, at the beginning of *Owd Sammy Twitcher's Visit tu 't Gret Exibishun e Darby*, Joseph Barlow Robinson informs readers that he aims to provide (3):

(3) (…) a fair representation of the thoughts and manner of speech of one of the many old farmers yet to be met with: jolly old chaps, with more genuine fun in them than half-a-dozen of the young ones growing up round them (1870b: 1).

This way, he establishes a link between the dialect he is about to depict and this specific identity. Similarly, instances of LD also use the dialect for characterization purposes. In them, country folk are the Derbyshire speakers, in contrast with the people coming from the city, who are presumed to be of a higher social class; example (4) may exemplify this:

(4) On the night of William Townend's homecoming, before he discovered himself to his fellow natives, he sat in the bar-parlour of the Golden Bull, posing like the mysterious stranger of fiction, who eventually shines as the wealthy son or brother of the ruined lord of the manor. He was a tall, black-bearded man, with sparkling eyes and bottle-shaped nose. *His well-cut clothes concealed in some measure the Peakland slouch, and his hands and waistcoat were embellished with costly jewellery*.

[…]

"Ay, I 've been here before. Time changes a man. I remember you ploughing against Tom Winterton at the Noe Valley Fair." (Murray Gilchrist [1897: 180-181], my emphasis).

William Townend was born in Derbyshire but moved to Canada after being rejected by Emma Bamber, to whom he intended to get married. Many years later, he returns to the county after becoming a wealthy city

man that has stopped using dialect, which testifies to the strong dialect awareness in 19th-century England. Just like William Townend, characters in the texts accommodate to the different social and linguistic settings by means of language, which points to a shift from the first to the second order of indexicality in that they are able to evaluate their speech in terms of correctness and appropriateness, modifying it accordingly. This can also be observed in (5) when German Ashford, a farmer characterised by his profuse use of dialect, avoids it when he meets the Squire, a high-class gentleman who speaks standard English:

(5) The old man himself, with one of the last queues left in England on one end of him, and short and blue stockings on the other, was sitting before a mass of papers at the table. After all, however, he was the squire, and German felt a certain "awe " as he entered.

[…]

"It were my sister's money," said German in a low voice; "she'd gived him every penny she had" (Verney 1868: 238-39).

Interestingly, accommodation goes in both directions since characters who tend to use the standard consciously decide to speak the dialect depending on their interlocutor. Example (6) shows how William Townend, the standard speaker in (4) above, deliberately chooses to use regional traits to his advantage when he meets Emma Bamber, his former fiancée, again:

(6) Old Maid Bamber came in at ten o'clock. (…) When she saw the boots she gave a wild and painful cry.

"Theer 's bin a ghoast here ! They 're Bill's, my lad Bill's. Nob'dy i' th' lond had feet that shiiape an' size!"

He stole behind her and caught her in his arms.

"Emma, wench," he said, *his speech losing the refinement which a broader life had given*. "I 've coom to ask yo' again. I 've never sin ony as I could care for but yo'!" (Murray Gilchrist [1897: 184-185], my emphasis).

Dialect speakers are not only characterised by means of language, but they are also linked to certain non-linguistic attributes and stereotypes. As shown

in (6), the Derbyshire dialect is regarded as a vulgar variety that is worth avoiding. Hence, native speakers are seen as unrefined, humble, uneducated people. This is clearly shown in *Stone Edge* (Verney 1868), where the absence of academic training is persistently associated with dialect users. After highlighting that "writing was a rare accomplishment" in the county (117), the narrator goes on to present how characters acknowledge their lack of formal education. Farmer German Ashford, for instance, laments in (7) the fact that he is not "booklearned" (122), a rare complaint in the county since, as the narrator emphasises, intellectual aspirations were usually present in higher social classes only:

(7) I bean't a learnin' nothin'; it's just muddlin' and milkin' and wabblin' i' th' mud arter plough tail. I'm like the little donkeys in the lane, I canna addle [earn] nought." *The burgher blood from his mother was stirring curiously in the lad*. "Roland would ha' learnt me to write and cipher, but feyther wouldn't let me nigh him. Well, good by, uncle, I must go; the minits runs as fast as rats down here (Verney [1868: 134], my emphasis).

Derbyshire dialect speakers are, thus, associated with humility and lack of formal education, yet they are also regarded as rude, boorish people, as described in (8):

(8) She lived in a little house beyond the Nether End of Milton, a quaint, pretty place, covered with ivy and Virginia creeper. It was her own property, and the desire of her heart was to keep up the reputation for good management which her mother, who was a "foreigner" from the Yorkshire Wolds, had acquired amongst *the rough-and-ready Peaklanders*. (Murray Gilchrist [1897: 179], my emphasis).

However, although the data in the LD sample trace an ideological link between the Derbyshire dialect and fairly negative social connotations, the DL representations point in a different direction. In the texts analysed, the dialect, far from being avoided, is encouraged both in speech and writing. In *Owd Sammy Twitcher's Visit tu 't Gret Exibishun e Darby* (Robinson 1870b), for example, the characters visit an exhibition in which the names of the rooms, the sculptures, the catalogues, etc. are all written in dialect. Indeed, we learn that "t'picter az fust towk [their] attenshun wor, *"Leyin daan t'Law""*, and then "annuther caw'd *"Hasses in a Shed""* (4), the dialectal equivalent to

"Horses in a Shed". It seems that the writer is trying to advocate and dignify the use of the dialect, somehow reacting against and challenging the ideas and perceptions about the variety circulated in the LD texts analysed. In fact, he pleads for the county, encouraging the "Darbysher foaks" to feel proud of it:

(9) Theers now uther kaanty con lick uz e meyin most things, an wee can mey ommast ivvery thing wee wanten. Theers gud stuff amung uz Darbysher foaks yet, an aar owd kaanty taan isna ta be sneezed at be a long chawk (Robinson 1870b: 20).

This ideological difference between both text types may be explained on account of the fact that, as noted by Beal (2004) and Wales (2006), due to the many social changes that took place in the period, attitudes towards dialect shifted in the 19th century, and its users started to feel proud of their localness and distinctiveness as regards their speech. As such, in the 1800s "dialect is by no means an object of shame, but cherished as an emblem of local identity" (Wales 2006: 129).

4. Concluding remarks

In this study, I have taken a preliminary linguistic and sociolinguistic approach to 19th-century Derbyshire dialect by means of the framework of enregisterment. As I have attempted to show, the analysis of literary renditions of this variety reveals that they are useful sources of information not only about the main linguistic forms and features that distinguished the dialect, but also about the sociocultural ideas linked with it.

The data analysed reveal that there is a common set of spellings suggesting dialectal pronunciation that are consistently associated with the dialect in both text types considered. The consistent use of forms including spellings such as <aa>, <u> or <o + n> for RP /aʊ/, /ʌ/ and /æ + n/, respectively, suggests that these traits were salient enough and had been sufficiently circulated so that a non-native audience could recognise such forms as characteristic of this dialect. The analysis has also shed light on the morphological and lexical features most commonly understood as Derbyshire, which include 2nd person pronoun *ye/yo*, non-standard past tense BE forms, *-na* negation, and the use of terms like the feminine pronoun *hoo*, *mun* 'must', *nowt* 'nothing', *sin* 'since' and *summat* 'something'. The fact

that all these features appear in instances of DL and LD alike testifies to the validity of both types of representation as linguistic portraits of the dialect since the features reflected in the LD texts, written for non-native users of the dialect, coincide with those employed in the DL material, which was produced by and for Derbyshire speakers. Although slight differences can be observed between the two types of dialect representation analysed, it seems evident that, by means of their use in writing, these linguistic traits were linked to a very specific type of speaker embodying certain non-linguistic attributes. As such, the Derbyshire people, characterised by their roughness, modesty and lack of literacy, were related to a recognizable linguistic repertoire which, in turn, became imbued with these socio-cultural connotations and came to index specific ideological features.

The production and circulation of DL and LD representing the Derbyshire dialect is in itself evidence of enregisterment since it points to the existence of third-order indexical links which indicate that the dialect was already regarded as a distinct variety within and outside the region where it was spoken. The circulation of the linguistic and cultural ideas described above through DL and, most of all, LD, allowed their propagation throughout the country, bringing the dialect in contact with larger audiences and helping to share the linkages between the variety and the values and identities associated with it. This way, collective linguistic and sociolinguistic ideas about the variety were created and incorporated to the public imagination, contributing, thus, to the enregisterment of the Derbyshire dialect.

However, extensive research still needs to be done in order to reach more comprehensive and representative results that may shed more light on the literary representation of this dialect. This would give further insight not only into the characteristics of the Derbyshire dialect, but also into the processes underlying its legitimation and enregisterment.

REFERENCES

Sources

Murray Gilchrist, R.
1897 *A Peakland Faggot. Tales of Milton Folk.* London: Grant Richards.
Robinson, J.B.
1870a *Owd Sammy Twitcher's Crismas Bowk for the Year 1870*. Derby.

1870b *Owd Sammy Twitcher's Visit tu 't Gret Exibishun e Darby.* Derby.

The Salamanca Corpus

2011– M.F. García-Bermejo Giner – P. Sánchez-García – J. Ruano-García – The DING Group, University of Salamanca (compilers), www.thesalamancacorpus.com/index.html, accessed December 2020

Verney, F.P.

1868 *Stone Edge*. London: Smith Elder and Cy.

Special Studies

Agha, A.

2003 "The social life of cultural value", *Language and Communication* 23, 231-273.

Beal, J.C.

2004 *English in Modern Times 1700-1945.* London: Arnold.

2009 "Enregisterment, commodification, and historical context: 'Geordie' versus 'Sheffieldish'", *American Speech* 84 (2), 138-156.

2017 "Northern English and enregisterment". In: J.C. Beal et al. (eds.) *Perspectives on Northern Englishes*. Berlin; Boston: De Gruyter, 17-39.

2019 "Enregisterment and historical sociolinguistics". In: S. Jansen et al. (eds.) *Processes of Change: Studies in Late Modern and Present-Day English*. Amsterdam: John Benjamins, 7-24.

Beal, J.C. – P. Cooper

2015 "The enregisterment of Northern English". In: R. Hickey (ed.) *Researching Northern English*. Amsterdam; Philadelphia: John Benjamins, 25-50.

Britain, D.

2007 "Grammatical variation in England". In: D. Britain (ed.) *Language in the British Isles*. Cambridge: Cambridge University Press, 75-104.

Clark, U.

2008 "The English West Midlands: Phonology". In: B. Kortmann et at. (eds.) *A Handbook of Varieties of English*. Berlin: Mouton De Gruyter, 134-163.

2013 "*'Er's from off*: The indexicalization and enregisterment of Black Country Dialect", *American Speech* 88 (4), 441-466.

Cooper, P.

2013 *Enregisterment in Historical Contexts: A Framework*. Unpublished PhD thesis, Sheffield: University of Sheffield.

2016 "'Deregisterment' and 'fossil forms': The cases of gan and mun in 'Yorkshire' dialect: The changes in what features constitute 'Yorkshire' dialect from the 19th century to the 21st", *English Today* 33 (1), 43-52.

2020 "Russian dolls and dialect literature: The enregisterment of nineteenth century 'Yorkshire' dialects". In: P. Honeybone et al.

(eds.) *Dialect Writing and the North of England*. Edinburgh: Edinburgh University Press.

Ellis, A.J.
1889 *The Existing Phonology of English Dialects Compared with that of West Saxon*. London: Trübner & Co.

García-Bermejo Giner, F.
1991 "Personal pronouns in Derbyshire as reflected in Sheridan Le Fanu's *Uncle Silas*". In: S.G. Fernández-Corugedo et al. (eds.) *Studia Patriciae Shaw Oblata*. Oviedo, 225-239.
1993 "Morphosyntactical variation in nineteenth-century Derbyshire". In: F.J. Ruiz de Mendoza Ibáñez et al. (eds.) *Actas XV Congreso AEDEAN*. Logroño: Colegio Universitario de La Rioja, 305-318.
1994 "Sources for the linguistic analysis of some nineteenth-century Midland dialects". In: W. Viereck (ed.) *Proceedings of the International Congress of Dialectologists*. Stuttgart: Franz Steiner, 232-245.

Halliwell, J.O.
1881 *A Dictionary of Archaic and Provincial Words, Obsolete Phrases, Proverbs, and Ancient Customs, from the Fourteenth Century*. London: John Russell Smith.

Ihalainen, O.
1994 "The dialects of England since 1776". In: R. Burchfield (ed.) *The Cambridge History of the English Language.* Vol 5: *English in Britain and Overseas: Origins and Development*. Cambridge: Cambridge University Press, 197-274.

Johnstone, B.
2009 "Pittsburghese shirts: Commodification and the enregisterment of an urban dialect", *American Speech* 84 (2), 157-175.
2013 *Speaking Pittsburghese: The Story of a Dialect*. Oxford: Oxford University Press.

Johnstone, B. et al
2006 "Mobility, indexicality, and the enregisterment of 'Pittsburghese'", *Journal of English Linguistics* 34 (2), 77-104.

Pegge, S.
1896 *Two Collections of Derbicisms Containing Words and Phrases in a Great Measure Peculiar to the Natives and Inhabitants of the County of Derby*. London: Oxford University Press.

Pietsch, L.
2005 "'Some do and some doesn't': Verbal concord variation in the north of the British Isles". In: B. Kortmann (ed.) *A Comparative Grammar of British English dialects: Agreement, Gender, Relative Clauses*. Berlin: Mouton De Gruyter, 125-209.

Ruano-García, J.
2012 "On the enregisterment of the Northern dialect in Early Modern English: An evaluation across literary text types". In: S. Martín

et al. (eds.) *At a Time of Crisis: English and American Studies in Spain*. Barcelona: AEDEAN, 376-383.

2020 "On the enregisterment of the Lancashire dialect in Late Modern English: Spelling in focus", *Journal of Historical Sociolinguistics* 6 (1), 1-38.

forthcoming "*Aw'm Lancashire, owd cock, and gradely hearty*: Enregistered Lancashire voices in the nineteenth-century theatre". In: D. Montini – I. Ranzato (eds.) *The Dialects of British English in Fictional Texts*. London: Routledge.

Silverstein, M.

1976 "Shifters, linguistic categories, and cultural description". In: K.H. Basso – H.A. Selby (eds.) *Meaning in Anthropology*. Albuquerque: University of New Mexico Press, 11-55.

Shorrocks, G.

1996 "Non-standard dialect literature and popular culture". In: J. Klemola – M. Kytö – M. Rissanen (eds.) *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt a. M.: Peter Lang, 385-411.

Upton, C. – J.D.A. Widdowson

2006 *An Atlas of English Dialects*. London; New York: Routledge.

Wakelin, M.F.

1977 *English Dialects: An Introduction*. London: The Athlone Press.

Wales, K.

2006 *Northern English. A Cultural and Social History*. Cambridge: Cambridge University Press.

Wright, J.

1896-1905 *English Dialect Dictionary*. Oxford: Henry Frowde.

1905 *English Dialect Grammar*. Oxford: Henry Frowde.

Address: PAULA SCHINTU, Departamento de Filología Inglesa, Universidad de Salamanca. Calle Placentinos 18, 37008, Salamanca, España.
ORCID code: orcid.org/0000-0002-2902-065X.

Ideologies of linguistic representation in Late Modern English: The case of James Fenimore Cooper*

Annalisa Scatà

Independent scholar

ABSTRACT

The aim of this article is to analyse James Fenimore Cooper's linguistic representations in his works with a particular focus on his most famous novel, *The Last of the Mohicans* (1826). In this novel Cooper depicted the typical contradictions of nineteenth-century American society both in the use of his own language and in the representation of Native American languages. In addition, by adding editorial footnotes to his novels, Cooper explained American customs and historical events to his British readers, but at the same time he supported the need to introduce new words in order to give an accurate representation of American reality.

The language created by Cooper in *The Last of the Mohicans* has often been debated by literary critics: many scholars accused Cooper of giving an idealized and romantic image of Native Americans, while others defended the author affirming that his representation was authentic and coherent with the historical period in which he lived. In fact, the way in which languages are presented bears witness to the ideology of the times, as Cooper created a linguistic hierarchy in which the Edenic language of the Delaware is presented as superior to the fallen and corrupt languages of the English and the French. On the other hand, by showing the destruction of Native American languages and cultures and by celebrating English as the only language understood by everyone, he seems to have implicitly suggested that the advancement of an Anglo-centric civilization was both advisable and inexorable.

Keywords: linguistic representations, ideology, editorial footnotes, Native American languages, linguistic hierarchy.

* This essay is based on the author's MA dissertation, which in 2018 was awarded the Gullì Prize by the Italian Society for North American Studies.

DOI: 10.25951/4361.

1. Introduction

The main purpose of this article is to examine James Fenimore Cooper's linguistic representations in *The Last of the Mohicans* (1826), on account of their value as emblematic of Cooper's attitude towards American English, which was distinctively ambivalent: he defended the Americans' need to introduce new words, but at the same time he condemned some linguistic innovations of theirs, while also showing an interestingly nuanced attitude towards native languages. As a matter of fact, his ambivalent attitude was related to the intellectual context of those years, characterised by the linguistic controversy between Samuel Johnson and Noah Webster and in which the development of a new 'American' language was a widely discussed matter. My investigation is based on a close reading of the texts in their historical context. A quantitative investigation of the data is beyond the scope of this study for various reasons: first of all, because my main interest is socio-historical, and secondly because the kind of representations under investigation are not amenable to the methods of corpus linguistics.

The article is structured as follows. In Section 2 I will focus on the linguistic ideas that Cooper expressed in the editorial footnotes that he added to his novels for their British edition. Then, in Section 3, I will examine the sources from which Cooper drew inspiration for his novel, in order to assess whether Cooper's Indians were idealized or authentic. This will enable me to focus on the linguistic hierarchy created by Cooper and on how he depicted the fight for the new world as a linguistic fight between Indian and European languages, where the Edenic language of the Delaware vanishes with the last of the Mohican, while the surviving tribes speak foreign and corrupt languages. Finally, in Section 4, some concluding remarks are presented.

2. Ambivalent attitudes to English and Native American languages

In order to understand Cooper's contradictions, it is important to take into consideration the historical context in which he lived. James Fenimore Cooper was born in 1789 and died in 1851, so his lifetime spans the so-called "Early Republic".[1] At that time, the development of a new language for

[1] "The Early Republic" (1780-1830) was a period of transition during which the Americans established a new government, faced the results of industrialisation, discussed social matters like slavery and extended their boundaries by conquering

the Americans was a widely discussed matter. In Cooper's opinion, it was essential for America to eliminate its mental dependence from England, but this wish for emancipation clashed with the complex problem of a shared language (Schachterle 2011: 37). In *Gleanings in Europe: England* (1837), Cooper commented on this matter writing that "of all the burdens, that of the mental dependence created by colonial subserviency, appears to be the most difficult to remove" (JFC *Gleanings* 1982: 233).

As a matter of fact, the creation of a national language was both a political and a social matter – see Simpson (1986). Conservative intellectuals, like John Witherspoon, condemned some "Americanisms", while some innovators defended American usage. Among the latter was Noah Webster, who promoted the introduction of distinctively American pronunciation, orthography and grammar, as more appropriate for a population that was socially different from the one in Britain. Cooper's main goal was to "create through a distinctive national language the mental independence from English opinion" (Schachterle 2011: 66). As we will see later in his footnotes, Cooper remarked on how French and Dutch contributed to the creation of American words, on how English words, that were obsolete in England, were used by the Americans, and how Native American vocabulary influenced the American one. Nevertheless, although Cooper could be described as a polyglot (he knew Latin and French very well), he didn't actually imagine a multilingual and transnational America; on the contrary, he promoted the creation of a language independent from England that could absorb new words as symbols of a rising national identity (Schachterle 2011: 66). In his works Cooper celebrated the diversity of American ethnic and linguistic communities, but at the same time he hoped that all those ethnic groups would adapt to American English, because the alternative was, as in the case of many Indian tribes and languages, extinction.

Cooper's ambivalence is also present in his representation of Indian languages. In the nineteenth-century, American and European ideas about Native American languages and cultures were closely connected to the image of the savage. First of all, the Indians were called 'savages' because they represented the opposite of 'civilized' Europeans. In addition, being savages was considered an initial stage in man's progress, whereas civilization was assumed to be a more advanced level. For Euro-American colonists this primordial condition of Native Americans must lead inexorably to

Western territories. In addition, in those years new technologies were introduced and uniquely American forms of art and literature were born (Bates 2015: xxxiii).

civilization and must be substituted by it. As Lucy Maddox affirmed, in that period "there were only two options for the Indians: to become civilized, or to become extinct" (Maddox 1991: 24). The Euro-Americans who were in favour of civilization obviously disagreed with those who were comfortable with the idea of extinction; as a result, those who, like Cooper, represented the Indians as noble savages opposed those who, like Francis Parkman, despised them.[2] However, most Euro-Americans believed that civilization and extinction were the only two options for the future of Native Americans, and hardly anyone was ready to admit that they were actually civilized, but in a different way. For most Euro-Americans, Indian languages provided another proof of their savagery, as reported by Maddox:

> White observers consistently concluded that because of the limitations of his or her language, the most complex intellectual maneuver any Indian (of whatever language group) could manage was the construction of a simple metaphor, or occasionally an analogy; the Indian could not speculate about things that have no visible form, nor comprehend notional ideas. (Maddox 1991: 24)

However, some scholars rejected this vision; one of them was John Heckewelder, who is one of the sources from which Cooper drew inspiration for his representation of Native Americans.[3] In his work *History, Manners, and Customs* (1818), Heckewelder reported his correspondence with Peter Duponceau, one of the greatest scholars of Indian languages of the time, and in this work they both describe these languages as complex social constructions. Their idea was in contrast with the stereotypical image of the savage, and Heckewelder confirmed that his goal was:

> To satisfy the world that the languages of the Indians are not so poor, so devoid of variety of expression, so inadequate to the communication

[2] Francis Parkman (1823-1893), was a famous American historian who wrote a seven-volume history of France and England in North America, covering the colonial period from the beginnings to 1763 (www.britannica.com/biography/Francis-Parkman, accessed September 2019).

[3] John Heckewelder (1743-1823) was a Moravian missionary in the United States. He had been adopted by and lived with the Lenápes (Delawares) for 49 years and he studied carefully the languages, manners, and customs of the Indians. He spent his last years writing numerous accounts of Native American life, notably his *Account of the History, Manners and Customs of the Indian Nations, etc.* (1818). (Johansen – Pritzker 2008: 695)

> even of abstract ideas, or in a word so *barbarous*, as has been generally imagined. (Heckewelder 1971: 125)

The contrasting views that characterized linguistic evaluations both in relation to European languages and to Native American ones emerge very clearly in Cooper's works and indeed in the editorial footnotes to his novels, which are carefully examined by Schachterle (2011).

Through an analysis of such footnotes it is possible to highlight how literary texts whose popularity is undoubted could and actually did contribute to the definition of linguistic perceptions that would become both pervasive and long-lasting on both sides of the Atlantic. It is in this framework that Cooper's notes become a valuable object of investigation for historical linguists. Although most of Cooper's novels focus on the Americans before, during and after the Revolution, in 1831, while Cooper was in Europe, he decided with his British editors, Colburn and Bentley, to review seven of his American novels for his growing British public, and to add explanatory notes. In these, Cooper supported the need to introduce new words in order to give an accurate representation of American reality, but at the same time he clarified American customs and historical events for his British readers. Moreover, he made significant comments pertaining to linguistic forms; such footnotes can be divided into two groups: those that defended and those that criticized American usage.

2.1. Defensive footnotes

According to Schachterle (2011), in order to explain the first group of notes, it is necessary to quote a comment from the novel *Satanstoe*, first published in 1845, in which Cooper says why mental dependence from England should be overcome – see excerpt (1):

(1) "Sleigh," as spelt, is purely an American word. It is derived from "slee," in Dutch; which is pronounced like "sleigh." Some persons contend that the Americans ought to use the old English words "sled," or "sledge." But these words do not precisely express the thing we possess. There is as much reason for calling a pleasure-conveyance by a name different from "sled," as there is for saying "coach" instead of "wagon." "Sleigh" will become English, ere long, as it is now American. Twenty million people not only can make a word, but they can make a language, if it be needed. (JFC *Satanstoe* 1990: 206)

In order to achieve cultural independence from England, Americans had to re-elaborate the languages inherited from the old world so that they could describe their new reality – a reality that was characterized by a different climate and topography.

In Chapter 5 of *The Politics of American English* (1986), David Simpson underlines how in *The Pioneers* (1823; revised British edition 1832) Cooper succeeded in representing the distinctiveness of American usage. The first footnote that appears in this novel concerns, once again, the word *sleigh*:

(2) Sleigh is the word used in every part of the United States to denote a traineau. It is of local use in the west of England, whence it is most probably derived by the Americans. The latter draw a distinction between a sled, or sledge, and a sleigh; the sleigh being shod with metal. Sleighs are also sub-divided into two-horse and one-horse sleighs. Of the latter, there are the cutter, with thills so arranged as to permit the horse to travel in the side track; the "pung," or "tow-pung," which is driven with a pole, and the "jumper," a rude construction used for temporary purposes, in the new countries. Many of the American sleighs are elegant, though the use of this mode of conveyance is much lessened with the melioration of the climate, consequent on the clearing of the forests. (JFC *Pioneers* 1980: 17)

In this comment Cooper highlights the fact that Americans were not creating new lexical items: they were just using English words that British readers in big cities had probably forgotten and which for this reason had become obsolete.

Among Cooper's many works, *The Pioneers* is the novel that contains most explanations both in the text and in the footnotes. In one of these notes Cooper refers to the origin of the word *Yankee* explaining that "in America the term Yankee is of local meaning. It is thought to be derived from the manner in which the Indians of New England pronounced the word English or Yengeese" (JFC *Pioneers* 1980: 53). The re-edited version of *The Prairie* (1827; revised British edition 1832), instead contains information on American customs that are supposed to be completely unfamiliar for British readers. One of the recurring characters in Cooper's novels, Natty Bumppo, is often described in this novel as a *trapper*, a lexical item that is clarified in the following note: "It is scarcely necessary to say, that this American word means one who takes his game in a trap. It is of general use on the frontiers. The beaver, an animal too sagacious to be easily killed, is oftener taken in this way than in any other" (JFC *Prairie* 1877: 17).

However, for Cooper no political or social reality elicited so much interest as Native Americans (Schachterle 2011: 49). In his representation of Native languages Cooper created two different linguistic worlds. On the one hand there are the Indians that white civilization has corrupted and who speak, according to Simpson, "an inelegant pidgin English unmarked by any evident memories of their own authentic, poetic locutions" (Simpson 1986: 205). On the other hand, there are the heroic Mohican who use their typical poetic expressions underlining the purity of Native American languages. The first note that appears in *The Last of the Mohicans* explains that Horican was the original name of Lake George, a lake that was renamed by both French and English colonists:[4]

(3) As each nation of the Indians had its language or its dialect, they usually gave different names to the same places, though nearly all of their appellations were descriptive of the object. Thus a literal translation of the name of this beautiful sheet of water, used by the tribe that dwelt on its banks, would be "The Tail of the Lake." Lake George, as it is vulgarly, and now, indeed, legally, called, forms a sort of tail to Lake Champlain, when viewed on the map. Hence, the name. (JFC *Mohicans* 1998: 16)

In this example Cooper reconstructs the name used by Native Americans probably by taking their cue from its physical traits (Schachterle 2011: 50). The author also gives another explanation of Native American customs in a note in which he shows how Indian tribes are called in different ways both by their friends and by their enemies.

Schachterle's analysis continues with another novel, *The Prairie*, where we can find an important note in which Cooper claimed that European and Native American languages were unified in an "American language" that was used by both, but was native for neither:

(4) The Americans and the Indians have adopted several words, which each believes peculiar to the language of the others. Thus "squaw", "papoose" or child, wigwam, &c. &c., though it is doubtful whether they belonged at all to any Indian dialect, are much used by both white and red men in their intercourse. Many words are derived from the French, in this species of Prairie nomaic. (JFC *Prairie* 1877: 488)

[4] See Section 3.2 for further analysis of the word "Horican".

According to Schachterle, this note represents Cooper's awareness of the fact that "cultural border-crossings between white men and red men were already beginning at the level of language" (Schachterle 2011: 52).

2.2 Critical footnotes

This second type of notes is found in the novels where Cooper used the authorial technique of the artificial editor. This strategy allowed him to present new narrators who often shared his own prejudices, but at the same time he could distance himself from their ideas. Among the footnotes in *Satanstoe* that ridicule American provincialisms there is an editorial note – quoted in (5) below – that shows how much Cooper was afraid that "American provincialism and exuberance sometimes produces linguistic changes that are awkward, comic, and too often grandiloquent" (Schachterle 2011: 55):

(5) It is northern [that is, Yankee] American, to call a small "lake" a "pond," a small "river" a "creek," even though it should be an "outlet," instead of an "inlet," &c. &c. It is a more difficult thing than is commonly supposed, to make two great nations, each of which is disposed to innovate, speak the same language with precise uniformity. The Manhattanese, who have probably fewer of the peculiarities of the inhabitants of a capital than the population of any other town in the world of four hundred thousand souls, the consequences of a rapid growth, and of a people who have come principally from the country, are much addicted to introducing new significations for words, which arise from their own provincial habits. In Manhattanese parlance, for instance, a "square" is a "park," or, even a "garden" is a "park." A promenade on the water, is a "battery!" It is a pity that, in this humour for change, they have not thought of altering the complex and imitative name of their town… – EDITOR. (JFC *Satanstoe* 1990: 384)

In this note Cooper mentions two great social groups: the Yankees and the Manhattanese and for him they are both guilty of innovations that he did not approve. For this reason, in his later works, Cooper tried to limit American changes that he considered exuberant. In a note to his novel *The Chainbearer*, of 1845, Cooper (1912) severely criticizes the pronunciation of the people of New England and their excessive use of Latinate words to indicate common objects.

In other novels Cooper restricted the right of linguistic creation only to the upper classes: in his view, New Englanders, the lower classes, and social climbers did not possess the proper authority to create and establish new usages. Cooper continued to support this idea in the chapter: "On Language" in *The American Democrat* (1838), where he underlined the most common *lacunae* in American English:

(6) The common faults of American language are an ambition of effect, a want of simplicity, and a turgid abuse of terms. To these may be added ambiguity of expression. (JFC *Democrat* 1931: 110)

Another interesting aspect of this work is its focus on the word *gentlemen*, that for Cooper means: "One elevated above the mass of society by his birth, manners, attainments, character and social condition" (JFC *Democrat* 1931: 112). Although Cooper never declared openly that *gentlemen* had the task of establishing linguistic usage, he firmly claimed that only those who belonged to this social class succeeded in avoiding the most common errors. For Cooper the duty of regulating and improving American English had to be given to "educated gentlemen of the middle states" (JFC *Gleanings* 1982: 28), a task that he himself attempted to carry out in his publications (Schachterle 2011: 68).

3. Linguistic representations in *The Last of the Mohicans*

3.1 Between realism and idealization

In *Cooper's Eloquent Indians* (1956), John T. Frederick starts his analysis underlining how the language created by Cooper in his novels has often been the object of heated debate. For example, in 1828, General Lewis Cass criticized Cooper in an article about the study of Indian languages. After having quoted about 20 figurative expressions from *The Last of the Mohicans* and *The Prairie*, Cass declared that "This is not the manner in which Indians talk, nor is it the manner in which any people talk" (Cass 1828: 374). Later, in relation to the language used by Cooper for his Native American characters, Cass concluded that in his opinion "[Cooper's] Uncas, and his Pawnee Hardheart… have no living prototype in our forests. They may wear leggins and moccasins, and be wrapped in a blanket or a buffalo skin, but they are civilized men, and not Indians" (Cass 1828: 376). Another critic, William Josiah

Snelling, who had first-hand experience of Native Americans, challenged the authenticity of the figurative language used in the autobiography of the Indian chief Black Hawk, for which he blamed Cooper's novels:[5]

> The only drawback upon our credence is the intermixture of courtly phrases, and the figures of speech, which our novelists are so fond of putting into the mouths of Indians […]. The term pale faces, often applied to the whites in this book, was, we think, never in the mouth of any American savage, excepting in the fanciful pages of Mr. Cooper. There are many more phrases and epithets of the like nature, and we only mention them, because we think it time that authors should cease to make Indians talk sentiment. (Snelling 1835: 69, 70).

Another famous critic who doubted the authenticity of Cooper's characters was Francis Parkman, who affirmed that "[Cooper's] Indian characters […] it must be granted, are for the most part either superficially or falsely drawn" (1852: 150). Parkman continued his accusation claiming that:

> The long conversations which he puts into their mouths, are as truthless as they are tiresome. Such as they are, however, they have been eagerly copied by a legion of the smaller poets and novel writers; so that, jointly with Thomas Campbell, Cooper is responsible for the fathering of those aboriginal heroes, lovers and sages, who have long formed a petty nuisance in our literature. (Parkman 1852: 150).

Also Georg Fridén reinforced the idea that Cooper's Indians were idealized: after comparing them with Byron's pirates and Ossian's Celtic heroes, he declared that "all these figures are but phases of the same romantic movement. Cooper saw his Indians in the light of romantic idealism. Cooper's Indian rhetoric is a poetic creation and not the speech of living men" (Fridén 1949: 55).

Nevertheless, Cooper himself proclaimed the authenticity of his representation in the preface to *The Last of the Mohicans*, where he wrote that "the reader who takes up these volumes in expectation of finding an imaginary and romantic picture of things which never had an existence will

[5] Snelling added that these figures of speech are "to be attributed to the bad taste of Black Hawk's amanuesis", who was probably influenced by Cooper's novels. (Snelling 1835: 69).

probably lay them aside, disappointed. The work is exactly what it professes to be in its title-page – a narrative" (JFC *Mohicans* 1998: 3). In addition, many scholars defended the author by placing his novel in the social context in which he lived and by considering the Indians' image that can be found in the sources that influenced Cooper. For example, James F. Beard stood in favour of Cooper's representation:

> Though Cooper seems never to have prepared a systematic list of readings, the extraordinary assimilation of information displayed in his fiction suggests that his knowledge of Indians was as full and authentic as discriminating study of the printed sources of his time allowed. (Beard 1983: xviii).

Even so, according to Frederick, it seems almost impossible to find a universal answer "to the broad question of whether Cooper's Indians are portrayed realistically or are idealized" (Frederick 1956: 1005), because it will probably always depend on a subjective judgement.

However, Cooper's figurative language can be examined objectively, and Frederick himself offers a comparison between the metaphorical expressions in Cooper's novels and those ascribed to the Indian speakers in the sources used by the author, in order to demonstrate that Cooper neither invented creatively nor copied from European authors; on the contrary, he "followed his sources with extraordinary fidelity" (Frederick 1956: 1005).

An essential piece of information for this type of study is the testimony of Cooper's daughter, Susan Fenimore Cooper. In *The Cooper Gallery* (1865), she guaranteed that her father tried to give a particular authenticity to the representation of his Indian characters and affirmed that "the earlier writers on these subjects, Heckwelder, Charlevoix, Penn, Smith, Elliott, Colden, were studied. The narratives of Lang, of Lewis and Clarke, of Mackenzie, were examined" (S.F. Cooper 1865: 129).

The list given by Susan Cooper was probably not complete, because, on the basis of the books that she mentioned, it could be supposed that he also read other works that were equally accessible and famous (Frederick 1956: 1006). Among such texts we can find the accounts of Alexander Henry, John Long and John Bradbury, the works of James Buchanan and Joseph Doddridge, and the official reports of Jedidiah Morse and Henry Rowe Schoolcraft. In all those works Cooper could find references to the figurative language of Native Americans, as in the work of Roger Williams, *A Key into the Language of North America*, where the author reported that "similitudes

greatly please them" and gave many examples of figure of speech that were typical of Indian speakers (Williams 1936: 132).[6]

Similarly, also John Eliot and Charlevoix commented on the Indians' figurative language affirming that "under such metaphorical language they usually expresse what eminent things they meane" (Eliot 1834: 44) and that "all these nations have some what of the Asiatic genius in their discourse, which gives it a figurative turn and expression" (Charlevoix 1923: I, 286). Charlevoix also observed that "they use a great many allegories and other figures" (1923: II, 76).

Colden, however, was critical of this abundant use of metaphorical expressions, affirming that "[the Indians'] Speeches abound with Metaphors, after the Manner of the Eastern Nations […] For the Indians having but few Words, and few complex Ideas, use many Metaphors in their discourse" (Colden 1922: 36). In addition, both Colden and Charlevoix complained about the inability of the interpreter to reproduce Native American figurative language with any accuracy. Heckewelder also provided examples of Indian use of figurative expressions; however, he found such usage excessive, while excusing it at the same time:

> The Indians are fond of metaphors. They are to their discourse what feathers and beads are to their persons, a gaudy but tasteless ornament. Yet we must not judge them too severely on that account […]. Even in enlightened Europe, many centuries have not elapsed since the best and most celebrated writers employed this figure in a profuse manner […] the immortal Shakespeare, himself, did not disdain it. (Heckewelder 1971: 137)

Despite the presence of these reliable accounts of the frequent use of metaphorical expressions in Indian languages, Cooper did not feel justified in inventing these expressions freely. For Frederick, if we compare the metaphorical expressions employed by Cooper and those attributed to Indian speakers in his sources, we can observe that more than half of them coincide, while the others are created by following models provided in the sources (Frederick 1956: 1009).

---

[6] The thirty-two chapters of this book provide a long list of recurring Indian expressions translated into English. These expressions refer to different themes, from religion to the natural and animal world, from relations to war and death. Some of these expressions are used by Cooper in his novels, e.g. *squaw* (Indian woman), *moccasin* (type of shoes) and *Manittó* (Gods).

As Cooper himself noticed, the Indian speaker "draws his metaphors from the clouds, the seasons, the birds, the beasts, and the vegetable" (JFC *Mohicans* 1998: 7). The most varied group of the expressions used by the author pertains to the animal world of the forests and the prairies, which played an essential part not only in Native American economy, but also in its culture. In particular, Cooper employed images of animals in metaphors and similes in order to represent human qualities or actions. In *The Prairie* an expert warrior is described as a "grizzly bear in combat" (JFC *Prairie* 1877: 365). This example relies on Colden's work, where the author quotes an exhortation of an Indian chief to his warriors, who are encouraged to behave like bears, because they must never give up in spite of all the difficulties. As a matter of fact, Native Americans respected bears on account of their tenacious bravery (Frederick 1956: 1009).

Another example can be found in the accounts of the Lewis and Clark expedition, in which it is reported that during a period of famine some Indians complained of having to "live like bears on roots and berries" (Frederick 1956: 1009). In addition to the bear and the more general "beast", Cooper's sources clarify his figurative references to other animals, such as the buffalo, the deer, the dog, the moose, the fox, the pig, the rabbit, the wolf, the puma and the marmot. The figure of the dog is undoubtedly the most frequent image employed by Cooper, especially to denigrate and express contempt, and it often appears in *The Last of the Mohicans*.[7] Among the most frequent expressions we can find: "dog of the palefaces […] go yell among the curs" and "paleface cur […] with tail between his legs", both taken from the novel *The Deerslayer* (1841) (JFC *Deerslayer* 1910: 327, 338). For these figures Cooper was inspired by the reports of Lewis and Clark, where we find the following simile: "White men are like dogs, the more they're beaten, the better they act" (Frederick 1956: 1009); and by those of General Cass, who reported the words of a Native American who described a white general who was retreating saying that he was "like a dog running off with his tail between his legs" (Cass 1826: 99n).

With regard to the ornithological field, Cooper allowed himself more freedom of imagination than in any other category. He preferred the image of the eagle to indicate ferocity, courage or a sharp eye, and he often used a particular type of bird, the wren, to describe the melodic voices of his heroines.[8] In addition, birds are a frequent metaphor of bearers of (often

[7] It appears six times in *The Last of the Mohicans*, four times in *The Deerslayer*, three times in *The Wept of W ish-ton-Wish*, and once each in *The Prairie* and *Satanstoe*.

[8] The figure of the eagle is used in all five of the *Leatherstocking Tales* and also in *The Wept of Wish-ton-Wish* and *The Redskins*, for a total of seventeen occurrences.

bad or false) news both in Cooper's novels and in his sources. Heckewelder described the expression "singing birds" as the typical figure to indicate "tale bearers or liars" (Heckewelder 1971: 138). In *The Last of the Mohicans* we can find expressions like "singing birds have opened their bills" – i.e., fake news is spreading – while Magua, the evil character in the novel, is described as "a singing bird" – that is, a liar (JFC *Mohicans* 1998: 347, 339).

Cooper also used some rhetorical figures taken from the vegetable world, like trees, flowers or leaves; the most recurring one is "as many as the leaves on the trees", often employed to refer to white settlers. From the natural world he also took the image of the clouds to represent mistakes, suspicions and misunderstandings, or the image of the sun to indicate friendship, peace and good faith (Frederick 1956: 1011).

The most common metaphorical expressions, present both in Cooper's novels and in his sources, are those in which a sense organ represents a mental ability and those in which a part of the body is used to describe someone's qualities or traits. These figures are also common in English, especially those relating to the heart (e.g. *good-hearted*, *a hearty welcome*, *put one's heart into sth*). Nevertheless, the most frequent figure is the one of the ears, that can indicate listening, comprehension or an agreement between people, like in "Your ears are ever open to slanderous reports" (Colden 1922: II, 107) or in "Why do you wish to stop my ears?" (JFC *Mohicans* 1998: 339).[9]

Other recurring expressions refer to war and fighting, such as *to bury the hatchet* and *to go on the warpath*, expressions that have since become part of ordinary English usage. Many figures also refer to blood and also in this case most of them are still in use (e.g. *hot blood*, *blood chilled by age*, *bloody hands*, *a spot stained with blood*); a very frequent simile based on this image compares white people's blood to *spring water*, as Native Americans thought that white people's blood was as pale as the colour of their skin. An example can be found in the reports of John Long, where a typical Indian expression is quoted about the white man's veins: in the Natives' view, they "run clear like the sea" (Long 1922: 140).

Another Native linguistic feature that Cooper borrows is the use of *woman* or *old woman* "in derogation and contempt" (Frederick 1956: 1013).[10] This usage is frequent in the Native American speeches reported by Mackenzie, Bartram, Heckewelder. Bradbury, Colden, Charlevoix and Long. Similarly, although less frequently, Native Americans also expressed

[9] The figure of the ears appears fifty-three times in eight novels.
[10] It appears thirty-two times in eight novels.

contempt through a comparison between adults and children; for example, John Eliot said that one Indian accused another of having asked "a papoose question" (Eliot 1834: 47) – i.e., a childish question, while Winslow reported the speech of a Native American who stated that white men "died crying, making sour faces, more like children than men" (Winslow 1624: 206). In addition, Schoolcraft quoted the words of an Indian chief who admitted that "sometimes the Indians have acted like children" (Schoolcraft 1825: 365).

According to Frederick, "the recurring charge that Cooper idealized and falsified his red men in this respect – that their eloquence was the product of his own imagination, or the effect of trans-Atlantic literary influence – is contrary to the facts" (Frederick 1956: 1017). Indeed, as this analysis has shown, Cooper never used his imagination freely in his representation of Native American languages, nor did he imitate the European romantic writers of his time. On the contrary, he acquired the necessary information to represent his Native Americans from the most reliable sources of his time and followed them closely in his linguistic representations. At the same, he employed such representations to reinforce the model of a linguistic hierarchy concerning both Native and Euro-American languages.

3.2 Cooper's fiction of language: The pure language of Native Americans

In relation to early nineteenth-century ideology concerning languages and their perceptions, Blakemore (1984) claims that in *The Last of the Mohicans* Cooper "imagined another language – a pure, unfallen language that would embody the new American reality" (1984: 21), but above all he imagined a prelapsarian language. Blakemore (1984) underlines how Cooper identified two types of languages: those that precede and those that follow the original sin. In the former group he places the Edenic language of the Delaware, while in the latter we find the corrupt Indian and European languages – in particular, French and English. In order to clarify this concept, we can refer to Steiner (1975), who summarizes the ancient theory according to which a pure language was in existence before the original sin, but after the expulsion of man from the Garden of Eden, it divided itself into many different languages as a result of man's imperfection, as shown in the biblical story of the tower of Babel.[11]

[11] According to this theory, the prelapsarian language enabled all men to understand one another and contained a divine syntax analogous to God's own diction. Being of

Among the sources mentioned in Section 3.1 above, John Heckewelder's *History, Manners, and Customs of the Indian Nations* (1818) is the text to which Cooper resorted most frequently, and it is from this source that Cooper took some of his theories about Native American languages. For example, Cooper described the Delaware as noble savages on the basis of Heckewelder's work, where Native Americans are presented as the original people and their language is considered the ancestor of all Indian languages. Another theory that reinforces this idea was outlined by Duponceau, who suggested that Delaware was the original language of Adam and Eve, "first taught to mankind by the great author of all perfection" (Heckewelder 1971: 406). Other ideas of Heckewelder's appear in *The Last of the Mohicans*, such as the superiority of the Delaware language over the corrupt language of the Huron and the sorrow for the impossibility of translating the beauties of that language into European languages. Cooper often stressed these concepts, but Blakemore argues that his support was futile after all: "his fiction of language is finally resolved in the writing out of a myth that inevitably ends with the destruction of the Indian world and the language that crystallized it" (Blakemore 1984: 22).

Nevertheless, in his novel Cooper seems to have an ambivalent attitude towards his own language, because it represents the legacy of an old and already fallen world. In addition, he blamed the end of the Indian world on the corruption of their languages by white settlers. As regards European languages, Cooper created a linguistic hierarchy in which the post-Edenic languages of the Euro-American world, in this case English and French, are considered inferior to the pure language of the Delaware. However, in this scheme he added that English is clearly a superior language among the post-Edenic ones, because it can be used as a mediator between the Delaware and the French.

In addition, in *The Last of the Mohicans* we can find a link between the expropriation of Native American lands and the expropriation of the languages of the peoples who lived there. For example, this connection is demonstrated in the 1850 preface, where Cooper comments on the Indian name "Horican", the original Indian name of "Lake George". In this passage the author informs readers of having extracted this name from some ancient maps of that area, where "it was ascertained that a tribe of Indians, called Les Horicans by the French, existed in the neighborhood of this beautiful

direct divine origin, this Edenic language was consistent with reality and words and meanings matched perfectly. (Steiner 1975: 58, 59).

sheet of water". Cooper explains that this name was first replaced with the French name *Lac du Saint Sacrement* and then with the English name "Lake George" – see (7) below:

(7) Its waters were so limpid as to have been exclusively selected by the Jesuit missionaries to perform the typical purification of baptism, and to obtain for it the title of lake "du Saint Sacrement." The less zealous English thought they conferred a sufficient honor on its unsullied fountains, when they bestowed the name of their reigning prince, the second of the house of Hanover. The two united to rob the untutored possessors of its wooded scenery of their native right to perpetuate its original appellation of "Horican." (JFC *Mohicans* 1998: 15-16)

Although Cooper criticized the original Indian name, saying it was unpronounceable, Blakemore (1984: 23) argues that the author was also very critical of the new names, because they represented the foreign corruption of the lake's intrinsic reality that was connected to the Indians. According to Blakemore, in this passage there is an implicit connection between the expropriation of the land and the deprivation of the Indians' right to name their land. Therefore, for Cooper, every time the colonists changed the name of the land, the land modified itself because it did not represent the Indian reality any longer. After having lost their land, they also lost their right to represent their reality and history, a right that was claimed by the white settlers.

Another theory discussed by Cooper in the 1826 preface is the link between the decline of the Indians' language and the end of their world. For the writer the confusion of Indian history was a result of the fragmentation of the primeval Native American language into different corrupt dialects that were made even more complicated by European names. For example, European colonists renamed the many Indian tribes in different ways, as the author explains:

(8) When it is remembered that the Dutch, the English, and the French, all gave appellations to the tribes that dwelt within the country which is the scene of this story, and that the Indians not only gave different names to their enemies, but frequently to themselves, the cause of the confusion will be understood"
(JFC *Mohicans* 1998: 8).

The introduction of post-Edenic names creates confusion in the historical reality of the Native Americans and Cooper represents this chaos of names

and languages in his novel. In Chapter 7, the author explains that "the Indians rarely use the same name when different tribes speak of each other" (JFC *Mohicans* 1998: 79). In Chapter 13, Cooper introduces the impossibility of communication among Indian tribes by describing a conversation between some Huron as "unintelligible" for the Delaware (JFC *Mohicans* 1998: 150). Later, in Chapter 19, the narrator summarises the themes that were discussed in the 1826 and 1831 prefaces:

(9) The confusion of nations, and even of tribes, to which Hawk-eye alluded, existed at that period in the fullest force. The great tie of language, and, of course of a common origin, was severed in many places; and it was one of its consequences, that the Delaware and the Mingo (as the people of the Six Nations were called) were found fighting in the same ranks, while the latter sought the scalp of the Huron, though believed to be the root of his own stock. The Delawares were even divided among themselves. (JFC *Mohicans* 1998: 224)

For his representation of the fragmented Indian language Cooper drew inspiration from the model of Babel. Nevertheless, although he describes the results of this division, he also affirms that the pure and incorrupt language of the Native Americans still existed in 1757. In order to represent Native American languages, the author constructs a double linguistic fiction: a corrupt language that is the result of the Indian languages' chaos and an original Indian language that is still spoken by the Delaware and the Mohican. For Cooper, Delaware is so superior and prestigious compared to the other languages that even Europeans cannot speak it. In the novel the superiority of this language is underlined in the scenes where Hawk-eye and the two Mohican communicate by using only the Delaware language. In these passages they move away from the other white people in order to discuss the situation in their language. For Blakemore "this physical separation is an objective correlative for the linguistic separation" (1984: 28), thus drawing attention to the fact that the main characters of this novel live in other linguistic and semantic spaces. When the characters have to discuss urgent and important matters, they use Delaware, creating the impression that this language is linked to important issues of life and death. Therefore, in the novel this language becomes the superior language of those who have been "initiated into the nuances of the forest and the mysteries of the Indian world" (Blakemore 1984: 29), distinguishing them from white "novices".

3.3 Linguistic hierarchy in *The Last of the Mohicans*

As already mentioned before, English is presented by Cooper as one of the post-Edenic languages that were inferior to the pure language of the Delaware; nevertheless, in his linguistic hierarchy, English is at an intermediate level between Delaware and French. As a matter of fact, he could not exaggerate his attacks against his own language, because he was writing for an English-speaking audience. For this reason, Cooper decided to compare English to a more corrupt language that symbolized the imperialistic ambitions of the Catholic French, described as "the white tempters in the garden of the new world" (Blakemore 1984: 30).

Throughout the novel French is presented as the demonic language that is associated with deceit and pedantry. The depiction of French and French Catholicism as corrupt and devious was linked to a deep-rooted ideology. Cooper probably took inspiration from John Milton's *Paradise Lost* (1667), because his description of French as the language of the Huron's manipulation and corruption (especially of Magua) recalls Satan's use of language in order to seduce and manipulate both the fallen and the unfallen in Milton's work (Blakemore 1984: 40). Although Cooper attacked this language implicitly, he never showed the superiority of English over French, but in some passages he just proclaimed English as the only authentic white language. In the novel, when Chingachgook asks Natty to speak with his "white brothers" (Heyward and the rest of the group), the latter answers: "That will I, and in English that the king needn't be ashamed to answer" (JFC *Mohicans* 1998: 42). According to Blakemore, although in this passage there are possible ironic nuances, it describes English as "the native language of the novel's white hero, and hence the esteemed white language of the novel" (Blakemore 1984: 30). During the scene in the Huron's cave, even though Hawk-eye describes English as "the genuine tongue of a white-skin" (JFC *Mohicans* 1998: 298), he cannot speak because his language would betray him; therefore, Heyward is obliged to speak French, or better, using Hawk-eye's words, he has to communicate with the Huron in their "jargon".

Natty's negative comments on French are frequent in *The Last of the Mohicans*; for example, in a scene he confuses French with the language of the dead because, finding himself in front of a French sentinel, he asks Heyward: "What says it? It speaks neither Indian nor English!" (JFC *Mohicans* 1998: 155). Natty probably recognized the sentinel's language; however, with this comment Cooper wanted to reinforce the inferiority of this language compared to English and Delaware. Also in the two conversations between

Heyward and Montcalm, French appears as the language of deceit and treason. At the beginning both characters speak French, but as soon as they try to confuse and deceive one another, "the discrepancy between what they say and what they mean contributes to the reader's sense that French is the language of rhetorical manipulation and trickery" (Blakemore 1984: 31). Colonel Munro also judges French negatively, complaining that it presents "a Jesuitical way of telling a man his misfortunes!" (JFC *Mohicans* 1998: 171). However, his judgement is more evident in the scene where he hesitates to listen to the French message written by Montcalm and says to Heyward: "Your mother was the only child of my bosom friend, Duncan; and I'll give you a hearing, though all the knights of St. Louis were in a body at the sally-port, with the French saint at their head, crying to speak a word under favor" (JFC *Mohicans* 1998: 178). This comment represents a rejection not only of the French message, but also of the French language in general, because it is characterized by that ambiguity, deceit and imperialistic ambition that the English novel purports to contrast with opposing values: simplicity, truth and honesty. Blakemore suggests that, even though Montcalm talks about chivalry and humanity, "his language is revealed to be vague and vacuous, full of glittering but hollow phrases" (Blakemore 1984: 31), as in the following example; while speaking to Heyward, Montcalm says: "Your commandant is a brave man, and well qualified to repel my assault. Mais, monsieur, is it not time to begin to take more counsel of humanity, and less of your courage? The one as strongly characterizes the hero as the other" (JFC *Mohicans* 1998: 174). With his critical comment on French, Cooper wanted to represent the French world view, "in which the specious words of French chivalry camouflage the sordid facts of French realpolitik" (Blakemore 1984: 32).

Furthermore, French is also the language that is borrowed by the devilish Huron, the "villains" of this novel. In the following passage Hawk-eye presents the Huron in this manner:

(10) I call them Iroquois, because to me every native, who speaks a foreign tongue, is accounted an enemy, though he may pretend to serve the king! If Webb wants faith and honesty in an Indian, let him bring out the tribes of the Delawares, and send these greedy and lying Mohawks and Oneidas, with their six nations of varlets, where in nature they belong, among the French! (JFC *Mohicans* 1998: 59)

In this further attack, Natty also condemns the tribes that employ French considering them as fallen and corrupt. Therefore, in *The Last of the Mohicans*,

language allows the reader to distinguish the good Indian tribes (Delaware and Mohican) from the bad ones (Huron). An appropriate example is represented by the evil character of the novel, Magua, who calls himself *Le Renard Subtil*: "'Tis the name his Canada fathers have given to Magua,' returned the runner, with an air that manifested his pride at the distinction" (JFC *Mohicans* 1998: 49). In this passage the corruption of Magua is reinforced by the fact that he voluntarily and proudly adopted a French name, abandoning his original name and therefore his Indian identity. Cooper continues his attack on French in the discussion between Heyward and Magua on the meaning of *Le Cerf Agile*, the French name given by Magua to Uncas:

(11) "'Le Cerf Agile' is not here?"
"I know not whom you call 'The Nimble Deer'," said Duncan gladly profiting by any excuse to create delay.
"Uncas," returned Magua, pronouncing the Delaware name with even greater difficulty than he spoke his English words. "'Bounding Elk' is what the white man says, when he calls to the young Mohican."
"Here is some confusion in names between us, Le Renard," said Duncan, hoping to provoke a discussion. "*Daim* is the French for deer, and *cerf* for stag; *elan* is the true term, when one would speak of an elk."
"Yes," muttered the Indian, in his native tongue; "the pale faces are prattling women! they have two words for each thing, while a red-skin will make the sound of his voice speak to him." Then, changing his language, he continued, adhering to the imperfect nomenclature of his provincial instructors. "The deer is swift, but weak; the elk is swift, but strong; and the son of 'Le Serpent' is 'Le Cerf Agile'".
(JFC *Mohicans* 1998: 104, 105)

In this scene Cooper implicitly shows the problem of Native American corruption on the part of European settlers, underlining ironically the fact that Magua has more difficulty in pronouncing the Indian name 'Uncas' than when he speaks English. Moreover, his speech is contradictory and disorganized: at first he attacks the European languages in his mother tongue, but immediately after that he starts again to speak in French, "adhering to the imperfect nomenclature of his provincial instructors". In addition, the debate about the meaning of the French words has the purpose of intentionally confusing Magua, in order to buy Heyward more time. Heyward's strategy is to start a conversation by deliberately mistranslating *Le Cerf Agile* as agile *deer*, although more precisely it means an agile male *stag*. This confusion is further increased by the corrupt French of Magua,

who thinks that Uncas' English epithet, *Bounding Elk*, is the same in the French translation, *Le Cerf Agile*. Another example of an intentionally wrong translation is given when Heyward deliberately fails to distinguish between the English name given by Hawk-eye to Uncas (Hawk-eye cannot speak French) and his French alias.

French is also the language used by both Native Americans and white people when they have to communicate, because neither group knows the mother tongue of the other and the white colonists rely on the Native Americans' knowledge of French. In the second part of the novel, when Heyward is in the Huron' village, their language is "unintelligible" for him; for this reason, he asks in French: "Do none of my brothers speak the French or the English?" (JFC *Mohicans* 1998: 266). In answer to this question a Huron warrior asks provocatively "in the language of the Canadas: 'When our Great Father speaks to his people, is it with the tongue of a Huron?'". Disguised as a French doctor, Heyward answers in an evasive manner, using the language of deceit and saying: "He knows no difference in his children, whether the colour of the skin be red, black, or white" (JFC *Mohicans* 1983: 266). According to Blakemore, Heyward moves the conversation cleverly "from a linguistic to a racial context in order to placate and distract the Hurons" (Blakemore 1984: 34). In addition, he continues his deception drawing the Huron' s attention to his own artificially painted skin – see (12):

(12) "When an Indian chief comes among his white fathers," returned Duncan, with great steadiness, "he lays aside his buffalo robe, to carry the shirt that is offered him. My brothers have given me paint and I wear it". (JFC *Mohicans* 1998: 267)

Heyward cleverly succeeds in overshadowing the question about language, "by shifting the category from language, which is 'internal' and intimate, to body paint – the 'external' sign of his deceptive disguise" (Blakemore 1984: 34), and by using French, the language of deceit, that contributes to the linguistic fragmentation of the Native Americans.

4. Concluding remarks

As suggested by Blakemore, in *The Last of the Mohicans* the Edenic village of the Delaware can be seen as a "linguistic microcosm" of the problems discussed in the novel: "chaotic communication, corrupt French, and the impossibility

of translation" (Blakemore 1984: 35). For Cooper, Delaware is a pure and natural language; nevertheless, it only appears in the novel when Uncas ends the chaos of foreign languages by affirming that he will talk "like his fathers with the tongue of a Delaware" (JFC *Mohicans* 1998: 346). However, Cooper does not provide a representation of the Delaware language: on the contrary, he emphasizes the death of this pure and ineffable language.

In addition, the author turns the Delaware's village into a model of Babel, where a language is destroyed by such corrupt, foreign languages as French, English and Huron. Already in 1826, with his representation of the end of a population and its languages and the resulting expansion of white civilization, Cooper predicted the future of America. In this linguistic chaos the author celebrates an Edenic Indian language and at the same time he regrets its destruction. However, at the end, he provides only one possible historical conclusion: "the prelapsarian Delaware tongue disappears with the last of the Mohican, while the extant fallen tribes speak only babel" (Blakemore 1984: 36). Finally, in this linguistic chaos, English imposes itself as the dominant language, understood by everyone, in a world where French and Dutch have been defeated and Native Americans have been decimated. If we read *The Last of the Mohicans* in this context we can affirm that Cooper depicted the fight for the new world also as a linguistic battle between Indian and European languages.

REFERENCES

Sources

Cooper, J.F.

1823 [1980] *The Pioneers; or the Sources of the Susquehanna, A Descriptive Tale.* Edited by J.F. Beard – L. Schachterle – K.M. Anderson. Albany, NY: SUNY Press.

1826 [1998] *The Last of the Mohicans. A Narrative of 1775.* Edited by J. McWilliams. Oxford: Oxford University Press.

1827 [1877] *The Prairie; a Tale.* New York: Hurd & Houghton.

1837 [1982] *Gleanings in Europe: England*. Albany, NY: SUNY Press.

1838 [1931] *The American Democrat.* New York: Alfred A. Knopf.

1841 [1910] *The Deerslayer; or The First War-Path.* Boston: Houghton, Mifflin.

1845 [1912] *The Chainbearer, or The Littlepage Manuscripts.* New York: Putnam.

1845 [1990] *Satanstoe, or The Littlepage Manuscript: A Tale of the Colony*. Albany, NY: SUNY Press.

Special studies

Bates, C.G.
2015 *The Early Republic and The Antebellum America: An Encyclopedia of Social, Political, Cultural, and Economic History.* New York: Routledge.
Beard, J.F.
1983 "Historical introduction". In: J.F. Cooper et al. *The Last of the Mohicans. A Narrative of 1757*. Albany, NY: SUNY Press, xv-xviii.
Blakemore, S.
1984 "Strange tongues: Cooper's fiction of language in *The Last of the Mohicans*", *Early American Literature* 19, 21-41.
Cass, L.
1826 "Indians of North America", *North American Review* XXII (January), 53-119.
1828 "Structure of the Indian languages", *North American Review* XXVI (April), 357-403.
Charlevoix, P.F.-X. de
1761 [1923] *The Journal of a Voyage to North America*. Edited by L.P. Kellog. Chicago: The Caxton Club.
Colden, C.
1727 [1922] *History of the Five Indian Nations.* New York: Allerton Book Co.
Cooper, S.F.
1865 *The Cooper Gallery; or Pages and Pictures from the Writings of James Fenimore Cooper*. Philadelphia: J.B. Lippincott.
Eliot, J.
1648 [1834] *The Clear Sun-shine of the Gospel Breaking forth upon the Indians in New-England.* Cambridge: Collection of Massachusetts Historical Society.
Encyclopedia Britannica
https://www.britannica.com/biography/Francis-Parkman, accessed September 2019
Frederick, J.T.
1956 "Cooper's eloquent Indians", *Proceedings of Modern Language Association* LXXI (December,) 1004-1017.
Fridén, G.
1949 *James Fenimore Cooper and Ossian*. Uppsala: Lundequistka Bokhandeln.
Heckewelder, J.G.E.
1818 [1971] *History, Manners and Customs of the Indian Nations.* Edited by W.C. Reichel. New York: Arno Press.
Johansen, B.E. – B.M. Pritzker (eds.)
2008 *Encyclopedia of American Indian History*. Vol. 1: *Chronological Essays. Issues in American Indian History. Events in American Indian History*. Santa Barbara, CA: ABC-CLIO.
Long, J.
1791 [1922] *John Long's Voyages and Travels in the Years 1768-1788.* Edited by M.M. Quaife. Chicago: R.R. Donnelley & Sons Company.

Maddox, L.
1991 *Removals: Nineteenth-Century American Literature and the Politics of Indian Affairs*. New York: Oxford University Press.
Parkman, F.
1852 "The works of James Fenimore Cooper", *North American Review* XXIV (January), 147-161.
Schachterle, L.
2011 "James Fenimore Cooper on the languages of the Americans: A note on the author's footnotes", *Nineteenth-Century Literature* 66 (1), 37-68.
Schoolcraft, H.R.
1825 *Travels in the Central Portion of the Mississippi Valley*. New York: Collins u.a.
Simpson, D.
1986 *The Politics of American English: 1776-1850*. New York: Oxford University Press.
Snelling, W.J.
1835 "Life of Black Hawk", review of *Life of Ma-Ka-Tai-Me-She-Kia-Kiak, or Black Hawk*, by Black Hawk, *North American Review* XL (January), 68-87.
Steiner, G.
1975 *After Babel: Aspects of Language and Translation*. London: Oxford University Press.
Williams, R.
1643 [1936] *A Key into the Language of North America*. Bedford, MA: Applewood Books.
Winslow, E.
1624 *Good News from New England, or, A True Relation of Things Very Remarkable at the Plantation of Plymouth in New England*. London: Printed by I.D. [John Dawson] for W. Bladen and J. Bellamie.

Address: Annalisa Scatà, Dipartimento di Lingue, Letterature e Culture Straniere, Università degli Studi di Bergamo, Piazza Rosate 2, 24129 Bergamo, Italy.
ORCID code: orcid.org/0000-0001-7609-0979.

Overlexicalization and semantic variation in the Early Modern English naming of Native Americans

Elisabetta Cecconi
University of Florence

ABSTRACT

This article takes as its focus the overlexicalization and semantic variation of reference terms for Native Americans in the period of the English exploration, settlement and colonization of Virginia (1584-1724). The corpus-assisted discourse analysis of pamphlets, first-hand accounts and letters taken from the *Virtual Jamestown Digital Archive* reveals that the lexical items used for naming the Natives undergo a process of amelioration and pejoration throughout the decades in relation to the changing historical and socio-cultural context in which they are used. An investigation of the definition and quotations of the same lexical items in the online version of the *OED* shows that although the dictionary attests the overlexicalization occurring for the Native Americans and reflects the most frequent lexico-syntactic patterns in which the words are found in the corpus, the choice of the quotations fails to account for instances of semantic variation and for the resulting ambivalent connotations of the terms.

Keywords: semantic variation, Native Americans, Early Modern English, corpus-assisted discourse analysis, *OED*.

1. Introduction

The proliferation of Early Modern English accounts of expedition and settlement in North America contributed to the construction of an ideology of colonization which was intended to elicit moral approval and financial support from the English readers at home (Borch 2004: 6; MacMillan 2013: 85). In order to this, authors were careful to construe the socio-cultural identity of the Native Americans in a way which could legitimize their own

DOI: 10.25951/4362.

right to settlement. Different – and at times ambivalent – representations of the Native Indians circulated in print, depending on the socio-economic and historical circumstances in which the texts were produced in the long course of the dispossession process (Jennings 1975; Fitzmaurice 2003; Moran 2007). When the Natives did not threaten the British interests but rather contributed to the process of profitable trade, they were represented as objects of curiosity and study, occasionally praised for their ingenuity and hospitality, more generally patronized for their cultural inferiority. Whenever they attempted to hinder the English expansion, they were represented as brutal savages and treacherous individuals.[1] The issue of the Native Americans and their treatment was highly controversial as it risked dividing public opinion. In linguistic terms, this anxiety was manifested in the abundance of quasi-synonymous reference terms for naming them. Such over-lexicalization indicates the contrasting concerns of society in the context of an emerging culture of colonization.

In this article, I shall conduct a corpus-assisted discourse analysis of first-hand accounts, letters and tracts written by explorers and settlers of Virginia in the period from 1584 to 1724 in order to identify the most common lexical items used to refer to the Native Americans and analyse their ideological bias in context. More precisely, I shall focus on features of semantic variation resulting from the changing socio-cultural and political variables throughout the colonization process, from the early years of exploration of the Newfoundland to the governmental policy of the state of Virginia at the beginning of the 18th century. Aspects of semantic variation will be assessed along the lines of pejoration and amelioration (Gramley – Pätzold 1992: 211). In the course of the analysis, I shall consider the extent to which the semantic changes found in the corpus are adequately reflected in the definition and examples in the *OED* of the same words during the same historical period. Indeed, the lexicographer's choice of quotations and sources unravels the way in which culture and society are represented both then and now (Simpson 2002; Brewer 2005).

2. Corpus and methodology

My data are taken from the *Virtual Jamestown Digital Archive* (hence *VJDA*). The corpus contains different text-types: letters, first-hand accounts and

[1] A similar ambivalent attitude towards the Native Americans appears in the 18th- and 19th-century documents. For an analysis of the representation of American Indians in the *Coruña Corpus of Historical English Texts*, see Dossena (2019).

tracts dated from the late 16th century to the beginning of the 18th century recounting the history of the colonization of Virginia. Although in many cases the idyllic representations of the prosperous life in the colony cannot be taken at face value, the texts are no less valuable as they reflect the logic of power and dominium at the basis of a Eurocentric view of the world through which social identities are shaped in discourse (Cecconi 2020). It is worth pointing out that many of these tracts were sponsored by the Virginia Company of London, beginning in 1606, and as a result they responded to the imperatives of profit and commercial interests usually camouflaged under the pretence of a Christian mission to the benefit of the pagan tribes.[2]

The texts will be examined according to the broad principles of corpus-assisted discourse analysis, theorized amongst others by Stubbs (1996, 2001) and Partington (2004, 2009). The methodology combines the usual qualitative approach to the analysis of the text with the quantitative analysis provided by Corpus Linguistics in the attempt to discover previously unnoticed regular patterns and link them to specific societal discourse practices. In this sense I follow Haarman and Lombardo's description of the approach as "a constant movement back and forth between data in the form of concordances, collocations and clusters on the one hand, and, on the other, the contextual information (i.e. the actual texts) retrievable by the software" (2009: 8). In order to achieve a better understanding of the ideological significance of certain lexical choices made by the author, my inquiry will extend beyond the textual context to include considerations concerning the wider socio-cultural and historical contextual matters within which the text has been produced (Pahta – Taavitsainen 2010: 551).

For the purpose of the analysis my database is divided into six sub-corpora covering about 25 years each and referred to as follows: Period 1 (1584-1599), Period 2 (1600-1624), Period 3 (1625-1649), Period 4 (1650-1674), Period 5 (1675-1699) and Period 6 (1700-1724). The segmentation of the corpus is intended to provide an insight into both distribution and semantic change of the lexical items throughout the six periods.

3. Analysis

My starting point was the elaboration of a wordlist from which I selected the first five most frequent words referring to the Native Americans, i.e. *Indian(s)*

[2] The Virginia Company Patent cites the natives as those who "live in darkness and miserable ignorance of the true knowledge and worship of God" and who are needful of "human civility and quiet government" (in MacMillan 2013: 85).

(1,074), *Sa(l)vage(s)* (211), *Inhabitant(s)* (173), *Heathen(s)* (77) and *Native(s)* (62).[3] Below is a graph which shows their quantitative distribution across the six sub-corpora and provides evidence of the complex relationship which exists between choice of descriptors, identity construction and ideological propaganda in specific historical periods.

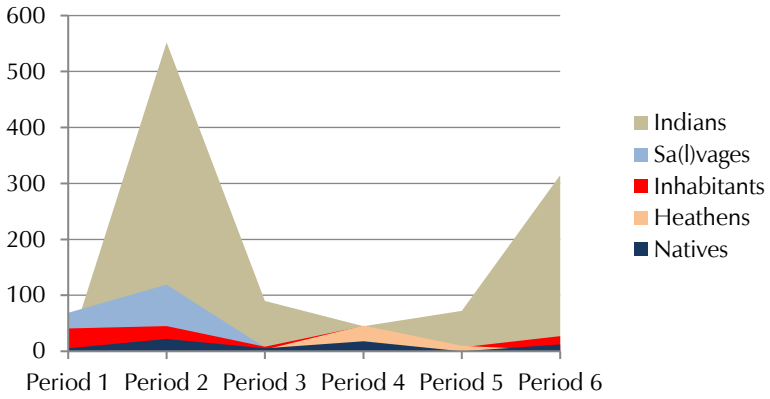

Table 1. Quantitative distribution of the reference terms for Native Americans in the *VJDA* from 1584 to 1724

Table 1 shows two peaks in the use of descriptors for the Native Americans. The first peak covers Period 2 and part of Period 3 when data register the highest frequency and variety of lexical items in the corpus. The decades at issue include the early phase of the settlement, followed by the two Anglo-Powhatan wars (1609-1614 and 1622-1632) ending in the Indians' defeat. In this early phase of exploration and settlement several competing descriptors were used by authors according to their ideological worldview or sensitivity.[4] The second peak occurs at the beginning of the 18th century when the General Assembly of Virginia established issues of law and order for the new colonial society and authors undertook retrospective accounts

[3] The computational analysis is carried out with the aid of the Sketch Engine software. Three major computational tools are used for the analysis: Wordlist, Word Sketch and Concordances.

[4] Contrasting representations of Native Indians coalesce in the pamphlets of the time. Barlowe (1584) used the noble savage motif to portray Native Indians as endowed with a natural nobility. Hariot (1588) portrayed Indians as inoffensive on account of their military weakness. Lane (1586) depicted them as uncivil and untrustworthy (see Moran 2007).

of the British colonization, highlighting the controversial relationship with the Natives.

The noun *Indian(s)* is the most frequent throughout the 6 sub-periods with a peak in 1600-1624 and in 1700-1724. *Sa(l)vages* comes second but its usage is restricted to the years of the early settlement (1584-1599), the starving time (1609-1610) and the first and second Anglo-Powhatan war (1609-1614 and 1622-1632). In Period 4 (1650-1674) there is a drop in the use of *Indians* which is accompanied by increasing occurrences of *Heathen*. The word appears to replace *sa(l)vages*, which falls into disuse by the mid-17th century. The descriptor *Heathen(s)* characterises the period of William Berkeley's government in Virginia (1641-1652 and 1660-1677), when his friendly policies towards Native Americans led to the rebellion of the English colonists headed by the wealthy landlord Nathaniel Bacon (1674-1676). The word *inhabitant(s)* is particularly frequent in the early period of the settlement when the English settlers come into contact with the Natives. Its use lessens in the mid-17th century due to a shift in reference for the word. As the English settlers established their colonies in North America, they became the new inhabitants and the noun changed its original reference from American Indians to English colonists. In Period 6, when the noun is frequently used, it mostly refers to the English people as the Inhabitants of Virginia or Maryland and it is unpremodified. It is only in tracts on the history of colonization that the word refers back to the Native Americans, and in several cases it is premodified by the adjective *primitive* to distinguish it from the new inhabitants. A similar – though less frequent – ambiguity appears to characterise the use of the word *native(s)* in Period 6 caused by the emergence of a new generation of English colonists who had been born in Virginia.

3.1 *Indian(s)*

The most frequent lexical item is *Indian(s)*. Table 2 on the next page shows the quantitative distribution of neutral/positive and negative semantics in the most frequent lexico-syntactic patterns (1584-1724).

Indian(s) occurs mostly as subject of neutral verbs which reveal the colonists' interest in their behaviour as a model on how to survive in the new land. The lexical verbs for which *Indian(s)* is encoded as subject are verbs of movement, action and saying: *come*, *say*, *use*, *see*, *make*, *tell*, *bring*, *set*, *give*, *call*. Some of these verbs (i.e. *give/bring*) entail positive connotations as they are meant to represent the American Indians' friendship and hospitality towards the newcomers.

(1) The Indians came presently down the River: they leaped on shore, and declared to the Governour, That they were subjects of a great Lord, whose name was Aquixo. (*Virginia richly valued*, 1609)

(2) Our provision now being within twentie dayes spent, the Indians brought us great store both of Corne and bread ready made. (*A True Relation of Such Occurrences and Accidents*, 1608)

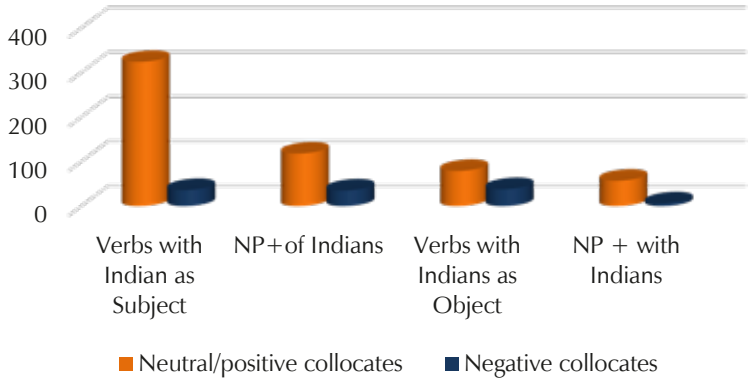

Table 2. Distribution of neutral/positive and negative collocates in the most frequent lexico-syntactic patterns of *Indians* in *VJDA*

Especially in the early phase of exploration and settlement the positive representation of the Indians was meant to reassure investors and travellers of their inoffensiveness and generosity, from which they would be able to make good profit. The second most frequent lexico-syntactic construction is [NP] *of* Indians, where the nouns occupying the [NP] slot vary according to the changing relationship between the Indians and the English settlers as the century progresses. Indeed, while the neutral lexemes in the [NP] slot (e.g. *canoes*, *number*, *cabins*, *language*, *conversion*, *corne*, *fashion*, *minds*) are quantitatively more salient than the negative ones, a closer inspection of the concordances reveals the ambivalent nature of the relationship between the two ethnic groups in moments of crisis. For example, in the context of the Anglo-Powhatan war of 1622 and the starving time which preceded it, the negative lexical items (e.g. *in the hands*, *attack*, *treachery*, *spye*, *conspiracy*, *cruelty/ies*, *dispight*, *spoile*, *ambush*) register an increase as the English settlers need to stress the treacherous character of the Indians in order to justify their Massacre and violent subjection (41%). In the next decades, negative semantics drops considerably in relation to the toleration policy adopted by

Governor Berkeley in the mid-17th century (12.5%) to peak again (68%) in the period of Bacon's rebellion (1675-1699), as shown in table 3.

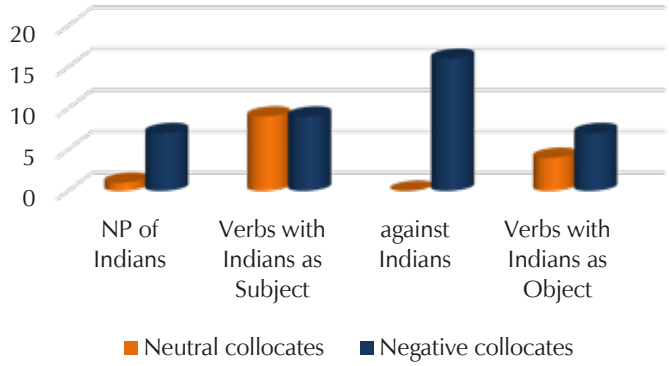

Table 3. The distribution of neutral/positive and negative collocates of *Indians* in the period of Bacon's Rebellion (Period 5)

The most frequent construction sees *Indian(s)* as subject of lexical verbs with negative meaning (e.g. *perpetrate*, *learn (not yet)*, *knock*, *torture*, *torment*, *kill*) or as subject of lexical verbs which – as a result of semantic prosody – acquire a negative connotation in context (*draw in*, *surprise*, *devise*). For the first time the construction *against the Indians* peaks as the pattern *with the Indians* drops. The construction [NP] *of the Indians* features a predominance of negative semantics, as documented by the collocates *cruelty/ies of*, *in the hands of*, *bloody proceedings of*, *destruction of*

(3) these Indians draw in others […] to their aides: which being conjoyned […] they dayly committed abundance of unguarded and unrevenged murthers upon the English, which they perpretrated in a most barbarous and horrid manner. Cruelties of the Indians By which means abundance of the Fronteare Plantations became eather depopulated by the Indian settlers, or deserted by the planters fears, who were compelled to forsake their abodes to find security for their lives; which they were not to part with in the hands of the Indians, but under the worst of torments. (*A Narrative of the Indian and Civil Wars in Virginia*, 1675-1676)

This diachronic change in the distribution of neutral and negative semantics of *Indian(s)* shows how its meaning is dependent on the ideological

imperatives of the time, oscillating between stereotypical representations of Natives as inoffensive and naive people in periods of relative stability and as savage enemies when fighting for supremacy (Dossena 2019: 16). In Period 6, when the colonization process is brought to completion, neutral terms peak again and negative semantics shrinks. Now there is a renewed interest in Indians and their geographical provenance, as the following clusters reveal: *American Indians*, *neighbour Indians*, *the East Indians*, *shore Indians*, *the native Indians*, *Canada Indians*, *Maryland Indians.* The most frequent colligational structure is *Indians* as subject with neutral and descriptive verbs referring to their nature, manner and habits (e.g. *have*, *be*, *use*, *do*, *make*, *call*, *take*). Neutral semantics also prevails in the second most frequent pattern [NP] *of Indians* (*use of*, *custom of*, *language of*, *manner of*, *fashion of*, *pastime of*), since authors are now describing Indians as objects of curiosity and study in their histories of the British Colonization.

In the *OED* the lexeme *Indian(s)* is defined as "a member of the aboriginal peoples of (any part of) the Americas; an American Indian" and is attested in the following quotations:

(4) "1576 H. Gilbert *Disc. Discov. New Passage Cataia* sig. f.ii^{v} Those Indians..came onely through our Northwest passage." (*OED*)

(5) "1612 Bacon *Ess.* (new ed.) 88 The Indians of the West have names for their particuler gods, thoughe they have no name for God." (*OED*)

(6) "1662 E. Stillingfleet *Origines Sacræ* iii. iv. §8 The tradition of the Flood is among the Indians, both in New France, Peru, and other parts." (*OED*)

The choice of the sources reflects the *OED*'s preference for literary and scholarly quotations (Brewer 2008: 120; Gulliver 2016: 75), the only exception being the work of the explorer, Humphrey Gilbert, who, however, never reached North America. The examples confirm the semantic preference of *Indians* for neutral and descriptive collocates referring to trade, movement and customs. In this regard, the appearance of the neutral lexical verbs *come* and *have names* in the *OED* is consistent with the quantitative salience of verbs of movement and saying in the *VJDA*. The quotations also feature two frequent colligational structures for the word *Indians*, i.e. Indians as subject and Indians as object. However, they fail to account for the quantitative and

qualitative significance of the pattern [NP] of *Indians*, whose salience emerged from the corpus-assisted analysis. As shown in the previous paragraph, the lexico-syntactic construction is interesting for investigating the ambivalent relationship between English settlers and Native Indians at different points in time. In this sense, the *OED* does not record the semantic change of the word but shows instead a one-dimensional representation of the *Indians*.

3.2 *Sa(l)vage(s)*

Sa(l)vage(s) is the second most frequent lexical item in the *VJDA* and occurs mostly as noun plural.

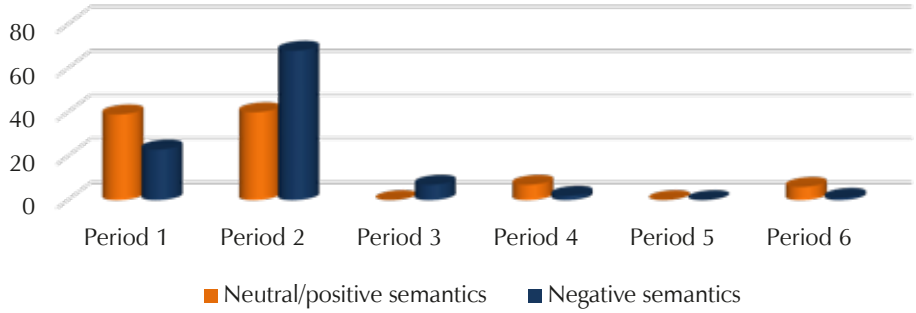

Table 4. Distribution of the neutral/positive and negative semantics for the word *Sa(l)vages* in the *VJDA*

The table shows the alternation of neutral and negative semantics across the six periods. In Period 1, *Indian(s)* has not yet appeared and the word *sa(l)vage(s)* is used as the main reference term for the Natives of the country. It is mostly found in neutral or even positive semantic networks due to the period of relative stability that the early settlers experienced in the new land:

(7) the ***Savages*** became so friendly that they often visited the English and dined with them. (*Two Tragical Events*, 1622)

Although *sa(l)vage(s)* underwent considerable alteration of meaning through the decades as the colonists' socio-economic needs evolved, one constant feature of its semantics was the 'uncivil' status of the referents. It was by means of the logic of the "Ignoble Savage" that the English colonists managed to justify the occupation of their territories (Jennings 1975; Williams 1992; Bickham 2005):

(8) such policy may be used by friendly signs, and courteous tokens towards them, as the savages may easily perceive (were their senses never so ***gross***) an assured friendship to be offered them. (*A True Report of the Late Discoveries*, 1584)

(9) there would be no labour under Heaven like this, to reduce them to civility. (*Virginia: More especially the South part thereof*, 1650)

The most frequent patterns in Period 1 are *sa(l)vages of* + [NP] (e.g. *Savages of the Maine*, *Savages of Moratok*, *Savages of the Ile*) and *sa(l)vages* + *inhabited*, *dwell*, *call* by which the English settlers describe the geographical distribution of the Native Americans in the regions. As already mentioned, the high frequency of neutral/positive semantics is consistent with the Virginia Company's instructions, according to which savages had to be described in relatively mild and tolerant terms in order to foster the commercial interests of investors in the colonization schemes (Jennings 1975: 74).[5]

The negative semantics of the word reaches its peak in Period 2 during the long Anglo-Powhatan war (63%). At that time the Powhatan was a community of more than 30 different tribes who occupied the lands of Tidewater Virginia from the Potomac River in the north to south of the James River, and parts of the Eastern Shore. The Chiefdom was ruled by Chief Powhatan, considered by the English settlers as the king of the savages (Gleach 1997). Three major lexico-syntactic patterns for *sa(l)vages* characterise the tracts of this period in which the word undergoes a process of pejoration beginning from the very moment in which the Native Indians unexpectedly attacked the English Captain Gabriel Archer.

The first pattern in order of frequency is [NP] *of Sa(l)vages* where the NP encodes two different semantic fields: a negative semantics of cruelty and deception (i.e. furie, treachery, danger, malice) and a positive semantics of news providers (information, relation).

(10) Our forces are now such as are able to tame the fury and treachery of the Savages: our Forts assure the Inhabitants, and frustrate all assailants. (*A True Declaration of the estate of the Colony in Virginia*, 1610)

[5] In the context of the failure of the Roanoke settlement, the quantitative findings in period 1 also reveal the emergence of a negative perception of the savages as untrustworthy and cruel people. In this regard, consider the quotation from Ralph Lane: "[it is] of mine opinion that we were betrayed by our owne Savages, and of purpose drawen foorth by them upon vaine hope to be in the ende starved"(*An Account of the particularities of the imployments of the English men left in Virginia*, 1586, contained in the *VJDA*).

(11) we may by this peace, come to discover the countrey better, both by our own travells, and by the relation of the Savages, as we grow in familiarity with them. (*A True Discourse of the present estate of Virginia*, 1614)

The second pattern is *Savages* + *be* + [predicative], e.g. *great thieves*, *lustie*, *nothing but hypocrisy and deceit*, *false and great hypocrites.*

(12) this conduct of the savages was nothing but hypocrisy and deceit, they only awaiting a favorable opportunity to kill out the English. (*Two Tragical Events*, 1622)

The third pattern is passive + *by Sa(l)vages*, with 7 occurrences of the cluster *were cutt off and slayne*:

(13) And those being Spente and devoured some weare inforced to searche the woodes and to feede upon Serpents and snakes and to digge the earthe for wylde and unknowne Rootes where many of our men weare Cutt off of and slayne by the Salvages. (*A True Relation by George Percy*, 1609-1612)

The last pattern features the collocation of *sa(l)vages* with words referring to animals:

(14) There came the Savages creeping on all foure, from the Hills like Beares. (*Observations gathered out of a Discourse of the Plantation of the Southerne Colonie in Virginia*, 1606)

(15) all the rest dancing about him, shouting, howling, and stamping against the ground, with many Anticke tricks and faces, making noise like so many Wolves and Devils. (*Observations*, 1606)

In Period 3 the word maintains the same negative connotations as in period 2, though its use drops to 7 occurrences.

(16) If God has not abated the Courages of the Savages in that moment of time, they so treacherously slew the English. (*A Perfect Description of Virginia*, 1649)

From Period 4 to Period 6 *sa(l)vages* maintains a low frequency and the proportion of negative semantics shrinks too (18%). This re-neutralization of

the word should be interpreted in light of the toleration policy adopted by the Governor of Virginia. In the concordances, savages are represented as possible helpers of the English colonists in the plantation and the common colligational structure features *sa(l)vages* as object of verbs such as *invite*, *encourage*, *show.*

(17) It will be good for you to incourage the Savages, when they finde any bottoms in the woods, to bring them to you, that you may get of the race, and seed to increase it. (*The Reformed Virginian Silkworm*, 1652)

(18) And let me tell you, being desirous that you may do all things with the least cost and labour to you, and to invite also the ***Savages*** to the work for their own gain. (*The Reformed Virginian Silkworm*, 1652)

The *OED* reports the two spelling variants of *savages* (*savage/salvage*) and its predominant usage as plural noun, which is consistent with my findings in the *VJDA*. Two major definitions are provided in the dictionary: that of savage as "a person living in a wild state; a member of a people regarded as primitive and uncivilized" and that of "a cruel or brutal person; (also) a person who is coarse, rough or uncouth". With respect to the first definition, a group of quotations from 16th-, 17th- and 18th-century authors and explorers attest the meaning of *sa(l)vages* as inhabitants of regions in Asia (W. Lithgow), North America (Ralph Lane, Swift) and the West Indies (J. Smith and E. Ward) and document their trade relationship with the English settlers as well as the settlers' reliance on their reports and guidance for survival:

(19) "1585 R. Lane *Let.* 12 Aug. in *Trans. & Coll. Amer. Antiquarian Soc.* (1860) 4 10 I leave to certyfye your honor of what lyekelyhuddes founde, or what the savvages reporte of better matters." (*OED*)

(20) "1612 J. Smith *Map of Virginia* ii. i. 3 Wee traded with the Salvages at Dominica." (*OED*)

(21) "1632 W. Lithgow *Totall Disc. Trav.* vi. 292 Some scattering Arabs, sold vs Water… Two of which Sauages our Captayne hyred, to guide vs." (*OED*)

(22) "1698 E. Ward *Trip to Jamaica* (ed. 3) 10 The next Morning the Salvages Man'd out a Fleet of their Deal Skimming-dishes." (*OED*)

(23) "1726 Swift Gulliver II. iv. ii. 18 I..took out some Toys, which Travellers usually carry for Presents to the Savage Indians of America." (*OED*)

The neutral representation of the savages as partners in trade and information providers reflects the findings in the *VJDA*, though it is curious that the choice of quoting Ralph Lane as source in the late 16th century is not consistent with the author's renowned scorn of the Native Americans and his preference for negative semantics (see note 5).

The remaining quotations are mostly taken from literary sources (Shakespeare, Flecknoe and Defoe) where the word usage loses its geographical reference to American Indians while retaining the meaning of pagan people and guides for explorers. There is only one quotation from Flecknoe's *Enigmaticall Characters* which attests the barbarity of the 'uncivil' but it is a literary source dated 1658 and can be considered only partially representative of the negative semantics of treacherousness and beastly violence attributed to the American Indians in the first half of the century.

(24) "1658 R. Flecknoe *Enigmaticall Characters* 67 Would tame fierce Lions, and civilize barbarousest Savages." (*OED*)

Regarding the second definition of *sa(l)vages* as cruel or brutal people, this usage is found no earlier than 1609 and again in a literary source (Shakespeare), whilst there are no quotations which refer to the treacherousness and fury of the savages of Virginia and North America from 1584 onwards.

(25) "1609 Shakespeare *Troilus & Cressida* v. iii. 51 Hect. Fie sauage, fie. Troy. Hector then 'tis warres." (*OED*)

(26) "1696 T. Comber *Disc. Offices* 114 But who would imagine, that our Christned Albion should breed such Salvages?" (*OED*)

(27) "1706 Ld. Godolphin Let. 22 Oct. in H.L. Snyder *Marlborough–Godolphin Corr.* (1975) II. 720 Some measures ought to be conceived for putting a stop to these savages." (*OED*)

By and large, the quotations in the *OED* reflect the neutral representation of *sa(l)vages* as partners in trade and guides for survival but apparently fail to account for the semantic change of the word which occurs in periods of crisis. Literary and theological sources predominate and even when first-hand accounts of explorers are quoted, the preference for neutral semantics obscures the ambivalent attitudes of the colonists towards 'the primitive and uncivil'.

3.3 *Inhabitant(s)*

The third most frequent lexical item is *inhabitant(s)*, commonly in the plural form. In Periods 1 and 2 the word refers to the Natives of the Newfoundland and occurs premodified by the adjectives *naturall* (10), *native* (3), *first* (2), *old*, *ignorant*. It is commonly encoded as subject of verbs of action and verbs of naming: *use* (17), *call* (13), *make* (12) and in the construction *inhabitants of* + [place] (16): (i.e. *the inhabitants of the Maine; of the all countrie*, *of that countie* (3), *of the countrie* (1), *of Secotan*, *of Virginia* (3), *of Mexico* (1)).

It is in Period 2 that for the first time the word appears in the pattern *inhabitants of the colony* in the Proceedings of the Virginia Assembly in 1619.[6] The cluster marks a turning point in the usage of the word, since from now on *inhabitant(s)* acquires a double referential meaning and can be applied both to the Native Americans and the English colonists. In Period 3 the two meanings coexist though the word usage drops from 45 to 8 occurrences. It is likely that – to avoid ambiguity – authors opt for other words to refer to the Native Americans at this stage (e.g. *Indians*, *Savages*). This shift of reference from Native Americans to the English colonists appears to be officialised in Periods 4 and 5 when the colonists are permanently settled in the territories and the word denotes their new legal status. The most frequent pattern is *inhabitants of* + [place] followed by issues of governance and land administration, e.g. *Inhabitants of Virginia* (5), *of Maryland* (3), *of the Colony/ie* (3), *of the side* (2), *of the North Side* (1), *of the Province* (1). In Period 6, the ambivalent referential meaning of the word re-emerges as authors recount the history of the colonization of Virginia.

In the *OED inhabitant* is defined as "one who inhabits; a human being or animal dwelling in a place; a permanent resident". It is also specified that in its early use the word was found only in the plural, the singular rarely occurring until late in the 16th century, as confirmed in the *VJDA* (3% of occurrences). The following quotations are reported for the period examined:

(28) "1588 R. Parke tr. J.G. de Mendoza *Comm. Notable Thinges* in tr. J.G. de Mendoza *Hist. Kingdome of China* 345 They did baptise certaine of the inhabitance." (*OED*)

[6] The quotation reads as follows: "Their fourth Petition is to beseech the Treasurer, Counsell and Company that they would be pleased to appoint a Sub-Treasurer here to collecte their rents, to the ende that the ***Inhabitants of*** this Colony be not tyed to an impossibility of paying the same yearly to the Treasurer in England". (*Proceedings of the Virginia Assembly*, 1619).

(29) "1593 *Tell-Trothes New-yeares Gift* (1876) 42 This citie… hath so dispersed her inhabitaunce into the other partes of the cuntrey." (*OED*)

(30) "1594 T. Bowes tr. P. de la Primaudaye *French Acad.* II. 408 If we consider both the house and the inhabitant, wee shall see that [etc.]." (*OED*)

(31) "1615 G. Sandys *Relation of Journey* 217 Frequented with Leopards, Bores, Iaccalls, and such like sauage inhabitants." (*OED*)

The quotations do not refer to the inhabitants of North America in the period of the discovery and settlement of Virginia.[7] Indeed, with the exception of Sandys's work, the 17th-century use of the word is not recorded. Moreover, in Sandys's *Relation of Journey*, the derogatory collocation *sauage inhabitants* does not find any correspondence in the *VJDA* where *inhabitants* usually maintains a neutral semantics even in moments of crisis. For example, during the Anglo-Powhatan war, the word is used to refer to the peaceful communities of Native Americans who were victims themselves of "the savages of Powhatan". The choice reveals a more positive attitude towards the Natives and assumes a more peaceful and cooperative relationship between them and the settlers (see example 10).

3.4 *Heathen*

The lexical item *Heathen(s)* has 76 occurrences: 64 as noun in the form *the Heathen*, 9 as adjective and 3 as plural noun in the phrase *the Heathens.*
The word is mostly used as adjective in the early phase of the settlement (Period 1 and 2) and takes on a negative semantic prosody as a result of its proximity to the word *devil*.

(32) but my opinion is that their heathen priests, who are the tools of the devil, were constantly working upon the credulity and ignorance of this people to make them believe that the English had come to exterminate them in the same way as the Spaniard (*Two Tragical Events*, 1622)

[7] Two of the quotations reported above are taken from texts translated from Spanish (de Mendoza) and from French (Primaudaye). The Romance languages very probably influenced the use of the word in the English translation.

(33) what is more excellent, more precious and more glorious, then to convert a heathen Nation from worshipping the divell to the saving knowledge, and true worship of God in Christ Jesus? (*A True Discourse* 1614)

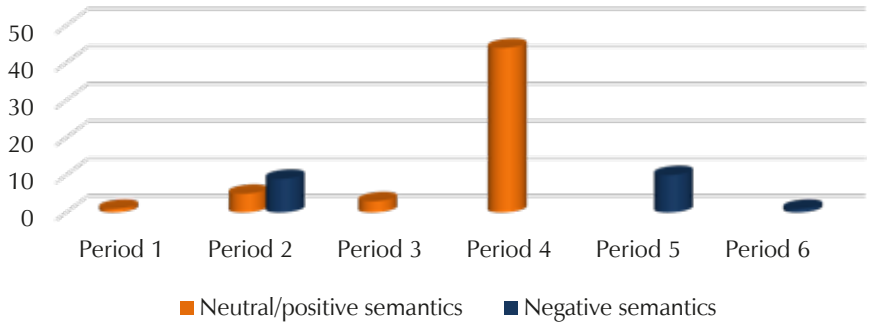

Table 5. Distribution of the neutral/positive and negative semantics of the word Heathen in the *VJDA*

Similar negative connotations characterise the occurrence of *Heathen(s)* as noun in proximity to words such as *Satan*, *ashamed*, *malice*. In Period 3, the years of Governor Berkley's toleration policy, the use of the word peaks and a process of amelioration changes its semantics. The Heathens are no longer represented as malicious believers in the devil but rather as victims of their own ignorance of God's existence. The most frequent pattern is [NP] + *of the Heathen* (25%) especially in the form *Conversion of the Heathen* (4) or *way of Converting the heathen* (1). This word usage has neutral meaning and reveals an attempt on the part of the authorities to establish uniformity of worship through conversion and moral exempla. While at that time the words *sa(l)vages* and *Indians* collocate with negative lexical items referring to the laziness, baseness and cowardice of the Natives, the choice of *Heathen* reveals a more compassionate and benevolent attitude.

The second most frequent construction contains *Heathen* as object with verbs such as *bring*, *win*, *gain*, *persuade* all referring to the conversion plan.

(34) so likewise it obstructs the hopefullest way they have, for the Conversion of the Heathen, which is, by winning the Heathen to bring in their Children to be taught and instructed in our Schooles. (*Virginia's Cure*, 1661)

(35) No hopes therefore of bringing the Heathen in love with the Christian Religion; whil'st so many evill and scandalous consequents attend the

Christians scatter'd manner of planting in that wildernesse. (*Virginia's Cure*, 1661)

There follows the pattern [modifier] + *Heathen*, where the adjectives have neutral or event positive meaning: *poor* (4), *discreet and sober* (2), *rational/l* (2) and *vertuous* (1).[8]

(36) and little hopes have the poor Heathen of redresse, whilst they see that Day so far neglected by the Christians. (*Virginia's Cure*, 1661)

(37) for if a sober discreet Heathen (and there are many such) should reply, Why hath not every Parish one of them, and Ministers belonging to them? why do no the Christians build their houses nearer them, that they may come oftner to them? why are they not better built? […] what defence could and ingenuous Christian make, which should not at once both shame himself and the Christians he would defend? (*Virginia's Cure*, 1661)

The quotations reveal a sense of pity towards the unbelievers and at the same time denounce the moral laxity of the Christians, which is detrimental to the desired peaceful coexistence with the Natives.

A drastic change in the semantics of the word occurs in period 4, when *Heathen* undergoes a process of pejoration apparently replacing the negatively connoted word *sa(l)vage(s)*. The occurrences of *Heathen* in the corpus are fewer in this period (10) but they all feature negative semantic prosody. The most frequent pattern is [NP] + *of the Heathen* which echoes the negative semantics of [NP] + *of the Indians*, *of the Savages*:

(38) Exposed to the Incursion, and murder of the Heathen (*The Declaration of the People against Sir W. Berkeley*, 1676)

(39) they judged too remiss in applying meanes to stop the fewrye of the Heathen (*An Account of our later Troubles in Virginia*, 1686)

(40) as the only man fitt in Verginia to put a stop to the bloody resolution of the Heathen (*A Narrative of the Indian and Civil Wars in Virginia In the Years 1675-1676*)

[8] The only negative premodifer appears in the noun phrase rigid heathen, referring to those heathens who are most difficult to convert as a result of their extremist position.

The second most frequent pattern is [premodifier] + *heathen* with derogatory connotations:

(41) for in a very short time they had, in a most inhumane manner, murthered no less than 60 innocent people no ways guilty of any actual injury done to these ill-discerning, brutish heathen. (*A Narrative of the Indian and Civil Wars in Virginia, In the Years 1675 and 1676*).

(42) For having in that unjust Gaine, betrayed and sold, His Maties: Countrie, and the Liberties of his Loyall Subjects to the Barbarous Heathen. (*The Declaration of the People*, 1676)

The process of pejoration in the semantics of the word is consistent with the political context of the time characterised by Bacon's fight against the Native Indians and by his harsh opposition to Berkley.

The *OED* records the usage of *heathen* as adjective from 971 as referring to "people holding religious beliefs of a sort that are considered unenlightened". For the period under consideration, the dictionary provides three quotations from theological works where *heathen* has neutral semantics and appears within noun phrases reflecting those found in the *VJDA*: *heathen priests* and *heathen-men*, the latter referring to the times of Ancient Greece and of Abraham. However, there is no attestation of the adjective *heathen* being used for the Native Americans at the time of the English settlement and of its negative semantics:

(43) "1563 W. Fulke *Goodle Gallerye Causes Meteors* ii. f. 13 Helena was of the Heathen men, taken as a Goddesse the daughter of Iupiter and Leda." (*OED*)

(44) "1627 R. Sanderson *Serm.* I. 263 Abimelech, an heathen-man, who had not the knowledge of the true God of heaven to direct him." (*OED*)

(45) "1706 M. Tindal *Rights Christian Church* 96 Made familiar to such Practices by the Heathen Priests." (*OED*)

The occurrence of *the Heathen* as collective noun is attested from 1000. There is only one quotation for the period examined and it is taken from John Milton's tragic drama *Samson Agonistes* "Spread his name Great among the Heathen round". Again, there is no attestation of the word being used

to refer to the Native Americans, nor to its ambivalent semantics, shifting from positive/neutral to negative according to the historical and political circumstances. The only quotation in the *OED* which refers to the Native Americans is dated 1736 and is taken from the *Works* of the theologian John Wesley. It provides evidence of the use of *Heathens* as plural noun, which is only occasionally found in my dataset:

(46) "1736 J. Wesley *Wks.* (1872) I. 25 My brother and I..went to pay our first visit in America to the poor Heathens." (*OED*)

The quotation is nonetheless consistent with the occurrences of *poor heathen* recorded in the *VJDA* at an earlier time (1650-1674) and documents the same compassionate attitude towards the Heathen which characterised the period of Berkley's government in Virginia.

3.5 *Native(s)*

Native(s) is the last lexical item to be examined. Its usage in the *VJDA* reflects both the general and specific definitions of the word reported in the *OED.* In its general meaning of "a person born in a specified place, region, or country" the word commonly refers to the English settlers and to England as the country of origin.

(47) Lord blesse England our sweet native countrey, save it from Popery, this land from heathenisme, & both from Atheisme. (*For the Colony in Virginea Britannia*, 1612).
(48) a very considerable number of nobility, clergy, and gentry, so circumstanc'd, did fly from their native country, as from a place infected with the plague [referring to the English Civil War]. (*A Voyage to Virginia*, 1649)

In its specific meaning of "a member of the indigenous ethnic group of a country or region, as distinguished from foreigners, esp. European colonists", the word refers to the Native Americans and oscillates between neutral and negative semantic prosody, although it is used most frequently with neutral prosody (72%).

(49) Every particular season (by the relation of the old inhabitants) hath his particular infirmity too: all which, if it had been our fortunes to have

seated upon some hill, accommodated with fresh springs and clear air, as do the natives of the country, we might have, I believe, well escaped. (*A True Reportory of the Wreck and Redemption of Sir Thomas Gates*, 1610)

(50) The devil has through the medium of the priests such an influence upon the natives that they only waited for a good opportunity to extirpate the foreigners. (*Two Tragical Events*, 1622)

The most common patterns for *natives* as plural noun referring to American Indians are *natives of* + [NP] (36%), and *natives* as object (31%).

(51) We never perceaved that the natives of the Countrey did voluntarily yeeld them selves subjects to our gracyous Sovraigne, nether that they took any pride in that title. (*The Tragical Relation of the Virginia Assembly*, 1624)

(52) He had not seen any native, or any thing in human shape, in all his round, not any other creature besides the fowls of the air, which he would, but could not, bring unto us. (*A Voyage to Virginia*, 1649)

In period 6 the *VJDA* registers the first occurrence of *native* as referring to an Englishman who was born and living in Virginia. The person concerned is the author of *The History of Virginia in Four Parts* (1722), Robert Beverly, who refers to himself as "A native and inhabitant of the place" and who – discussing the condition of servants in the colony – explains that they "become as free in all respects and as much entitled to the liberties and privileges of the country, as any of the inhabitants and natives are, if such servants were not aliens". Although these two occurrences of *natives* as referring to Virginians remain isolated in the corpus, they still give an insight into the increasing ambiguity in the referential meaning of the word from the second half of the century. It was just at that time that the term *Virginians* began to be used to designate the new generation of English colonists who had been born in the Newfoundland, as distinct from *our natives*, i.e. the English settlers who had been born in England and then moved to Virginia and from the American Indians who were simply named *natives* or – in Period 6 – *native Indians*. The relexicalization from unpremodified *natives* into *native Indians* (41% of occurrences) appears consistent with the necessity to specify the referential meaning of the word in the complex ethnical scenario.

(53) This, and a great deal more, was the natural production of that country, which the native Indians enjoyed, without the curse of industry. (*The History of Virginia in Four Parts*, 1720)

In the *OED* there are no quotations recording the usage of the cluster *native Indians* in the context of North America. Nonetheless the relexicalization process from *natives* to *native(s)* + [Nationality] is attested for other populations, as the following quotation from a translated work shows:

(54) "1687 A. Lovell tr. J. de Thévenot *Trav. into Levant* i. 59 The native Turks are honest People, and love honest People." (*OED*)

The specific meaning of *native(s)* is recorded in several quotations from geographical books, essays, histories and fiction:

(55) "1603 R. Johnson in tr. G. Botero *Hist. Descr. Worlde* 153 He committed no lesse an error in suffering the Natiues to keepe their possessions and to inhabit all their townes." (*OED*)

(56) "1631 J. Smith *Advts. Planters New-Eng.* iv. 10 More [land] to spare than all the natives of those Countries can use and culturate." (*OED*)

(57) "1652 P. Heylyn *Cosmographie* iv. ii. sig. Oooo3^{v} Inhabited by the Natives only, though the Portugals did sometimes endeavour a Plantation in it." (*OED*)

(58) "1695 W. Temple *Introd. Hist. Eng.* (1699) 5 The North-East part of Scotland was by the Natives called *Cal Dun*." (*OED*)

(59) "1725 D. Defoe *New Voyage round World* i. 2 The Stories of their Engagements, when they have had any Scuffle either with Natives, or European Enemies." (*OED*)

Smith's tract (1631) is the only source to attest the use of *natives* as referring to the American Indians of Virginia. The word occurs in the pattern *natives of* + [NP], which reflects its quantitative salience in the *VJDA*. Interestingly, in the quotations the word *natives* oscillates between negative connotations (i.e. "suffering the Natives", "scuffle either with Natives") and neutral semantics (i.e. "inhabited by the Natives only"; "by the Natives called Cal Dun"), just as the Natives are represented ambivalently in the *VJDA*.

4. Conclusion

The corpus-assisted discourse analysis has revealed the semantic variation of lexical items referring to the American Indians in the period from 1584 to 1724 in the VJDA and has assessed it along the lines of pejoration and amelioration in relation to aspects of the historical and socio-cultural context of exploration and settlement. The results reflect the changing relationship which existed between English colonists and Native Americans over the decades and account for the ambivalent attitudes of the English authors in the representation of the Other. As the British Empire expanded, a similar tendency to overlexicalization is traceable in other new colonies in the rest of the world. Studies on 19th century travel journals, for example, document the usage of the same referents (Indian(s), Savage(s), Native(s), Inhabitant(s)) for labelling indigenous peoples in Australia and India (Shvanyukova 2020, Samson 2020). Interestingly, while the referents used by the male explorer to describe his encounter with the Australian aborigines remain emotionally neutral (Shvanyukova 2020), those used by Victorian female travellers for the natives of India (Samson 2020) echo the negative semantic prosody which I found in my data.

In light of the controversial role played by the Native Americans in the history of British colonization, the *OED* documents the overlexicalization which characterised the Early Modern English accounts of the Native Americans, through the attestation of *Indian(s)*, *Sa(l)vage(s)*, *Inhabitant(s)*, *Heathen(s)* and *Native(s)*. However, few of the quotations reported refer to the exploration, settlement and colonization of Virginia and North America (13%) and only 28% of the examples examined come from tracts and pamphlets written by explorers and future settlers travelling eastward and westward. In line with the *OED*'s traditional preference for the great works of literature, history and philosophy, the bulk of the quotations for the period under investigation (71%) come from literary sources (Shakespeare, Milton, Swift, Defoe), philosophical and religious tracts (Bacon, Comber, Sanderson) and histories translated from foreign authors (de Mendoza, de Thévenot, Botero). By and large the definition of each word and the choice of the quotations reflect its general and specific meanings, its spelling variants and its most frequent lexico-syntactic patterns in the *VJDA*. However, some major semantic changes which affect the words in relation to the evolving relationship between the American Indians and the English settlers are not fully attested and the resulting semantic ambivalence of the terms is often not noted in favour of a one-dimensional representation of their usage. There

is also no clear evidence of the referential shift of the word *inhabitant(s)* and the referential ambiguity of *native(s)* in the context of the new generation of English colonists born in the occupied territories. Nonetheless, the choice of quotations for *natives* attests the ambivalent connotations of the word in line with the complex and ideologically-biased process of identity construction documented in the corpus.

REFERENCES

Sources

Oxford English Dictionary Online (OED online)
http://www.oed.com, accessed September 2019

Virtual Jamestown Digital Archive
http://www.virtualjamestown.org/, accessed September 2019

Special studies

Bickam, T.
2005 *Savages Within the Empire. Representations of American Indians in Eighteenth Century Britain.* Oxford: Clarendon Press.

Borch, M.F.
2004 *Conciliation – Compulsion – Conversion. British Attitudes Towards Indigenous People 1763-1814.* Amsterdam; New York: Rodopi.

Brewer, C.
2005 "Examining the OED", http://oed.hertford.ox.ac.uk, accessed September 2019
2008 "Johnson, Webster and the Oxford English Dictionary". In: H. Momma – M. Matto (eds.) *A Companion to the History of English Language.* Oxford: Blackwell, 113-122.

Cecconi, E.
2020 "British colonial ideology in the language of pamphlets on Virginia (1584-1624)". In: N. Brownlees (ed.) *The Language of Discovery, Exploration and Settlement.* Newcastle u. T.: Cambridge Scholars Publishing, 8-24.

Dossena, M.
2019 "'There were always Indians passing to and fro'. Notes on the representation of Native Americans in CHET documents". In: I. Moskowich et al. (eds.) *Writing History in Late Modern English: Explorations of the Coruña Corpus.* Amsterdam: John Benjamins, 22-39.

Fitzmaurice, A.
2003 *Humanism and America. An Intellectual History of English Colonization, 1500-1625.* Cambridge: Cambridge University Press.
Gleach, F.W.
1997 *Powhatan's World and Colonial Virginia: A Conflict of Cultures.* Lincoln: University of Nebraska Press.
Gramley, S. – K. Pätzhold
1992 *A Survey of Modern English.* New York: Routledge.
Gulliver, P.
2016 *The Making of the Oxford English Dictionary.* Oxford: Oxford University Press.
Haarman, L. – L. Lombardo
2009 "Introduction". In: L. Haarman– L. Lombardo (eds.) *Evaluation and Stance in War News: A Linguistic Analysis of American, British and Italian Television News Reporting of the 2003 Iraqi War.* London: Continuum, 1-26.
Jennings, F.
1975 *The Invasion of America. Indians, Colonialism and the Cant of Conquest.* Chapel Hill: University of North Carolina Press.
MacMillan, K.
2013 "'Bound by our regal office': Empire, sovereignty, and the American colonies in the seventeenth century". In: S. Foster (ed.) *British North America in the Seventeenth and Eighteenth Centuries.* Oxford: Oxford University Press, 72-97.
Moran, M.G.
2007 *Inventing Virginia. Sir Walter Raleigh and the Rhetoric of Colonization, 1584-1590.* New York: Peter Lang.
Pahta, P. – I. Taavitsainen
2010 "Scientific discourse". In: A.H. Jucker – I. Taavitsainen (eds.) *Historical Pragmatics.* Berlin; New York: Mouton de Gruyter, 549-586.
Partington, A.
2004 "Corpora and discourse, a most congruous beast". In: A. Partington – J. Morley – L. Haarman (eds.) *Corpora and Discourse.* Bern: Peter Lang, 11-20.
2009 "Evaluating evaluation and some concluding thoughts on CADS". In: J. Morley – P. Barley (eds.) *Corpus-Assisted Discourse Studies on the Iraq Conflict: Wording the War.* London: Routledge, 261-303.
Samson, C.
2020 "Victorian women discovering and representing India in their travel journals". In: N. Brownlees (ed.) *The Language of Discovery, Exploration and Settlement.* Newcastle u. T.: Cambridge Scholars Publishing, 128-144.
Shvanyukova, P.
2020 "The language of 'messy' encounters: Matthew Flinders's depictions of Australian Aborigines (1801-1803)". In: N. Brownlees (ed.)

The Language of Discovery, Exploration and Settlement. Newcastle u. T.: Cambridge Scholars Publishing, 179-194.

Simpson, J.A.
2002 "The revolution of English lexicography", *Dictionaries* 23, 1-15.

Stubbs, M.
1996 *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture.* Oxford: Blackwell.
2001 *Words and Phrases: Corpus Studies of Lexical Semantics.* Oxford: Blackwell.

Williams, R.A. Jr
1992 *The American Indian in Western Legal Thought. The Discourses of Conquest.* New York: Oxford University Press.

Address: Elisabetta Cecconi, Scuola di Scienze Politiche, University of Florence, via delle Pandette 32, 50127 Firenze.
ORCID code: orcid.org/0000-0002-7819-0559.

A Cultural Linguistics approach to the "discovery of childhood" in sixteenth and seventeenth century Britain

Paloma Tejada Caller

Universidad Complutense of Madrid

ABSTRACT

This article addresses childhood as a culturally constructed life-stage in sixteenth and seventeenth century Britain from a Cultural Linguistics or Ethnolinguistics approach. Intended as a language-centred contribution to the interdisciplinary field of Age Studies, this lexicographical study intends to prove that a new modern intersubjective conceptualization of childhood emerges in English during the period under consideration, confirming what historians have called the "discovery of childhood". Using the *Oxford English Dictionary* and the *Historical Thesaurus of the Oxford English Dictionary* as data sources, a corpus of 103 new words and word senses appearing during the Early Modern English period and associated with the five 'child' meanings listed as [PERSON] was compiled and analysed through a purposely-created chart of 24 parameters. Quantitative and qualitative results obtained verify the corpus as a coherent set of terms providing cumulative evidence of a clear change in the perception and conceptualization of childhood, arguably organized around what we have called *anchoring words*.

Keywords: Age Studies, Cultural Linguistics, lexicography, childhood, life-stage, Early Modern English.

1. Introduction

This study sets out to verify whether there is *linguistic* evidence of any "discovery of childhood" in Britain during the sixteenth and seventeenth centuries, as apparently advocated by social and historical scientists. That is, the study aims at the identification and linguistic reconstruction of the

DOI: 10.25951/4363.

subjective, but not individual, system of internalized structures, schemes of perception, conception and possibly action, common to members of the Early Modern English society regarding childhood. More globally, falling back on the idea that until now linguistic analysis has been very much underutilised in social and humanistic studies, this contribution intends to expand Cultural Linguistics tenets and methodology to the field of Age Studies (AS).[1]

Three ideas lie behind the theoretical framework of this contribution. First, the need to contribute linguistic thinking to the burgeoning field of AS; second, the need to concentrate on the apparently neglected notion of childhood as an age stage. And third, the claim that lexicographical approaches may prove adequate to reveal socially and culturally shared schemes of perception. In this introduction I will briefly address each of these guidelines.

As is increasingly better-known, AS constitute an interdisciplinary field where the cross-fertilization of concepts and methods from previously distinct research traditions has given way to a newly shared sensitivity about age and age-stages, revealing additional gaps of research (Charise 2014; Segal 2014). Oddly enough, linguistics has been slow in responding to the age-challenge. Some have even advocated that linguistics has been age-blind up to recently (Coupland 2009). Thus, it seems timely to take on an ageing lens in the humanities, making age and age-stages a focus of attention.

From the 1990s AS have essentially made of age an analytical category, opposing the idea that chronological age is age. From this apparently simple theoretical tenet, the new tradition set out to fight against predefined taxonomies of social types, which portray a generally assumed coherence and stability of age categories through times and places. This means that age has become increasingly understood as a cultural notion implying the assumption that the conceptualization, definition, even the existence of life stages like childhood, adolescence, adulthood or later life, – to mention just the most commonly accepted ones -, varies across time, space and populations. Any age stage seems ultimately dependent on cultural schemes, responsible for the number of life-stage categories recognised

[1] We hold to the short designation of *Age studies* (*AS*), as opposed to *Ageing Studies*, for two basic reasons. First, it requires no disambiguation as for the interpretation of life as an ageing process, involving different stages from birth to death. And second, the term *Age studies*, introduced by Gullette in 1993, is used by the Modern Language Association and in the humanities in preference to "ageing studies" (Cole – Ray 2010: 17).

within a community, their boundaries and duration and the prestige granted to them. In other words, understanding age as a cultural concept implies the assumption that there exist socially shared patterns of knowledge that should be unearthed. However, due to the interdisciplinary nature of AS and the merging of both social and humanistic traditions within a common arena, discussion on how to reach this goal has reinstated the divide between disciplinary approaches to social and human issues as for the validity of their explanations (Goffman 1983; Hagestad – Dannefer 2001: 4; Biggs et al. 2003; Katz 2014, among others). For advocators of the humanities, empirical explanations seem stereotyping and all too sweeping, whereas for empirical scientists the humanities often turn anecdotal and hence non-significant, particularly because of an alleged over-emphasis on individual micro-interactions that seems to render the social norms and values invisible. It is indeed to reduce this divide that linguistics may be called into play. Following Nikander (2009), who regards linguistics as a suitable middle ground in the theoretical and analytical debate, it is also our contention that this discipline may provide a potential tool to curb the estrangement between the supposedly scientific and the irreducibly humanistic grasp in the cultural construction of childhood patterns (Tejada 2019). As will be expounded below, contemporary linguistics may bring life-stage studies at least two gains. First, a theoretical approach preventing research from falling into a methodological individualism, and second, a fact-based argumentation, in as far as it is linguistic data, free of common research biases, that lie at the base of hypotheses on cultural conceptualisations and critical attitudes (Tejada 2019).

Regarding the need to focus on childhood, suffice it to say that social research has singled it out as the "forgotten category" in the life-course (Frijhoff 2012: 23), or the "silent" and "obscure" years (King 2007: 389; see also Thomas 2004). Moreover, there seems to be scholarly agreement that it was only in the 1960s that the child was *discovered* in social discourse; that is, it was then that the child became newly perceived as an autonomous object of historiography (King 2007; Fass 2009; Dekker 2012; Frijhoff 2012, among others). More specifically, historians seem to have discovered childhood some sixty years ago very much due to Phillipe Ariès' ground-breaking essay, originally released in 1960 and later translated into English as *Centuries of Childhood: A Social History of Family Life* (1962). As is well-known, Ariès' work sparked off a heated and productive debate, which, despite its intensity, or because of it, gave way to an outpouring of studies on the history of childhood.

Most important for us is the fact that childhood may be a modern invention. In his controverted book, Ariès argued that childhood was discovered within the sixteenth and seventeenth centuries in Western Europe. The child was then discovered in social practice, becoming separately identified and perceived as an autonomous human being. Ariès attributes this discovery to a series of social and economic changes in Europe, expanding from the middle classes down through the social scale from the sixteenth to the twentieth century: namely, the rise of privacy in the household; changes in education[2], and very importantly, a more codified system of apprenticeship. According to Ariès, by the eighteenth century the mentalities had changed[3]. Moreover, from the sixteenth century onwards a change in the way people felt about children seems to have been brought about, their perception steadily sliding from neutrality to a higher valuation. Drawing on these premises, scholars in the social sciences have insightfully concluded that childhood "as distinct from biological immaturity is neither a natural nor a universal feature", though "it appears as a specific and cultural component of many societies." (Prout – James 1990: 8). Notwithstanding the above, this judicious and insightful debate has scarcely attracted the attention of linguists so far. It is probably the alleged linguistic blindness on age that has resulted in a clear lack of research on a dynamic definition and meaning of childhood in cultural and historical terms.

The third idea guiding this study concerns the fact that lexicographical and semantic research has been recently re-valued. In English linguistics at least, investigation in this field has been progressively re-gaining ground for various reasons, among which two are remarkable and should be mentioned here. First, the release of the new online edition of the *Oxford English Dictionary* (*OED*) in combination with the *Historical Thesaurus of the OED* (*HTOED*) has contributed much to the reappraisal of these studies, as Busse (2012), Coleman (2012), Crystal (2014) and others have widely argued. Second and more important is the existence of a scholarly sustained tradition stemming from Sapir (1921, 1957), and ultimately deriving into Cultural Linguistics, stressing the idea that "vocabulary is

[2] Even at the risk of oversimplification, education moved from impersonal and stative wisdom into some kind of schooling, which opened the way to a new stage in child rearing, that of the so-called "intrusive parent" in deMause's (1974) terms.

[3] Despite its contentious nature, this is not the place to discuss Ariès' theory. Suffice it to say that most scholars currently agree on the rightness and accuracy of his overall idea, recognising that he tried to reconstruct the way people thought about the idea of childhood, rather than how individual children were reared or treated. See King (2007: 271), Kline (2008), Lowe (2009: 66), Mawhinney (2015: 22), among others.

a very sensitive index of the culture of a people" (Sapir, 1957: 34, 36; Brinton – Closs Traugott 2005; Bartmiński 2009), and that the lexis is one of the most conclusive bases for investigating linguistic worldview. According to this conceptual model, vocabulary would constitute the institutionalization of social meaning, words representing access-nodes to shared knowledge. That is, relatively stable lexical units, together with grammatical relations, are considered to be relic, codified traces of historical experiences and common understanding, influencing the thought patterns of their speakers. Consequently, the vocabulary of a language may be viewed as a body of accumulated innovations and meanings which have been salient for the community at a certain historical stage. These principles, most conspicuously followed by the Lublin school of Ethnolinguistics (see Bartmiński 2016), constitute not only a powerful analytical framework, but also a methodological advantage for social and cultural researchers. Being a bottom-up approach, analyses are unlinked to the researcher's intuitions or guided input. Explorations move from linguistic categories to speakers, rather than the reverse, thus escaping the risk of departing from predefined, previously established identities of language-users. Examining cultural conceptualizations rather than speakers, we may claim that language-centred approaches to AS represent an improved theoretical framework for the grasping of "cultural cognition", defined as the collective cognition that characterises a cultural group. In our case, it reveals itself as a useful scheme to unearth intersubjective patterns of Early Modern English childhood. In sum, this contribution claims that lexicographical studies within the Cultural Linguistics tradition are adequate to explore whether sixteenth- and seventeenth-century English shows "new categories, schemas, conceptual metaphors and propensities for certain perspectives" on childhood reflecting the cultural cognition of those who spoke the language at the time (Tomasello 1999: 169).

2. Methodology

As defined in the *OED*, the word *childhood* is both the state and the stage of being a child.[4] However, looking up the dictionary definitions of *child* is of little use, definitions being limited in their explanatory power and ultimately imprecise. On the one hand, the child is defined quantitatively, in

[4] "Childhood": a) "The state of being a child"; b) "The stage of life or period during which one is a child." Cf."childhood, n." OED Online, www.oed.com/view/Entry/31631, accessed September 2019.

relative terms of age, *child* and *youth* being synonymous expressions. On the other, we are provided with very little qualitative characterization. Hence, in order to reconstruct a qualitative assessment of childhood during Early Modern English, the opposite path should be walked; that is, from meaning to words, as allowed by the new *HTOED*.

Drawing on the premise that the Early Modern English conceptual category 'child' may be reconstructed using a lexicographical *HTOED* approach, a table of synonyms associated with the senses of 'child' as PERSON was produced, namely, ['child' as INFANT], ['child' as CHILD], ['child' as YOUNG MAN] and ['child' as GIRL], to which ['child' as FETUS] was added, considering the first *OED* definition for 'child' with reference to state or age.[5] Interpreted as markers of meaning profiles rather than sociolinguistic patterns of variation, it was assumed that, examined together, these synonyms would be revelatory about the thinking of a society (Kay 2010). Hence, after a global assessment of 'child' synonyms registered in the *HTOED* from Old English to the sixteenth century, a corpus of 103 new words and word senses (i.e. extended meanings of pre-existing words) appearing during the Early Modern English period and associated with the five 'child' meanings mentioned above was collected.[6] Subsequently, a chart of twenty-four study parameters was devised (Table 1) including, stylistic and axiological information, register connotations captured from cross-reference definitions of the terms, and further details, to obtain quantitative and qualitative results.

Also, in order to identify the degree of lexical and semantic innovation occurring during the Early Modern English stage, the period was divided into eight 25-year subcategories (Table 2).[7]

---

[5] "The unborn or the newly born". Cf. "child, n." OED *Online*, www.oed.com/view/Entry/31619, accessed September 2019.

[6] Listed in alphabetical order, the corpus comprised the following inventory of synonyms: *bantling*, *bonne*, *boy*, *brat*, *bratchet*, *bratling*, *breed*, *bud*, *budling*, *butter-print*, *callow*, *cherub*, *child*, *child in arms*, *childling*, *chit*, *chitterling*, *chrisom*, *chrisomer*, *cockerel*, *cockling*, *codling*, *cub*, *dandiprat*, *dandling*, *demy*, *doveling*, *embryo*, *eyas-musket*, *fairy*, *feture*, *filly*, *flosculet*, *geniture*, *gorrell*, *gossoon*, *halflang*, *hans-in-kelder*, *hensour*, *hobbledehoy*, *hoppet*, *imp*, *kid*, *kinchin*, *kinchin-mort*, *kitling*, *lad*, *ladykin*, *lambkin*, *lap-child*, *little girl*, *loneling*, *loon*, *lullaby-cheat*, *maggie*, *maid*, *man-boy*, *minx*, *miss*, *mistress minx*, *muchacho*, *ninny*, *piccaninny*, *piggy*, *pigsney*, *pledge*, *prill*, *puppet*, *ragazzo*, *schoolboy*, *schoolgirl*, *shapeling*, *shaver*, *skipper*, *slut*, *snipper-snapper*, *spaught*, *spear*, *sprig*, *spring*, *squall*, *stranger*, *stubbed boy*, *stubble boy*, *tadpole*, *tenderling*, *tendril*, *two-year-old*, *urchin*, *urchin*, *vriester*, *wag*, *wean*, *whelpling*, *woman-child*, *womb-infant*, *youngster*, *youngster*, *younker*, *younker*, *younkerkin.* Note that terms classified under two different *HTOED* nodes appear twice.

[7] Throughout the article a number in brackets may appear together with a given term, indicating period of creation according to the *OED*, as for example, in *bantling* [4].

Table 1. Parameters of analysis

Parameters	
Child senses	new word
	new word sense
	period
	embryo/ fetus
	infant/baby
	child
	girl
	young man
	meaning
Currency	currency: obsolete
	last quotation
	last quotation after 1900
Zoosemy and nature metaphors	plants
	animals
Axiological nuances	diminutive/ hypocoristic
	contempt/ depreciative
	affection/endearment
	irony/playful
	negative overtone
	positive overtone
Stylistic labels	slang
	coll
	regional/ cant
	gendered
	uncertain/ unknown/ obscure etymology

Table 2. Time-periods for analysis

Periods		Periods	
1	1500-1525	5	1601-1625
2	1526-1550	6	1626-1650
3	1551-1575	7	1651-1675
4	1576-1600	8	1676-1700

For the sake of clarity, Table 3 illustrates the distribution of terms per time period. For a qualified assessment of each term, the analysis required not only a close reading of definitions and dictionary labels, but also examination of quotations as well as frequent cross-reference acts through the dictionary and the thesaurus.

Table 3. List of new words/ new word senses per period and child meaning

Period	Embryo/ fetus	Infant/ baby	Baby girl	Child	Girl	Young man
1				*brat*	*woman-child*	*hensour* *younker*
2	*feture*			*kitling* *younkerkin*	*bonne* *urchin*	*gorrell* *hobbledehoy* *lad* *spear*
3		*chrisomer*		*urchin*	*kinchin-mort*	*cockerel* *loon* *spring* *wag*
4	*embryo*	*chrisom* *tenderling*		*bantling* *bratchet* *breed* *budling* *lambkin* *loneling* *pledge* *ragazzo* *schoolboy* *tadpole* *two-year-old*	*dandiprat* *minx* *mistress minx* *prill*	*boy* *cockling* *demy* *imp* *muchacho* *pigsney* *shaver* *snipper-snapper* *spaught* *stubble boy* *youngster*
5	*womb-infant*		*child*	*bud* *butter-print* *chit* *dandling* *doveling* *eyas-musket* *piggy* *puppet* *whelpling* *younker*	*child* *filly* *little girl* *maggie* *squall* *tendril*	*codling* *cub* *skipper*

6	*hans-in-kelder*	*childling* *flosculet*		*fairy* *ninny* *youngster*	*ladykin*	*man-boy*
7	*geniture* *shapeling*	*bratling* *child in arms* *lullaby-cheat* *stranger*	*maid*	*chitterling* *lap-child* *piccaninny*	*miss* *schoolgirl* *slut* *vriester*	*callow* *halflang* *sprig*
8		*hoppet*		*cherub* *kid* *kinchin* *wean*		*gossoon* *stubbed boy*

3. Results and discussion

To assess whether Early Modern English reveals a changing construct of childhood, it is requisite to refer briefly to the previously existing situation, before turning to specific data and results. A look at the synonyms for 'child' meanings registered in the *HTOED* up to the sixteenth century would confirm that English exhibited a very limited number of terms for 'child', words encompassing little emotion or metaphor. At the time, childhood would be recognised as the first of a two-step ladder of life, whereby children would be at an early stage, as opposed to "non-children", or adults. These two steps were apparently defined around two basic external axes of youth and stature. Children would, thus, be either living persons or things that had lived for a relatively *short* time, and people of a *short stature* as opposed to fully grown human beings. Alongside this external characterization, childhood would be mainly associated to the idea of offspring and identified in terms of static, unaged social roles. Up to the sixteenth century terms like *page*, *groom* or *servant* are given as synonyms of 'child' in the *HTOED*.

According to our data, this system proves to change during the Early Modern English period, confirming a conceptual transformation. During the sixteenth and seventeenth centuries a much more refined picture of childhood is obtained, whereby childhood is no longer a vague first step in the ladder of life, but a stage with a well-profiled delineation of substages. As will be argued, during the sixteenth and seventeenth centuries there seems to arise a cognitive and social separate perception of fetuses, infants, young children and youths,

leading to an awareness of children as individuals. Two of these categories obtain lexical prominence during the period: namely, the ungendered 'child' as [CHILD], constrained to a context of intimacy and affection, and that of 'child' as [YOUNG MAN/ BOY], a predominantly male later childhood, evaluated against broader social norms. Further in this shift, perception seems to move from an external to an internal or subjective assessment, new terms exhibiting a clear growth in emotional colouring and imagery.

3.1 Cumulative evidence for a new perception of childhood

A global analysis of data leads to the characterization of the corpus around three main features, which should be read in terms of cumulative evidence for the new perception summarized in the previous paragraph. First, during the period under study, a significant growth of new words and word senses applied to the notion of childhood is perceived. If we compare the rate of innovation per century affecting 'child' synonyms from OE to the nineteenth century, the intensification of novelty during the sixteenth and seventeenth centuries becomes obvious. (Figure 1).[8]

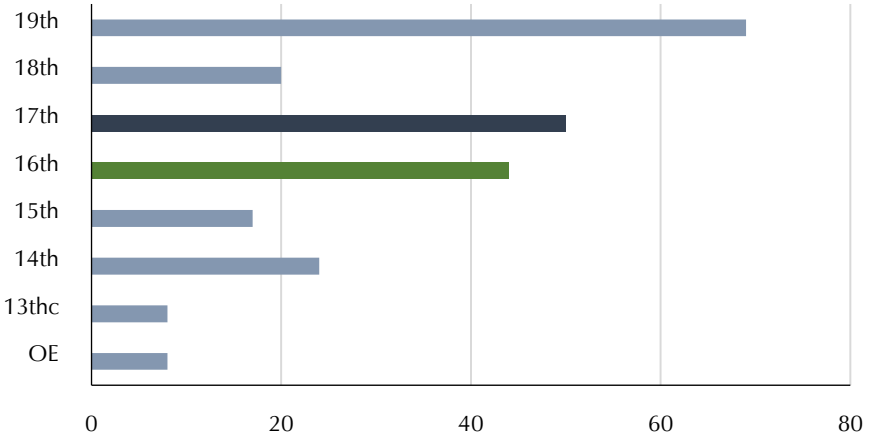

Figure 1. Lexical and semantic innovation for 'child' (all meanings) from OE to the 19th c.

[8] As may be observed, the greatest lexical activity and linguistic awareness seemingly occurs in the nineteenth century, following the Victorian concern and social discourse on childhood. As for Renaissance lexical and semantic innovation on the construal of child images, the years 1576-1625 (periods 4 and 5) constitute a most active stage, something very much in line with scholarly observations on Early Modern lexical productivity (See Nevalainen 2000: 336, and references thereby).

However, it is notable that a complex distribution of lexical and semantic innovation is obtained when narrowing the scope of analysis. During the period under scrutiny the rate of innovation varies widely across child meanings, as shown in Figure 2. Moreover, there seems to be a clear phasing of innovation affecting the two most conspicuous categories of childhood: period 4 points to a peak in the innovative construal of the [YOUNG MAN/BOY], whereas the 'child' as [CHILD] seems to have been more consistently discovered at a later stage (1576-1625).

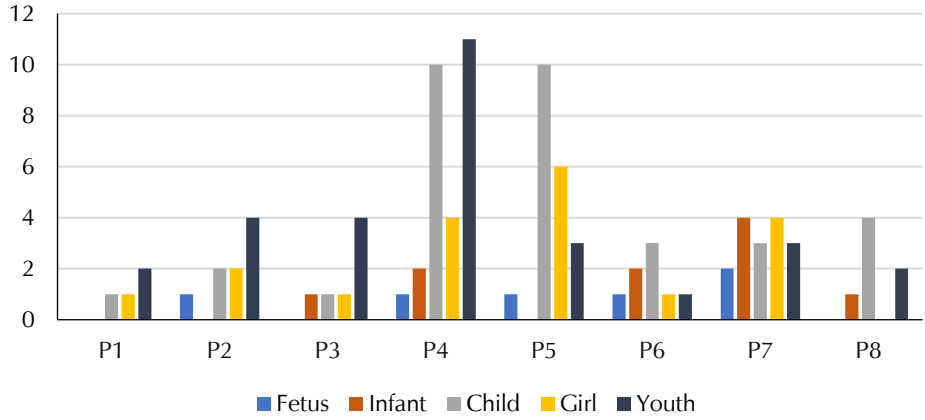

Figure 2. Lexical and semantic innovation per child meaning and time-period (16th c. and 17th c.)

A second feature defining the corpus is that almost 80% of the words could be classified as "expressives". This term has been taken from typological scholarship (Tufvesson 2007; Steriopolo 2016) to include diminutives, nicknames, metaphors, sound-symbolism, alliterative terms and terms indicating emotions (such as endearment/affection or contempt), attitudes (i.e. jocular uses, judgements of approval or disapproval) and evaluations (whether of an aesthetic, behavioural, moral kind).[9] Particularly significant in our corpus is the rate of diminutives. During this period the creation of diminutives with either a denotative or emotional meaning is to be highlighted, this feature affecting 25.2% of the terms.[10] Also, the corpus

[9] According to typologists, expressives may constitute "a distinct class of words denoting sensory, emotional or other types of perceptions of the speaker, in relation to a particular phenomenon." (Tufvesson 2007) This definition is in line with other more intricate definitions of expressive or evaluative meaning less adequate to our purposes.

[10] Cf. *bratchet*, *bratling*, *budling*, *childling*, *chitterling*, *chrisomer*, *cockerel*, *cockling*, *demy*, *doveling*, *flosculet*, *hoppet*, *kinchin*, *kitling*, *ladykin*, *lambkin*, *maggie*, *minx*, *ninny*, *piggy*,

demonstrates a significant growth in metaphors and imagery, mirroring the interests of the speech community. Noteworthy is the almost exclusive use of plants (12% of the terms) and animals (20% of the terms) for the symbolic representation of childhood. As evidenced in Table 4, figurative synonyms span from hedgehogs, to foals, pigs or doves and from buds to apples.

Table 4. Animal and plant metaphors per period and child meaning

Animal metaphors	Period	Fetus	Infant	Child	Girl	Young man
Animal metaphors	1					
	2			*kitling*	*urchin*	*gorrell*
	3			*urchin*		*cockerel*
	4			*tadpole* *lambkin* *Bratchet*		*pigsney* *cockling*
	5			*eyas-musket* *doveling* *whelpling* *chit* *piggy*	*filly*	*cub*
	7			*chitterling*		*callow*
	8			*kid*		
Plant metaphors	2					*spear*
	3					*spring*
	4			*budling* *bantling*		*imp* *stubble boy*
	5			*bud*	*tendril*	*codling*
	6		*flosculet*			
	7					*sprig*
	8					*stubbed boy*

Third and perhaps most substantial, the corpus stands out as a heterogeneous combination of non-standard and unstable terms. Specifically, the corpus provides ample evidence of a) terms tinged with colloquial and informal

pikaninny, *prill*, *shapeling*, *tenderling*, *whelpling*, *younkerkin*. Note that, according to the *OED*, the suffix *-ling* in *loneling*, *bantling* is not a diminutive marker, but a morpheme meaning "concerned with".

connotations, words belonging to jargons, slang, children's talk (e.g. *butter-print*, *piggy*, *kinchin*, *pigsney*), regional terms (e.g. *wean*, *hoppet*, *maggie*), fanciful formations and rhyming slang (e.g. *snipper-snapper*, *hobbledehoy*), etc.; b) words proving heterogenous origins, and frequently exhibiting obscure etymologies, which might suggest not only local innovation, but also the adoption and adaptation of terms introduced in the language through oral interaction, through "street exchange", rather than as literate or literary creations;[11] and c) terms revealing a tendency for quick obsolescence.[12] Apparently, Early Modern English words for childhood did not have a lasting effect on the language, being either rejected or ignored by the later speech community. As a matter of fact, the vocabulary of childhood seems to have been renewed in the nineteenth century. This reveals a most interesting result, taking into account the period under consideration. A high proportion of words defining childhood during the sixteenth and seventeenth centuries appear to be popular and emotional formations. Most remarkably, the lexis of childhood, deeply connected with subjectivity and intimacy, seems to be a new fashion reflecting individual needs rather than social institutionalization from above, a phenomenon most likely to have occurred in the nineteenth century (see footnote 8). Although further research is needed, we are probably witnessing an incipient moment of vocabulary creation on the edges of language, where no attitudes of correctness seem yet to be monitoring behaviour.[13]

3.2 Childhood substages

Moving on to particular details in our analysis, a careful scrutiny of the corpus reveals that at the end of the period there appeared descriptive terms to denote childhood substages. These terms I have called "anchoring words". That is, by

[11] Among the words of uncertain/ obscure/ unknown origin, one may mention: *boy*, *brat*, *bratling*, *callow*, *codling*, *cub*, *dandiprat*, *filly*, *gorrell*, *gossoon*, *hensour*, *hobbledehoy*, *kinchin-mort*, *lad*, *loon*, *minx*, *ninny*, *piccaninny*, *pigsney*, *puppet*, *prill*, *slut*, *snipper-snapper*, *spaught*, *sprig*, *squall*, *stubble boy*, *tendril.* On the need to revise *OED* etymologies see Sayers (2016).

[12] It is significant that 40% of the terms in the corpus are marked as *obsolete* in the *OED* and a further 37% were apparently no longer in use after 1900. Among the terms marked as obsolete in the *OED*, one may mention: *bonne*, *breed*, *budling*, *butterprint*, *callow*, *chrisomer*, *codling*, *dandiprat*, *dandling*, *demy*, *feture*, *flosculet*, *geniture*, *gorrell*, *hans-in-kelder*, *hensour*, *spear*, *imp*, *kitling*, *lap-child*, *loneling*, *maggie*, *maid*, *mistress minx*, *ninny*, *prill*, *shapeling*, *spaught*, *spring*, *squall*, *stubbed-boy*, *tendril*, *urchin*, *vriester*, *wag*, *womb-infant*, *younkerkin.*

[13] On standardization and popular formations, see Marchand (1969), as cited in Nevalainen (2000: 431).

the eighteenth-century English has apparently lexicalized the stereotypes of childhood, around which the conceptualization of the whole stage revolves. By this time, boundaries have become established between a *womb-infant* and a *child-in-arms*; or between these and the *lap-child*, the *man-boy* and the *woman-child*. With this idea in mind, we can reconstruct the cultural schemes behind each of the childhood substages and the differences established in the sixteenth and seventeenth centuries with respect to the previous period.

3.2.1 The womb-infant

The first interesting result is that the unborn child becomes separately perceived from the rest of the subcategories. One should not forget that the divide between abortion and infanticide, between the born and the unborn child, had not always been clear (King 2007: 391). According to our data, from 1576 onwards there seems to be a growing number of words for embryos. Divided into the specialized and the familiar stylistic registers (cf. *feture* vs *hans-in-kelder*), the new perception of embryos is channelled through popular metaphors of either smallness, bringing children closer to animals and plants as living entities getting shaped, as in *shapeling*, or confinement, as in *womb-infant*, whereby the defining term *womb* would be interpreted as a container, an image also elicited in the more complex *hans-in-kelder*, implying a second metaphorical level.[14]

3.2.2 The child-in-arms

The baby constitutes a second newly-perceived childhood stereotype, as the *child-in-arms*. The data in our corpus show it as a late perception. It is only from the seventeenth century onwards that a more explicit appreciation of babies is proved, through the creation of an increasing number of new synonyms (cf. Table 3 above). In relative terms, this category exhibits the largest degree of lexical innovation (i.e. of new words). Considering that 78% of the total amount of synonyms for this category of 'child' as [INFANT] are new words, it may clearly be claimed to be an invention. More specifically, during the years under consideration, babyhood seems to move from a religious to a more expressive and evaluative perception. At the beginning of the period the baby is identified as a *chrisom* or a *chrisomer* (periods 3 and 4) suggesting religious innocence symbolized through the white robes at baptism. Then, a more physical and descriptive definition is obtained:

[14] As for *Hans-in-kelder (Jack in the cellar)*, see Nevalainen (2000: 422) on phrasal nouns as a pattern of Early Modern English innovation.

the baby is understood as a small and fragile being (as in *tenderling* [4], *or childling* [6]), often through the use of denotational diminutives evaluating size), to reach a final phase of lexical innovation, comprising metaphorical perceptions, such as *lullaby-cheat* [7] or *stranger* [7], that would suggest a newly recognised need for affection and care in these entities.

3.2.3 The lap-child

Moving on to the next substage, the *lap-child* stands out as one of the two most salient categories of Early Modern English childhood, stereotypically recognised also as *the two-year-old* [4]*;* This third category is apparently discovered during periods 4 and 5. Our results suggest that the language then experiences an explosion of new words and word meanings for this substage, managed through a wide variety of resources, from borrowings through to new derivatives and meaning extensions of existing words. In absolute terms, this substage scores the largest number (34%) of new terms and new word-meanings, of which 52% are new coinages. Furthermore, the *lap-child* is confirmed as a gender-neutral category, 77% of the terms in the corpus being ungendered.[15] Moreover, almost no term in our collection is descriptive in nature, most being expressives and more particularly diminutives and animal metaphors (cf. Table 3 above). Rather than evaluating size, diminutives in this category seem to convey qualitative evaluation, responding thereby to their allegedly prototypical use (Ponsonnet 2018).[16] As for zoometaphors that are used as synonyms for the *lap-child*, most of them express positive evaluation, as suggested by terms like *kitling*, *doveling* or *lambkin*, among others.[17] However, the language-specific nature of this resource would require further research.

From a semantic point of view, this third childhood substage represents a newly discovered world in the sphere of familiar routines (cf. *dandling* [5], *lap-child* [7]). Members of this category are construed as small beings (*kinchin*, *wean*, *lap-child*), with some tinges of under-development or rudiment; lovable

[15] Save for a minority of terms, leaning towards either the girl-side (*fairy*, *puppet*), or the boy-side (*youngster*, *younker*, *younkerkin*). The term *urchin*, initially applied to female children with a pejorative moral meaning, extends to "raggedly, or untidily clothed" boys in period 3, and is endowed with a more aesthetic nuance.

[16] According to Ponsonnet 2018 (who draws on previous studies by Wierzbicka 1984 and Jurafsky 1996), diminutives are usually anchored in intimacy across languages, and express milder emotions of endearment, familiarity, affection or approval, and positive judgements as opposed to augmentatives.

[17] Cf. Sakalauskaite (2010: 17), for a definition of zoometaphors as metaphors in which the behaviour, emotion, or appearance of an animal is a reference to those of humans. See also Kiełtyka – Kleparski (2005).

and funny creatures (as evidenced by *kitling*, *eyes-musket*, *lambkin*, *doveling*, *budling*, *bud* or *kinchin*), who are also moveable and, therefore, troublesome, cheeky (*brat*, *chitterling)* or dependent, as in *lap-child*. They may even be cast as a burden (*piggy*), calling on scant resources; or ill-tempered and roguish, as suggested by *urchin.* Particularly significant in this lexical subgroup is the term *pledge* [4], newly depicting the *lap-child* as a token of mutual love and duty between the parents. A late perception of the *lap-child* as beautiful and innocent is also to be noted (*fairy* [6], *cherub* [8]). Peripherally, one may add that a marginal scheme based on the idea that alien is 'ugly' or 'low status' is also perceived in *tadpole* and *ragazzo.*[18]

All in all, our results show that the *lap-child* stereotype (or the 'child' as [CHILD], as in the *HTOED* node) is built in positive terms of lightness, vivacity, playfulness and tenderness, and as beings deserving caresses and delicacy (*fairy*). However, the emotional load in these terms leads us to introduce a caveat here. Whereas few words or word extensions are negatively loaded,[19] it cannot be ignored that the denotative value of smallness easily slides into the idea of insignificance or lack of importance, as in *bratchet.* Likewise, the presence of a subset of unstable terms, apparently fluctuating towards the notions of manipulation and mild or more serious contempt, such as *puppet*, *urchin*, *chitterling*, may be of note.

3.2.4 The man-boy

The fourth *HTOED* subcategory of 'child' as [YOUNG MAN], prototypically labelled as the *man-boy*, may be said to be the most complex and the second most conspicuous in the corpus. It comes forth earlier in the period and scores 30% of the total sum of new creations. In the construal of the *man-boy*, a change of focus is appreciated, the adult being now what constitutes the standard for assessment. Our results confirm a physical or metaphorical evaluation of members in this category according to aesthetic norms prevailing in a broader social context, in an outer sphere. Early Modern English boys seem to be construed by the adult community as the non-alike, and essentially evaluated along two dimensions: physical development and behaviour. Throughout the period the *man-boy* stereotype is steadily

[18] Cf. Sakalauskaite (2010) on the cross-linguistic uses and negative connotations of tadpoles and toads. During the period *tadpole* is applied to black infants. As for *ragazzo*, it was a synonym for "servant" or "pageboy".

[19] Cf., however, *bantling* (with "bastard" connotations); *whelpling* (apparently used for vile creatures "acting on impulses that most resemble animals"); *butter-print* (understood as "burden"), or the already mentioned *ragazzo* and *tadpole*.

described as undeveloped, as a "rudiment of an adult" (e.g. *demy*, *halflang*, *spaught*, *codling*, *cub*, *callow*). Through terms denoting plants rather than animals (e.g. *spear*, *spring*, *imp*, *stubbed*, *sprig*), boys are judged half-tall, thin, raw, shapeless or clumsy, features easily interpreted as "defective", expressing some kind of disapproval or subjective negative evaluation, bordering contempt. Notwithstanding the above, the *man-boy* is likewise distinguished as lively and vigorous (*youngster* [4]), something not too surprising at a time when size and strength were more important than age (Mintz 1993). As for social behaviour, both positive and negative judgments are brought up in the corpus. From the very beginning, the *man-boy* might be reckoned as gay and fashionable, even before lively or strong (*younker* [1], *hensour* [1]). Alongside, more negative profiles are found. Youngsters were apparently glimpsed as mischievous (*wag* [3]); idle (*loon* [3]); boastful (*cockerel* [3], *cockling* [4]); conceited or arrogant (*snipper-snapper* [4]).

As mentioned above, the analysis of synonyms recorded for this category certainly proves complex, given that there seems to be reasonable evidence to defend a multifaceted scheme of perception. The lack of univocal undertones in many of the terms in this subset apparently allows for a two-fold view of the *man-boy*: that of the adult confronting men-to-be, just mentioned, and a second peer-based perception. That is, it might be the case that some of the items in the category could be recognised as the lexicalized expression of an in-group grasp; some kind of horizontal assessment that would add to the adult's vertical view. According to this assumption, terms could be used either by adults or youngsters with different connotations. Positive uses could express adult's endearment or peer camaraderie, depending on the context. In turn, negatively-loaded expressions might be used descriptively or contemptuously by adults, or rather undergo a process of "reversal of judgement", whereby bad behaviour, toughness, etc. (as in *snipper-snapper*, *wag*, *loon*, etc.), would turn into pride of group or identity markers rather than contempt (Eckert 2003). However, that would require both a contextualized, discursive interpretation of positive and negative undertones suggested in definitions and labels, and the consideration of other significant features that seem to concur in this subclass of terms. The hypothesis, thus, remains open to conjecture.

3.2.5 The woman-child

The analysis of the corpus provides evidence of a last (and uncertain) category of Early Modern English childhood: that of 'child' as GIRL, very expressively stereotyped as a *woman-child* [1] or a *little girl* [5]. It is worth recalling that the

word *girl* entered the language in 1375 meaning "young woman", and only from 1400 was it used to refer to a female child. Contrary to what happened with the *man-boy*, the *woman-child* does not seem to be a discovery in the Early Modern English period, probably because she had never been perceived as a child in the past, and she was *not yet* perceived as a child during this stage. Save for a certain increase of attention recorded in period 5, the relatively low degree of lexical and semantic innovation affecting the category (18%), may be a clear indicator of this claim. Moreover, in accordance with our results, the *woman-child* lacks a clear profile as a life-stage. There is still a blurred distinction between *girl* and *woman*. Though the category comprises some new descriptive terms, like *woman-child* [1], *ladykin* [6], *little girl* [5] (pointing to their short stature), these come mixed up with other synonyms pointing rather to the girl social condition of being unmarried and susceptible to being wooed or courted (Cf. *miss* [7], *vriester* [7]*)*. Moreover, the existence of comparatively few diminutives in the corpus points to the lack of endearment nuances in the perception of this category. It is true that words like *tendril* [5], *filly* [5] and *ladykin* [6] are described as terms of affection. However, it is to note that some -if not all- of them may constitute pseudo-euphemistic terms of abuse, or social and moral indicators. More particularly, it may be the case that *filly*, an instance of equine terms, would fit the interpretation by Borkowska – Kleparski (2007: 43), following Kiełtyka (2005), who identify the use of "mare" images as frequently used for contemptible women. Similarly, the presence of nicknames in the corpus, such as *prill* [4] or *maggie* [5], frequently classified as devices to express affection, may be argued to be socially and morally loaded, close to "cheeky", "rustic", "low", "flirtatious", "roguish", as suggested by dictionary cross-references. The *woman-child* is, thus, a clear case of semantic imbalance in the proportion of positive-negative undertones. According to our results, words in this *HTOED* sub-corpus seem to experience early processes of pejoration, an otherwise well-researched topic in the literature, as shown in Hughes (2000), Borkowska – Kleparski (2007) or Łozowski (2015), among others.

3.3 An age-blurred childhood

As will have been noticed, nothing has been said thus far regarding age-boundaries separating duly identified categories. In fact, our corpus definition of childhood remains age-blurred, but for the term *two-year-old*, given as a synonym for the 'child' as [CHILD]. This keeps in line with social studies, where the possibility of stages being tightly linked to a certain age or age-

range is still a line of enquiry. And furthermore, there seems to be a lack of agreement in the literature as for childhood age-values (Orme 2001; Clarke 2004; Cunningham 2006; Dekker 2012; Frijhoff 2012). As shown in our corpus, qualifying indexes other than age are used as substage categorizing factors, such as the child's walking ability, their size, and degree of dependency, the home routines involved in their caring and conspicuous gendering issues. In view of the foregoing, and considering the prominence of both the *lap-child* and the *man-boy* as the two most-profusely portrayed stages during the period, one might conclude that Early Modern English childhood could initially be depicted as a space comprising two separate spheres. On the one hand, the home, a private sphere for children up to the *lap-child*, where boys and girls would be treated the same, parents developing emotional links towards them[20]. On the other, a more public sphere, inhabited by would-be adults, where children would be judged accordingly. In this second and outer sphere, institutions like schooling or apprenticeship would act as strong gendering agents, as the stereotyping labels of *man-boy* and *woman-girl* suggest.[21]

4. Concluding remarks

Our results confirm that Cultural Linguistics can contribute interesting insights to AS. The lexicographical and semantic approach utilised here has proved illuminating in the reconstruction of Early Modern English cultural schemes responsible for the concept of childhood, allowing us to contend that childhood emerged as a cultural component of English society during that period. Confirming advances in social studies, the corpus provides cumulative evidence that during the Early Modern period childhood is recognised as a distinct stage of life, children becoming visible

[20] Though the corpus does not give any hints on the issue, this would apparently affect the commoner family rather than the nobility, where children would usually be left under the care of non-related women (see Orme 2001).

[21] The urge to put further research effort into this later childhood stage, at the cross-roads of contemporary childhood, adolescence and youth, proves binding. Merging our tentative conclusions with the historical evidence available, age 7 would apparently mark a gradual move from infancy to later childhood. At this vague marker of a wider range of 7 to 12 years, children would be sent away as apprentices and committed to tutors of whatever sort, which would result in an increasing separation between the spheres of home, education and work, especially for males. Running up to the age of 20 or over, these youngsters would be deemed old enough to protect themselves, while still dependent on adults. (Cf. Ben-Amos 1994; Dekker 2012; Frijhoff 2012; Mawhinney 2015)

as separate autonomous beings, beyond social role and external assessment of appearance, as was apparently the case prior to the fifteenth century. Moreover, the heightened lexical activity of the period points to a new structuring of patterns and meanings involved in the process of growing up. However, considering that the written language is conservative in nature, this should have begun earlier than the sixteenth century. Interestingly, the singular composition of the corpus suggests that the concept of childhood is at an incipient moment of perception and creation. Early Modern English exhibits no fixed terms yet for the new construct, and lexical innovation apparently reflects individual needs rather than the social institutionalization and the linguistic normalization of childhood. Closely connected with this, our results suggest that Early Modern childhood is not a prestigious stage overall, since undertones of unimportance, underdevelopment, burden or lack of a polished nature come up too often. However, a more detailed approach to this issue is required.

During the Early Modern English period five childhood identities or stereotypes seem to emerge, both socially and biographically, the most important divide being that between early and late childhood: or between the family and an outer domain. The focus of innovation seems to lie on two particularly salient categories: the *lap-child* and the *man-boy*, the girl not yet having been discovered as a child. As for the time-based definition of childhood, it remains age-blurred, other indexes being used for the qualification of main childhood substages.

Given that lexicographical sources offer few hints on issues of class, ethnicity, educational levels, etc., the social dimension of Early Modern English childhood has not been addressed. This limitation of research should be counterbalanced in the future. Likewise, a natural progression of this work should attend to the cultural construal of the broad and complex category of later childhood along the time-axis.

REFERENCES

Sources

Historical Thesaurus of the Oxford English Dictionary (*HTOED*)
https://www.oed.com/public/htoed/loginpage, accessed March 2019
Oxford English Dictionary (*OED* online)
http://www.oed.com/, accessed March 2019

Special studies

Ariès, P.
1962 *Centuries of Childhood: A Social History of Family Life*. Translated by R. Baldick. New York: Jonathan Cape.
Bartmiński, J.
2009 *Aspects of Cognitive Ethnolinguistics*. Edited by J. Zinken. London; Oakville, CT: Equinox.
2016 "Ethnolinguistics in the year 2016", *Ethnolinguistic* 28, 9-31.
Ben-Amos, I.K.
1994 *Adolescence and Youth in Early Modern England*. New Haven, Conn.: Yale University Press.
Biggs, S. – J. Hendricks – A. Lowenstein
2003 *The Need for Theory: Critical Approaches to Social Gerontology*. New York: Baywood.
Borkowska, P. – G.A. Kleparski
2007 "It befalls words to fall down: Pejoration as a type of semantic change", *Studia Anglica Resoviensia* 4, 33-50.
Brinton, L.J. – E. Closs Traugott
2005 *Lexicalization and Language Change*. New York: Cambridge University Press
Busse, B.
2012 "A celebration of words and ideas: The stylistic potential of the *Historical Thesaurus of the Oxford English Dictionary*", *Language and Literature* 21 (1), 84-92.
Charise, A.
2014 "The future is certain: Manifesting age, culture, Humanities", *Age, Culture, Humanities* 1, 11-16.
Clarke, J.
2004 "Histories of childhood". In: D. Wyse (ed.) *Childhood Studies: An Introduction*. Malden: Blackwell Publishers, 3-12.
Cole, T.R. – R.E. Ray
2010 "Introduction: The humanistic study of aging past and present, or why gerontology still needs interpretive inquiry". In: T.R. Cole – R.E. Ray – R. Kastenbaum (eds.) *A Guide to Humanistic Studies in Aging: What Does It Mean to Grow Old?* Baltimore: The Johns Hopkins University Press, 1-29.
Coleman, J.
2012 "Using dictionaries and thesauruses as evidence". In: T. Nevalainen – E. Closs Traugott (eds.) *The Oxford Handbook of The History of English*. Oxford: Oxford University Press, 98-110.
Coupland, J.
2009 "Discourse, identity and change in mid-to-late life: Interdisciplinary perspectives on language and ageing", *Ageing & Society* 29, 849-861.

Crystal, D.
2014 *Words in Time and Place: Exploring Language Through* the Historical Thesaurus of the Oxford English Dictionary. Oxford: Oxford University Press
Cunningham, H.
2006 *The Invention of Childhood*. London: BBC Books.
Dekker, J.J.H. et al.
2012 "Discoveries of childhood in history: An introduction", *Paedagogica Historica* 48 (1), 1-9.
DeMause, L.
1974 "The evolution of childhood". In: L. DeMause (ed.) *The History of Childhood*. New York: Psychohistory Press, 1-73.
Eckert, P.
2003 "Language and adolescent peer groups", *Journal of Language and Social Psychology* 22 (1), 112-118.
Fass, P. (ed.).
2009 *The Routledge History of Childhood in the Western World*. Abingdon: Routledge.
Frijhoff, W.
2012 "Historian's discovery of childhood", *Paedagogica Historica* 48 (1), 11-29.
Goffman, E.
1983 "The interaction order", *American Sociological Review* 48 (1), 1-17.
Gullette, M.M.
1993 "Creativity, aging, gender: A study of their intersections, 1910-1935". In: A. Wyatt-Brown – J. Rossen (eds.) *Aging and Gender in Literature. Studies in Creativity*. Charlottesville, VA.: University Press of Virginia, 19-48.
Hagestad, G.O. – D. Dannefer
2001 "Concepts and theories of ageing: Beyond microfication in social science approaches." In: R.H. Binstock – L.K. George (eds). *Handbook of Aging and the Social Sciences* (5th edn.). New York: Academic Press, 3-21.
Hughes, G.
2000 *A History of English Words*. Oxford: Blackwell Publishers.
Jurafsky, D.
1996 "Universal tendencies in the semantics of the diminutive", *Language* 72 (3), 533-578.
Katz, S.
2014 "What is Age Studies?", *Age, Culture, Humanities* 1, 17-23.
Kay, C.
2010 "*Historical Thesaurus of the Oxford English Dictionary*: Introduction". In: M. Adams (ed.) *Cunning Passages, Contrived Corridors: Unexpected Essays in the History of Lexicography.* Monza: Polimetrica, 255-270.

Kiełtyka, R.
2005 "The axiological-cognitive analysis of the evaluative developments in the domain of EQUIDAE: A pilot study", *Studia Anglica Resoviensia* 3, 59-75.

Kiełtyka, R. – G.A. Kleparski
2005 "The scope of English zoosemy: The case of domesticated animals", *Studia Anglica Resoviensia* 3, 76-87.

King, M.
2007 "Concepts of childhood: What we know and where we might go", *Renaissance Quarterly* 60 (2), 371-407.

Kline, D.T.
2008 Review of A. Classen (ed.) *Childhood in the Middle Ages and the Renaissance: The Results of a Paradigm Shift in the History of Mentality*. Berlin: Walter de Gruyter, 2005, http://www.h-net.org/reviews/showrev.php?id=22822, accessed March 2019

Lowe, R.
2004 "Childhood through the ages". In: T. Maynard – N. Thomas (eds.) *An Introduction to Early Childhood Studies*. London: SAGE, 65-75.

Łozowski, P.
2015 "Towards a lexicographic account of GIRL: Forms, meanings and values", *Journal of Language and Cultural Education* 3 (3), 16-29.

Marchand, H.
1969 *The Categories and Types of Present-day English Word-Formation: A Synchronic-Diachronic Approach* (2nd end.). Munich: C.H. Beck.

Mawhinney, S.E.
2015 *Coming of Age: Youth in England 1400-1600*. Unpublished PhD thesis, York: University of York.

Mintz, S.
1993 "Life stages". In: M.K. Cayton – E.J. Gorn – P.W. Williams (eds.) *Encyclopedia of American Social History*. Vol 3. New York: Scribner, 2011-2022.

Nevalainen, T.
2000 "Early Modern English lexis and semantics". In: R. Lass (ed.) *The Cambridge History of the English Language.* Vol. 3: *1476-1776*. Cambridge: Cambridge University Press, 332-458.

Nikander, P.
2009 "Doing change and continuity: Age identity and the micro–macro divide", *Ageing & Society* 29 (6), 861-879.

Orme, N.
2001 *Medieval Children*. New Haven, Conn.; London: Yale University Press.

Ponsonnet, M.
2018 "A preliminary typology of emotional connotations in morphological diminutives and augmentatives". In: M. Ponsonnet – M. Vuillermet

(eds.) *Morphology and Emotions across the World's Languages*. Amsterdam: John Benjamins, 17-50.

Prout, A. – A. James
1990 *Constructing and Reconstructing Childhood.* London: Routledge.

Sakalauskaite, A.
2010 *Zoometaphors in English, German, and Lithuanian: A Corpus Study*. Unpublished PhD thesis, Berkeley, CA: University of California.

Sapir, E.
1921 *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.
1957 "Language". In: D.G. Mandelbaum (ed.) *Culture, Language and Personality: Selected Essays*. Berkeley, CA: University of California Press, 1-44.

Sayers, W.
2016 "English etymologies from the popular register", *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 133 (4), 171-181.

Segal, L.
2014 "The coming of Age Studies", *AgeCultureHumanities* 1, 31-34.

Steriopolo, O.
2016 "Expressives across languages", *Annual Review of Linguistic*s 2, 293-324.

Tejada Caller, P.
2019 "Speakers as cultural insiders. Or what linguists may contribute to the idea of a 'normal life'". Plenary talk given at *The 6th IABA-Europe Conference "Knowing the Self: Auto/Biographical Narratives and the History of Knowledge"*. Madrid, 19-21 June 2019.

Thomas, N.
2004 "Sociology of childhood". In: T. Maynard – N. Thomas (eds.) *An Introduction to Early Childhood Studies*. London: SAGE, 75-85.

Tomasello, M.
1999 *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.

Tufvesson, S.
2007 "Expressives". In: A. Majid (ed.) *Field Manual Volume 10*. Nijmegen: Max Planck Institute for Psycholinguistics, 53-58.

Wierzbicka, A.
1984 "Diminutives and depreciatives: Semantic representation for derivational categories", *Quaderni di Semantica* 5, 123-130.

Address: PALOMA TEJADA CALLER, Department of English Studies, Facultad de Filología, Universidad Complutense de Madrid, Plaza Menéndez Pelayo s/n, 28040 Madrid, España.
ORCID code: orcid.org/0000-0002-1201-9839.

Maria Edgeworth's choice of auxiliary verb in perfect tenses

Nuria Calvo Cortés
Universidad Complutense Madrid

ABSTRACT

The present study concentrates on Maria Edgeworth's use of auxiliary verbs (*have* or *be*) in combination with the participle *gone* in perfect tenses and the possible reasons for the choice of one or the other. A corpus containing most of her novels and moral tales was compiled and all the examples in which *gone* appeared were extracted and later analysed manually. The results show a clear preference for *be*, even in her later works, dating from when *have* was already well-established in the language as the auxiliary verb for perfect tenses. As for the reasons, while Maria Edgeworth was thought to have been influenced by the Irish English variety in her choice, a comparison between her use of auxiliaries in the narrative sections of her works and in the voice of the characters might indicate a possible manipulation by editors or an intention to show differences in speech. Similarly, it is also likely that some of the components of motion situations may have motivated her choices.

Keywords: Late Modern English, perfect tenses, choice of auxiliary, motion situations, editors' manipulation.

1. Introduction

Maria Edgeworth wrote her novels and moral tales between the end of the 18th century and the first half of the 19th century. She lived most of her life in Ireland, and her writings often reflect the variety of English spoken in Ireland at that time (Hollingworth 1997; Ó Gallchóir 2005; Manly 2007).

The period when Edgeworth wrote coincided with the time in which the verb *have* became more frequently used in combination with past participles to create perfect tenses (Rydén – Brorström 1987). Until then, *be*

DOI: 10.25951/4364.

had been the most common auxiliary verb in such structures, particularly when combined with intransitive verbs of motion.

Although perfect tenses and their evolution in British English have been analysed (cf. 2.1), Edgeworth has not been included in those studies, probably due to her being classified as an Anglo-Irish writer. Similarly, most of these studies have not taken into account the semantic components of motion situations despite the fact that most of the verbs taking longer to be used in combination with *have* are verbs that refer to some kind of motion (e.g. *go*, *come*, *become* or *grow*). From a cognitive linguistics perspective, a typical motion situation involves the presence of a *figure* that moves along a *path* in relation to a *ground* (cf. 3.1). For instance, in *The girl ran into the park*, *The girl* is the *figure*, *into* lexicalises the *path* and the *ground* is represented by *the park*. Semantics and syntax are connected and some of the complementation patterns observed in previous studies (cf. 2.1) represent the *path* and the *ground*. However, whereas the semantic components are always present, they may not be reflected in the surface representation of the sentence. This can be observed in *The girl ran in*, where the *ground* is not present in the surface structure, but semantically it is obvious that the girl has moved into a particular *ground*, which is probably omitted because it can be inferred from the context.

The present study focuses on the analysis of the choice of auxiliary verb in perfect structures with the participle *gone* in Maria Edgeworth's main works. In order to observe if she might have been influenced by Irish English, whether her choices might have been motivated by the components of motion situations, or if a possible intervention of editors may explain the auxiliaries found in her works, a corpus-based study was carried out. The results were compared to equivalent usage by other British female writers of the same period, since they were supposed to be more conservative than male authors in the introduction of *have* in this particular grammatical structure. The high number of similarities among the writers leads to the conclusion that Edgeworth was probably following the same pattern as other female writers. This might have been due to the presence of specific components of motion situations, rather than to the influence of the Irish English variety of the language. The differences between the results found in the narrator's and the characters' voices also point to the possibility that the editors intervened and changed some of the auxiliary verbs in the narrator's voice to conform to the usage that was becoming common at the time, that is, the increase in the use of *have* as opposed to *be* in these structures. Alternatively, Edgeworth may have wanted to reflect these differences herself in order to show different types of speech.

The study is limited to the verb *go*, as it is one of the verbs that was used in combination with *be* in perfect structures for a longer time (Kytö 1997). Also, it is present in both physical and abstract contexts. Even today in some English dialects it is still fairly common to use expressions such as *he is gone* with a physical locative meaning as well as a metaphorical one (Anderwald 2014).

2. Previous studies

2.1 Perfect tenses in Late Modern English

Perfect tenses in Late Modern English became the subject of study in the 1980s and 1990s coinciding with the time when this specific period of the English language attracted the attention of researchers (Beal 2004).

Rydén and Brorström (1987) were the first to offer a comprehensive description of the structure of perfect tenses with mutative verbs in the 18th and 19th centuries, including in their study two types of texts, letters and comedies, as well as numerous authors of both centuries. They offer a detailed description of the uses of both auxiliaries during the period analysed and, as regards the verb *go*, they show that whereas in the 18th century *be* is used more frequently than *have* in expresions such as *she is gone somewhere*, in the first half of the 19th century the uses of both auxiliaries seem to be balanced, and it is in the second half of the 19th century when *have* is more commonly used, and examples as *they have gone somewhere* become the norm. Rydén and Brorström (1987) also identify differences in the usage of the two auxiliaries depending on the type of sentence or complementation present in the sentence. In some ways, the categories that they establish are fairly similar to the ones used in the present study, with the difference that here the components of motion events are considered. In addition, despite the comprehensive study in terms of number of authors and tokens analysed, it is restricted to two types of genre, which as they indicate "probably come reasonably close to the "real" speech of the day" (Rydén – Brorström 1987: 13). They come to several conclusions regarding the choice of auxiliary verb. The use of *have* appears to be conditioned by specific contexts, such as those of some hypothetical sentences, iterative/durative contexts and some verbal forms. They also acknowledge that there are some authors who seem to follow their own patterns. Furthermore, when comparing the works written by men and by women, in general, women tended to use *be* more

often than men did so, and for a longer time, even in the 19th century, when *have* had become fully established as the auxiliary verb used in combination with the past participle form of the verbs.

Kytö also concentrates on the study of perfect tenses with intransitive verbs of "transition or change" (1997: 17). Her treatment covers a long period in the history of the English language, from Middle to Modern English, and the mid-18th century is suggested to be "the final turn-over in the history of the paradigm" (Kytö 1997: 32). The study presents a detailed analysis of both the external and the internal factors influencing the choice of auxiliary in perfect tenses. They include differences in text type, gender, sentence type and complementation in general. Some of these aspects had already been discussed in Rydén – Brorström (1987), but as in that previous study, the components of motion situations are not considered. Finally, the author also acknowledges the need for further studies to rule out other possible influences.

The 21st century has witnessed additional studies of this issue. One such is Sorace (2000). English is among the European languages analysed, and the focus is on the choice of auxiliary possibly being motivated by the differences shown by the types of intransitive verbs used. A number of varieties of English are studied by Werner et al. (2016), who conclude that the auxiliary verb *be* is still commonly used by speakers of some of these varieties. McFadden (2017) dates the disappearance of *be* as a perfect auxiliary to the beginning of the 19th century, slightly earlier than do Rydén – Brorström (1987). However, all these studies emphasise that further investigation is required.

More recently, the perfect tenses of Jane Austen's novels together with those of her letters have been analysed (Calvo Cortés 2019). Although the present enquiry takes previous work into consideration, it also adds two new aspects in connection with the choice of auxiliary. First, that choice might have been influenced by editor manipulation since there is a considerable difference between the use of these auxiliaries in the novels and in the letters. Whereas in the letters *be* is used in 87.06% of the cases, in the novels this figure is reduced to 71.59%, which is still higher in comparison with the use of *have*, but not as high. At the same time, the letters display a more personal way of writing, which could have made them less likely to have undergone such alteration. The second aspect refers to the influence of the components of motion events (cf. 3.1) on the choice of auxiliary. In this respect, the analysis of Jane Austen's work suggests that it is the different types of *ground* that seem to condition the type of auxiliary used whereas the *figure* does not appear to have much influence on this.

Finally, other British women writers of the 18th century have also been studied in relation to their use of these two auxiliaries to form perfect tenses and the possible connection of their choices with the components of motion situations (Calvo Cortés 2020). This study shows variation among the different writers although some similarities can be observed. For instance, there seems to be a preference for non-human *figures* to combine with *have* more often than with *be* together with verbs such as *fall*. Similarly, the type of *ground* appears to condition the type of auxiliary. In addition, most of these authors show differences between the narrator's and the character's voice in relation to their use of *have* or *be*, as *have* is the preferred auxiliary verb of the narrator, as opposed to *be*, which tends to be more frequently used by the characters.

2.2 Irish English and perfect tenses

Regarding Irish English, much has been discussed about perfect tenses in this variety of the English language, including the extent to which the Irish language might have influenced the perfect tense forms found in the English variety spoken in Ireland (Harris 1993; Filppula 1999; Siemund 2004; Hickey 2000, 2007; Ronan 2012; McCafferty 2014). These studies refer to the variety of forms used for perfect tenses in Irish English, and the different meanings that they express depending on the contexts where they are used.

In Irish there is no equivalent to the English verb *have*, therefore, the perfect aspect structures with this auxiliary verb could only have come from English. On the other hand, in Irish there are structures with the verb *be* expressing meanings similar to those of some of the English perfective structures (for instance, indicating either an absent state, e.g. *Tá sé imithe* ('Is he/it gone'), or a more dynamic or resultative state, e.g. *Tá sé imithe abhaile* ('Is he/it gone home'), which might imply the influence of Irish on the English variety spoken in Ireland. Nevertheless, it is not clear to what extent this influence contributed to the permanence of the perfective structures with *be* in Irish English, since in Early Modern English *be* was the most common auxiliary verb used in these structures, and there are other European languages, e.g. German, which have maintained the distinction between the auxiliaries depending on the lexical verb that accompanies them.

All in all, most of these studies conclude in some manner that the perfect tenses with *be*, as used in Ireland, probably came from English but they would have been reinforced by the Irish language (Filppula 1999).

2.3 Maria Edgeworth and Irish English

Despite having been born in England, Maria Edgeworth spent most of her life in Ireland. This has caused her to be classified as an Anglo-Irish writer, who has been regarded as a clear example of a writer who used vernacular Irish English. According to Ó Gallchóir (2005: 132), she was a "faithful and innovative recorder of vernacular speech", and Hollingworth (1997) also indicates that she used the vernacular language at least in some of her writings. In addition, she has been considered a defender of the richness of the Irish language, as Manly (2007: 160) explains: for Maria Edgeworth, Irish speakers "from the highest to the lowest, in daily conversation […] employ a superfluity of wit and metaphor […] astonishing and unintelligible to a majority of […] English yeomen', and the Irish language was 'one of pristine purity".

3. The present study

Due to her interest in the vernacular Irish English speech, Maria Edgeworth was expected to reflect in her writings the language that the people around her would have spoken in Ireland at the time. Consequently, regarding perfective structures, she was supposed to offer the possibility to discover if Irish English was different from the English spoken in England in the same period, as well as from the English used by some of her contemporary English female writers.

Since the verb *go* is a clear example of a motion verb, probably the most basic one signifying its main semantic concept (cf. 3.1), it seemed to be the logical starting point for the present study.

3.1 Features of a basic motion situation

Before providing a detailed description of this investigation, it is essential to clarify the concept of motion situation. So far, the analyses of perfect structures do not seem to have considered the possible influence of motion situations, with the exception of the mentioned studies of Jane Austen (Calvo Cortés 2019), and of 18th-century British women writers (Calvo Cortés 2020). Given the fact that the verbs that resisted the use of *have* for a longer time are verbs that encode some semantic content of motion, an analysis of these structures is vital.

After Talmy's description of lexicalization patterns (1987), motion situations became prominent in cognitive linguistics studies. These studies have mainly concentrated on the description of the elements of motion situations (Talmy 1975, 1987, 2000), in their connection with metaphorical contexts (Ramscar et al. 2009), and in the classification of languages in the world depending on the representation of the different components of a motion situation in the surface structure (Matsumoto 1996; Slobin 1996; Talmy 2000).

A basic motion situation (cf. Fig. 1) involves four main semantic concepts. They are the *figure*, or object that moves or is located somewhere; the *ground*, the place where the *figure* is moving to or where it is located; the *path*, which refers to the trajectory of the movement; and the motion, which can be represented in the syntactic structure either by a *deep Be-located* verb (e.g. *be*) or a *deep Move* verb such as *come* (Talmy 2000).

The girl (*figure*) is going (*motion*) to (*path*) the park (*ground*)

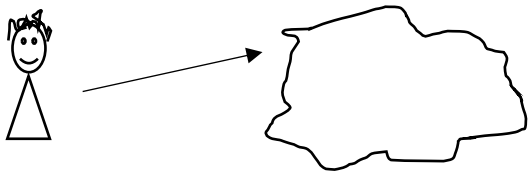

Figure 1. Basic motion situation

The verb analysed in the present study, *gone* (base form *go*), is a *deep Move* verb. This means that the situations in which it is present are motion situations or motion events, which implies the presence of the other components that have been described above. The analysis of these components in the examples found will be essential in order to understand the choice of the auxiliary.

Gone is a very basic verb of motion in the sense that it does not lexicalise any element involved in the motion event apart from motion. In other words, the *figure*, the *ground* and the *path*, as well as any other possible co-events, such as *manner*, are realised by other words or phrases, as long as they are present in the syntactic representation. This explains, for example, the need to use a prepositional phrase after *go* to indicate the *path* and the *ground* as in *he has gone into the room*, whereas a verb such as *enter* does not

require the same kind of complement phrase since it incorporates the *path* in the semantic content of the verb, and only the *ground* requires to be represented by a separate component, a noun phrase in this case, as in *he has entered the room*.

Apart from the component of motion, which is clearly present in the instances analysed here, the components of *figure* and *ground* will be those considered in the present study as they may be responsible for the variation shown in these structures. However, the *path* component does not seem to influence the changes in these types of structures unless it is combined with the *ground* in one lexical element, as for instance in *aboard* (Calvo Cortés 2014), therefore, this component will be mentioned only in relation to the *ground*.

3.2 Research questions and hypotheses

The choice of auxiliary verb used in perfect tenses found in Maria Edgeworth's work may have been motivated by various reasons. In the present paper, four questions and hypotheses are suggested and they will become the focus of the analysis. It is probable that not just one but a combination of several reasons contributed to Edgeworth's choices.

First, the probable intervention of the editors. Differences found in other writers (Calvo Cortés 2020) regarding the use of the two auxiliaries in the narrator's or the characters' voices were also thought possibly to be present in Maria Edgeworth's works. Although access to the original manuscripts would be required to confirm this, if notable differences are observed in this respect, the editors' influence may be considered as a possible cause of differences.

Second, the potential variation according to the type of verbal forms in which the auxiliaries appear. Previous studies have pointed at these differences (cf. 2.1). It is expected that Maria Edgeworth's use of *be* and *have* will also vary according to whether the auxiliary verb is in the infinitive, gerund, past or present.

Third, the likely effect of the different components of motion events. As the semantics and syntax of motion events may have contributed to the choice of auxiliary made by other writers (Calvo Cortés 2019, 2020), Maria Edgeworth might have been influenced by this in a similar way. This implies that the structures in which *be* or *have* are present will be connected to different types of *figures* and/or *grounds* present in each structure.

Fourth, the possible influence of the Irish English variety. As mentioned above, in Irish there is no verb *have* equivalent to the English verb *have*

used in perfect tenses, whereas structures similar to the ones with *be* are present and they are commonly used to indicate an absent state, or a more dynamic or resultative state (cf. 2.2). Even though there is no agreement on the extent of the influence of the Irish language on Irish English, it may have contributed to the use of *be*-perfect structures for a longer time in Irish English. Consequently, it is not unreasonable to think that this might have influenced Maria Edgeworth's writings, which means that it was expected to find more examples with *be* in the corpus.

3.3 Method

In order to carry out the present study, a corpus of works by Maria Edgeworth was compiled. This compilation of moral tales and novels was obtained from texts available in *the Guttenberg Project*, and the total number of words in the corpus is 1,813,408.

The *Antconc* software programme was used to extract all the tokens containing the participle form *gone*, 540 in total. The first step was to discard the examples that were not suitable for the study. They included: all the tokens where the participle verb appeared on its own, that is without an auxiliary verb, and those amounted to 66; the examples that included *'s*, which were 22, since it was not possible to determine if those instances of *'s* represented *has* or *is*, as they could be either; and finally, the 4 occurrences of the combination *have been gone*, since it incorporates both auxiliaries and the meaning could differ.

All the remaining examples, 446, were then analysed manually, and a Fisher Exact test was carried out to calculate the p value and therefore, the significance of the different results. A Fisher Exact test was chosen as opposed to a Chi-square test because some of the data are 5 or fewer, which impedes the use of the latter. The analysis started with the division of the examples into two groups, the ones with *be* and those with *have*. A distribution was also established taking into account whether the perfect structures were found in the narration of the texts or in the sections corresponding to the voice of the characters. This was followed by an analysis of the different verbal forms used, which included present perfect tenses (*have/has gone* and *am/is/are gone*), past perfect tenses (*had gone* and *was/were gone*), an infinitive (*have gone* and *be gone*) and *-ing* forms (i.e. *having gone* or *being gone*).

Next, since the components of motion events were expected to influence the choice of auxiliaries, the *figure* and the *ground* present in the extracted examples were all analysed and grouped according to separate

categories. As regards the *figure*, the examples were divided into those containing a human *figure* (e.g. *he* in *He had gone directly to the inn*) and those containing a non-human *figure* (e.g. *things* in *things have gone so far*). As for the *ground*, it was important to distinguish at first if the *ground* was present in the surface structure (e.g. *he was gone to a neighbouring town*, where *neighbouring town* is the *ground*) or not (e.g. *he is gone*, where there is no explicit *ground* in the sentence, although it might be implied in the context). Secondly, the examples that contained an explicit *ground* displayed a wide variety of *grounds*. However, a major distinction was established between those that were physical, i.e. *into the church* in *He had gone into the church*, and those that were metaphorical, i.e. *mad* in *He has gone mad*.

Finally, a comparison was established between Maria Edgeworth's use of *be* and *have* in perfect structures with *gone* and those of four contemporaneous English female writers, namely Jane Austen, Fanny Burney, Elizabeth Inchbald and Ann Radcliffe.

3.4 Results and discussion

The results included in this section refer to the four main aspects analysed. Firstly, the general results of the presence of both auxiliaries in the total examples are quantified and presented. This section also includes the results of the distribution of examples according to the voice, that is, the narrator's or the characters'. Secondly, the combinations of the two auxiliary verbs with *gone* are compared considering the different types of verbal forms used, namely present, past, infinitive or present participle (*-ing*). Thirdly, the *figure* and *ground* elements present in each of the examples are described. Finally, the results found in Maria Edgeworth's corpus are compared to the ones found in some of her contemporary female writers.

3.4.1 General results and voice

The data analysed in the corpus show a clear preference for the use of the auxiliary verb *be* as opposed to *have*. A total of 141 of the 446 examples analysed, that is, 31.61%, include *have* as the auxiliary verb, whereas this figure increases dramatically to 68.39%, that is 305 examples, in the case of the auxiliary verb *be*.

Although *be* is the predominant verb in the whole corpus, the use of the two auxiliaries displays measurable variation when the variable of voice is considered. As shown in Table 1, despite the higher number of tokens with *be* in both the narrator's and the characters' voices, it is in the latter where

the use of *be* surpasses amply that of *have* (79% versus 21%). In contrast, the difference between the two auxiliaries is much reduced in the narrator's voice, as *be* is present in 58% of the examples as opposed to 42% in the case of *have*. The statistical analysis shows a p value of 0.0001, which means that this result is significant, as it is < 0.05. Therefore, it can be stated that the presence of the two auxiliary verbs in the different voices is not random, and either the editors or Maria Edgeworth herself would have opted for indicating this difference. The editors may have wanted to change the *be*-perfect forms for *have*-perfects to reflect the tendency of the time whereas Edgeworth may have intended to show differences between the characters and the narrators, with the characters' speech reflecting a variety closer to Irish English.

Table 1. Distribution of examples according to voice

	Be	*Have*
Narrator's voice	58% (128)[1]	42% (93)
Character's voice	79% (177)	21% (48)

[1] Raw figures are all presented in parentheses in all the tables.

3.4.2 Verbal forms

Since the corpus used for the present analysis contains tales and novels, many of the verbs used are in past perfect tenses and present perfect tenses. Past perfect tenses are usually embedded in the narration, while present perfect tenses are the ones more often used by the characters in their own speech.

In addition, two other verbal forms were present in the corpus, the perfect infinitive, *be gone* or *have gone*, and the present participial form, *having gone*. As regards the perfect infinitive form, in most cases it appears following a modal verb, i.e. *She could have gone earlier*.

As shown in Table 2, some differences can be observed in the use of both auxiliaries in combination with one type of structure or another. The first result that stands out is that the forms with *have* are scattered and used in all the contexts. However, the distribution is not balanced. If the uses of the infinitive and the participial form are not considered to begin with, there seems to be a fair discrepancy in the use of *have* in present and past tenses, as *have* appears more often in past than in present form. On the contrary, when the results of *be* are observed, there seems to be a much more even distribution between the uses in present and past tenses.

In addition, although both *be* and *have* are used in the infinitive form, *have* is used almost twice as frequently as *be* in perfect infinitive forms, and *being gone* is never used in the corpus, whereas just a few examples are present with *having gone* (cf. Table 2). These results align with those of previous studies (Rydén – Brorström 1987; Kytö 1997) which have also shown that *have* tends to be the preferred verb in the infinitive form, as opposed to *be*. Furthermore, they are significant since the p value after the statistical test is applied is 0.0001, that is < 0.05. This means that the choice of auxiliary is not random, but in this case motivated by the verbal form.

Table 2. Distribution of examples according to verbal forms

	Be	*Have*
Present perfect tense	44.59% (136)	31.21% (44)
Past perfect tense	45.90% (140)	48.94% (69)
Infinitive (have/be + gone)	9.51% (29)	17.02% (24)
-ing (having/being + gone)	0%	2.83% (4)

3.4.3 Figure and Ground

3.4.3.1 Figure

All motion events include a *figure* that moves or is located somewhere, as described in 3.1. This *figure* coincides with the syntactic subject of the sentence in all the examples found, which is the most common function of the *figure*. Due to the fact that in English the subject always has to be present in the syntactic structure, the *figure* is also always present in the surface structure of the motion events analysed here.

The idea of motion could imply that either a *figure* moves by itself or is moved by somebody else. However, in all the examples analysed, the *figure* is the one that moves, independently of it being human or non-human, for instance *a letter*. Because of this, it was expected that most of the *figures* would be human, since non-human *figures* do not tend to physically move by themselves, unless they are animals or a metaphorical meaning is implied, as in *a letter has gone in the post*.

As can be seen in Table 3, even though the human *figures* dominate in combination with both auxiliary verbs, when a non-human *figure* is present the use of *have* is slightly preferred to the use of *be*. This difference is not significant, as shown by the p value of 0.1907, which implies that this element of the motion situation does not affect the type of auxiliary. However,

when considered together with the rest of the results, it may contribute to indicating a possible pattern. As shown in 3.4.3.2, the fact that metaphorical *grounds* seem to favour the use of *have* could also explain the higher presence of non-human *figures* with *have*, because, unlike human *figures*, non-human *figures* often move in a metaphorical way, and they can even be metaphorical themselves (e.g. *swearing has gone out of fashion*).

Table 3. Distribution of human and non-human *figures*

	Be	*Have*
Human *figure*	78% (238)	72% (102)
Non-human *figure*	22% (67)	28% (39)

3.4.3.2 Ground

As opposed to the component of the *figure*, the *ground* is not always present in the surface structure of the sentence in English, that is, in the syntactic representation of the motion event (e.g. in *The girl is gone* there is no indication of where the motion event ends). However, it is always present in the deep structure, the semantic content of the sentence (e.g. although in *The girl is gone* there is no 'external' ground, in the deep structure a location can be inferred). Usually, the context contributes to the understanding of what the *ground* is if it has been left out of the surface structure, but at the same time a variety of interpretations may arise when the *ground* is not explicitly present.

In addition, when the *ground* is present, it is represented by many different words or phrases, sometimes with a physical meaning (e.g. a place) and sometimes with a metaphorical one (e.g. a state of being, such as *red* in *he has gone red*, implying embarrassment). Furthermore, there is such a range of subtleties of meanings that a further distribution proved essential.

The categories established were as follows: *no explicit ground*, when there was no *ground* present in the surface structure (e.g. *he is gone*); *physical location* (e.g. *Mr Henry was gone to Amsterdam*); *to action*, which includes a verb, but it implies that there is a physical place where that particular action takes place (e.g. *the old gentleman was gone to dine*); *intransitive preposition*, that is, when the *ground* is represented by a preposition functioning on its own rather than accompanied by a complement, and it incorporates both *path* and *ground* (e.g. *all the reports that have gone abroad*); *intransitive preposition + prepositional phrase*, in which as well as the intransitive preposition, there is a prepositional phrase referring to the *ground* more specifically (e.g. *Mr Hope and the band is gone up to the castle*), and these are examples incorporating a complex *path*,

as it is expressed both in the intransitive preposition and in the preposition introducing the complement *ground*; *intransitive preposition + to action*, which is similar to the previous category but incorporates an action rather than a physical place (e.g. *he had gone out to take a solitary walk*); *noun phrase*, where the complement is not preceded by any preposition, but it refers to some kind of *ground* (e.g. *they have gone a single step*); *preposition + metaphorical location*, which functions as *ground* but it is abstract or metaphorical (e.g. *he has gone through certain mysterious preparations*); the word *far* or any form containing it, such as *farther*, *so far* or *too far* (e.g. *she had gone too far*), and this could have been incorporated into the category of intransitive prepositions, as the term *far* can have different functions, however, since it was only used with *have* it was considered important to keep it as a separate category, in order to see if it could contribute to a better understanding of the choice of auxiliary; *to a state*, in which the *ground* is represented by an abstract concept referring to *a state of being* (e.g. *Zadig should have gone blind*).

Table 4 shows that *have* is used in a wider variety of contexts as regards the type of *ground* present, whereas *be* is much more restricted to certain *grounds*. Not only does the type of *ground* appear to condition the type of auxiliary, but the lack of explicit *ground* in the surface structure also seems to show a preference for *be*. Contrastingly, *be* is not used in combination with metaphorical locations, with the word *far*, and when the *ground* refers to a *state of being* (e.g. *something must have gone wrong*). On the contrary, the most frequent combination with *have* is the category of *preposition + metaphorical location*, followed by the *intransitive preposition* category.

Table 4. Distribution of different types of *grounds* with both auxiliaries

Type of ground	*Be*	*Have*
No explicit ground	77.05% (235)	7.10% (10)
Physical location	9.18% (28)	17.02% (24)
To action	5.57% (17)	7.10% (10)
Intransitive preposition	4.59% (14)	16.31% (23)
Intransitive preposition + prepositional phrase	2.62% (8)	3.54% (5)
Intransitive preposition + to action	0.66% (2)	3.54% (5)
Noun phrase	0.33% (1)	7.10% (10)
Preposition + metaphorical location	0%	26.24% (37)
Far	0%	9.22% (13)
To a state	0%	2.83% (4)

As shown in Table 4, the use of *be* is mainly restricted to those examples in which the *ground* is not explicitly represented in the surface structure, representing 75.96% of the total. The *physical location* is the second most common *ground* in combination with *be*, although the percentage is appreciably low compared to the first category. The rest of the combinations present lower percentages compared to the results with *have*.

As opposed to the results of the *figure* element, which did not prove to be significant, the results of the *ground* component are different. As shown by the p value of 0.0000, the differences between the uses of *be* and *have* in relation to the types of *grounds* that have been identified in the present study are significant, in other words, they are not combined randomly with one auxiliary or the other.

3.4.4 Comparison with some contemporaneous female English writers

In general, Maria Edgeworth seems to have had a preference for *be* as the auxiliary of perfect tenses, as shown in 3.4.1. As can be observed in Table 5, the equivalent usage patterns found in works of other contemporaneous female writers are in agreement with Edgeworth's, in the sense that *be* tends to be the preferred verb in all of them. Previous studies have shown that this also occurs with other intransitive verbs of motion (Calvo Cortés 2019, 2020). However, the case of *gone* is particularly interesting as it prevails in combination with *be* for a longer time. For this reason, and since this verb is the focus of the present study, a summary of how Edgeworth's results compare to those of some of these other writers is provided here.

The figures presented in Table 5 include both the raw data and the normalised data since there is a discrepancy in the size of the five corpora used. The percentages of uses of both auxiliaries provide a clear insight into the differences between them. They also help in observing the similarities and/or differences between the usage of the two auxiliaries and the five writers. The normalised data indicate that, in general, Austen uses perfect tenses more than the other authors, whereas the other four writers display very similar total figures regarding their use of perfect forms.

Table 5 shows that Maria Edgeworth's results are very similar to those found in Jane Austen in relation to the distribution of the two auxiliaries in perfect tenses. Both authors coincided in time although Edgeworth lived longer and, therefore, continued writing, and published her last novel, *Helen* (the last work included in the present corpus), in 1834. Nevertheless, both writers had works published or republished after their deaths and different later editions may have reflected changes in the novels, including

adaptations of grammatical features, such as changing *be* for *have* in perfect tenses. Access to the original manuscripts would provide clarification on this and could confirm if the editors had indeed manipulated the grammar or not. Since these original documents have not been accessed, this remains a speculation.

Table 5. Distribution of auxiliaries in Maria Edgeworth and in other contemporaneous female English writers

	Be	*Have*
Maria Edgeworth	68.39% (305) [1.68][1]	31.61% (141) [0.77]
Jane Austen	71.59% (247) [3.16]	28.41% (98) [1.25]
Fanny Burney	82.43% (122) [2.38]	17.57% (26) [0.51]
Elizabeth Inchbald	78.57% (33) [2.15]	21.43% (9) [0.58]
Ann Radcliffe	80.39% (82) [2.28]	19.61% (20) [0.56]

[1] Normalised data by 10,000 words is presented in square brackets.

The results from the three other writers included in this comparison, Burney, Inchbald and Radcliffe, reveal an even higher preference for *be*. All the chosen novels written by these women were published between 1778 and 1798. Although a period of 20 or 30 years may not have been too long in other contexts, in this particular situation it is a factor to be considered, since this is precisely the period when the use of *have* (as opposed to *be*) starts to dominate the perfective structures. This could explain the slightly higher percentage of *be* in these three authors when compared to Edgeworth and Austen. However, differences are slight and the five writers can be said to have behaved similarly in relation to their choices of auxiliaries.

4. Conclusions

The present paper has attempted to further our understanding of the uses of auxiliary verbs in perfect tenses by Maria Edgeworth. Her general preference for *be* as an auxiliary can be explained due to certain factors. Similarly, her choice of *have* in specific contexts may be connected to her conceptualization of motion situations.

The intervention of editors in the changes of auxiliary verbs cannot be demonstrated by analysing the examples in the printed works alone. However, the discrepancy in the data found in the analysis of the voices in

Maria Edgeworth's works may provide some insights into this issue. The fact that there are many more instances of *be* than *have* in the characters' voices as opposed to a more balanced result in the narrator's voice could indicate manipulation of the texts by the editors, since they would be more likely to maintain the direct speech of the characters, which may reflect more variation and features of the vernacular language. The significance of the results in this respect contributes to the supposition that if the editors did not change the auxiliaries, Maria Edgeworth likely chose *be*-perfects on more occasions in some characters' speech with a particular intention.

When verbal tenses are analysed, *be* is the preferred choice in the present perfect tense, while a more balanced percentage can be observed in the past perfect tense. In the statistical tests, the differences shown by the two auxiliaries have been demonstrated not to be random but possibly motivated in part by the type of verbal forms used.

Although it has been thought that Edgeworth's usage may have been influenced by Irish English, this does not seem to have been true of her choice of auxiliary, as her uses match those of other female writers of the time, who were likely virtually unfamiliar with this variety of English. All these writers prefer the usage of *be* to *have* when combined with the participial form *gone*.

The fact that Maria Edgeworth's results regarding auxiliary choice are closer to Jane Austen's than to the ones of Burney, Inchbald and Radcliffe could be due to the relative proximity of their lifespans. In other words, their writings were published between two and three decades later than the novels chosen for analysis written by the latter three writers. Nevertheless, the similarity of results with Austen's may simply have been due to chance.

One of the most influential factors seems to be the components of motion situations. Although the *figure* does not appear to play an important role in the selection of auxiliary, the non-human *figures* tend to slightly prefer *have* as the auxiliary of perfect tenses, which might indicate that abstract *figures* or *figures* that do not typically move by themselves combine with *have* better. However, no significant results were found for this variable. What seems to condition the type of auxiliary used most is the *ground*. Whenever no explicit *ground* is present in the surface structure, *be* is usually the preferred option, whereas when metaphorical locations are implied in the *ground*, *have* is the chosen verb. Previous studies (e.g. Kytö 1997) had associated *have* with the presence of a complement, which can be understood similarly to the presence of the *ground* in the surface structure. However, this study suggests that not only do complements influence the choice, but more specifically that the presence of metaphorical or abstract semantic content

is what determines which auxiliary is used in many situations, since both metaphorical *figures* and *grounds* tend to prefer *have* in perfect structures, even if only slightly in the case of the *figures*.

Similarly, it remains unknown to what extent Edgeworth might have been influenced by her reading of English literature, and despite wanting to reflect the Irish English variety in her texts, other influences may also have been present. Her use of either auxiliary may have been motivated by some of the factors suggested here, but she seems to have been a fairly conservative author in this respect.

Further studies are required to understand the uses of both auxiliaries. They include perfect structures with other verbs in Maria Edgeworth, other Irish writers (both female and male), and English male writers of the same time. Additionally, in order to rule out the influence and manipulation of editors, other types of texts need to be examined, for instance private correspondence, and when possible in the original manuscripts.

REFERENCES

Sources

Project Gutenberg

https://www.gutenberg.org/ebooks/8826, *Tales and Novels* by M. Edgeworth. Vol. 1: *Moral Tales*, accessed December 2020

https://www.gutenberg.org/ebooks/8720, *Tales and Novels* by M. Edgeworth. Vol. 2: *Popular Tales*, accessed December 2020

https://www.gutenberg.org/ebooks/9455, *Tales and Novels* by M. Edgeworth. Vol. 3: *Belinda*, accessed December 2020

https://www.gutenberg.org/ebooks/9439, *Tales and Novels* by M. Edgeworth. Vol. 4: *Castle Rackrent; An Essay on Irish Bulls; An Essay on the Noble Science of Self-Justification*; *Ennui*; and *The Dun*, accessed December 2020

https://www.gutenberg.org/ebooks/9414, *Tales and Novels* by M. Edgeworth. Vol. 5: *Tales of a Fashionable Life*, accessed December 2020

https://www.gutenberg.org/ebooks/9620, *Tales and Novels* by M. Edgeworth. Vol. 6: *The Absentee; Madame de Fleury; Emilie de Coulanges;* and *The Modern Griselda*, accessed December 2020

https://www.gutenberg.org/ebooks/8937, *Tales and Novels* by M. Edgeworth. Vol. 7: *Patronage*, part 1, accessed December 2020

https://www.gutenberg.org/ebooks/9321, *Tales and Novels* by M. Edgeworth. Vol. 8: *Patronage*, concluded; *Comic Dramas*; *Leonora*; and *Letters*, accessed December 2020

https://www.gutenberg.org/ebooks/9107, *Tales and Novels* by M. Edgeworth. Vol. 9: *Harrington*; *Thoughts on Bores*; and *Ormond*, accessed December 2020

https://www.gutenberg.org/ebooks/8531, *Tales and Novels* by M. Edgeworth. Vol. 10: *Helen*, accessed December 2020

Special Studies

Anderwald, L.

2014 "The decline of the be-perfect, linguistic relativity, and grammar writing in the nineteenth century". In: M. Hundt (ed.) *Late Modern English Syntax*. Cambridge: Cambridge University Press, 13-37.

Beal, J.C.

2004 *English in Modern Times*. London: Arnold.

Calvo Cortés, N.

2014 "A corpus-based study of gradual meaning change in Late Modern English". In: A. Alcaraz Sintes – S. Valera Hernández (eds.) *Diachrony and Synchrony in English Corpus Linguistics*. Bern: Peter Lang, 25-54.

2019 "Conservatism or the influence of the semantics of motion situation in the choice of perfect auxiliaries in Jane Austen's letters and novels". In: C. Claridge – B. Bös (eds.) *Developments in English Historical Morpho-Syntax*. Amsterdam: John Benjamins Publishing, 175-198.

2020 "Variations from letter-writing manuals: Humble petitions signed by women in Late Modern London". In: A.H. Jucker – I. Taavitsainen (eds.) *Manners, Norms and Transgressions in the History of English. Literary and Linguistic Approaches*. Amsterdam: John Benjamins, 184-212.

Filppula, M.

1999 *The Grammar of Irish English. Language in Hibernian Style.* London: Routledge.

Harris, J.

1993 "The grammar of Irish English". In: J. Milroy – L. Milroy (eds.) *Real English: The Grammar of English Dialects in the British Isles*. London: Longman, 139-186.

Hickey, R.

2000 "Models for describing aspect in Irish English". In: H. Tristram (ed.) *The Celtic Englishes II*. Heidelberg: Carl Winter, 97-116.

2007 *Irish English. Its History and Present-day Forms*. Cambridge: University Press.

Hollingworth, B.

1997 *Maria Edgeworth's Irish Writing. Language, History and Politics.* London: Macmillan Press Ltd.

Kytö, M.
1997 "Be/have + past participle: The choice of the auxiliary with intransitives from Late Middle to Modern English". In: M. Rissanen – M. Kytö – K. Heikkonen (eds.) *English in Transition: Corpus-based Studies in Linguistic Variation and Genre Styles*. Berlin: Mouton de Gruyter, 17-84.

Manly, S.
2007 *Language, Custom and Nation in the 1790s. Locke, Tooke, Wordsworth, Edgeworth*. Hampshire: Burlinton Ashgate.

Matsumoto, Y.
1996 "Subjective motion and English and Japanese verbs", *Cognitive Linguistics* 7, 138-226.

McCafferty, K.
2014 "[W]ell, are you not got over thinking about going to Ireland yet: The Be- perfect in eighteenth and nineteenth century Irish English". In: M. Hundt (ed.) *Late Modern English Syntax*. Cambridge: Cambridge University Press, 13-37.

McFadden, T.
2017 "On the disappearance of the Be perfect in Late Modern English", *Acta Linguistica Hafniensia* 49 (2), 159-175.

Ó Gallchóir, C.
2005 *Maria Edgeworth: Women, Enlightenment and Nation*. Dublin: University College Dublin.

Ramscar, M. – T. Matlock – L. Boroditsky
2009 "Time, motion and meaning: the experiential basis of abstract thought". In: K.S. Mix – L.B. Smith – M. Gasser (eds.) *The Spatial Foundations of Language and Cognition*. Oxford: Oxford University Press, 67-82.

Ronan, P.
2012 "The development of periphrastic tense and aspect periphrasis in Irish and Welsh". In: L. Brems et al. (eds.) *Grammaticalization and Language Change: Origins, Criteria and Outcomes*. Amsterdam: John Benjamins Publishing, 226-246.

Rydén, M. – S. Brorstöm
1987 *The Be/Have Variation with Intransitives in English*. Stockholm: Almqvist & Wiksell International.

Siemund, P.
2004 "Substrate, superstrate and universals. Perfect constructions in Irish English". In: B. Kortmann (ed.) *Dialectology Meets Typology*. Berlin: Mouton de Gruyter, 401-434.

Slobin, D.I.
1996 "Two ways to travel: Verbs of motion in English and Spanish". In: M. Shibatani – S.A. Thompson (eds.) *Grammatical Constructions*. Oxford: Clarendon Press, 195-219.

Sorace, A.
2000 "Gradients in auxiliary selection with intransitive verbs", *Language* 76, 859-890.
Talmy, L.
1975 "Semantics and syntax of motion". In: J.P. Kimball (ed.) *Syntax and Semantics.* Vol. 4. New York: Academic Press, 181-238.
1987 "Lexicalization patterns: Semantic structure in lexical forms". In: T. Schopen (ed.) *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, 57-149.
2000 *Towards a Cognitive Semantics.* 2 vols. Cambridge, MA: MIT Press.
Werner, S – E. Seoane – C. Suárez-Gómez (eds.)
2016 *Re-assessing the Present Perfect*. Berlin: Mouton de Gruyter.

Address: Nuria Calvo Cortés, Departamento de Estudios Ingleses: lingüística y literatura, Facultad de Filología, edificio A, Plaza Menendez Pelayo s/n, 28040 Madrid, España.
ORCID code: orcid.org/0000-0003-1405-8943.

Terms of reference and discursive representations: A case study with Saddam Hussein in the 1991 Gulf War

Sonja Kuosmanen
University of Helsinki

ABSTRACT

This study investigates connections between terms of reference and discursive frames using U.S. newspaper reports on the 1991 Gulf War and Iraqi President Saddam Hussein as a case study. Combining results from a quantitative and qualitative analysis of different terms of reference, three discursive frames were identified: The Statesman, The Madman and Our Saddam. The most common discursive frame was of a foreign head of state. Newspapers also included representations that both demonized Saddam Hussein and discussed him in the context of Middle Eastern in-groups. Previous research has highlighted the demonization of Saddam Hussein in news reports during the war, but the findings of this study suggest that an overtly demonizing discursive frame was a minority view. However, its use nonetheless showed lasting impact beyond the end of the military operation. Press reports thus showed more varied and ambivalent representations than previous analyses may have suggested.

Keywords: term of reference, naming, corpus, Persian Gulf War, Saddam Hussein.

1. Introduction

For many decades, Iraq and its then-leader President Saddam Hussein represented a controversial topic in U.S. foreign policy discussions. The United States, leading a coalition of other nations, fought wars against Iraq both in the 1990s and the 2000s. Saddam Hussein was a central figure in each of these conflicts. This study focuses on the Persian Gulf War as a case study to investigate the variety and prevalence of discursive representations for an enemy 'Other'. While now several decades past, this conflict continues

DOI: 10.25951/4365.

to serve as a case study for the examination of media manipulation and propaganda. Researchers have highlighted the media's role in amplifying negative representations of Saddam Hussein and Iraq in both the Persian Gulf and Iraq Wars (presented in sections 2 and 3). Using a sizable corpus of news reports from during and after the U.S.-led military operation against Iraq in 1991, this study investigates different terms of reference and discursive representations for Saddam Hussein as well as their prevalence and continuation after the war itself. Did press reports in fact offer a full-throated support for the demonization of Saddam Hussein during the war? Did these demonizing or other representations remain in use after the military operation?

Specifically, this study investigates contemporary U.S. newspaper reports for diachronic changes in the use of specific types of terms of reference for Saddam Hussein as well as whether – and the extent to which – these terms of reference can be associated with specific discursive representations. To accomplish this, the analysis uses a corpus of U.S. newspaper reports from January to July 1991. Corpus analytical tools are utilized to investigate the use of different terms of reference and changes over time, while a systemic-functional study of a smaller sample of examples expands the analysis into categories that implicate specific kinds of processes and participants. Findings from the analysis are drawn together in discursive frames that show different representations for Saddam Hussein, associated with certain terms of reference and systemic-functional categories.

The analysis identifies distinct discursive frames used for the Iraqi president in newspaper reports. These discursive frames are titled The Statesman, The Madman, and Our Saddam, depicting Saddam Hussein as, respectively, the legitimate though not always benevolent leader of a nation, a dangerous and volatile individual, and the man represented in the words and views of Iraqi and other regional voices. Of the three discursive frames, The Statesman is the most common. The analyzed newspapers showed clear differences in their use of various terms of reference and the identified discursive frames could also be associated with specific terms of reference. The term "Saddam" in particular was associated with two somewhat contradictory frames, Our Saddam and The Madman.

The analysis ultimately presents an ambivalent and conflicting picture of press reporting. The comparatively lower frequency of The Madman discursive frame suggests that the demonization of Saddam Hussein in the media, which has been highlighted in previous research, was not a dominant practice in contemporary press reports. However, while used at a lower rate of frequency, The Madman discursive frame was nonetheless maintained

and even somewhat strengthened during the aftermath of the war. This suggests that this frame had a lasting effect in the way Saddam Hussein was represented in U.S. media, which may well have also carried over into the later Iraq War.

2. Historical context

In August 1990, following a period of escalating tensions over oil drilling rights, Iraq invaded the small neighboring state of Kuwait. Following the invasion, throughout the fall of 1990, international leaders attempted to mediate the situation with Iraq while the UN Security Council passed several resolutions condemning the invasion. In the United States, both the administration of President H.W. Bush and wealthy Kuwaiti interest groups undertook a campaign to sway American public opinion in favor of a U.S. military operation against Iraq (Kellner 2004: 137-144). The Bush administration sought to cast Saddam Hussein as the villain in a struggle between the forces of good and evil (Peer – Chestnut 1995: 89-91).

The UN Security Council ultimately set the deadline of January 15, 1990, for Iraq's withdrawal from Kuwait. When that deadline passed, an international coalition led by the United States launched Operation Desert Storm to drive Iraq out of Kuwait. The five-week operation resulted in the removal of Iraqi forces from Kuwait but did not oust Saddam Hussein as Iraq's leader (Mazarr et al. 1993.) In addition to a military confrontation, however, the operation was a carefully planned media spectacle: "[w]hen the US began military action against Iraq on January 16, 1991," Kellner (2004: 144) writes, "the mainstream media became a conduit for Bush administration and Pentagon policies", framing the war as an exciting adventure and shunning dissenting opinions.

During Operation Desert Storm, this framing was aided by strict controls imposed on the press coverage of the Gulf War (Nohrstedt 1992: 119-120), including limited access to the conflict area, placements in so-called pools that offered access to the war front and troops, and censorship of reports. Journalists who did not follow these guidelines risked being deported from the conflict area. A comparison of television broadcasts and newspaper reporting prior to the U.S. military operation found that newspaper reports on the conflict were comparatively more critical (Peer – Chestnut 1995). The coverage of the military operation itself, to audiences both in the United States (Kellner 2004) and around the world (Mowlana et al. 1992), was dominated by Western and especially U.S. media. The media

coverage showed a U.S. high-tech war machine and sympathetic soldiers (Kellner 2004: 147-148) while demonizing Iraq and its leader through images, headlines, and even editorial cartoons (Artz – Pollock 1995). In an examination of newspaper editorials from the first weeks of the military operation, Hackett and Zhao (1994) found that anti-war voices were also represented to some extent, though editorials nonetheless emphasized viewpoints favorable to the Bush administration.

In mid-February 1990, while the U.S. military operation was still underway, U.S. President George Bush urged the Iraqi people to overthrow their leader. However, when Kurdish and Shi'ite insurrections in Northern and Southern Iraq, respectively, rose in response, the United States failed to provide support and the insurrections were crushed by Saddam Hussein, who was able to reassert his power over Iraq after the war (Kellner 2004: 150; Atkinson 1993: 488-489). Thus, while achieving his immediate objectives for the war, President Bush's framing of the conflict as a "moral crusade" ultimately left the struggle unresolved and the villain in power (Atkinson 1993: 497).

It has now been several decades since the Persian Gulf War, and the United States has since engaged in other more recent wars, including another war against Iraq in 2003. However, the Persian Gulf War was chosen here as a case study for several reasons. Firstly, the conflict has been extensively investigated, so it offers a baseline for further investigation and indeed continues to serve as a relevant example (for another recent study, see, for example, Oddo 2018). Secondly, unlike in the later Iraq War, the United States did not have a pre-established hostile relationship with Saddam Hussein prior to the Persian Gulf conflict; the two were rather uneasy allies in their opposition to Iran (Oddo 2018: 41-42). Examinations of the Iraq War have revealed the presence of pre-established frames of a rogue country and its leader (see, for example, Abid – Manan 2016; Oddo 2011), at least some of which can be assumed to have originated from the Persian Gulf War. Thirdly and finally, this case study can broaden our understanding of journalistic practices and press reactions in times of conflict and war.

3. Theoretical and methodological frameworks

3.1 Theoretical framework

The focus of this study is the language of the media, and specifically newspaper articles. For written or printed media discourse, its interactional nature is less obvious than in face-to-face communication. In written discourse, "shared

meanings, knowledge of the language, knowledge of the world, and other beliefs must be taken into account in such a characterization of discourse meaning" (Van Dijk 1988: 9). While the traditional model of communication involves a speaker or sender, a message, and the hearer or receiver, an actual newspaper article is handled by multiple individuals and undergoes many edits before finally being put into print (Bell 1991: 34-35).

Characteristics of news reports and quality press include long, complex sentences, many nominalizations ('disruption' instead of 'they disrupted'), formal jargon borrowed from sources such as officials and policymakers, and syntactic structures that are rare in other discourse forms (such as the inverted declarative sentence structure: 'something happened, someone declared') (Van Dijk 1988: 10-11). Due to journalists' reliance on both spoken and written second-hand sources, the news story can be viewed as a layered whole of embedded texts within texts (Bell 1991: 50-51). Sourcing and constructions of news items is often closely linked to the actions and opinions of powerful social groups, as items are also selected and composed based on a conception of the intended target audience (Richardson 2007: 1).

This study examines terms of reference – essentially practices of naming – for Saddam Hussein in newspaper reports. In previous studies, naming has been examined in the context of public discourses such as news as well as elsewhere, including in the context of sexist discourses (see, for example, Page 2003; Mills 2003). Naming is a powerful ideological choice (Clark 1992: 209). Expressions used to refer to a participant are intertwined with social values, but the connection between the chosen expression and the intended meaning is also context-dependant (Fowler 1991: 99). Clayman (2010) examined the strategic use of address terms in news broadcast interviews and found that address terms were used as a strategic tool to signal actions such as soliciting attention or to cast the interview in certain ways. Page (2003) showed that the patterns of naming choices for the same individual can be linked to different and even contradictory ways of discursive representation and that these representations can be explored through the naming practices used in newspaper reports.

Journalism is a tool and channel for societal influence (Richardson 2007: 180). The media can have profound impact on how certain actors and events are represented through the compounding effects of repetition of images and concepts (Gerbner et al. 1986). To investigate these compounding representations, extensive corpora of media texts have been used to examine the representation of actors and events in newspaper reports. Baker et al. (2013), for example, studied British newspaper reports to investigate representations of Muslims and Islam and found that while newspapers

did to some degree contribute to negative stereotypes, they offered varied and often ambivalent representations. Branum and Charteris-Black (2015) compared different representations of the Edward Snowden affair in the British press and found clear differences in both ideological position and style among newspapers, including in naming strategies.

Because of its role in shaping societal discourses, in times of societal tension such as war, journalism is both vulnerable to and the target of external influence (Richardson 2007: 185-186). The Gulf War is one example of a conflict where media coverage of a war not only echoed an administration's (Bush's) positions but also demonized an enemy in both domestic (U.S.) press reports (Artz – Pollock 1995) and elsewhere (see, for example, Martín Rojo 1995). In the run-up to the war, changes in press reporting extended also to terms of reference in news articles, which changed to highlight the prestige of President Bush and the delegitimacy of Saddam Hussein (Kuosmanen 2019). Both in text and image, Iraqi interests were attributed to the individual ambitions of Saddam Hussein, who in turn was portrayed as a dangerous, uncivilized, and irrational individual (Artz – Pollock 1995). Similar patterns were later found in the Iraq War (see, for example, Lule 2004; Steuter – Wills 2010; Popp – Mendelson 2010). The perspectives of Iraqi civilians have often been neglected (for one perspective on Iraqi women in particular, see Al-Ali 2011).

In both wars, the president and the White House made great efforts to demonize the Iraqi president in their appeals to the media and the public (Oddo 2011; Abid – Manan 2016; Oddo 2018). Hart and Fuoli (2020), however, also found that in order to be effective, political leaders have to supplement their pro-war appeals with cited evidence or design their message to address existing favorable attitudes among the public. Thus, as one example of such supportive messaging, Oddo (2018) traced the fabrication of a story of atrocities by Iraqi soldiers that was shared extensively in the media with the aid of the White House.

3.2 Methodological framework

To investigate newspaper reports, this study uses two analytical tools: corpus methodology and systemic-functional analysis. This section will present these two tools and discuss some analytical challenges arising from applying these tools to newspaper reports.

Discourse analysis has been increasingly combined with corpus methodology in a variety of different approaches. Baker (2006) and Mautner

(2009) among others have discussed the benefits and limitations of using corpus linguistics as an analytical tool in discourse analysis. Branum and Charteris-Black (2015) and Baker et al. (2013) are among those who have used a combination of discourse analysis and corpus methodology to study the representation of actors and events in newspaper reporting, conducting comparative as well as diachronic analysis.

Systemic functional linguistics (SFL) provides an additional layer of both qualitative and quantitative analysis. Based on the work of Michael Halliday (see, for example, Halliday – Matthiessen 2004), the systemic-functional analysis used in this study focuses on experiential meaning. Experiential meaning is explored through a system called Transitivity that centers on a process type, participants associated with that process, and additional circumstantial information. There are multiple process types, each with associated participant categories: material, mental, behavioral, and existential processes (for one overview of Transitivity and its grammatical categories, see Eggins 2004: 206-253). In addition, the material process, for example, identifies participants that are implicated by the process. In the subsequent analysis, terms referring to categories of systemic functional grammar will be capitalized to denote the use of these terms in their systemic-functional meaning.

Transitivity has been presented as an analytical tool for discourse analysis by Fowler (1991: 70-76), for example, and used by, among others, Clark (1992) and Page (2003) in their studies of sexist naming practices in British press reports. Its value is in its "facility to analyse the same event in different ways" and to link discursive choices to ideological significance (Fowler 1991: 71). The specific methods of analysis and materials used in this study are described in more detail in the following section.

4. Material

The material used in the study consists of a corpus of U.S. newspaper articles on the Gulf War conflict compiled from January-July 1991. The corpus has been divided into sub-corpora, one for each month of articles included. Table 1 shows the detailed composition of the corpus, including the number of articles, word types (unique words) and word tokens (number of overall words) for the whole corpus and each sub-corpus. As Table 1 shows, *The New York Times* and *The Washington Post* together represent over 80% of the total material in the corpus, while *The Wall Street Journal* makes up a smaller

portion of the overall corpus. Numbers for word types and word tokens were obtained using AntConc (Anthony 2011). Table 1 shows the composition of the corpus in total and for each sub-corpus per number of newspaper articles included, word types (unique word tokens) and total word tokens (total number of words). Table 2 shows the overall composition of the corpus divided by newspaper.

Table 1. Composition of the corpus per sub-corpus

Sub-corpus	Articles	Word types	Word tokens
Jan-91	1,085	34,278	1,083,690
Feb-91	734	25,325	690,110
Mar-91	515	25,021	497,440
Apr-91	350	19,218	339,620
May-91	202	15,560	201,446
Jun-91	119	11,114	102,081
Jul-91	114	10,299	106,882
Corpus total	**3,119**	**54,803**	**3,021,269**

Table 2. Composition of the corpus per newspaper

Newspaper	Articles	Word types	Word tokens
The New York Times (*NYT*)	1,385	33,614	1,256,689
The Washington Post (*WP*)	1,193	36,786	1,219,413
The Wall Street Journal (*WSJ*)	541	25,644	545,167

The articles included in the corpus were retrieved from the Proquest Database of Historical Newspapers using the search words "Iraq" and "Kuwait" for articles published between January 1 and July 31, 1991. Search results were limited to articles, and other newspaper genres such as editorials and advertisements were excluded. The articles, which are stored in the database as scanned pdf files of newspaper clippings, were processed using an optical recognition program and saved as text files. All articles retrieved by the search that referenced the ongoing conflict were included in the corpus. The number of articles and words is thus also representative of the volume of reporting on the Gulf War in these newspapers between January and July 1991.

As mentioned above, the study involves a quantitative analysis of concordances and frequencies for specific terms of reference for Saddam

Hussein. For the systemic functional analysis, a sample of 20 random concordances from each sub-corpus was selected, resulting in a total of 140 concordances for SFL analysis. For each concordance, the entire sentence within which the searched term of reference was located was retrieved for scrutiny. Concordances were distributed among different categories of terms of reference as well as newspapers based on the frequencies to ensure a balanced representation for each sub-corpus, newspaper, and term of reference. Table 3 shows the overall number of concordances per term of reference and newspaper; detailed frequencies per term of reference and newspaper based on which this distribution was done are presented later in Table 4 in the Analysis section.

Table 3. Distribution of concordances chosen for systemic-functional analysis

Concordances per term of reference	Honorific + Hussein	Saddam Hussein	Saddam	Concordances Total
	28	58	54	140
Concordance per newspaper	NYT	WP	WSJ	
	60	60	20	

The specific concordances were selected from each sub-corpus using a random number generator, based on which the corresponding number was selected from the list of AntConc concordance search results. Thus, while concordances for each newspaper and term of reference were allocated based on frequencies in the corpus, the specific concordances themselves were randomized. A systemic-functional analysis of these concordances was completed using Transitivity.

The news texts included in the analysis are both syntactically complex and, due to the types of topics covered here, full of figurative expressions describing societal and political events that are both abstract and complicated. Thus, analyzing the samples chosen for the systemic functional analysis poses several challenges. Firstly, the types of processes described in press reports can be difficult to categorize according to the various Process Types. Example 1 demonstrates that the categorization of Processes can sometimes offer multiple options and interpretations:

(1) [President Saddam Hussein's army]$^{\text{Actor/Sayer}}$ [threatened to attack] $^{\text{Process:Material/Verbal}}$ [with chemical and biological weapons]$^{\text{Circumstance}}$… (*NYT*, 17 July 1991)

The Process can be categorized as either a Material (acting in a threatening manner) or Verbal (issuing a threatening statement) Process. Neither the immediate context of the clause nor the rest of the paragraph provide additional clues, so either interpretation can be considered equally valid.

Secondly, the same elements and clauses can be implicated in multiple layers of Processes and Participant roles. As demonstrated in example 2, the same actor ("President Hussein") can be implicated in multiple layers of Processes; as the Participant (Actor) in a Material Process that is at the same time a Participant (Phenomenon) for a Mental Process that in turn is a Participant (Verbiage) for a Verbal Process.

(2) [Mr. Vorontsov, the Soviet representative]Sayer, [said]Process:Verbal [[Moscow]Senser [had "reason to believe"]Process:Mental [[President Hussein]Actor [was ready to withdraw]Process:Material [unconditionally] Circumstance [in a very short time frame.]Circumstance]Phenomenon]Verbiage (*NYT*, 26 February 1991)

5. Analysis

This section presents the results of the analysis and proceeds in two steps. First, as a starting point, a corpus analysis presents the concordances and frequencies for different terms of reference for Saddam Hussein. Second, a qualitative analysis using the systemic-functional framework of Transitivity is conducted on 140 sample concordances.

5.1 Corpus analysis

The first step in the analysis was to complete a concordance search for three identified types of term of reference for Saddam Hussein. These three types were: an honorific (*President* or *Mr.* combined as *Honorific + Hussein*) accompanied by the surname *Hussein*; a reference with a first name and last name (*Saddam Hussein*); and a reference with the first name only (*Saddam*). Some concordances for *Saddam Hussein* also have a pre-or post-modifying "President [Saddam Hussein] of Iraq" or "President of Iraq [Saddam Hussein]". These were included in the *Saddam Hussein* category as such instances typically served an identifying or introductory function when Hussein was first mentioned in the articles. While this is not an exhaustive

list of terms used to refer to Saddam Hussein, these were identified as most frequently used based on a sample reading of the newspaper articles[1].

The concordance for each term of reference as well as its normalized frequencies are included in Table 4. Additionally, percentages are provided for each newspaper to show how these three terms of reference are distributed within concordances for that specific newspaper.

Table 4. Frequencies and newspaper distribution for terms of reference for Saddam Hussein

Sub-corpus	Honorific + Hussein	Saddam Hussein	Saddam	Freq*
1	2	3	4	5
Jan-91	**2.87 (N=311)**	**11.99 (N=1,299)**	**12.56 (N=1,361)**	**Sub-corpus**
	30.13% (N=307)	58.68% (N=598)	11.19% (N=114)	NYT
	0.06% (N=1)	30.43% (N=471)	69.51% (N=1,076)	WP
	0.74% (N=3)	56.93% (N=230)	42.33% (N=171)	WSJ
Feb-91	**3.65 (N=252)**	**13.81 (N=953)**	**12.59 (N=869)**	**Sub-corpus**
	31.81% (N=251)	56.27% (N=444)	11.91% (N=94)	NYT
	0.10% (N=1)	30.95% (N=303)	68.95% (N=675)	WP
	0.00% (N=0)	67.32% (N=206)	32.68% (N=100)	WSJ
Mar-91	**3.22 (N=160)**	**9.67 (N=481)**	**10.61 (N=528)**	**Sub-corpus**
	37.21% (N=160)	49.30% N=212)	13.49% (N=58)	NYT
	0.00% (N=0)	28.89% (N=154)	71.11% (N=379)	WP
	0.00% (N=0)	55.83% (N=115)	44.17% (N=91)	WSJ
Apr-91	**5.54 (N=188)**	**13.34 (N=453)**	**10.19 (N=346)**	**Sub-corpus**
	38.09% (N=187)	50.10% (N=246)	11.81% (N=58)	NYT
	0.24% (N=1)	34.15% (N=140)	65.61% (N=269)	WP
	0.00% (N=0)	77.91% (N=67)	22.09% (N=19)	WSJ
May-91	**2.13 (N=43)**	**9.03 (N=182)**	**10.18 (N=205)**	**Sub-corpus**
	29.25% (N=43)	62.59% (N=92)	8.16% (N=12)	NYT
	0.00% (N=0)	29.85% (N=80)	70.15% (N=188)	WP
	0.00% (N=0)	66.67% (N=10)	33.33% (N=5)	WSJ

[1] One example of a term of reference not included was the use of last name only ("Hussein"). This term of reference was excluded because of its comparative rarity (0,74 per 10,000 words, N = 225), and because of the challenge in identifying instances referring specifically to Saddam Hussein rather than to other similarly named individuals, such as King Hussein of Jordan.

1	2	3	4	5
Jun-91	**0.78 (N=8)**	**6.96 (N=71)**	**6.17 (N=63)**	**Sub-corpus**
	26.67% (N=8)	60.00% (N=18)	13.33% (N=4)	NYT
	0.00% (N=0)	41.00% (N=42)	58.00% (N=58)	WP
	0.00% (N=0)	91.67% (N=11)	8.33% (N=1)	WSJ
Jul-91	**0.94 (N=10)**	**8.33 (N=89)**	**14.97 (N=160)**	**Sub-corpus**
	13.16% (N=10)	46.05% (N=35)	40.79% (N=31)	NYT
	0.00% (N=0)	24.71% (N=42)	75.29% (N=128)	WP
	0.00% (N=0)	92.31% (N=12)	7.69% (N=1)	WSJ
Total Corpus	**3.22 (N=972)**	**11.68 (N=3,528)**	**11.69 (N=3,532)**	**Corpus**
	32.39% (N=659)	55.16% (N=1,645)	12.44% (N=371)	NYT
	0.07% (N=3)	30.74% (N=1,232)	69.19% (N=2,773)	WP
	0.29% (N=3)	62.48% (N=651)	37.24% (N=388)	WSJ

* Frequencies reported per 10,000 words (raw frequencies in parentheses). Percentages for each newspaper represent the distribution of terms of reference within concordances for that newspaper.

As Table 4 shows, *Saddam* and *Saddam Hussein* are the two most common terms of reference while the honorific term of reference is used for the most part only by the *NYT*. *Saddam Hussein* is most common in January 1991, during the run-up to and the commencement of Operation Desert Storm, and in April 1991, when terms for Iraq's future were being negotiated in the aftermath of the military operation and when the honorific term of reference was also most frequent. Combined, these normalized frequencies suggest that the focus on Saddam Hussein was particularly high in January, April, and June; waning in February and May; and at its lowest in June. There is a comparative rise in normalized frequencies for the July sub-corpus after a downward trend. A reading of the sample concordances selected for systemic-functional analysis suggests that in July, the focus of the newspaper reporting had moved partly from reporting current developments to re-examination and re-evaluation of the events of the past 12 months. This may have also prompted a renewed focus on Saddam Hussein and his actions.

It is noteworthy that a foreign head of state is referred to by his first name only consistently and frequently in all three newspapers. The *WP* uses *Saddam* most frequently, while the *NYT* and the *WSJ* generally show a stronger preference for *Saddam Hussein*, and *Honorific + Hussein* is almost exclusively used by the *NYT*. However, the *NYT* also abruptly increases its use of *Saddam* in July 1991. Different newspapers have different guidelines on how to use terms of reference: the *NYT*, for example, explicitly advises

the use of "courtesy titles", with a main title for government officials when first introduced and a Mr., Mrs., or Miss for the following references (Siegal – Connolly 2015: 79). Thus, some differences in the types of references used can be explained by editorial practices.

Occasionally, different types of terms of reference are mixed even within the same clause, as shown in example (3). This example uses the first name only reference *Saddam* for the Iraqi President but an honorific for his U.S. counterpart. This type of reference shows an evaluation of prestige for these two national leaders through an implicit juxtaposition. In other cases, some concordances contain two types of terms of reference for Saddam Hussein within the same clause, as in example (4).

(3) Mr. Bush will struggle to keep the two subjects separate by stressing the extent and gravity of Saddam's cheating. (*NYT*, 31 July 1991)

(4) Like Saddam Hussein it seems to believe that an open Israel-Iraq confrontation will wreck the coalition; that by involving the hated Israelis, Saddam will succeed in separating the Arab partners from their American and European allies… (*WSJ*, 21 January 1991)

Step two of the analysis examines three discursive frames, discussed in connection with different terms of reference and the systemic-functional analysis of the sample concordances.

5.2 Discursive frames

While the three types of terms of reference to some extent imply different levels of prestige and evaluation, a qualitative analysis of the ways in which they are used was also needed. For this purpose, the 140 concordances collected for systemic-functional analysis were also used to investigate the contexts in which these terms appear. The main results of this phase of the analysis are the three discursive frames: The Statesman, The Madman, and Our Saddam. This section focuses on discussing these three frames and their relationship to specific terms of reference and systemic-functional categories.

As Page (2003: 563) notes, the choice of names in news reports is not a simplistic measure but rather a starting point for closer examination. The selected concordances were categorized in terms of these three frames in two separate ways: first, with the term of reference for Saddam Hussein visible in each concordance; and second, a blind categorization of the

concordances with the terms of reference hidden from view (replaced by the word 'referent'). This two-stage analysis was made to ensure that the specific term of reference included in the concordance did not overdetermine the choice of discursive frame – assigning every instance of "Saddam" to The Madman frame, for example. In an otherwise similar sentence, the choice of a U.S. newspaper to use *Saddam* rather than *Saddam Hussein* or *President Hussein* can potentially already be a signal, particularly when this choice departs from the newspaper's editorial practices on naming.

Two of the discursive frames, The Statesman and The Madman, can be viewed as the opposing ends of the same spectrum and many concordances could be placed in a grey area between the two extremes. As the leader of the Republic of Iraq, Saddam Hussein has at his disposal the various powers of the state. What separates the discursive frames of The Statesman from The Madman is whether the Iraqi president's use of this power is represented as legitimate. Additionally, there is a subset of examples that represents Saddam Hussein through the words of Iraqi, Kuwaiti, and other Middle Eastern voices – these examples are discussed under the discursive frame of Our Saddam. As an indication of the frequency of these frames, approximately half of the analyzed concordances were categorized as belonging to The Statesman discursive frame, one quarter to The Madman frame, and one quarter to the Our Saddam frame. However, some concordances also overlapped into two categories.

5.2.1 The Statesman

The discursive frame of The Statesman contains instances where Saddam Hussein is represented as the leader of Iraq who, while not always engaged in positive or constructive actions, is afforded the legitimacy of head of state and military leader. However, as previously mentioned, the frames of The Statesman and The Madman represent a spectrum along which many concordances can be placed. The concordances included under The Statesman frame are most clearly associated with the *Honorific + Hussein* and *Saddam Hussein* terms of reference. However, in an analysis of the concordances in which the term of reference was masked as 'referent', the term of reference "Saddam" was also slightly more often included in this discursive frame.

Material Processes are found commonly across the discursive frames. However, what does vary are the types of actions represented. For this discursive frame, there are several examples where Saddam Hussein is portrayed as either attempting a diplomatic resolution or navigating the

boundaries of his situation, as in examples (5) and (6). In other cases, he is engaged in actions that may not be similarly constructive but are nonetheless the actions of a head of state, such as breaking diplomatic relations or refusing to comply with a U.N. resolution. Within this discursive frame, the Iraqi president also appears most often as a Participant in the Actor role, as the one engaged in and initiating actions, in comparison to any other discursive frame.

(5) [Thus [in settling the Kuwait issue]Circumstance]Verbiage, [the sources]Sayer [say]Process:Verbal, [[Mr. Hussein]Actor [is seeking to make]Process:Material [the best deal he can…]Goal]Verbiage (*NYT*, 08 January 1991)

(6) [Mr. Vorontsov, the Soviet representative]Sayer, [said]Process:Verbal [[Moscow]Senser [had "reason to believe"]Process:Mental [[President Hussein]Actor [was ready to withdraw]Process:Material [unconditionally in a very short time frame]Circumstance.]Phenomenon]Verbiage (*NYT*, 26 February 1991)

In connection with Verbal Processes, examples for this discursive frame are most likely to include the Iraqi president in a Participant role as the Sayer – the Participant saying or communicating. The Sayer role is most common in the concordances for the honorific term of reference, as in examples (7) and (8), though other terms of reference are also found in Sayer roles, as demonstrated by example (9). In some cases, as in example (10), Saddam Hussein is in a Sayer role while also being embedded within the Verbiage of another Verbal Process, meaning that his reported words are further reported by another source.

(7) [In their 90-minute discussion]Circumstance [today]Circumstance, [Mr. Arafat and Mr. Hussein]Sayer [reportedly]Circumstance [touched on]Process:Verbal [the scheduled Friday meeting of European Community foreign ministers on the Gulf crisis]Verbiage… (*NYT*, 03 January 1991)

(8) [In a televised pep talk he gave to a group of leading Mosul citizens on Saturday]Circumstance, [President Hussein]Sayer [seemed to concede] Process:Verbal [the sensitivity of the issue]Verbiage. (*NYT*, 07 May 1991)

(9) [On Friday]Circumstance, [Saddam]Sayer [promised] [the United Nations] Receiver that [[he]Sayer [would permit]Process:Verbal [U.N. inspectors]

Receiver [''prompt and unimpeded access" to locations designated for inspection]Verbiage]Verbiage. (*WP*, 07 July 1991)

(10) [U.S. reports]Sayer [indicate]Process:Verbal that [[Saddam Hussein]Sayer [has ordered]Process:Verbal [the killing of as many as 20 officers in the upper military hierarchy]Verbiage.]Verbiage (*WSJ*, 26 July 1991)

5.2.2 The Madman

The Madman discursive frame represents instances in which Saddam Hussein and his actions are represented as illegitimate – through the use of terms such as "dictator" or "regime" – or the Iraqi president personally and his actions as criminal and worthy of suspicion. In comparison to The Stateman discursive frame, this is the opposing end of a spectrum of legitimacy. Within The Madman discursive frame, terms of reference *Saddam* and *Saddam Hussein* are particularly prevalent. In addition, in the analysis of masked 'referent' concordances, a few examples of *Honorific + Hussein* were also included in this category, but they were nonetheless comparatively underrepresented in this discursive frame.

Examples (11) and (12) represent instances categorized in The Statesman and The Madman discursive frames, respectively, and illustrate the basis on which certain concordances were placed in either category. The two statements share a clear similarity: in both cases, Saddam Hussein is breaking with an agreement or directive. However, in example (12), it is further implied that he is unreliable and unworthy of trust, and thus this example was categorized in The Madman frame.

(11) And it can – if Saddam Hussein would simply comply unconditionally with all the resolutions of the United Nations. (*NYT*, 16 February 1991)

(12) Since Saddam flouted the original agreement, it would be reckless to rely on his unsupported word. (*WP*, 05 May 1991)

Similarly to the concordances as a whole, The Madman discursive frame includes many Material Processes. However, in comparison to The Statesman, this particular frame is associated with immoral and even criminal activities such as massacring, flouting agreements, carrying out a genocide, attacking, getting away with murder, as well as with a variety of Participant Roles. These roles range from Actor (the participant completing the action) as in example (13), to Goal (the participant affected by the action) in example (14)

and Recipient or Beneficiary (the participant gaining or benefiting from the action). Goal Participant roles are found in connection with the terms of reference Saddam and Saddam Hussein both within The Statesman and The Madman frames, with many connected to how other events have acted upon the Iraqi leader to affect his fate and future specifically as the leader of Iraq.

(13) [Worse still]Circumstance, [Saddam]Actor [has hidden away]Process:Material [an arsenal of weapons of mass destruction as well as the missiles to deliver them]Goal… (*NYT*, 31 July, 1991)

(14) [Although [Saddam]Goal [has]Process:Material [clearly]Circumstance [been weakened]Process:Material by [the war]Actor]Verbiage, [the official]Sayer [said] Process:Verbal, … (*WP*, 28 February 1991)

There are only a few Verbal Processes in this discursive frame, and in these concordances, Saddam Hussein is typically placed within the functional elements of Verbiage or the Target (the Participant of whom others are speaking). Notably, among the Verbal Processes there are several cases of President Bush specifically being quoted as describing his Iraqi counterpart in unflattering terms. Examples (15), (16), and (17) illustrate this:

(15) [Mr. Bush]Sayer [will struggle to keep]Process:Verbal [the two subjects] Target [separate]Verbiage by [stressing the extent and gravity of Saddam's cheating]Verbiage. (*NYT*, 31 July 1991)

(16) [He]Actor/Sayer [campaigned]Process:Material [like Richard Nixon]Circumstance and [talked]Process:Verbal] [like Lyndon Johnson]Circumstance, [saying] Process:Verbal [publicly]Circumstance that [[Saddam Hussein]Carrier [was] Process:Relational ["worse than Hitler,"]Attribute]Verbiage … (*NYT*, 16 June 1991)

(17) ["Everyone knows that the man was cheating and lying,"]Verbiage [Mr. Bush]Sayer [said]Process:Verbal [of Mr. Hussein's effort to conceal his nuclear abilities]Target. (*NYT*, 02 July 1991)

In contrast to the sample concordances as a whole, there are more Mental Processes than Verbal Processes in the concordances categorized in The Madman discursive frame. In these instances, Saddam Hussein often

appears in the Participant role of a Phenomenon (the represented thought, perception or feeling). The Mental Processes in these instances serve a somewhat similar function to Verbal Processes, representing evaluations of Saddam Hussein by others who have then, in turn, relayed this evaluation to a third party such as a reporter. These evaluations are generally not positive, as they include unsettledness and mistrust, as in example (18), as well as predictions about the Iraqi president's future actions, as in example (19).

(18) [MASSOUD BARZANI]^{Senser} … [mistrusts]$^{Process:Mental}$ [Saddam's pledges]^{Phenomenon} (*WP* 03 May 1991)

(19) [The White House]^{Senser} [expects]$^{Process:Mental}$ [Saddam to continue his cat-and-mouse game of divulging only as much as he thinks he must to avoid getting whacked by U.S. air power]^{Phenomenon}. (*WP* 14 July 1991)

Finally, it is worth noting that of the concordances categorized in The Madman discursive frame, slightly more than half were in the final three months of the corpus (May, June, and July) in comparison to the first four. What this shows is that at the very least this representation of Saddam Hussein was not overemphasized during the military operation – quite the opposite, in fact. The Madman discursive frame was maintained and possibly even strengthened after the military operation itself was concluded. The rise and subsequent oppression of the Shi'ite and Kurdish uprisings in Iraq may have also played a role. The selected concordances from the last three months of the corpus show that newspaper reports were as concerned with reviewing past events as with reporting current development.

5.2.3 Our Saddam

The third and final discursive frame discussed in this study is the Our Saddam frame. This discursive frame is closely connected with the term of reference "Saddam", and to a lesser extent with that of "Saddam Hussein". In comparison to The Madman frame, however, the Our Saddam frame is the representation of the Iraqi leader in the words of other regional actors. Some of these actors express views supportive of Saddam Hussein while others are critical or antagonistic, but overall this discursive frame is represented as a Middle-Eastern point of view that is separate from a U.S.-based perspective. Concordances included in the Our Saddam frame are distributed quite evenly among newspapers and months and include

mainly Material and Verbal Processes as well as some examples of Relational Processes. However, there is concentration of examples in May 1991, when there are multiple examples of this discursive frame, with several sources evaluating the future of Saddam Hussein as the leader of Iraq.

The first-name reference is used in both direct and indirect reported speech and by, among others, an Iraqi Shiite dissident, a Kurdish resident, and the Turkish president. Verbal Processes are connected with either man-on-the-street type interviews on the one hand or officials and political leaders on the other hand: in example (20), the Sayer is Iraqi foreign minister Tariq Aziz, while in example (21) it is a Palestinian merchant. The sole Kuwaiti voice is that of a soldier who after example (22) goes on to say that the Kuwaitis have "a bigger heart".

(20) [Aziz]Sayer [disputed]Process:Verbal [reports that the [raid]Actor [killed] Process:Material [some relatives and officials close to Saddam who may have taken refuge there]Goal.]Verbiage (*WP*, 08 May 1991)

(21) ["[No one]Actor [forced]Process:Material [Saddam]Goal [to withdraw] Process:Material,"]Verbiage [said]Process:Verbal [George Thalji, a merchant]Sayer. (*NYT*, 04 March 1991)

(22) ["[Saddam Hussein]Carrier [has]Process:Relational [a big army]Attribute,"] [says]Process:Verbal [Sgt Mohammed Rasheed]Sayer. (*WSJ*, 21 February 1991)

In comparison to The Madman and The Statesman discursive frames, here the usage of "Saddam" is by a perceived in-group – Middle Eastern voices, leaders and citizens, talking about one of their own, regardless of their approval of Saddam Hussein's actions. This type of usage is also implicitly acknowledged in example (23) by a U.S. analyst discussing what could be said by people in the region and what by U.S. government actors:

(23) ["[Some of the things we might like to say about Saddam]Carrier [were far better left]Process:Relational [to the Middle Easterners]Attribute,"]Verbiage [the analyst]Sayer [said]Process:Verbal. (*WP*, 09 April 1991)

Additionally, a specific sub-group in this discursive frame, including examples of Relational Processes, is connected to the Iraqi president's familial and tribal ties, as in example (24):

(24) [At odds with the Majids]Attribute [are]Process:Relational [Saddam's half-brothers, the Ibrahims]Carrier, who [share]Process:Relational [a common mother with Saddam]Attribute. (*WSJ*, 01 March 1991)

The conclusion will draw together the findings from the analysis and discuss their implications.

6. Conclusion

The analysis identified three distinct discursive frames for Saddam Hussein: The Statesman, The Madman, and Our Saddam. The Statesman is the most common of these discursive frames, while The Madman and Our Saddam frames were comparatively rarer. While The Madman frame is closely associated with the *Saddam* term of reference, the same term is also used in connection with the Our Saddam discursive frame, demonstrating that the same term of reference can function in different and even contradictory ways depending on the specific context in which it is used. The three analyzed newspapers also show marked differences in their use of the different terms of address, each using a somewhat different distribution of the analyzed terms of reference.

The comparatively largest share of The Statesman discursive frame and the smaller share of The Madman frame suggests that while representations of Saddam Hussein as a dangerous individual were to some extent highlighted, this was not the dominant practice in news reports during the analyzed time period. While previous studies have emphasized the demonization of Saddam Hussein in the Persian Gulf conflict (see, for example, Artz – Pollock 1995; Martín Rojo 1995), the results of this study echo the findings by Baker et al. (2013: 225) in their examination of Muslims and Islam: that newspapers did not generally employ extreme negative and generalizing stereotypes but rather presented a more ambivalent and subtle image. The efforts of President George H.W. Bush in promoting military action is well documented in the case of the Persian Gulf War (Kellner 2004; Oddo 2018). However, the findings in this study may be evidence in support of Hart and Fuoli (2020), who found that political leaders' appeals for war may have a limited persuasive impact on the media and the public without cited evidence or supportive societal attitudes.

It is, however, also notable that instances of The Madman discursive frame were not limited to the active war operation in the early half of the analyzed timeframe. Rather, the discursive frame was maintained and even

strengthened after the end of Operation Desert Storm. This indicates that the delegitimizing effects of The Madman discursive frame had a lasting impact beyond the end of the military operation itself. Again, keeping in mind the findings of Hart and Fuoli (2020), the later arguments for military action against Saddam Hussein made by President George W. Bush in the build-up for the Iraq War (as explored by, among others, Oddo 2011; Abid – Manan 2016) would have benefited from the discursive frames established previously in the Persian Gulf War.

REFERENCES

Sources

ProQuest Historical Newspapers™
https://about.proquest.com/products-services/pq-hist-news.html, accessed December 2018

Special studies

Abid, R.Z. – S.A.Manan
2016 "Constructing the 'self' and the 'other' in Bush's political discourse before and after the Iraq war (2002-2008)", *Journal of Language and Politics* 15 (6), 710-726.

Al-Ali, N.
2011 "A feminist perspective on the Iraq War", *Works and Days* 29, 99-114.

Anthony, L.
2011 *AntConc* (Version 4.3.3w). Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software, accessed December 2018

Artz, L.W. – M.A. Pollock
1995 "Limiting the options: Anti-Arab images in U.S. media coverage of the Persian Gulf War crisis". In: Y.R. Kamalipour (ed.) *The U.S. Media and the Middle East – Image and Perception*. Westport, Conn.: Praeger, 119-136.

Atkinson, R.
1993 *Crusade: The Untold Story of the Persian Gulf War*. New York: Houghton Mifflin Company.

Baker, P.
2006 *Using Corpora in Discourse Analysis.* London: Continuum.

Baker, P. – G. Costas – T. McEnery
2013 *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge: Cambridge University Press.

Bell, A.
1991 *The Language of News Media*. Oxford: Blackwell.
Branum, J. – J. Charteris-Black
2015 "The Edward Snowden affair: A corpus study of the British press", *Discourse & Communication* 9 (2), 199-220.
Clark, K.
1992 "The linguistics of blame: Representations of women in *The Sun*'s reporting of crimes of sexual violence". In: M. Toolan (ed.) *Language, Text and Context: Essays in Stylistics*. London: Routledge, 208-226.
Clayman, S.E.
2010 "Address terms in the service of other actions: The case of news interview talk", *Discourse & Communication* 4 (2), 161-183.
Eggins, S.
2004 *An Introduction to Systemic Functional Linguistics* (2nd edn.). London; New York: Continuum.
Fowler, R.
1991 *Language in the News: Discourse and Ideology in the Press*. London: Routledge.
Gerbner, G. et al.
1986 "Living with television: The dynamics of the cultivation process". In: J. Bryant – D. Zillman (eds.) *Perspectives on Media Effects*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 17-40.
Hackett, R.A. – Y. Zhao
1994 "Challenging a master narrative: Peace protest and opinion/editorial discourse in the US press during the Gulf War", *Discourse & Society* 5 (4), 509-541.
Halliday, M.A.K. – C.M.I.M. Matthiessen
2004 *An Introduction to Functional Grammar* (3rd edn.). London: Edward Arnold.
Hart, C. – M. Fuoli
2020 "Objectification strategies outperform subjectification strategies in military interventionist discourses", *Journal of Pragmatics* 162, 17-28.
Kellner, D.
2004 "The Persian Gulf TV war revisited". In: S. Allan – B. Zelizer (eds.) *Reporting War: Journalism in Wartime.* London; New York: Routledge, 136-154.
Kuosmanen, S.
2019 "Terms of reference and objectivity in US press reports in the Gulf War in 1990", *Journalism*, https://doi.org/10.1177/1464884919847310, accessed December 2018
Lule, J.
2004 "War and its metaphors: News language and the prelude to war in Iraq, 2003", *Journalism Studies* 5 (2), 179-190.

Mautner, G.
2009 "Checks and balances: How corpus linguistics can contribute to CDA". In: R. Wodak – M. Meyer (eds.) *Methods of Critical Discourse Analysis* (2nd end.). London: Sage, 122-143.
Martín Rojo, L.
1995 "Division and rejection: From the personification of the Gulf conflict to the demonization of Saddam Hussein", *Discourse & Society* 6 (1), 49-80.
Mazarr, M.J. – D.M. Snider – J.A. Jr. Blackwell
1993 *Desert Storm: The Gulf War and What We Learned.* Boulder, Colo.: Westview Press.
Mills, S.
2003 "Caught between sexism, anti-sexism and 'political correctness': Feminist women's negotiations with naming practices", *Discourse & Society* 14 (1), 87-110.
Mowlana, H. – G. Gerbner – H.I. Schiller (eds.)
1992 *Triumph of the Image: The Media's War in the Persian Gulf – A Global Perspective*. Boulder, Colo.: Westview Press.
Nohrstedt, S.A.
1992 "Ruling by pooling". In: H. Mowlana – G. Gerbner – H.I. Schiller (eds.) *Triumph of the Image: The Media's War in the Persian Gulf – A Global Perspective*. Boulder, Colo.: Westview Press, 118-127.
Oddo, J.
2011 "War legitimation discourse: Representing 'Us' and 'Them' in four presidential addresses", *Discourse & Society* 22 (3), 287-314.
2018 *The Discourse of Propaganda: Case Studies from the Persian Gulf War and the War on Terror*. University Park: Pennsylvania State University Press.
Page, R.E.
2003 "'Cherie: lawyer, wife, mum': Contradictory patterns of representation in media reports of Cherie Booth/Blair", *Discourse & Society* 14 (5), 559-579.
Peer, L. – B. Chestnut
1995 "Deciphering media independence: The Gulf War debate in television and newspaper news", *Political Communication* 12 (1), 81-95.
Popp, R.K. – A.L. Mendelson
2010 "'X'-ing out Enemies: *Time Magazine*, visual discourse, and the war in Iraq", *Journalism* 11 (2), 203-221.
Richardson, J.E.
2007 *Analysing Newspapers*: *An Approach from Critical Discourse Analysis.* Basingstoke: Palgrave Macmillan.
Siegal, A.M. – W.G. Connolly
2015 *The* New York Times *Manual of Style and Usage* (5th edn., revised and updated by P.C. Corbett et al.). New York: Three Rivers Press.

Steuter, E. – D. Wills
2010 "'The vermin have struck again': Dehumanizing the enemy in post 9/11 media representations", *Media, War & Conflict* 3 (2), 152-167.
Van Dijk, T.A.
1988 *News Analysis: Case Studies of International and National News in the Media*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Address: SONJA KUOSMANEN, English Philology, Faculty of Arts, PO Box 24, FI-00014 University of Helsinki, Finland.
ORCID code: orcid.org/0000-0003-2477-0803.

Making sure everybody is on the same page: interactional communication strategies in BELF encounters

Valeria Franceschi
University of Verona

ABSTRACT

Communication in business often comes with high stakes. To achieve successful communication and positive business outcomes in international and intercultural settings, it is necessary for participants to convey their message effectively and accurately. The success of BELF interactions relies heavily on cooperative behavior, which is essential to meaning negotiation and to the prevention or management of communication problems. Participants in international business interactions highlight the importance of "asking clarifying questions and of checking, double-checking, confirming, and reconfirming" (Louhiala-Salminen – Kankaanranta 2011: 256). This paper aims at exploring how BELF users interact with other participants to ensure intelligibility and comprehension. Specifically, the use of appeals for assistance and responses will be investigated, the use of which has already been attested in ELF. A qualitative approach will be adopted to analyze naturally-occurring BELF data drawn from business conversations and meetings in the Professional Business and Professional Organizational subsections in the VOICE corpus.

Keywords: BELF, English as a Lingua Franca, Communication Strategies, business communication, intercultural communication.

1. Introduction

Professional communication in English has become more and more common in recent years, due to increasing workforce mobility and advancements in technology that foster international business relations. Such business interactions are often carried out by professionals who come from different

DOI: 10.25951/4366.

cultural and linguistic backgrounds and who adopt English as a shared language of communication. English in its lingua franca role in multilingual business contexts has been defined as BELF (English as a Business Lingua Franca), highlighting the domain of use and the speakers' membership in "the global business community" (Kankaanranta – Louhiala-Salminen 2010) in addition to being ELF speakers.

As business interactions are often task-oriented, with specific and sometimes sensitive goals to be achieved within a time-frame, it is paramount that the finer details of the topics discussed in the conversation are clear to all participating parties (Louhiala-Salminen – Kankaanranta 2011; Palmer-Silveira 2013).

For this reason, the success of BELF interactions relies heavily on cooperative and listener-oriented behavior, in order to negotiate meaning, prevent or manage any communication problem(s) that may arise. In a 2011 study, professionals highlighted the importance of "asking clarifying questions and of checking, double-checking, confirming, and reconfirming" (Louhiala-Salminen – Kankaanranta 2011: 256). In Ehrenreich's 2010 study, interviewees similarly remarked that "comprehension checks, asking for clarification and repetition, attention to facial expression" (2010: 422) are important strategies to ensure effective communication.

This paper aims at expanding current knowledge on the use of Communication Strategies (CSs) in ELF in work encounters, focusing specifically on the types of strategies identified by professionals themselves as paramount to effective communication. The study focuses on face-to-face conversations and meetings within the Professional Business and Professional Organizational sections of the VOICE corpus. The investigation adopts mainly a qualitative approach, with a focus on the function and effectiveness of the CSSs in these BELF encounters.

1.1 Communication Strategies

Studies on Communication Strategies started in the late 1970s in relation to Second Language Acquisition (SLA) (e.g. Selinker 1972; Tarone 1977; Canale – Swain 1980). Originally, non-native speakers were thought to adopt these strategies to compensate for their lack of proficiency in the target language. This 'deficit' perspective is indeed embedded in multiple definitions of CSs, as in Corder's, which views a CS as "a systematic technique employed by a speaker to express his meaning when faced with some difficulty" (1983: 16) and Bialystok's, which emphasizes "[manipulating] a limited linguistic

system in order to promote communication" (1983: 102). Faerch and Kasper similarly mention that CSs are used by learners to "overcome problems" and "create the conditions for intake" (1983: 36). In early studies on CSs, the focus remained firmly on the use of such strategies to solve "gaps", "breakdown" and "problems" in communication (Coupland et al. 1991: 3) and to compensate for failure in communication (Canale – Swain 1980: 30).

However, the use of CSs is not limited to learners trying to compensate for their imperfect knowledge of the language. On the contrary, CSs are found to be pervasive in native speaker use as well, regardless of the presence of non-native speakers in the conversation (Firth – Wagner 1997). These types of strategies exist for both L1 and L2 speakers, and they are considered an essential characteristic of effective communicators (Savignon 1997: 47; Widdowson 2003); in Tarone's words, CSs are "tools used in a joint negotiation of meaning where both interlocutors are attempting to agree as to a communicative goal" (Tarone 1980: 420), whose "main function […] in the interactional approach is to assist both communicators to agree on and convey the meaning in an interactional situation" (Jamshidnejad 2011: 3758). The effective use of CSs also involves the ability, on the part of participants, to respond to "a variety of changing and often unexpected interpersonal conditions" and to "meet the demands of ongoing communication" through the selection of the most appropriate strategy in any given situation (Savignon 1997: 44), independent of their native or non-native status.

For these reasons, study of CSs is relevant in ELF, as they constitute an important aspect of speakers' communicative competence, which they possess and deploy as they see fit alongside other components to reach mutual understanding. As the ELF approach neither adopts a deficit perspective nor sees the ELF user as a learner striving towards – but never quite reaching – native-like competence, SLA frameworks are not appropriate for the analysis of ELF data (Björkman 2014: 124). CS research in ELF is indeed "achievement rather than problem-oriented" (Kaur 2009: 41).

1.2 Communication Strategies in ELF and BELF

CSs are considered in ELF as part of "normal pragmatic practice" (Widdowson 2003), employed as a means of facilitating and co-constructing understanding (Kaur 2009; Björkman 2011, 2014). They appear to be frequently employed by ELF users, as both pre-emptive and retroactive moves, in order to ensure that communication is successfully achieved. Indeed, studies have shown that ELF speakers do not make use exclusively of compensatory strategies to

the purpose of solving instances of non-understanding or misunderstanding. On the contrary,

> [e]ven in sequences when there appear not to be overt displays of misunderstanding, the non-native participants are found to employ various interactional practices to check, monitor and clarify understanding. These include clarification and confirmation requests, which are in turn receipted by moves to repeat, rephrase or explain some prior utterance (Kaur 2009: 46).

Indeed, ELF users appear to display a high degree of mutual cooperation, being highly aware of the many variables that may come into play in an ELF interaction and of their own – and their interlocutor's – varying language competence and different communicative styles. Awareness of these asymmetries, as Björkman calls them (2014: 124) appears to lead ELF speakers to adopt a pro-active attitude (Mauranen 2006b; Kaur 2009). Indeed, the achievement of understanding in ELF communication is a joint enterprise, where participants in ELF interactions work together through the use of various strategies to guarantee the success of the communicative event, assessing and adapting their use of the language at each turn according to the needs of the interaction. ELF speakers appear to make use of explicitation strategies (Mauranen 2006a), both speaker-initiated and other-initiated, that have the purpose of preventing non-understandings and misunderstandings from occurring. As will be seen in our data as well, mutual cooperation is not limited to the use of CSs to ensure successful communication, but also to the maintenance of face. In Kirkpatrick's words, "the overarching goal in this type of lingua franca conversation is to ensure communication on the one hand, while preserving the face of the participants on the other" (2008: 33), suggesting that it is important in ELF conversation to make sure participants are up-to-date and feel involved in the conversation.

To date, most studies on CSs in ELF have primarily examined the functions and use of certain communication strategies (cf. e.g. Firth 1996; Wagner – Firth 1997; Meierkord 2000; Lichtkoppler 2007; Cogo 2009; Bjørge 2010), with fewer studies having attempted to propose a systematic taxonomy of CSs occurring in natural ELF data (Björkman 2014). Research on the topic spans different geographical contexts, such as Asia (Kirkpatrick 2007, 2008), as well as different domains of use: larger studies were carried out in relation to academic ELF (cf. Kaur 2009; Björkman 2011, 2013, 2014; Mauranen 2006b, 2012). Investigations have also been carried out on leisure

interactions (Vettorel 2019) and BELF (Pitzl 2010). Other studies looking at CSs in BELF have focused primarily on code-switching and plurilingual practices, which have been shown to play a role in meaning negotiation, establishing interpersonal relationships, expressing L1 or LN cultural affiliation, or projecting professional identity (Ehrenreich 2010; Cogo 2012, 2016; Franceschi 2017).

In BELF communication users appear to be "strategically competent", using both verbal and non-verbal strategies to achieve their communicative goals. Indeed, "applying appropriate communicative techniques" is one of the three main skills BELF users should possess to be successful communicators, alongside "the use of appropriate business related concepts" and "using the corresponding specific terminology in English" (Palmer-Silveira 2013: 9-10). As a prevalently goal-oriented type of communication that may have high stakes, "[t]he dynamism [of BELF] entails that strategic skills, such as ability to ask for clarifications, make questions, repeat utterances, and paraphrase (see Mauranen 2006b), gain in importance for successful communication" (Kankaanranta – Louhiala-Salminen 2013: 28).

2. Methodology and data selection

In light of previous studies on BELF, this paper intends to look at how participants in BELF ensure that comprehension is maintained throughout the interaction by actively checking with their interlocutors, either pre-emptively or by signaling lack of understanding. An additional point of investigation is related to responses to appeals for assistance, as the study also aims at identifying instances of lack of response and the reasons they occur.

In order to investigate the use of interactional communicative strategies in BELF encounters, a choice was made to study naturally-occurring spoken data in professional environments where participants do not share a common language, effectively characterizing them as ELF speakers. The corpus selected for this purpose was the *Vienna-Oxford International Corpus of English* (VOICE), which includes about a million words of ELF spoken data in various contexts of communication and different types of interaction. For the scope of this enquiry, it was decided to concentrate on the Professional subset of data in the corpus, specifically on the Professional Business (PB) and Professional Organizational (PO) sections of the corpus. Within these sections, only the conversation and meeting subsections were selected for

analysis, as the nature of these communicative events is highly interactional. The subcorpus contains a total of circa 291,000 words.

Analysis of the data will be carried out from a qualitative perspective building on existing CSs theory. The qualitative approach will contribute to shedding light on the use of each individual strategy and identify which strategies are favored by the participants to the interaction. The qualitative aspect of this investigation will look at selected occurrences from the corpus, in order to show how speakers make effective use of the strategies.

The focus of this study will be mainly on three macrosets of strategies, that is, appeals, comprehension checks and responses, as illustrated by Dörnyei and Scott in their 1997 comprehensive taxonomy of CSs (191-2). These strategies were then adapted and updated according to the linguistic behaviors adopted by participants in the communicative events, where they employed strategies aimed at ensuring that the message has been shared effectively and accurately, either self-initiated or other-initiated.

This taxonomy was adopted for convenience; however, CSs included in the 'appeals' macrocategory were attested by multiple scholars as frequent strategies used in ELF interactions in multiple contexts, although sometimes referred to with different terminology: confirmation checks, clarifications, repetitions (Mauranen 2006b; Kaur 2009); requests for repetition and clarification (Kirkpatrick 2007); repetitions (Lichtkoppler 2007); requests for confirmation and understanding (Kaur 2010, 2011); overt questions, clarification requests (Björkman 2014). Other terminology was borrowed by different scholars for those strategies that were not included in Dörnyei and Scott's list.

1. Appeals
    a. Direct and indirect appeals
    b. Request for repetition
    c. Request for clarification
    d. Request for confirmation
2. Comprehension checks
3. Responses
    a. Repetition
    b. Response to direct appeal with other-repair
    c. Response (rephrasing, expanding, definition, exemplification, etc.)
    d. Lexical anticipation in response to hesitation
    e. Repair after request for confirmation

4. Other
    a. Other-initiated word replacement
    b. Statement of incorrectness followed by replacement / paraphrase / extension / definition / exemplification

The CSs were identified through careful analysis of the previous and following turns. As a strategy's surface form and its pragmatic function may not necessarily match, paying close attention to the context is paramount to determine the function of the strategy (Ohta 2005; Foster – Ohta's 2005; Kaur 2009; Jamshidnejad 2011; Björkman 2014).

The classification of certain appeals was not always straightforward, due to their ambiguity. Certain appeals may be either requests for repetition or for clarification, as in the case of open class repair initiators (Drew 1997) such as "sorry?" "huh?", which are used as signals of non-understanding but do not indicate where the problem lies. In such cases, it is up to the next speaker to decide whether to use repetition or attempt clarification through paraphrase or extension. In Kaur's data, it appears that "[t]he ELF speaker is more likely to repeat a prior turn when faced with an open class repair initiator as this involves straightforward recycling of existing material" (Kaur 2009: 142). This appeared to be consistent with the data in this study, as most responses provided in these ambiguous cases were repetitions. The participants that had produced the appeal appeared to be satisfied with the response received in the vast majority of instances, either explicitly signaling comprehension or by continuing the conversation. In these cases, the appeal was marked as a request for repetition. Where repetition was not enough to ensure comprehension, further clarification was prompted in an additional move, as will be seen in the data.

Some instances of CS use were not categorized at all, i.e. where unclear speech was transcribed and it was not possible to reconstruct the context, or function of the CS.

3. Results

A summary of the number of instances for each of the strategies selected and investigated may be seen in Table 1 below. It should be noted that, as described above, classification was not always straightforward, and in some cases, identified strategies were not added to the count where unclear speech or gaps in the transcription made it impossible to ascertain their

function. Nevertheless, a breakdown of frequencies will help disclose BELF users' preferences when employing CSs.

Table 1. Frequency of the CSs identified in the VOICE subcorpus

Strategy type	No. of occurrences
APPEALS	
Direct and indirect apppeals	11
Indirect appeals	10
Requests for repetition	73
Requests for clarification	35
Requests for confirmation	72
Comprehension checks	38
RESPONSES	
Repetition	67
Response to direct appeal with other-repair	5
Response (rephrasing, expanding, definition, exemplification, confirmation, etc.)	112
Lexical anticipation in response to hesitation	6
Repair after request for confirmation	1
No response to appeal/ request	18
OTHER	
Other-initiated word replacement	21
Statement of incorrectness and replacement / paraphrase / extension	7

As can be seen, requests for repetition and confirmation appear to be the most frequent CSs in the subcorpus. This may be explained by the fact that these strategies are immediate and require a lower cognitive load for the speaker: indeed, requests for repetitions are usually expressed via one-word minimal queries, and requests for confirmation consist in repetition of the trigger with rising intonation. As a consequence, repetition (67) and confirmation (52 out of the 112 responses) are similarly the most common responses to appeals and requests, with more elaborated responses coming into play only when further clarification is needed or expressly requested.

There are, however, multiple factors that play a role in strategy selection, including other-oriented face-saving behaviors, as will be discussed in the following subsections.

3.1 Appeals

In the categories of appeals, speakers either signal an existing problem or what they perceive might be a potential problem (Mauranen 2006b; Kaur 2011; Björkman 2014) in their own understanding of previous turns (request for repetition, clarification or confirmation) or in their own use of the language. In the latter case they might ask explicitly for help (direct appeal), or their behavior, verbal or non-verbal (for instance pauses or hesitations) may be an attempt to "elicit help from the interlocutor" (Dörnyei – Scott 1997: 191).

3.1.1 Direct or indirect appeal

These types of appeals are usually responded to via a response to a direct appeal or through lexical anticipation (Kirkpatrick 2007), where another speaker completes the previous utterance by providing what they think is the term the original speaker intended, as in extract 1 below.

(1)
200	S2:	and that they have to er treat their father in special way because if they er don't don't catch him er catch his erm
201	S4:	mhm
202	S4:	attention or attraction yeah yeah yeah yeah
203	S2:	attention (PBcon594)

In the passage above, S4 notices S2's difficulty in completing her turn and provides a couple of potential options before S2 can retrieve the word she was looking for. The intended meaning is then accepted and confirmed by S2 through repetition. The strategy used here is one of anticipation (Kirkpatrick 2007). This type of behavior, in addition to ensuring comprehension, may also be oriented at preserving S2's face, with S4 intervening helpfully when she displays difficulty in her turn.

In direct appeals, the speaker asks explicitly for assistance with a word they cannot express, as in example 2 below:

(2)
101	S1:	= and these er are (.) actually part of the fish these are (1) er what do you call the (1) eggs of fish (.)
102	S7:	the eggs of fish?
103	S1:	yah
104	S7:	caviar
105	S1:	well if it's a very

106 SS: @@ <1> @@ </1>
107 S1: <1> e:r EXquisite er </1> egg from a very exquisite fish <2> but wh- wh- wh- (.) wh- wh- </2> er er (1)
108 S7: <2> but then it's a fish the: the: eggs of fish (.) i guess </2>
109 S7: <3> (i don't know) </3> (POcon534)

S1 makes a direct appeal, asking for the specific word used to define "fish eggs", that is, roe. S7 asks for confirmation, and, once he receives it from S1, proceeds to respond to the appeal by offering what he thinks may be an appropriate word. However, S1 does not appear entirely satisfied, so he adds further information to highlight the difference between S7's suggestion and his intended meaning. The two participants do not come to a decision as to the most appropriate word for fish egg, but the conversation is not disrupted and mutual intelligibility is achieved, as both participants agree on the meaning of S1's utterance.

Direct and indirect appeals were not common in the data; indirect appeals through hesitation occurred slightly more often; it is, however, unclear whether speakers would have used a direct appeal or attempted a rephrasing if the interlocutor had not provided a solution via lexical anticipation. It is hard to speculate as to why direct appeals appear not to be a commonly-used CS in this dataset: it may be due to the speakers trying to make themselves understood through other strategies such as rephrasing and simplification before disrupting the flow of the conversation with a direct question. In multiple instances in the data, rather than producing a direct appeal, speakers signal their linguistic "insecurity" by pronouncing marked terms in a rising intonation, as if to check the accuracy/correctness of their linguistic choice with the others, and leaving room for other-initiated word replacement or meaning negotiation sequences.

3.1.2 Request for repetition

Requests for repetition appeared to be fairly common in the data, alongside requests for confirmation. Most commonly, as will be seen in example 4, requests occur via minimal queries/open class repair initiators such as "sorry?", "huh?", "what?", "pardon?", "pardon me?", all identified in the data, rather than more explicit requests for repetition or signals of non-understanding requiring participant intervention.

In example 3, other types of appeal occur throughout the passage, probably due to high levels of background noise that may have impaired message comprehension.

(3) 378 S4: funding is e:r (1) certainly part of e:r (1)
379 S6: (right one)
380 S4: the administrative aspect (1)
381 S2: sorry? (.)
382 S4: the <1> funding </1> and the financial aspects
383 S6: <1> funding </1>
384 S2: yes mhm mhm that's right (7) {S2 is writing (7)}
385 S7: contracts and agreements (.)
386 S4: sorry? (.)
387 S3: <2> i- </2>
388 S2: <2> con</2><3>tracts agreements </3> (.)
389 SX-f: <3> contracts (and) agreements </3>
390 SX-f: <un> x xx </un>
391 S4: agreements ah (.) {S2 starts writing} (POmtg315)

In line 381, non-understanding is signaled through the use of the minimal query "sorry?". S4 and S6 both provide a response through repetition and expansion, which is then followed by S2's signal of understanding. A second signal of non-understanding is carried out just two lines later by S4; response occurs via repetition, which seems to satisfy S4, who in turn confirms comprehension through repetition/echoing. Both appeals in this case were interpreted as requests for repetition, except for S4, who interpreted S2's minimal query as a request for clarification.

3.1.3 Request for clarification

In the following passages, more overt requests for clarification are made, where open repair initiators are usually substituted by wh- questions such as "what is/are…?", "What do you mean by/with…?", sometimes followed by the repetition of the problematic item. At times, requests for clarification may not be formulated clearly enough to be perceived as such, making additional moves necessary to reach mutual understanding, as can be seen in the example below:

(4) 1724 S1: the action (.) o- er DONE is a key account management program we asked er [first name2] (.) to set up e:r some program (.) with imp- (.) the the important er (.) key ACCOUNT (.) to develop some ANNUAL program (.) for ou-
1725 S4: what is an annual program? (.)

1726 S1: it's a YEARLY (.) based (.) <4><soft> (program) </soft></4>

1727 S4: <4><L1ger> ja ja {yes yes} </L1ger></4> (i mean) in regards of <5> assortment </5> in regards of placement or in regards of e:rm activities? (.) (PBmtg3)

In (4), S4 asks for clarification of the expression "annual program". Assuming S4 does not know the meaning of the world "annual", he substitutes the trigger with a word similar to "year", supposing it would be more familiar to S4 and thus easier to understand. However, S4 signals in German, his L1, that he did in fact understand the form of the expression ("*ja ja*"), but what he was interested in was the aspect, or aspects, S1's company meant to schedule in the program. The conversation, not reported here in its entirety, continues until S4 is satisfied that there is shared understanding on the contents of the annual program. This once again suggests the importance, in professional conversations, of obtaining a precise understanding of what the interlocutor means, especially when different companies/institutions may have slightly different interpretations or practices related to the same expression.

3.1.4 Request for confirmation

Another common appeal in the data were requests for confirmation. Often, participants repeated part of the previous speaker's turn with a rising intonation, in order to check whether their own understanding was correct (Kaur 2009: 112). In some cases, a repetition with a questioning intonation may not be a request of confirmation, but may instead express surprise (Corsaro 1977: 190), or encouragement to continue and/or provide further information (Foster – Ohta 2005: 407). Such occurrences were detected in the current dataset as well.

In the majority of cases of confirmation requests, a confirmation is produced, but in one case, where the participant heard or understood incorrectly, a repair was provided, as shown in the following example:

(5) 2146 S8: <fast> so how many how many </fast> er lower decks is on the <spel> a </spel> three: three hundred er three three THIRTY for the: (.) for the cargo? six er bellies er for the cargo? o:r less. (.) for the connections. (.) as well. (.) right NOW you're on the (backside you're) right? (.)

2147 S2: you mea- you mean the: ex <3> europe flights? makes </3>

2148 S8: <3> hongkong er hongkong </3> tokyo. (2)
2149 S2: oh <4> the the </4> TOKYO how MANY flights? (.)
2150 S8: <4> hongko- hongkong </4>
2151 S8: mhm
2152 S2: <8> e:r </8>
2153 S1: <8> no:</8> how many pallets per er er er does fit on a on on a <5> eight three thirties </5>
2154 S2: <5> in the three thirties </5><9> three thirties </9>
2155 S8: <5><un> xxx </un></5><9><un> xx </un> didn't </9> know what <6> they've e:r </6>
2156 S1: <6> three thirties </6> yeah yeah
2157 S8: between hongkong and japan right?
2158 S2: yeah?
2159 S8: okay. h- how many positions you keep for for cargo?
2160 S2: i believe it was THREE <7> three </7>
2161 S8: <7> three </7> for for cargo? =
2162 S2: = lower lower deck <10> pallets </10> (.)
2163 S8: <10> mhm </10>
2164 S2: the: the <spel> p m cs </spel>
2165 S8: yah yah yah (PBmtg300)

In the passage above, we see S2 asking for confirmation that he has understood S8's question. S8 provides a repair, indicating the route he was referring to. S2 contributes a repetition of part of S8's turn to signal comprehension and asks again for confirmation that he has understood his question. S8 repeats the name of the second city (2150), probably to ensure that S2 is aware both cities are involved, but then the conversation stalls, with S2 hesitating because of S8's lack of uptake of his request for confirmation. S1 then intervenes, signaling S2's lack of comprehension and rephrasing S8's original question (2153). In the following line, S2 provides a repair, correcting the aircraft type. S1 signals comprehension via repetition, and then S8 takes the floor again, reminding S2 that the question refers to the Hong Kong – Tokyo route, to ensure a relevant answer. S2's confirmation, in rising intonation, appears to encourage S8 to continue. When S8 finally receives his answer, he asks for confirmation that he had understood correctly, followed by S2's move, providing additional information until S8 signals comprehension ("yah yah yah").

The passage above also constitutes an example of collaborative negotiation, where three speakers participate in the process of achieving

mutual comprehension. The face of all participants is maintained, with speakers taking the necessary actions to achieve communication without leaving room for awkwardness. An example of this focus could be considered S1's intervention in line 2153, where he compensates for S8's lack of uptake and provides S2 with the means to answer the original question.

3.2 Comprehension checks

Comprehension checks are a pro-active move "on the part of the speaker to establish that his or her prior utterance has been sufficiently understood by the recipient" (Kaur 2009: 182). Comprehension checks in the corpus were prevalently expressed through single words or non-verbal sounds produced with a rising intonation, such as "okay?", "you know?", "yes?/yeah?", "huh?". Only occasionally the comprehension check would be more articulated, as in the examples below.

(6) a) 2146 S1: […] there's a (.) character called (.) digimon (.) <5> digimon </5> is <8> japa</8>nese er <9> you know </9> di- digimon? (.) (PBmtg3)

b) 366 S7: yeah okay? (.) i i didn't realize that about the belgians because that's hh that's the FRENCH system <fast> you know what i <1> mean</1></fast> (.) (POcon549)

In the corpus data, comprehension checks are rarely followed by an explicit signal of comprehension. Usually, the conversation continues without an explicit request for clarification or repetition, which is interpreted by the speaker as a sign that comprehension has been achieved. In some cases, comprehension checks are followed by signals of non-understanding, which triggered instances of co-construction, as can be seen clearly in (7) below: non-understanding is here due to an instance of codeswitching to refer to a locally-connoted concept. This type of meaning co-construction has been attested in analysis of leisure communicative events (cf. Vettorel 2019) in the VOICE corpus, where students discussed the local food and festivities of their own home cultures.

(7) 207 S4: then the: (.) organizer was saying everything was done but not the papers (.) not yet this written fact (.) hh then (.) it was the eight of july (.) the idea it is to do e:rm <smacks lips> er a session during a <L1fre> braderie {sale} </

L1fre> you know <L1fre> braderie {sale} </L1fre> it is sort of sale can you say sale?

208 S3: mhm

209 S4: e:r in this

210 S1: maybe you <7> should explain that a <LNfre> braderie {sale} </LNfre></7> a <LNfre> braderie {sale} </LNfre> for us is <8> for </8> you it's different than it's for [S2] <9> because </9> i i talked to her about this and she thought that it's <10> all about </10> selling furniture (1)

211 S3: <7> but can you explain a bit mo:re </7>

212 S3: <8> yeah </8>

213 S10: <9> no no </9>

214 S10: <10> we don't know </10>

215 S3: er

216 S10: <1> [S4] [S4] who is </1>

217 S4: <1> no no </1> no no no no erm it is you know when you have the the sale er in july you have a lot of <2> sale </2> in in in er shops (.)

218 S1: <2> yeah </2>

219 S4: okay?

220 S3: mhm

221 S4: and you have -s this also in january (.) (POmtg439)

S4 is clearly aware that her interlocutors may not be familiar with the French word *braderie* (line 207). Her comprehension check is combined with a translation, mitigated by the vague expression "sort of", which marks a certain degree of insecurity about the accuracy of the chosen translation. Indeed, she checks with the others that "sale" is an acceptable option: "can you say sale?" (line 207). Both S1 and S3 ask for additional clarification, with S1 specifying that more information regarding *braderie* is necessary, due to a different understanding of the word on the part of different speakers. S10 also reinforces the request for clarification by claiming he and his colleagues are not familiar with the concept. S4 then responds to the requests by providing additional information about the event, producing another comprehension check at line 219. S3's non-lexical vocalization is interpreted as a sign of encouragement and S4 continues her explanation. The conversation, not reported here for reasons of space, continues with further explanations and exemplifications by S4.

3.3 Responses

As seen in the previous sections, appeals and requests are generally attended to by other participants, who provide an appropriate response to the produced request to solve an existing or potential problem. Requests for confirmation are generally followed by confirmations – or, in one case, a repair –, while requests for repetitions are usually responded to with the repetition of the relevant segment of the trigger utterance. Should repetition be insufficient to restore mutual intelligibility, additional information may be elicited, with a request for clarification (cf. example 4) – or offered by the speaker.

When answering requests for clarification, speakers may rephrase the word – or utterance – to extend the context (Dörnyei – Scott 1997), or again, provide an exemplification or a definition of the problematic item.

(8) 2278 S1: okay (4) er NOW er talking about TARGET . (2) er (.) kids are (no more) (.) GULLIBLE {word is used in the presentation material} than adults. per<5>haps even less so </5> they're
2279 S5: <5><un> xxxxx </un></5>
2280 S4: excuse me er (.) GULLIBLE i've never heard that word. what does that mean? (.)
2281 S5: hm
2282 SX-1: <pvc> gullabry <ipa> ˈgʌləbri </ipa> </pvc> (more like) <6> that?</6>
2283 S4: <6> @ </6>
2284 S5: @@ (.)
2285 S2: <L1kor> x [first name5] xx?</L1kor>
2286 S4: gullible?
2287 S1: gullible gullible (1) yeah gullible means (2) not english word (.) (but) like er GREEDY ? (.)
[…]
2319 S1: <2> O:H </2> yeah i think <3> (it) EASY to be influenced </3> […] (PBmtg3)

When S4 hears a word he is not familiar with ("gullible"), he asks for clarification by explicitly asking for the meaning of the word. SX-1 attempts to solve the problem by producing an alternative – although non-standard – pronunciation of the word. It is hard to provide an interpretation for this

move, but S1 might have thought S4 could have recognized the word if pronounced differently. Alternatively, given S4 and S5's laughter, SX-1's turn might be an inside joke incomprehensible to outside listeners. In line 2286, S4 brings the attention back to the issue with a repetition of the problematic item with a rising intonation. After some hesitation, S1 suggests a possible answer via paraphrasing. However, the answer provided is not correct, creating a situation of erroneous shared knowledge. The conversation on the topic, however, continues with S1 signaling uncertainty about his response and other participants intervening in the conversation, in another example of meaning negotiation. In the end, S1 provides another definition of the word, this time the correct one, ensuring that agreement on the meaning of the word and comprehension of the presentation material are achieved.

While in many cases it is possible to reach shared understanding with an appeal/request and answer sequence, sometimes multiple turns may be necessary to ensure that communication has been effectively achieved for all parties involved. An example may be seen in the extract below: S8 requests confirmation of her understanding of S2's explanation by offering a paraphrase of the element on which she wants confirmation, preceded by a meta-cognitive expression making her request explicit (line 912, "this means…").

(9)	912	S8:	this means that <un> xxx </un> a kind of a set of regulations?
	913	S2:	mhm
	914	S8:	you they have to follow or to apply? =
	915	S2:	= they (.) CAN follow yeah
	916	S8:	if they: wish to have a a: label er [org2] label (.) or not.
	917	S2:	hm:<4> no?</4>
	918	S1:	<4> maybe it's not a </4> set of regulation but <5> rather </5> er (.)
	919	S2:	<5> no?</5>
	920	S1:	a <6> manual </6> of good practice.
	921	SS:	<6><un> xxx </un></6>
	922	S2:	practi- <7> exactly.</7>
	923	S1:	<7> is something </7><8> which is </8> serves well the purpose. =
	[…]		
	936	S7:	sort of recommendations rather than <1> regulations </1>

937 SS: <1> recommendations </1>
938 S1: <2> rather than re</2>gulations
939 S8: <2> recommendation huh?</2>
940 S2: yah (POmtg314)

In this case the first request for confirmation is immediately followed by a second one in line 914. Reacting to S2's "mhm", which may have been interpreted as a sign of lack of comprehension, S8 attempts to clarify her own move by adding to his first question. This seems to be successful, as S2 answers, using sentence stress to mark that the regulations are not compulsory – "they (.) CAN follow" rather than "have to follow". S8 continues with her own interpretation of S2's explanation and appears not to have caught on to S2's message, who then explicitly says S8's understanding was not accurate, and S1 intervenes, providing an alternative definition to S8's proposal. S2 accepts S1's definition ("exactly"), and S1 continues paraphrasing S2's previous explanation to enhance clarity, the explanation interspersed by other speakers' positive backchannels. S7 takes the floor, providing other-initiated word replacement that is then taken up by other participants via repetition. S8 then appears to ask once again for confirmation of the negotiated word 'recommendation', which S2 confirms in the following turn. In this extract, a misunderstanding emerges via a request for confirmation and is solved through the joint work of multiple participants, mutually co-constructing meaning and reaching understanding on a main aspect of their project.

3.4 Other-initiated word replacement and repairs

In addition to checking for other people's comprehension or making requests to ensure that their own understanding is correct, participants in BELF interactions may ensure the success of an interaction by repairing a speaker's turn when the linguistic element employed may be inaccurate or ambiguous, and potentially hindering effective communication and comprehension of important business points, as in the following example:

(10) 1839 S2: maybe yours is a: little bigger than
1840 S1: wider
1841 S4: <8> mhm </8>
1842 S5: <8> mhm </8>
1843 S2: erm <6> yah mhm yah mhm yah </6>

1844 S5: <6> yah </6> this <7> was </7> the problem <1> (because) </1> it was <2> too </2> wide
1845 S2: <7> mhm </7>
1846 S2: <1> yeah </1>
1847 S2: <2> mhm </2>
1848 S2: yeah wide (PBmtg3)

Participants in the passage are discussing the size of a product display item. When S2 suggests the issue is that the product was too big (line 1839), S1 provides an other-initiated word replacement, specifying that the issue was with the width, which is then confirmed by S5 in line 1844. S2 accepts the repair through repetition of the correct form. In this case, while the term employed by S2 was correct in a broad sense, S2 still felt it was necessary to provide a repair as it would be important to know exactly where the problem was to avoid potential further issues with the product. In other cases, a word replacement may not be as important to the interaction as in the instance above. In another event (POmtg315), the use of "work market", despite being understandable, triggers another-initiated word replacement with a speaker suggesting the more common collocation "labor market". The use of a more technical collocation may be preferred in very task-oriented meetings such as this one, as it may later appear in a document shared with other groups and institutions.

3.5 Unattended appeals for assistance

The analysis of the corpus showed that the vast majority of appeals and requests in the dataset were addressed by other participants. However, it was also true that in a small number of instances, such appeals or requests appeared not to be reacted to by any of the other participants in the communicative event. Looking at the individual occurrences, it could be noted that in most of these cases, the lack of uptake could be justified in several ways; it was only in very few instances that signals of non-understanding were simply left unattended.

In some instances, the appeal for assistance was followed by a signal of comprehension, clarifying that intelligibility had not actually been lost. In other cases, the use of the <soft> tag could mean that the would-be recipient(s) did not hear the request. In the extract below another situation is represented, where stopping the flow of the conversation to attend to an appeal might not have been deemed necessary or appropriate by the recipients.

(11) 567 S5: er but then you <6> won't </6> see
568 S3: <6> huh?</6>
569 S3: huh?
570 S5: <7><un> xxx </un></7>
571 S9: <7> this is fine </7> (.) yah
572 S5: that's fine (.) (POmtg315)

In this conversation participants are taking notes on the board. S5 points out that the latest element added is not visible. S3 signals lack of understanding through a non-lexical minimal query ("huh?"). S9 and S5 seem to fix the visibility issue and signal agreement. S3's appeal may have been left unattended because it was only relevant to the writing process and not to the content. In yet other instances, requests for repetition appear to go unattended because the utterances triggering the request do not contain any relevant information. Participants may then decide not to disrupt the flow of the conversation to address the queries, as doing so would not be beneficial or enhance communicative effectiveness.

4. Concluding remarks

The need to come to a mutual agreement and ensure that all parties understand the language used in the same way is certainly not exclusive to ELF interactions, but it may be more sensitive in BELF contexts. Participants are indeed aware that their linguistic levels may be different and that the implications of the same word or expression may be influenced by multiple factors, such as a) individual proficiency in English, b) linguistic and cultural background, or c) business culture. As a result, they similarly appear to be mindful of the potential risks of misunderstanding these conditions entail, and react by acting accordingly, employing requests and confirmation checks to prevent these risks (Mauranen 2006b; Kaur 2009; Björkman 2014). One of the major aspects of BELF, and one of the elements that sets it apart from other uses of ELF, is related to the knowledge of the "shared professional area of expertise involving special concepts and terminology, as used by the relevant discourse community" (Kankaanranta – Planken 2010: 391) and its related vocabulary. This is in line with what an interviewee stated in Kankaanranta and Planken's study, where "sharing the jargon and content" (394) was prioritized over grammatical correctness. As illustrated in 3.4,

"work market" might have been perfectly understandable to all participants, but an other-initiated word replacement is used in favor of a more technical term, which may be considered more appropriate to use in minutes or any documents stemming from the meeting.

In these data, requests are generally attended to by participants, who either repeat, clarify or confirm the marked item, ensuring that all parties are on the same page. Participants tend to show a tendency to employ minimal queries ("sorry?", "huh?", repetition of the word/expression in a rising intonation) when performing their requests, which may make it difficult, not only from a methodological point of view but for the interlocutors themselves, to discern whether the request requires only a confirmation or repetition as an answer, or a clarification of any type. In most cases, speakers appear to provide either a repetition or a confirmation, waiting for additional requests on the part of the asker before giving additional information and further interrupting the flow of the conversation with an explanation sequence. It is possible that the interlocutor may expect a more elaborate request for clarification (e.g. "what do you mean with servers") and interpret this type of minimal query as repetition/confirmation by default. In addition, interlocutors may want to preserve the 'askers'' face by assuming first that the disruption in the conversation is due to a lack of or imperfect hearing rather than a lack of knowledge of the problematic item.

The corpus showed that appeals and requests are rarely left unattended, and even when this occurs, the event can usually be justified: the request might not need to be addressed directly, as the problematic item or turn may not be relevant to the conversation at hand, as shown in example 12. In other examples in the corpus, lack of hearing due to the request being uttered in a soft voice may be behind an unattended request. In a different case, not reported here, a non-verbal gesture like a nod, not transcribed in the turn, may result in an apparent unattended request. On the other hand, comprehension checks appear to go mostly unanswered, with silence being interpreted as confirmation of understanding, or followed by non-lexical vocalizations such as "mhm", "hm".

It may, therefore, be said that the participants in the conversations analyzed here show strong cooperative and face-saving behavior in BELF encounters while they engage actively as listeners in those conversations and respond to each appeal and request, with multiple participants employing as many turns as necessary to reach mutual understanding and achieve their common professional goal.

Acknowledgements

This paper is supported by the PRIN 2015 Prot. 2015REZ4EZ –"English as a Lingua Franca in domain-specific contexts of intercultural communication: a Cognitive-functional Model for the analysis of ELF accommodation strategies in unequal migration contexts, digital-media virtual environments, and multicultural ELF classrooms"

REFERENCES

Sources

Vienna-Oxford International Corpus of English (VOICE)
2013 B. Seidlhofer – S. Majewski – R. Osimk-Teasdale – M.-L. Pitzl – M. Radeka – N. Dorn – L. Lippert (compilers), version POS Online 2.0, http://voice.univie.ac.at/pos/, accessed October 2018

Special studies

Bialystok, E.
1983 "Some factors in the selection and implementation of communication strategies". In: C. Faerch – G. Kasper (eds.) *Strategies in Interlanguage Communication*. London: Longman, 100-118.
Bjørge, A.K.
2010 "Conflict or cooperation: The use of backchanneling in ELF negotiations", *English for Specific Purposes* 29 (3), 191-203.
Björkman, B.
2011 "Pragmatic strategies in English as an academic lingua franca: Ways of achieving communicative effectiveness", *Journal of Pragmatics* 43 (4), 950-964.
2013 *English as an Academic Lingua Franca: An Investigation of Form and Communicative Effectiveness*. Berlin; Boston: De Gruyter Mouton.
2014 "An analysis of polyadic English as a lingua franca (ELF) speech: A communicative strategies framework", *Journal of Pragmatics* 66, 122-138.
Canale, M. – M. Swain
1980 "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics* 1, 1-47.
Cogo, A.
2009 "Accommodating difference in ELF conversations: A study of pragmatic strategies". In: A. Mauranen – E. Ranta (eds.) *English as*

a Lingua Franca: Studies and Findings. Newcastle u. T.: Cambridge Scholars Publishing, 254-273.

2012 "ELF and superdiversity", *Journal of English as a Lingua Franca* 1 (2), 287-313.

2016 "'They all take the risk and make the effort': Intercultural accommodation and multilingualism in a BELF community of practice". In: L. Lopriore – E. Grazzi (eds.) *Intercultural Communication: New Perspectives from ELF*. Rome: Roma Tre-Press, 365-383.

Corder, S.P.

1983 "Strategies of communication". In: C. Faerch – G. Kasper (eds.) *Strategies in Interlanguage Communication*. London: Longman, 15-19.

Corsaro, W.A.

1977 "The clarification request as a feature of adult interactive styles with young children", *Language in Society* 6 (2), 183-207.

Coupland, N. – H. Giles – J.M. Wiemann

1991 *Miscommunication and Problematic Talk*. London: Sage.

Drew, P.

1997 "'Open' class repair initiators in response to sequential sources of troubles in conversation", *Journal of Pragmatics* 28, 69-101.

Dörnyei, Z. – M.L. Scott

1997 "Communication strategies in a second language: Definitions and taxonomies", *Language Learn*ing 47 (1), 173-210.

Ehrenreich, S.

2010 "English as a business lingua franca in a German multinational corporation", *Journal of Business Communication* 47 (4), 408-431.

Faerch, C. – G. Kasper

1983 "Plans and strategies in foreign language communication". In: C. Faerch – G. Kasper (eds.) *Strategies in Interlanguage Communication*. London: Longman, 20-60.

Firth, A.

1996 "The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis", *Journal of Pragmatics* 26 (2), 237-259.

Firth, A. – J. Wagner

1997 "On discourse, communication and (some) fundamental concepts in SLA research", *Modern Language Journal* 81, 285-300.

Foster, P. – A.S. Ohta

2005 "Negotiation for meaning and peer assistance in second language classrooms", *Applied Linguistics* 26 (3), 402-430.

Franceschi, V.

2017 "Plurilingual resources as an asset in ELF business interactions", *Journal of English as a Lingua Franca* 6 (1), 57-81.

Jamshidnejad, A.
2011 "Functional approach to communication strategies: An analysis of language learners' performance in interactional discourse", *Journal of Pragmatics* 43, 3757-3769.
Kankaanranta, A. – L. Louhiala-Salminen
2010 "'English? – Oh, it's just work!': A study of BELF users' perceptions", *English for Specific Purposes* 29 (3), 204-209.
2013 "'What language does global business speak?' – The concept and development of BELF", *Ibérica* 26, 17-34.
Kankaanranta, A. – B. Planken
2010 "BELF competence as business knowledge of internationally operating business professionals", *Journal of Business Communication* 47 (4), 380-407.
Kaur, J.
2009 *English as a Lingua Franca. Co-constructing Understanding*. Saarbrücken: VDM-Verlag Müller.
2010 "Achieving mutual understanding in World Englishes", *World Englishes* 29 (2), 192-208.
2011 "Raising explicitness through self-repair in English as a lingua franca", *Journal of Pragmatics* 43 (11), 2704-2715.
Kirkpatrick, A.
2007 "The communicative strategies of ASEAN speakers of English as a lingua franca". In: D. Prescott (ed.) *English in Southeast Asia: Varieties, Literacies and Literatures*. Newcastle u. T.: Cambridge Scholars Publishing, 118-137.
2008 "English as the official working language of the Association of Southeast Asian Nations (ASEAN): Features and strategies", *English Today* 24 (2), 27-34.
Lichtkoppler, J.
2007 "'Male. Male.' – 'Male?' – 'The sex is male.' The role of repetition in English as a lingua franca conversations", *Vienna English Working Papers* 16 (1), 39-65.
Louhiala-Salminen, L. – A. Kankaanranta
2011 "Professional communication in a global business context: The notion of Global Communicative Competence", *IEEE Transactions on Professional Communication* 54 (3), 244-262.
Mauranen, A.
2006a "A rich domain of ELF – the ELFA corpus of academic discourse", *Nordic Journal of English Stud*ies 5 (2), 145-159.
2006b "Signaling and preventing misunderstanding in English as lingua franca communication", *International Journal of the Sociology of Language* 177, 123-150.
2012 *Exploring ELF: Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press.

Meierkord, C.
2000 "Interpreting successful lingua franca interaction. An analysis of non-native-/non-native small talk conversations in English", *Linguistik Online* 5 (1), https://bop.unibe.ch/linguistik-online/article/view/1013/1673

Ohta, A.S.
2005 "Confirmation checks: A discourse analytic reanalysis", *Japanese Language and Literature* 39, 383-412.

Palmer-Silveira, J.C.
2013 "The need for successful communication in intercultural and international business settings: Analytic and comparative studies, new trends and methodologies", *Ibérica* 26, 9-16.

Pitzl, M.-L.
2010 *English as a Lingua Franca in International Business: Resolving Miscommunication and Reaching Shared Understanding*. Saarbrücken: VDM-Verlag Müller.

Savignon, S.
1997 *Communicative Competence. Theory and Practice*. New York: McGraw-Hill.

Selinker, L.
1972 "Interlanguage", *IRAL* 10, 219-231.

Tarone, E.
1977 "Conscious communication strategies in interlanguage: A progress report". In: H. Douglas Brown – C.A. Yorio – R.H. Crymes (eds.) *On TESOL'77: Teaching and Learning English as a Second Language: Trends in Research and Practice: Selected Papers from the Eleventh Annual Convention of Teachers of English to Speakers of Other Languages, Miami, Florida, April 26-May 1, 1977.* Washington D.C.: TESOL, 194-203.
1980 "Communication strategies, foreigner talk, and repair in interlanguage", *Language Learning* 30 (2), 417-431.

Vettorel, P.
2019 "Communication strategies and co-construction of meaning in ELF communication: Drawing on multilingual resources pools", *Journal of English as a Lingua Franca* 8 (2), 179-210.

VOICE Project
2007 "Mark-up conventions". VOICE Transcription Conventions [2.1]. http://www.univie.ac.at/voice/documents/VOICE_mark-up_conventions_v2-1.pdf, accessed October 2018

Wagner, J. – A. Firth
1997 "Communication strategies at work". In: G. Kasper – E. Kellerman (eds.) *Communication Strategies: Psycholinguistic and Sociolinguistic Perspectives.* London: Longman, 323-344.

Widdowson, H.
2003 *Defining Issues in English Language Teaching*. Oxford: Oxford University Press.

APPENDIX 1

Selected VOICE transcriptions conventions

SX-f	unidentified speaker, identified gender
<L1ger></L1ger>	tag signaling that a word is produced in the speaker's L1. The language is specified.
<LNfre></LNfre>	tag signaling that a word is produced in an LN. The language is specified.
?	rising intonation
e:r	lengthened sound
(.)	brief pause
(1)	longer pause, timed in seconds
<pvc><pvc>	variation from native norms in terms of phonology, morphology or lexis. May indicate creations that do not exist in ENL.
<ipa></ipa>	phonetic transcription in case of significant variation in pronunciation
@	laughter
<1> </1>	overlapping speech (because) uncertain transcription
<un>xxx</un>	unintelligible speech
{words}	translation of non-English speech and contextual events
CAPS	emphasis of a syllable or a word
=	continuation/completion of speech by another speaker without pauses
<soft> </soft>	soft voice
<fast></fast>	fast pronunciation
<spel></spel>	words spelled out by the speaker
[org23]	anonymization of organization names
[S4] / [first name 2]	anonymization of participants / non-participants
di-	the hyphen indicates that a word is not produced in its entirety.

Address: VALERIA FRANCESCHI, Department of Foreign Languages and Literatures, University of Verona, Lungadige Porta Vittoria, 41 37129 Verona, Italy.
ORCID code: orcid.org/0000-0003-0224-049X.

Legal English as a lingua franca in academia: The strategic use of repetitions in lectures

Patrizia Anesa
University of Bergamo

ABSTRACT

This paper investigates the functions of repetitions in ELF academic discourse, exploring the complex nature of their realizations. The data presented originate from a corpus of Law lectures in which English is used as a lingua franca. Special attention is given to the identification of the specificities and the functions of repetitions which may typify these events from a disciplinary perspective.

The main hypothesis is that the realization of repetitions in an event which lies at the crossroads between legal and academic discourse is influenced by the lecturers' cultural allegiance to the use of normative and semantic repetition. Data show that repetition displays a tendency, on the one hand, to enhance formal correctness, and, on the other hand, to favor intelligibility and mutual understanding. The discussion in this paper contributes to an ongoing debate on the extent of the use of English in English-medium tertiary-education settings at a global level, and more specifically, on the use of specific pragmatic devices in this context.

Keywords: legal English, legal English as a lingua franca, EAP, Law lectures, repetition.

1. Introduction

Given the widespread use of English in academia across fields and countries, studies focusing on the use of English as a lingua franca (ELF)[1] in tertiary-education settings represent a particularly vibrant field for investigation (e.g. Björkman 2011; Jenkins 2014; Kaur 2011; Mauranen 2016; Murata 2018).

[1] Here ELF is intended as a "site of complex language contact" (Mauranen 2018: 107), which involves language users rather than language learners (Mauranen 2018: 113).

DOI: 10.25951/4367.

This work focuses on lectures, which epitomize a central site of knowledge-disseminating practice. In particular, Law lectures constitute the privileged locus of analysis. Indeed, the usage of ELF in academia is becoming increasingly important even in Law studies, in which the adoption of a lingua franca has traditionally been hindered by the fact that the specificities of a legal system or culture are generally considered inseparable from the usage of a given national language.

Within ELF studies focusing on academic settings, the investigation of pragmatic devices has received considerable attention. In particular, research has confirmed that ELF users in academic contexts show a high pragmatic competence and employ a variety of strategies, such as repetition, to achieve specific objectives, namely signaling misunderstanding (e.g. repetition of problematic items) or preventing misunderstanding (e.g. confirmation checks, interactive repair, and self-repair) (Cogo 2009; Cogo – Dewey 2006; Mauranen 2006; Taguchi – Ishihara 2018). In this respect, both other- and self-repetition represent proactive tools which contribute to improving understanding and mutual intelligibility in ELF (Cogo 2009; Kennedy 2017; Lichtkoppler 2007).

The use of repetition in ELF academic interaction varies considerably in relation to the type of communicative situation. For instance, in conversations between faculty members and students, Lichtkoppler (2007) identifies different categories related to the functions of repetition, including time-gaining, utterance-developing, prominence-providing, ensuring accuracy of understanding, showing listenership, cohesion[2], and borrowing. In the case of lectures, however, self-repetition generally predominates. For instance, self-repair aiming at rephrasing content, wording, or grammar, is employed in order to secure comprehension (Mauranen 2006) or to increase explicitness (Kaur 2011).

Clearly, repetitions may also have a negative impact on the fulfilment of given communicative objectives. Some of these aspects are addressed in Relevance Theory and, within this framework, repetitive, redundant utterances are believed to go against the principle of optimal relevance, since such language features require extra processing effort (Sperber – Wilson

[2] For a discussion of repetitions and their role in establishing cohesive relations in the text see de Beaugrande – Dressler (1981) and Halliday – Hasan (1976), with particular reference to the concept of reiteration. In particular, reiteration represents a broad form of cohesion which involves "the repetition of a lexical item, or the occurrence of a synonym of some kind, in the context of reference; that is, where the two occurrences have the same referent" (Halliday – Hasan 1976: 318-319).

1986). Another drawback of repetition is pointed out by de Beaugrande and Dressler (1981), who claim that excessive use of lexical recurrence might lower the informativity of the text.

This study focuses on the functions of repetition, by observing its usage in a circumscribed context, i.e. Law lectures, in which ELF is employed. More specifically, this analysis aims to gain a finer understanding of how repetition works in ELF Law lectures by answering the following research questions: 1) What types of repetition are identifiable? 2) What purpose do they serve in the Law lectures under investigation? 3) Does the specificity of the discipline affect the way repetitions are employed?

Thus, the present work attempts to verify whether the pragmatic functions that repetition assumes in ELF lectures *tout court* also emerge in ELF Law lectures or whether disciplinary specificities become manifest. It is plausible to assume that repetition is a common pragmatic device in lectures regardless of the type of discipline involved. However, in the field of Law, repetition ontologically plays a crucial role in that it is a constitutive element of the law itself. In this respect, we can reason with Philippopoulos-Mihalopoulos (2011: 49) that "only through law's obsessive normative repeating can justice spring forth". Indeed, normative repetition is not simply a form of recollection, but is an immanent and instrumental feature of the law, one often deemed necessary in order for the law to be faultless and unambiguous in its verbal realizations. Nevertheless, it should not be forgotten that the process of normative repetition also implies that every repetition entails variation, a form of transcendence from the original item. Thus, it is not mere duplication, but can have specific legal value as it is placed within specific legal practices.

Starting from this theoretical premise, it can be hypothesized that lecturers discussing the law use repetition in a way which is influenced by the awareness of the necessity of repetition as a constitutive element of the law. Thus, normative repetition may translate into a semiotic and semantic repetition which is employed in the discussion of the law itself.

2. Legal English in academia

Traditionally, legal language has been divided into subcategories including different genres. For instance, Kurzon distinguishes between 'language of the law' i.e. the language used "in documents that lay down the law", and 'legal language', i.e. the language "used when people talk about the law",

but with a specific usage, as happens in the case of judges' opinions, legal textbooks, or lawyers' speeches in court (Kurzon 1997: 120). In this interpretation, the notion of legal language seems to be a prerogative of the legal profession. Conversely, Trosborg's (1995) well-known categorization includes five subcategories: language of the law; language of the courtroom; language in textbooks; lawyers' speech; and people talking about the law. Thus, Trosborg paved the way for a more inclusive conceptualization of legal language, given that situations in which people talk about the law (and so meaning professionals, but also other potential participants) are taken into account.

This work focuses on the use of Legal English in the academic context. Particular attention is devoted to Law lectures, which may be seen as belonging to a particular subcategory of 'people talking about the law'. The field of analysis is even more specific in that the emphasis is on those communicative events where English assumes the contours of a lingua franca, in that speakers and audience are predominantly non-native users of the language.

Considerable research has been carried out on the features of English-medium lectures. Insights have been offered, *inter alia*, on cultural diversity, focusing for instance on genre-related issues (e.g. Thompson 1994), cultural matters (e.g. Flowerdew – Miller 1995), interaction dynamics (e.g. Bligh 2000), and pragmatic considerations (e.g. Björkman 2011; Crawford Camiciottoli 2004; Molino 2015). As early as 1967, Garfield stated that English should be employed by researchers as the language of academia because of its accessibility (Garfield 1967). Undeniably, the appropriateness of ELF in the academic sphere has been subject to lively debate between those who see it as a fundamental tool for international communication and those who fear homologation and lack of diversity.[3] Without entering into the debate on the role of English-medium lectures in tertiary education, the objective of this analysis is to focus on a specific type of academic ELF event. This will be done by adopting an approach which can help us to reflect on the complexity of the processes of knowledge dissemination and negotiation of meaning in occurrences which lie at the crossroads between the academic and legal worlds.

[3] For a discussion of the ideological implications of the spread of English as the language of academia, and a reflection of its potential effects in terms of marginalization or obliteration of existing differences, see Gotti (2012).

3. Research framework

The material for analysis is based on fifteen Law lectures given in English. All communicative events were transcribed broadly drawing on MICASE conventions; however, in the inevitable trade-off between accuracy and readability, the latter aspect was privileged, given that prosodic features are not investigated in this work. The transcripts have been coded by two coders[4] using *QDA Miner Lite* (Provalis Research 2011) in order to conduct a qualitative analysis of the main discursive devices emerging. Among the different codes identified, repetition (and its subcodes) will be described in the analytical section.

The choice to include Law lectures exclusively is to allow a focus on one specific academic field. Although it is plausible to assume that similar pragmatic features emerge in lectures dealing with other disciplines, the homogeneity of the corpus is deemed necessary in order to gain a finer understanding of the specificities of a sociolect in a given context. In this respect, it should be kept in mind that different disciplines imply "a certain degree of interdisciplinary diversity and a degree of intradisciplinary homogeneity" (Hyland 2000: 10), and the notion of "disciplinary culture" (cf. Bondi 2005: 6) is the result of a dialectic process which develops across disciplinary variation.

The corpus under investigation consists of lectures dealing with contract law, corporate crimes, ethics, and arbitration. All the events took place between 2012 and 2016 in Italy, Malaysia, and Brazil. This choice derives from practical reasons related to the accessibility of the material, as well as from the desire to include samples of lectures from different continents, in line with the cross-national nature of legal English in academia, which is postulated in this paper.

Some participants are present in more than one recording and the nationalities involved are Italian, Malaysian, Indian, French, Spanish, German, and Brazilian. Table 1 summarizes the corpus details.

This analysis focuses on an area where academic and legal communication intersect, in that it is based on the observation of lectures given in the area of Law, primarily delivered to postgraduate students and legal training experts. Most of the speakers recorded have an advanced level of English, although the extensive range in levels of proficiency available in

[4] The final level of inter-coder agreement was 85%. In the case of discrepancies, a third coder was consulted.

ELF studies may be substantial and is one of its intrinsic features. Indeed, ELF is inherently characterized by enormous diversity (Mauranen 2007, 2016), even within specific subject fields such as Law.

Table 1. Corpus details

Lecture code	Number of speakers involved	Approximate number of participants	Approximate number of words	Length (minutes)
L1	1	90	5365	35
L2	1	100	4025	31
L3	1	100	4128	28
L4	1	100	7016	50
L5	1	100	3984	27
L6	1	50	10045	45
L7	2	20	13615	65
L8	8	100	14150	60
L9	7	110	12613	51
L10	4	125	12398	45
L11	1	82	6840	23
L12	1	84	8252	38
L13	1	50	9088	48
L14	3	76	9539	33
L15	5	24	8076	44

The style of the lectures investigated is heterogeneous and varies from formal to more "conversational" (Dudley-Evans 1994), and thus from purely monologic to more dialogic forms. In this respect, we can reason with Bamford and Bondi (2005: IX) that "[a]lthough all texts are interactive, some are more interactive than others". The majority of the events analyzed show a predominance of a monologic style, and only in few cases is the audience actively involved in the interaction. Generally, lecturers lead the conversation, but they may elicit answers or comments from the audience, and in the corpus under investigation this happens in six lectures (out of the fifteen). All the same, in this paper, interaction is not seen exclusively as the presence of dialogic elements but, more broadly, as a process displaying devices which favor the negotiation of meaning (see Crawford Camiciottoli 2004; Sinclair – Coulthard 1992), regardless of the distribution of turns.

4. Interactional dynamics in ELF Law lectures

Pragmatic strategies are essential to negotiate meaning in interaction (De Bartolo 2014: 453) and in ELF they can enhance comprehension and intelligibility (Cogo – Dewey 2012), and the co-construction of meaning.

Among the several elements observed within an interactional framework we can find repetitions, reformulations, clarifications, comprehension checks, and clarification requests, all of which can be adopted in order to improve communication success and to signal cooperation and involvement (cf. Mauranen 2006). Obviously, different interactional devices are used in combination. Without attempting to describe all possible functions related to these linguistic practices, the aim is to focus on one recurrent phenomenon, that of repetition, which is investigated as an effective linguistic resource which contributes to communicative success in ELF. Thus, it is a tool which may be seen to function as an enhancer of all levels of understanding.[5]

4.1 Forms of repetition

Different classifications of repetition have been offered, and Table 2 presents an overview of the main categories according to the following criteria: speaker, exactness, time-lapse, and intentionality.

As regards speakers, repetition is classified as same- or other-speaker repetition (Schegloff 1996: 177). While other-speaker repetition as a strategy for negotiating meaning has often been investigated (Cogo 2009; Mauranen 2006; Watterson 2008), monologic repetition has received less attention

Table 2. Types of repetition

Criterion	Type
Speaker	Same Other
Exactness	Exact Quasi-exact Reformulation
Time-lapse	Immediate Delayed
Intentionality	Spontaneous Deliberate

[5] The multifaceted idea of understanding, which comprises understanding accent, propositional content, and pragmatic sense, is to some extent aligned with the components of Smith's model of understanding, namely intelligibility, comprehensibility, and interpretability. According to Smith's seminal work, which developed within the field of cross-cultural communication, intelligibility relates merely to the recognition of words and utterances, comprehensibility involves the understanding of the meaning of words and utterances, and interpretability goes further and includes the implied and pragmatic meaning (Smith 1992: 88).

(as in this case processes of co-construction of meaning are less evidently verbalized) and it represents the privileged (but not exclusive) object of this study.

The level of exactness will also be taken into account. In this work, following a well-established tradition, reformulation is treated as a form of repetition displaying lower exactness (e.g. Johnstone et al. 1994). In this respect, reformulation, or paraphrasing, inherently implies the repetition of a certain concept, however using a different linguistic formulation. It thus differs from repetition in terms of structure, although it may be seen as being placed at the end of a continuum ranging from exact repetition to quasi-exact repetition to reformulation, which may potentially be intended as a semantic form of repetition.

Repetition has been described in relation to the time which elapses between the production of an element and the repeated one. The focus here is predominantly on immediate repetition, although the delayed type is also investigated. Repetition is also distinguished by spontaneous/unintentional vs deliberate forms (Biber et al. 1999). The former includes unintentional repeats and reformulations, while the latter is mainly based on rhetorical devices, for instance with the aim to increase clarity or involvement.

4.2 Functions of repetition

Studies on repetition have pointed out its ability to improve communication (Kaur 2012), its spontaneity (Tannen 1987), and its necessity within the conversational flow (Johnstone 1987). Repetition represents a commonly-used rhetorical device which mainly fulfils emphatic purposes or is an organizing tool. The functions of repetitions are diverse (as illustrated in seminal studies such as Norrick 1987 and Schegloff 1996) and vary in case of self- or other-repetition, as illustrated in Table 3 (broadly adapted from Norrick 1987).

The functions lying behind the use of repetitions are innumerable (see e.g. Tannen 1987), both in the case of native speakers (NSs) and non-native speakers (NNSs), and it should be pointed out that this list is not exhaustive. Additionally, the same repetition device can clearly fulfil a variety of objectives simultaneously.

One objective is to enhance understanding as repetition favors "semantically less dense discourse" (Tannen 1987: 582). Along the same lines, Lichtkoppler (2007) also stresses that repetition can favor mutual understanding among participants. The function related to the improvement of correctness is also applicable to ELF interactions, especially in the case

of other-repetition repair (see e.g. Lichtkoppler 2007: 59). Repetition is often noticeable in sequences where understanding needs to be enhanced, checked, or confirmed. The co-construction of understanding can take place both in self- and other-repetition: in the first case the speaker assumes that the repetition device can facilitate the recipient's understanding, while in other-repetition the interlocutor can use it to check understanding or prompt the repetition he/she may need. In this respect, as Kaur notes, "[i]t is perhaps easier to ascribe the function of promoting recipient understanding to dialogic repetition given the negotiated nature of the phenomenon" (Kaur 2012: 597). Indeed, the development of the exchange can facilitate the identification of the function assumed by a repetition device. As will be seen (in Section 5), this function plays an important role in the events analyzed.

Table 3. Functions of repetition

Self-repetition	Other-repetition
hold the floor; gain planning time; bridge interruption; enhance textual coherence; facilitate understanding; improve correctness; …	show attention; show participation; signal receipt; signal agreement/disagreement; initiate repair; check understanding; improve correctness; …

Other-repetition can often express alignment and solidarity. However, self-repetition can also be an implicit form of solidarity, especially if used in order to facilitate comprehension on the part of the interlocutor. Another function of repetition in conversation is the guaranteeing of a certain rhythm between turns from a prosodic perspective. However, in the case of lectures, the assigning of turns is not spontaneous in that it is the main speaker who generally leads the conversation and decides the intervention of other interlocutors. As shall be seen, this function is not manifest in the cases analyzed.

5. Repetition in lectures

In this study the main object of analysis is represented by those segments which include the repetition of one or more elements. The focus is on exact repetition and repetition with little variation (no more than two words

reformulated within a segment). These include parallel phrasing, key word repetition and repetition of lexical bundles, as well as repetition of repaired segments.

5.1 Self-repetition or other-repetition

As mentioned above, repetition can assume the form of self- or other-repetition. In particular, self-initiated repetition is important in pre-empting potential communicative problems. Lectures generally present self-repetition, given their predominantly monologic nature, but cases of other-repetition are also present when the lecturers are particularly active in eliciting interaction, as illustrated in the case below:

(1) S1: Are you saying that taxing Jordan say at three percent tax rate for good causes to feed the hungry *is theft*?
S3: I think it's unjust, yes I do believe *it's theft*, but perhaps it is necessary to condone that theft.
S1: But *it's theft*. [S3: Yes]. [SS LAUGH] Why *is it theft* (DEIDENTIFIED)? <AUDIO DISTURBANCE>
S3: Because *It's theft* because ehm, at least in my opinion and by the libertarian opinion he earned that money fairly ehm, and it belongs to him and so to take it from him is by definition theft.

The term *theft* identifies a specific crime within a definite legal apparatus. The discussion in this case focuses on its definition and thus no speaker deviates from the use of the same term. The presence of lexical repetitions is in line with the preservation of the established terminology and the degree of precision that could be lost with the implementation of other linguistic choices. It also allows compliance with the requirements for precision and unambiguity.

Both self-repetition and other-repetition can be immediate and delayed (see Table 2). In the following excerpt, delayed repetition is used within the logical reasoning to demonstrate a thesis syllogistically:

(2) So taxation, actually, *is morally equivalent to forced labor* because forced labor, involves the taking of my leisure, my time, my efforts... just as taxation takes the earnings that I make with my labor... and so, for Nozick and for the libertarians taxation for redistribution is theft as (DEIDENTIFIED) says. But not only theft it is morally equivalent

to laying claim to certain hours of a person's life... and labor so: *it's morally equivalent to forced labor*.

The moral equivalence between taxation and forced labor is expressed at the beginning of the paragraph, and then reiterated at the end of the logical argumentation, in order to reinforce the validity of the statement.

5.2 Parallel phrasing

Parallel phrasing is defined by Norrick (1987: 254) as a form of repetition that is typically encountered in listing elements. The parallelism makes use of not-exact repetition, where one or more elements vary within the same construction.

(3) *every person*, *every market*, *every industry*, *every organization* has a back story

(4) *There's a thing called the Rome Convention that* talks about private international law as you all call it here, we call it conflict of laws in the United States. But *there's this thing called the Rome Convention that* says that wuh wuh what law do I recognize as governing law and is enforceable?

(5) *There was a concept of arbitration*, *there was a concept of mediation*, *there was a concept of conciliation* that existed in most of the civilized world, by the way what was the civilized world, a thousand years ago?

The replication of a specific pattern is often employed rhetorically in academic events. Parallel phrasing is generally used for emphatic purposes, while also favoring explicitness and understanding. Indeed, it allows the speech to be less dense from an informational perspective and also attempts to overcome the difficulties inherent in the explanation of complex legal concepts.

5.3 Keyword repetition and repetition of lexical bundles

Exact or quasi-exact repetition of keywords, especially in the case of specific legal terms, emerges frequently in the corpus analyzed, as exemplified by the passage below:

(6) And, *Sharia* to the extent not inconsistent therewith there are different ways of drafting this. But something else that's very useful because the courts are unclear themselves as to what *Sharia* means. *Sharia* to, and you know this, *Sharia* to a Malaysian is a little bit different than *Sharia* to a Saudi. It's a little bit different than *Sharia* to a Pakistani, mean this is my own experience. As a professor in the United States once said, we should use plural, *Sharias*. He's overstating it of course. He says there are five *Sharias*.

The substitution of the term *Sharia* with a synonym appears impossible, given the conceptual importance of this keyword within the lecture. Similarly, the use of a pronoun may impede clarity. Indeed, the law is often characterized by lexical precision and a Law lecture cannot abstain from the preservation of this feature. The following excerpt also shows the need to express this form of accuracy through the repetition of specific lexical bundles and phrases:

(7) if the laws of Malaysia are the proper law to be applied by the arbitral tribunal, then the reference to *the Sharia Advisory Council or Sharia expert* is a must. Also not expressively articulated in article eleven para one the arbitral tribunal is well-advised to consult with and seek the assistance of the parties and their council in formulating the questions and issues to be referred to *the Sharia Advisory Council or the Sharia expert*. The same applies to the preparation of the reference including the relevant facts to the extent that they are not in disputes or have been agreed by the parties as stipulated in rule eleven para three. Rule eleven para two speak of a decision of *the Sharia Advisory Council or the Sharia expert*.

The use of the phrase *Sharia Advisory Council or Sharia expert* emerges recurrently, and the same lexical bundles are used without any modification, as even minor changes could lead to conceptual inaccuracies. Although the repetition of the same segment, and in particular the multiple repetition of the word *Sharia*, may appear redundant and could potentially be considered implicit, it is instead explicitly verbalized in all cases, in order to conform with the need for clarity and precision.

Legal vocabulary is traditionally described as stable and inflexible, and such features are clearly visible in 'the language of law'. The genre of a Law lecture, being inherently hybrid, potentially presents a higher level of flexibility and dynamicity from a lexical perspective. However, it displays

a tendency to preserve technical precision and accuracy. In this respect, the reiteration of a specific item is often merely a natural process but, at the same time, it can contribute to enhancing clarity, as happens in the case below:

(8) […] its enforcement under *the New York Convention*. *The New York Convention* is the most successful convention.

Synonymy varies in its susceptibility to alteration; however, as a tendency, technical terms are generally repeated rather than paraphrased. Indeed, repetition may assume the form of a formulaic device which forestalls possible ambiguity. Moreover, the juxtaposition of the same expression can help the listener to follow the logical pattern envisaged by the lecturer.

5.4 Repaired repetition

Repaired repetition is defined as repetition with a modification, which is found immediately after the pronunciation of the first segment. Similarly, the "reformulation repetition" is described by Murata (1995: 353) as a tool typical of NNSs which is used to offer a more accurate utterance. In the locutionary process the speaker tries to find the exact word, thus partially repeating certain segments:

(9) the Jewish in terms of religion has long tradition than than Islam in terms of legal and uh uh uh *regulation framework regulatory framework* for for for their community

A segment is repeated in another form devoid of deviation from those patterns which are deemed correct, or in order to guarantee exactness, which is generally considered as an indispensable requirement in legal discourse.

6. Conclusions

The aim of this paper was to observe the multifaceted role of repetition in an event which is an expression of legal and academic discourse and, more specifically, to describe how it is employed in ELF Law lectures by adopting a qualitative approach. Consequently, from a theoretical perspective, this study also attempted to frame the Law lecture not only as a locus of analysis which represents an instantiation of academic discourse but, drawing on

the concept of interdiscursivity, as a genre which lies at the interconnection between the academic and the legal world and is based on their mutual hybridization in terms of discursive practices.

A communicative event such as the lecture locates participants at the heart of knowledge dissemination practices. Even in the case of monologic events, the constitutive relationship between the lecture and the audience is an integral element of the event, which is interactional by its nature. This analysis suggests that repetition represents an interactional device which enhances the construction of negotiated meaning, and is used especially when dealing with specifically legal concepts. Moreover, repetition is a characteristic of legal language and assumes a series of functions, such as guaranteeing lexical precision or improving clarity, which may be reflected in the usage of repeated elements.

More specifically, strategies such as parallel phrasing, keyword repetition, and repaired repetition can be usefully employed to emphasize the significance of a certain lexical element, or to pre-empt potential misunderstanding. The repetitions identified in these events are generally verbatim, and this choice contributes to preserving lexical accuracy and precision. Keyword repetition, in particular, often refers to precise legal concepts whose meaning is explicitly contextualized within a certain legal system. Parallel phrasing, instead, usually regards less specific legal terms, and the parallelism serves to improve clarity, as well as for rhetorical purposes, and to produce a certain prosodic rhythm. The repetition of a repaired segment may be used to introduce a concept and specify the exact term when lexical precision is needed, pursuant to the desire to provide accurate speech.

Repetition in ELF Law lectures is in line with Mauranen's (2006: 146) idea that ELF is characterized by "considerable effort invested in preventing misunderstanding". Especially in other-repetition, the final aim is not only to signal non-understanding but rather to confirm understanding in order to favor the continuation of the discussion. It is a multiple dialogic repetition which also allows the negotiation of meaning. As Schegloff (1997: 527) points out, "many repeats are not used for repair, but to implement other actions, and particular, desirable ones". Therefore, the smooth running of the conversation, and the involvement of the interlocutors, are some of the objectives that repetition fulfils. However, in the events analyzed, self-repetition is predominant and seems to indicate a tendency towards formal correctness rather than efficiency.

Irrespective of the academic discipline, repetition is a common feature of lectures at large, be it for informative, educational, or persuasive purposes. However, its role in Law lectures in which English is used as the lingua franca seems to be deserving of particular attention. Indeed, when the construction of meaning happens across languages, legal systems, and professional cultures, the sapient use of discursive devices drawn from a user's pragmatic repertoire, of which repetition constitutes only one example, may prove essential to improve understanding and guarantee accuracy.

Given the variability and the contextuality that characterize the use of English, and in order to limit the risk of oversimplification, the analysis was limited to a specific context of usage, the Law lecture. Further avenues for research could adopt a more contrastive stance in order to gain a deeper understanding of the disciplinary divergences as regards the usage of pragmatic devices in lectures.

From a broader perspective, we may argue that, fundamentally, the law may be international, but it is not universal, and neither is the use of English as a legal lingua franca in academia, which needs constant problematization. Thus, even in the case of internationally recognized practices, the complexities of specific pragmatic dynamics cannot be ignored, and should also be analyzed from a cross-cultural perspective. In this respect, quantitative analyses could also be conducted in order to explore variations in the use of pragmatic devices across countries.

On a final note, it should be pointed out that the impact of the discussion on the nature of English on teaching practice has largely been marginal in legal academic endeavors, but its conceptualization as an international language calls for a reflection on its implication in teaching and training. Most materials focus on Legal English by adopting an Anglocentric standpoint, often with exclusive reference to legal systems used in countries such as the US, or England and Wales, while only a few available teaching resources are slightly more in line with the nature of English as an international language (see Campos 2010). Moving beyond the idea of Legal English as the language of Anglophone countries, and adopting a more encompassing analytical lens, the investigation of settings in which Legal English is used a lingua franca can contribute to a more nuanced understanding of the function that English assumes in legal communication in different contexts on a global scale.

REFERENCES

Sources

Data set of Law lectures (Italy, Malaysia, Brazil. 2012-2016). Collected by P. Anesa.

Special studies

Bamford, J. – M. Bondi
2005 "Introduction". In: J. Bamford – M. Bondi (eds.) *Dialogue within Discourse Communities: Metadiscursive Perspectives on Academic Genres*. Tübingen: De Gruyter, VII-XXIV.

Bondi, M.
2005 "Metadiscursive practices in academic discourse: Variation across genres and disciplines". In: J. Bamford – M. Bondi (eds.) *Dialogue within Discourse Communities: Metadiscursive Perspectives on Academic Genres*. Tübingen: De Gruyter, 3-30.

Biber, D. et al.
1999 *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Björkman, B.
2011 "Pragmatic strategies in English as an academic lingua franca: Ways of achieving communicative effectiveness?", *Journal of Pragmatics* 43, 950-964.

Bligh, D.
2000 *What's the Use of Lectures?* San Francisco: Jossey-Bass.

Campos, M.Á.
2010 "Going beyond the obvious in English for legal purposes: A few remarks on international legal English as a lingua franca in Europe". In: L.Á. López – R. Crespo Jiménez (eds.) *Professional English in the European Context: The EHEA Challenge*. Bern: Peter Lang, 175-194.

Cogo, A.
2009 "Accommodating difference in ELF conversations: A study of pragmatic strategies". In: A. Mauranen – E. Ranta (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle: Cambridge Scholars, 254-273.

Cogo, A. – M. Dewey
2006 "Efficiency in ELF communication: From pragmatic motives to lexico-grammatical innovations". *Nordic Journal of English Studies* 5 (2), 59-93.
2012 *Analysing English as a Lingua Franca. A Corpus-driven Investigation*. London: Continuum International Publishing Group.

Crawford Camiciottoli, B.
2004 "Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study", *Journal of English for Academic Purposes* 3 (1), 39-54.
De Bartolo, A.M.
2014 "Pragmatic strategies and negotiation of meaning in ELF talk", *EL.LE* 3 (3), 453-464.
de Beaugrande, R. – W. Dressler
1981 *Introduction to Text Linguistics.* London: Longman.
Dudley-Evans, T.
1994 "Variation of the discourse patterns favoured by different disciplines and their pedagogical implications". In: J. Flowerdew (ed.) *Academic Listening*. Cambridge: Cambridge University Press, 146-158.
Flowerdew, J. – L. Miller
1995 "On the notion of culture in L2 lectures", *TESOL Quarterly* 29 (2), 345-373.
Garfield, E.
1967 "English – an international language for science?", *Current Contents* 26, 19-20.
Gotti, M.
2012 "Cross-cultural aspects of academic discourse", *Brno Studies in English* 38 (2), 59-78.
Halliday, M.A.K. – R. Hasan
1976 *Cohesion in English.* London: Longman.
Hyland, K.
2000 *Disciplinary Discourses. Social Interactions in Academic Writing*. London: Longman.
Jenkins, J.
2014 *English as a Lingua Franca in the International University. The Politics of Academic English Language Policy*. Abingdon: Routledge.
Johnstone, B.
1987 "An introduction", *Text* 7 (3), 205-214.
Johnstone, B. et al.
1994 "Repetition in discourse: A dialogue". In: B. Johnstone (ed.) *Repetition in Discourse: Interdisciplinary Perspectives*. Norwood, NJ: Ablex, 1-20.
Kaur, J.
2011 "Raising explicitness through self-repair in English as a lingua franca", *Journal of Pragmatics* 43 (11), 2704-2715.
2012 "Saying it again: Enhancing clarity in English as a lingua franca (ELF) talk through self-repetition", *Text & Talk* 32 (5), 593-613.
Kennedy, S.
2017 "Using stimulated recall to explore the use of communication strategies in English lingua franca interactions", *Journal of English as a Lingua Franca* 6 (1), 1-27.

Kurzon, D.
1997 "Legal language: Varieties, genres, registers, discourses", *International Journal of Applied Linguistics* 7 (2), 119-139.
Lichtkoppler, J.
2007 "'Male. Male.' – 'Male?' – 'The sex is male.' The role of repetition in English as a lingua franca conversations", *Vienna English Working Papers* 16 (1), 39-65.
Mauranen, A.
2006 "Signaling and preventing misunderstanding in English as lingua franca communication", *International Journal of the Sociology of Language* 177, 123-150.
2007 "Hybrid voices: English as the lingua franca of academics". In: K. Flottum (ed.) *Language and Discipline Perspectives on Academic Discourse*. Newcastle: Cambridge Scholars, 243-259.
2016 "What is going on in academic ELF? Findings and implications". In: P. Vettorel (ed.) *New Frontiers in Teaching and Learning English*. Newcastle: Cambridge Scholars, 31-54.
2018 "Second language acquisition, world Englishes, and English as a lingua franca (ELF)", *World Englishes* 37 (1), 106-119.
Molino, A.
2015 "Checking comprehension in English-medium lectures in technical and scientific fields", *The European English Messenger* 24 (1), 22-29.
Murata, K.
1995 "Repetitions: A cross-cultural study", *World Englishes* 14 (3), 343-356.
Murata, K. (ed.)
2018 *English-Medium Instruction from an English as a Lingua Franca Perspective. Exploring the Higher Education Context*. Abingdon: Routledge.
Norrick, N.
1987 "Functions of repetition in conversation", *Text* 7, 245-264.
Philippopoulos-Mihalopoulos, A.
2011 "Repetition or the awnings of justice". In: O. Ben-Dor (ed.) *Law and Art: Justice, Ethics and Aesthetics.* New York: Routledge, 45-58.
Provalis Research
2011 *QDA Miner Lite: Qualitative and Mixed-Method Software*, https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/
Schegloff, E.
1996 "Confirming allusions: Toward an empirical account of action", *American Journal of Sociology* 102, 161-216.
1997 "Practices and actions: Boundary cases of other-initiated repair", *Discourse Processes* 23, 499-545.

Sinclair, J. – M. Coulthard
1992 "Towards an analysis of discourse". In: M. Coulthard (ed.) *Advances in Spoken Discourse Analysis*. London: Routledge, 1-34.
Smith, L.
1992 "Spread of English and issues of intelligibility". In: B.B. Kachru (ed.) *The Other Tongue* (2nd edn). Urbana: University of Illinois Press, 75-90.
Sperber, D. – D. Wilson
1986 *Relevance: Communication and Cognition.* Oxford: Basil Blackwell.
Taguchi, N. – N. Ishihara
2018 "The pragmatics of English as a lingua franca: Research and pedagogy in the era of globalization", *Annual Review of Applied Linguistics* 37, 80-101.
Tannen, D.
1987 "Repetition in conversation: Toward a poetics of talk", *Language* 63 (3), 574-605.
Thompson, S.
1994 "Frameworks and contexts: A genre-based approach to analysing lecture introductions", *English for Specific Purposes* 13 (2), 171-186.
Trosborg, A.
1995 "Introduction", *Journal of Pragmatics* 23 (1): Special Issue on "Laying down the law – Discourse analysis of legal institutions", 1-5.
Watterson, M.
2008 "Repair of non-understanding in English in international communication", *World Englishes* 27 (3-4), 378-406.

Address: Patrizia Anesa, Room 206, University of Bergamo, Piazza Rosate 2, Bergamo 24129, Italy.
ORCID code: orcid.org/0000-0002-0795-745X.

Stefan Dollinger, *Creating Canadian English*: *The Professor, the Mountaineer, and a National Variety of English*, Cambridge: Cambridge University Press, 2019a, xviii + 283 pp.

(Reviewed by Carol Percy, University of Toronto, Canada)

The first chapter of *Creating Canadian English* presents distinctive perspectives on the monograph's concept and on the chapter's title: "What is Canadian English?" Dollinger acknowledges English as one of at least 263 other languages spoken in Canada, at least 60 of which are Indigenous. He epitomizes the effect of English on First Nations languages by showing how colonial toponyms like "Chatham and Discovery Islands" obliterate Indigenous heritage and cultural-linguistic knowledge. Immigrant languages other than English appear in a chronological chart of immigration waves. Canadian and other English dialects feature in more detail in the contexts of settlement history and new dialect formation theory. And we meet American- and British-born scholars among the "Big Six" who edited the *Dictionary of Canadianisms on Historical Principles* (*DCHP*), one of four dictionaries published for Gage between 1962 and 1967, Canada's Centennial year. Their perspectives and editing arguably "created" Canadian English. Dollinger's book focuses almost solely on these texts and these men. But this introductory perspective on the origins of Canadian English is the first of many episodes likely to engage readers who are intrigued by the intellectual and the logistical challenges of codifying colonial varieties of English.

Dollinger's monograph officially begins in 1940, when Queen's University professor Henry Alexander mentioned Canadian English amid chapters on British and American English in *The Story of our Language*. With Alexander as its anchor, the well-researched second chapter on "The Heritage of Canadian English" features American projects to interconnect the men who codified Canadian English for Gage. From 1946, the American autodidact Charles Lovell distinguished Canadian vocabulary from the Americanisms he accumulated for the University of Chicago Press. The title's "mountaineer", Lovell combined his love of hiking and lexicography with first-hand fieldwork on words like *hobo*. Lovell died in 1960, just after

DOI: 10.25951/4368.

securing Canada Council funding for a full year of work on the *DCHP*. The *DCHP* would be steered by the "professor": Walter Avis received his PhD from the University of Michigan after studying at Queen's with Alexander. Canadianizing the spelling of an American dictionary, conducting early research, and training scholars like Avis, Alexander is Dollinger's "godfather" of Canadian English.

Lexicography produces dictionaries and heart failure: Avis's 1979 death opens chapter 3, "Avis Pulls It Off". Dollinger reconstructs the background for Avis's training in the 1950s: the *Middle English Dictionary* at Ann Arbor and (in nearby Wisconsin) the future *Dictionary of American Regional English*. And he illustrates Avis's employment at the Royal Military College, with pan-Canadian undergraduate dialects and excessive administration. Emerging through the story is the Canadian Linguistic Association (1954) and its Lexicographical Committee (1957), and an eventual agreement by Gage (1959) to fund the *DCHP* as well as three-tiered school dictionaries. Dollinger draws on obituaries, family calendars, and especially letters between Lovell and Avis to map the realities of long-distance dictionary composition. He perceptively interprets events like the first-name basis between Charlie and Wally (1955) and the activities of other editors, volunteers, and family members, including women. Like Matthew H. Scargill, the founding chairman of the lexicographical committee and the coordinator of the volunteer reader programme (1957-59), some left less evidence of their labour.

Perspectives and methods in compiling the *DCHP* dominate chapter 4, "The 'Technology': Slips, Slips, and More Slips". Duplications, errors, and omissions on *quotation* or *citation slips* (Canadians say both) underscore the difficulties of coordinating reading programs and classifying Canadianisms. A late addition was the word *canoe* – not unique to Canada but certainly Canadian. The *DCHP-1* editors' assumptions about belonging and lexicography pervade records for potential entries like *Canuck*, *Indian residential school*, *Ojibwa*, and *Stikeen* – the latter First Nation names that were ultimately excluded. Their spellings here and in the records are the product of standardization – itself a settler conception. And the absence of *Anishnabe* reflects how slowly settler Canadians recognized Indigenous groups' preferred terms. As the *DCHP-2*'s editor-in-chief, Dollinger deploys his analytical skills: even non-*Canuck*s should appreciate his reasoning whereby Hawaiian *kanaka* 'man, person' became the ultimate self-designator of a Canadian, via the east coast of North America and whaling ships.

The marketing of regional dictionaries features in chapter 5, "1967 – Excitement and Hype". The *DCHP* is sketched alongside its fellow pioneering dictionary of a regional English, Cassidy and Le Page's *Dictionary of Jamaican English* (1967), and compared to perpetually edited dictionaries of dominant varieties like the *Oxford English Dictionary*, sucking money but providing data for the formerly profitable genre of the desk dictionary. Royalty statements and estimates of postage and advertisements underscore the challenges of producing new titles for a small market amid American and British competition. The rise of the Gage series of school dictionaries extends the story well past 1967. Dollinger contextualizes the 1983 rebranding of the former Senior dictionary as the *Gage Canadian Dictionary*, which dominated the market until 1998. The challenge of codifying non-dominant varieties is epitomized by determining criteria for Canadianisms: as the editor of *DCHP-2* (2017), Dollinger's six criteria included cultural significance (such as *eh*) distinguished from negative legacy (such as *residential school*).

Chapter 6, "Riding the Wave of Success", perversely concerns the declining interest in dictionaries by academic linguists and the lack of revision of the *DCHP* by its surviving editors. A 1973 edition was abridged but not updated with newly topical words like *eh*. A trade book was published but soon forgotten. Other projects were never finished or published, for personal or political reasons: Dollinger relentlessly exhumes abandoned manuscripts – one for a nearly-complete bilingual dictionary, aborted before the 1980 Quebec referendum on separatism. Dollinger sees the death of Avis as "the last nail in the coffin of a revision project" (2019a: 145) and "a serious setback for an academic field that had just found its first firm footing" (2019a: 67). Instead, variationist and corpus linguistics diverted the next generation of professional academics from the vocabulary and standards that typically interest the public. Canadian vocabulary was nevertheless collected by the.*Oxford English Dictionary*: between 1968 and 1978 Douglas Leechman was paid by editor Robert Burchfield for his contributions to the *OED* supplement.

The supposed Americanization of Canadian English along with the entry of Oxford UP into the Canadian market sets the scene for chapter 7, "A Global Village and a National Dictionary War". Dollinger explains the implications of the American and British genealogies of the Canadian dictionaries published respectively by Gage (1997), ITP Nelson (1997), and Oxford (1998). Readers interested in Canadian pronunciation will find some information here: perceptions of the Canadian-American relationship are indexed by pronunciations of words like *news* and *Iraq*. But the bulk of the

chapter contextualizes the appearance of the *Canadian Oxford Dictionary* (*COD*) in the rise of corpus lexicography and the history of the *OED*, and elaborates on the challenge of codifying and marketing lexical Canadianisms. To what extent should encyclopedic and regional terms be included in a national desk dictionary? For Dollinger, the success of the *COD* and its second edition (2004) reflects the outreach of its editor and the prestige of England in a former colony. The later failure of all three dictionaries reflects not simply the internet but specifically Oxford UP's retreat from this small market after monopolizing it.

The problems involved in codifying changing social attitudes concerns chapter 8, "Decolonizing *DCHP-1* and *DCHP-2*". The existence of countries like Canada reflects settlers' oppression and extermination of original inhabitants. The *DCHP-1* itself was (as Dollinger observes) an anti-colonial project in its codification of distinctively local lexical norms. But the existence of *residential schools* and the testimony of its survivors betray the Canadian state's sustained erasure of Indigenous languages and cultures. Further research in the citation files suggests that it was only in late 1967 that settler journalists and thus lexicographers might be expected to recognize residential schools as sites of cultural genocide. For numerous headwords including *go Indian* and *good Indian*, quotations and definitions emphasize negative attitudes, unrecognized with usage labels. As the editor of *DCHP-2*, Dollinger retained but flagged such evidence of the negative legacy of *DCHP-1* and consulted widely when writing new entries – and this book. And although proper names were not added to the *DCHP-2*, in the last chapter we learn that the third most common semantic domain is the "Aboriginal".

Readers of *Creating Canadian English* will anticipate its author's answer to the question posed by the final chapter: "Is There Really a Canadian English?" Readers of media should wonder how much we can conclude from journalists' repeated rediscoveries of *Canadian English* – and whether the decline in references to that phrase might reflect the ongoing closure of newspapers. And the opinion of the public (even university students) about whether Canadian English is distinctive likely reflects sociopolitical ideologies rather than linguistic realities. Academic linguists were among those *In Search of the Standard in Canadian English*, edited by W.C. Lougheed in 1986, though in chapter six Dollinger identified only J.K. (Jack) Chambers as a member of the new generation's "variationist camp". But to anyone with internet access, the existence of the *DCHP-2* (as of this monograph) will prove the distinctiveness and the development of Canadianisms.

Creating Canadian English is an outstanding work of research. It brings to light and interprets much unpublished material. Touching on pronunciation and spelling, it focuses mostly on vocabulary – a subject of particular interest to the general public. Its interconnected stories describe the craft of lexicography and the scholarly lives of the editors whose dedication resulted in the inevitably imperfect codification of ever-changing Canadian English. Keen and occasionally confused readers will want an even richer bibliography and index – an index that includes figures like Jaan Lilles, who argued against the existence of Canadian English, and all substantial references to culturally significant words like *eh* or tricky ones like *Indian*. But with its multiple theoretical perspectives, *Creating Canadian English* will also appeal to non-Canadian scholars of fields like World Englishes, English linguistics, European lexicography, and "pluricentric" languages more generally – another area of expertise for Dollinger (2019b).

REFERENCES

Avis, W.S. (chief editor), C. Crate, P. Drysdale, D. Leechman, M.H. Scargill, C.J. Lovell (eds.)
1967 *A Dictionary of Canadianisms on Historical Principles*. Toronto: W.J. Gage.

Dollinger, S.
2019b *The Pluricentricity Debate: On Austrian German and Other Germanic Standard Varieties*. New York: Routledge.

Dollinger, S. (chief editor) and M. Fee (ed.)
2017 *The Dictionary of Canadianisms on Historical Principles, Second Edition*. With the assistance of B. Ford, A. Gaylie and G. Lim. Vancouver: University of British Columbia. www.dchp.ca/dchp2/

Lougheed, W.C. (ed.)
1986 *In Search of the Standard in Canadian English*. Strathy Language Unit Occasional Papers, 1. Kingston, ON: Queen's University. https://qspace.library.queensu.ca/handle/1974/23743.

Address: Carol Percy, Department of English, University of Toronto, St. George Campus, 170 St. George Street, Toronto, Ontario, Canada.
ORCID code: orcid.org/0000-0002-9871-1745.