



**‘Jaysus, keep talking like that and
you’ll fit right in’**

**An Investigation of Oral Irish English
in Contemporary Irish Fiction**

Ana María Terrazas-Calero

A THESIS SUBMITTED TO MARY IMMACULATE COLLEGE
(UNIVERSITY OF LIMERICK)
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Supervisors

Dr. Anne O’Keeffe and Dr. Carolina P. Amador-Moreno

(Submitted to Mary Immaculate College, University of Limerick, May, 2022)

DEDICATION

In loving memory of my late grandmother, Cesárea Hidalgo Cáceres.

You always asked with all kinds of expletives when I would finish the thesis.

It took time, and I wish you were here to see it, but I hope I made you proud because *¡¡por fin
he acabado la p*** tesis, abuela!!*

DECLARATION

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

Signed:

A handwritten signature in black ink, appearing to read 'Ana María Terrazas-Calero', written over a light grey rectangular background.

Ana María Terrazas-Calero

ABSTRACT

This project is an interdisciplinary and comparative investigation of the reproduction of linguistic features of Irish English (IrE) present in *contemporary* IrE fiction. To do this, a corpus of over 1 million words comprising 16 works of fiction published in the Republic of Ireland by 8 authors was compiled: the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE).

The goal of this thesis, therefore, is to determine 1) which are the most frequently reproduced features of IrE orality in contemporary IrE fiction, 1a) how realistic is their fictional portrayal when contrasted against real spoken uses, 2) what does the use of the most frequently reproduced features in the corpus encode with regard to speaker identity, and 3) in what manner may modern Irishness be encoded through the reproduction of pragmatic items in fiction. Utilizing a variety of interdisciplinary methodologies, including Corpus Stylistics, Corpus Linguistics, Sociolinguistic, and Pragmatic techniques, the thesis identifies signature linguistic features that are thought to be representative of IrE in the corpus via quantitative and qualitative, comparative corpus analysis. To evaluate the level of realism inherent in the fictional rendition, the findings are contrasted against the *Limerick Corpus of Irish English* and the *BNC2014*. A second corpus comprising books by one of the CoFIrE authors, i.e. Paul Howard, was also compiled. Thus, the *Ross O'Carroll-Kelly Corpus* (CoROCK) was created given this series' reputation for being a chronicler of modern Ireland and because of the high frequency of IrE orality reproduction these books were found to contribute to CoFIrE. Two case studies on non-standard, non-traditionally IrE high frequency intensifiers are conducted on CoROCK to better answer the research questions regarding the potential indexation of modern Irishness through speech reproduction in fiction. Finally, by evaluating the type of speaker identity these features may index when used in *contemporary* fiction, this thesis determines the type of modern Irishness that appears to be encoded through fictional speech representations.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr. Anne O’Keeffe and Dr. Carolina P. Amador-Moreno for their incredible support throughout all these years. I do not think I would have embarked on this PhD project if it had not been for Dr. Carolina P. Amador-Moreno. Not only was this your original idea but you have for many years been a role model to me. Muchas gracias por tu continuo apoyo y confianza. No sabes cómo te lo agradezco. I also have no words to express how much of an inspirational source Dr. Anne O’Keeffe has also been, taking me under her wing at Mary Immaculate College, always being a phone-call away in case I needed support. I value both of you very much both personally and professionally. Thank you!

Thank you to Mary Immaculate College for awarding me the *Postgraduate Studentship Award* which allowed me to begin my doctoral studies in Ireland. My thanks also to the *Irish Research Council* without whose three years of funding through their *Postgraduate Government of Ireland Scholarship* I would not have been able to conduct this investigation and disseminate my work through international conferences.

My gratitude goes out to Dr Joan O’Sullivan, at Mary Immaculate College, for her invaluable friendship and for her trust in me and my teaching abilities, always asking me to participate in the MA in Applied Linguistics classes. You will never know how much I appreciate that! Special thanks also to the Spanish Department at the University of Limerick where I was able to hone my foreign language teaching skills for several years.

I would like to thank all the authors who agreed to have their books included in my corpus. Special thanks goes out to Paul Howard for not only agreeing to have his novels added, but for sending as many books as I needed, for accepting to participate in a public reading hosted at Mary Immaculate College, and for basically putting up with me every time I wish to add to the *Ross O’Carroll-Kelly Corpus*!

My most personal “thank yous” go out to Ana M^a Martín Sánchez. My soul sister. No te haces una idea de lo que tu amistad significa para mí y de cuánto me has ayudado todos estos años. Thank you to Mariano Paz for taking a chance on me, for educating me on ‘five-star-rated’ cinematic pieces, y por todos los recuerdos preciosos que no olvidaré. My gratitude to Teresa Cuixeres as well. You’re the most caring, most loving person I have ever met. Gracias por todo el apoyo que me diste y que continúas dándome. ¡Eres la mejor! Thank you as well to John O’Grady for your friendship and support throughout all these years. To Roman for always helping me no matter what and for introducing me to Russian traditions. Thank you to Nancy Ávila, whose positivity and constant support always bright up my day. ¡Eres una heroína!

The biggest ‘thank you’ goes to my father, Antonio Julián, for his unwavering support. For all the rides to and from the airport when I was headed out or returning from a conference. For putting up with me, especially in these last few years. For simply having my back no matter what and for teaching me to never give up. Eres el mejor padre del mundo. Finally, I would also like to dedicate this thesis to my mother who, I am sure, would have supported me in everything she could.

LIST OF FIGURES

Chapter 3

- Figure 3.1. Diagram of the *top-bottom* approach used to annotate CoFIrE.
Figure 3.2. Example of a search for tag <ML> across CoFIrE using *Wordsmith Tools* (Scott, 2012).
Figure 3.3. Screen capture of *Table Books* in CoFIrE MS Access database.
Figure 3.4. First 15 items on *Table Features* in CoFIrE MS Access database.
Figure 3.5. Screen capture of *Table Characters* in CoFIrE MS Access database.

Chapter 4

- Figure 4.1. Percentage of annotated items per text in CoFIrE.
Figure 4.2. Most recurrent pragmatic items/occurrences in CoFIrE.

Chapter 5

- Figure 5.1. Raw occurrences of GO per *tense/aspect* in CoFIrE.
Figure 5.2. GOES (pro)nominal colligates in CoFIrE.
Figure 5.3. Raw number of occurrences of *pronominal spectrum+GO* colligational patterns in CoFIrE.
Figure 5.4. Discourse-pragmatic range of GO in CoFIrE.
Figure 5.5. GO distribution of occurrence for *age, class, and location* in CoFIrE normalized by 100,000 words.
Figure 5.6. Raw distribution of GO by *tense* across the 100-item random *LCIE* and *BNC2014* samples.
Figure 5.7. Raw distribution by *subject pronoun+GOES* colligational pattern in randomized *LCIE/BNC2014* samples.
Figure 5.8. Distribution across *grammatical subject+GO* patterns in CoFIrE and *LCIE/BNC2014* samples with data normalized to 100.
Figure 5.9. Distribution of GO per function with CoFIrE and *LCIE/BNC2014* data normalized to x100 words.
Figure 5.10. Raw distribution of GO by social class in 100-item *LCIE/BNC2014* samples.
Figure 5.11. Age and gender distribution of BE LIKE in *LCIE/BNC2014* samples normalized to x100.
Figure 5.13. Distribution of BE LIKE by grammatical person across CoFIrE and *LCIE/BNC2014* randomized samples with data normalized by 100 words.
Figure 5.13. Distribution of BE LIKE by 3rd-person subject pronouns across CoFIrE and *LCIE/BNC2014* randomized samples with data normalized by 100 words.
Figure 5.14. Raw functional distribution of BE LIKE across CoFIrE.

Chapter 6

- Figure 6.1. Sample image of the Excel file illustrating the classification of intensifying *fucking* in CoFIrE.
Figure 6.2. Distribution of InF by gender in *LCIE/BNC2014* 100-random sample compared with CoFIrE (normalized to 100 words).
Figure 6.3. Distribution of InF across age cohorts in *LCIE/BNC2014* (100 random samples) with CoFIrE's data normalized to 100 words.
Figure 6.4. Distribution of InF occurrences by social class in CoFIrE.

Figure 6.5. Distribution of InF by social class across CoFIrE (data normalized to 100 words) and *LCIE/BNC2014* samples following NRSSG occupation-based classification outlined in section (3.8.2).

Figure 6.6. Distribution of orthographic renditions of InF by county in CoFIrE.

Figure 6.7. Distribution of InF by colligational pattern preference in *LCIE/BNC2014* 100 random item samples, with CoFIrE data normalized to 100 words.

Figure 6.8. Functional distribution of InF across CoFIrE and *LCIE/BNC2014* randomize samples.

Figure 6.9. Gender distribution of InF across CoFIrE and *LCIE/BNC2014* randomized samples.

Figure 6.10. InF (Negative) emotion breakdown across CoFIrE and the *LCIE/BNC2014* randomized samples.

Chapter 7

Figure 7.1. Brinton's (2017, 9) updated list of DPM features based on Aijmer & Simon-Vandenberg's (2011) categories.

Figure 7.2. Sample image of the Excel file illustrating the classification of clause-medial *likes* in CoFIrE.

Figure 7.3. Raw functional distribution of *like* across the *LCIE/BNC2014* randomized samples.

Figure 7.4. Age distribution of medial *like* across CoFIrE and *LCIE/BNC2014* samples with all data normalized to by 100 words.

Figure 7.5. Class distribution of medial *like* normalized by x100 words across CoFIrE and *LCIE/BNC2014* samples.

Figure 7.6. Raw regional distribution of medial *like* in CoFIrE.

Figure 7.7. Contrastive pragmatic function distribution of medial *like* across CoFIrE and *LCIE/BNC2014* samples normalized by 100 words.

Figure 7.8. Pragmatic distribution of medial *like* across CoFIrE and *LCIE/BNC2014* samples normalized to by 100 words.

Figure 7.9. Pragmatic functionality of focuser medial *like+high rising terminal* pattern in CoFIrE.

Chapter 8

Figure 8.1. Screen capture of frequency list run on CoROCK-SO with *Wordsmith Tools* (Scott, 2012).

Figure 8.2. Screen capture of a concordance line search for SO in CoROCK-SO.

Figure 8.3. Screen capture of MS Excel documenting the classification of occurrences of Non-lexical *totally* in CoROCK-T.

Figure 8.4. Distribution of NISo occurrences in CoROCK-SO by year of publication.

Figure 8.5. Colligational pattern distribution of NISo across CoROCK-SO by year of publication.

Figure 8.6. Distribution of NISo+Noun colligational pattern variation in CoROCK-SO by year of publication.

Figure 8.7. Longitudinal analysis of NISO orthographic prosodic representation across CoROCK-SO by year of publication.

Figure 8.8. Distribution of *NISo+Intensifier Boosters* distribution across CoROCK-SO by year of publication.

Figure 8.9. Gender distribution of occurrences of *NISo+Oh my God* pattern across CoROCK texts by year of publication.

Figure 8.10. Distribution of the conveyance of *positive/negative* connotations through NISo across CoROCK-SO by year of publication.

Figure 8.11. Raw distribution of *positive/negative* emotion catalogue conveyed through NISo in CoROCK-SO.

Figure 8.12. Longitudinal distribution of NLT occurrences across CoROCK-T texts by year of publication.

Figure 8.13. Distribution of NLT prosodic representation in CoROCK-T by year of publication.

Figures 8.14. and 8.15. Raw distribution of NISo and NLT by gender across CoROCK-SO and CoROCK-T respectively.

LIST OF TABLES

Chapter 2

Table 2.1. Broad research topics and subtopics in CS studies of fictional IrE orality in fiction.

Table 2.2. Summary of (mostly CL-based) studies of fictional IrE portrayal dealing with subtopics (b) and (c).

Chapter 3

Table 3.1. Restrictive parameters in CoFIRE.

Table 3.2. Academic sources consulted to create CoFIRE's *sampling frame*.

Table 3.3. CoFIRE authors, texts, book codes, and approximate word count/book.

Table 3.4. Word count averages/book in CoFIRE.

Table 3.5. Summary of annotation system developed for and implemented in CoFIRE.

Table 3.6. Pragmatic sub-tags applied to CoFIRE texts.

Table 3.7. Illustration of the social stratification system and occupations in CoFIRE.

Table 3.8. National Readership Survey's social grade occupation-based classification in the UK.

Table 3.9. CoROCK books, year of publication, book codes, and approximate word counts.

Chapter 4

Table 4.1. Annotated items and frequency of production in CoFIRE.

Table 4.2. Number of characters and frequency of item occurrence by social class in CoFIRE.

Table 4.3. Annotated item occurrence raw distribution and normalized per hundred words (PHW) within the *upper-class* in CoFIRE.

Chapter 5

Table 5.1. Raw non-standard quotative occurrence count in CoFIRE.

Table 5.2. Percentage of GO and BE LIKE production per CoFIRE author.

Table 5.3. Distribution of GO by *lexical form* in CoFIRE.

Table 5.4. Raw and percentage distribution of subject pronoun + *goes* colligation in CoFIRE.

Table 5.5. Gender distribution of GO in CoFIRE raw and normalized by x100 words.

Table 5.6. Distribution of GO by *lexical form* across CoFIRE (by %) and raw 100-item randomized *LCIE/BNC2014* samples.

Table 5.7. Raw age and gender distribution of GO across *LCIE/BNC2014* samples.

Table 5.8. Distribution of BE LIKE by form and occurrences in *LCIE* sample.

Table 5.9. Contrastive tense distribution of BE LIKE normalized to x100 across CoFIRE and *LCIE/NC2014* randomized samples.

Chapter 6

Table 6.1. Distribution and percentage of InF occurrences in CoFIRE and *LCIE/BNC2014* randomized samples.

Table 6.2. Gender distribution of InF in CoFIRE.

Table 6.3. Distribution of phonetically represented variants of InF in CoFIRE.

Table 6.4. Cases of InF in disagreement with Mackenzie's 'predictions' with checkmarks indicating their appearance in CoFIRE and/or *LCIE/BNC2014* randomized samples.

Table 6.5. Distribution of *Standard* and *Non-Standard* InF *Tmesis* across CoFIRE and 100-item *LCIE/BNC2014* samples, all normalized to 100 words.

Table 6.6. Classification of InF *Tmesis* across CoFIRE and the *LCIE/BNC2014* samples.

Table 6.7. Percentage of distribution of InF by gender in CoFIRE.

Table 6.8. Pragmatic functions of InF as used by narrator *Ross O'Carroll-Kelly* in CoFIRE.

Table 6.9. *Negative* emotion breakdown and distribution of InF when produced by narrator *Ross O'Carroll-Kelly* in CoFIRE.

Chapter 7

Table 7.1. Contrastive clausal-position preference of *like* across CoFIrE and *LCIE/BNC2014* samples with all data normalized to by 100 words.

Table 7.2. Raw gender distribution of medial *like* across CoFIrE and *LCIE/BNC2014* samples.

Table 7.3. Raw distribution of Southside-produced clause-medial *like* by author in CoFIrE.

Table 7.4. Corresponding attributes between the American/Irish English stereotypes.

Table 7.5. Qualitative classification of all occurrences of clause-medial *like* in CoFIrE.

Table 7.6. Raw distribution of the focuser medial *like+high rising terminal* pattern by age/gender in CoFIrE

Chapter 8

Table 8.1. Rank order of SO and NISo occurrences across CoROCK-SO.

Table 8.2. NISo contexts of occurrence in CoROCK-SO. All emphases in the original.

Table 8.3. NISo+Noun Phrase pattern variation in CoROCK-SO.

Table 8.4. Distribution of NISo+intensifying booster items across the CoROCK-SO texts.

Table 8.5. Raw distribution of NISo by pragmatic connotation in CoROCK-SO.

Table 8.6. Raw distribution of occurrences of NLT in CoROCK-T.

Table 8.7. Non-gradable NLT colligational pattern preference across CoROCK-T.

Table 8.8. Anaphoric *standalone* NLT functions across CoROCK-T.

Table 8.9. Raw NLT colligational pattern distribution across CoROCK-T texts by year of publication.

Table 8.10. NLT+Intensifier Boosters distribution across CoROCK-T texts.

Table 8.11. Distribution of NLT and NLT+intensifying booster items occurrences across the CoROCK-T texts by year of publication.

Table 8.12. Distribution of NLT connotation conveyance in CoROCK-T.

Table 8.13. NLT Connotation conveyance distribution across CoROCK-T books.

Table 8.14. NLT pragmatic functionality and distribution of occurrences in CoROCK-T.

Table 8.15. Distribution of occurrences by NLT *Negative Speaker Emphasis / Negative Emphatic Agreement* emotion conveyance in CoROCK-T.

Table 8.16. Distribution of NISo & NLT occurrences by age and gender across CoROCK.

8.17. Raw distribution of occurrences of NISo & NLT by gender across CoROCK-SO and CoROCK-T.

Chapter 9

Table 9.1. Associated Valley-Girl items present in the *Southside Socialite-Rugby Jock/D4 Head* repertoires in CoFIrE and CoROCK.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	v
LIST OF TABLES	vii
1. FICTIONALIZING LANGUAGE TO CONVEY NATURAL ORALITY AND IDENTITY	1
1.1. Fictional (literary) dialects as mirrors of natural orality	1
1.2. Research Questions	4
1.3. Locating the Study	5
1.4. Rationale for this study	9
1.4.1. Why focus on pragmatic IrE elements?	16
1.5. Overview of chapters	20
2. CORPUS STYLISTICS AND THE STUDY OF FICTIONAL (LITERARY) DIALECTS	23
2.1. Introduction to Corpus Stylistics	23
2.2. How to do Corpus Stylistics?: <i>Corpora</i> in Corpus Stylistics	25
2.3. Criteria for designing and building corpora	28
2.3.1. Representativeness	28
2.3.2. Sampling	30
2.3.3. Balance	32
2.3.4. Corpus analytical tools in Corpus Stylistics	34
2.4. Corpus Stylistics Studies in Fiction	35
2.4.1. Corpus Stylistics Approaches	36
2.4.2. Approaches within <i>inter-textual</i> , synchronic CS analyses of fictional (literary) dialect-dialogue	40
2.5. Studying Irish English in Fiction	43
2.5.1. CS Investigations of Fictional IrE: Research Trends & Existing Studies	45
2.6. Concluding Remarks	52
3. METHODOLOGY	55
3.1. Rationale for creating the <i>Corpus of Contemporary Fictionalized Irish English</i>	55
3.2. Designing the <i>Corpus of Fictionalized Irish English</i>	56
3.2.1. CoFIrE design criteria: Establishing the ‘variety’	57
3.2.2. CoFIrE design criteria: Other parameters	59
3.2.3. Developing CoFIrE’s sampling frame and sampling technique and limitations of CoFIrE	63
3.3. The <i>Corpus of Fictionalized Irish English</i>	68
3.4. CoFIrE authors and books	72
3.4.1. Kevin Barry	72

3.4.2.	Claire Keegan	73
3.4.3.	Éilís Ní Dhuibhne	73
3.4.4.	Nuala Ní Chonchúir	75
3.4.5.	Roddy Doyle	76
3.4.6.	Donal Ruane	77
3.4.7.	Donal Ryan	78
3.4.8.	Paul Howard	79
3.5.	CoFIrE Analytic Process: Text preparation	81
3.6.	Data Annotation: Designing the CoFIrE Annotation System	82
3.6.1.	CoFIrE's Annotation System	86
3.7.	Processing and managing the CoFIrE Metadata	90
3.8.	Corpora used to compare and contrast the CoFIrE fictional data	96
3.8.1.	Method of comparison/contrast with LCIE and BNC2014	98
3.8.2.	Sociolinguistic cross examination	99
3.9.	The <i>Ross O'Carroll-Kelly Corpus</i> and subcorpora	100
3.10.	Concluding remarks	102
4.	OVERVIEW OF GENERAL COFIRE FINDINGS	104
4.1.	Summary of features and frequency of occurrence in CoFIrE	104
4.2.	Pragmatic prominence in CoFIrE	106
4.3.	Top pragmatic items in CoFIrE	109
4.4.	Annotated item production by social class in CoFIrE	110
4.5.	Concluding remarks	112
5.	NON-STANDARD QUOTATIVE <i>GO</i>	114
5.1.	Non-standard quoting devices	114
5.2.	CoFIrE's Non-Standard Quotative Repertoire	117
5.3.	<i>GO</i> in the literature	120
5.3.1.	<i>GO</i> in IrE	123
5.4.	Methodology of analysis	124
5.5.	Intralinguistic variables of <i>GO</i> in CoFIrE: <i>Tense</i>	125
5.5.1.	<i>Lexical form</i> and <i>grammatical subject</i> of <i>GO</i> in CoFIrE	127
5.6.	Pragmatic Functions of <i>GO</i> in CoFIrE	132
5.6.1.	Pragmatic Subfunctions of <i>GO</i> in CoFIrE	133
5.7.	The Social Markings of <i>GO</i> in CoFIrE	136
5.8.	Contrasting the CoFIrE data against the reference corpora	139
5.8.1.	<i>Tense/Aspect</i> contrast between CoFIrE and <i>LCIE/BNC2014</i> samples	139
5.8.2.	<i>Lexical form</i> and <i>grammatical subject</i> in CoFIrE and <i>LCIE/BNC2014</i> samples	141
5.8.3.	Pragmatic Functions across CoFIrE and the contrastive samples	146
5.8.4.	Social markings of <i>GO</i> across CoFIrE and contrastive samples	149
5.9.	The remaining CoFIrE quotative repertoire: <i>BE LIKE</i>	151
5.9.1.	The social markings of <i>BE LIKE</i> in CoFIrE	153
5.9.2.	Social markings of <i>BE LIKE</i> in <i>LCIE</i> and <i>BNC2014</i> samples	153
5.9.3.	<i>BE LIKE Tense</i> in CoFIrE and contrastive samples	155
5.9.4.	<i>Grammatical person</i> + <i>BE LIKE</i> across CoFIrE and contrastive samples	156

5.9.5.	Pragmatic uses of BE LIKE in CoFIrE and in contrastive samples	158
5.10.	Concluding remarks	160
6.	TABOO LANGUAGE: <i>FUCKING</i>	162
6.1.	The pragmatic functionality of taboo language	162
6.1.1.	The multifunctionality of taboo language in fictional dialogue	165
6.2.	Intensifying <i>Fucking</i>	166
6.2.1.	Intensifying <i>fucking</i> in the literature	169
6.2.2.	Intensifying <i>fucking</i> in spoken and fictionalized IrE	170
6.3.	Method of analysis	171
6.4.	Intensifying <i>Fucking</i> in CoFIrE	174
6.5.	The sociolinguistics of Intensifying <i>Fucking</i> in CoFIrE	175
6.5.1.	Gender and Intensifying <i>Fucking</i> across CoFIrE in contrastive samples	176
6.5.2.	Age and Intensifying <i>Fucking</i> across CoFIrE and contrastive samples	178
6.5.3.	Social class and Intensifying <i>Fucking</i> across CoFIrE and contrastive samples	179
6.5.4.	Geographic location indexation through Intensifying <i>Fucking</i> in CoFIrE	181
6.6.	Intensifying <i>Fucking</i> Colligational Pattern Preference across CoFIrE and contrastive samples	185
6.6.1.	<i>Tmesis</i> Pattern Preference across CoFIrE and contrastive samples	190
6.7.	The pragmatics of Intensifying <i>Fucking</i> across CoFIrE and contrastive samples	191
6.8.	Intensifying <i>Fucking</i> and the homodiegetic narrator in CoFIrE: a <i>Ross O'Carroll-Kelly</i> case study	196
6.9.	Concluding remarks	198
7.	DISCOURSE PRAGMATIC MARKER <i>LIKE</i>	201
7.1.	<i>Like</i> in the literature	204
7.1.1.	<i>Like</i> in IrE	207
7.2.	Methodology of analysis	209
7.3.	Frequency of occurrence of <i>like</i> in CoFIrE and contrastive samples	210
7.4.	Clausal positional preference of <i>like</i> in CoFIrE and contrastive samples	211
7.5.	Sociolinguistic distribution of clause-medial <i>like</i> in CoFIrE and contrastive samples	213
7.5.1.	Clause-medial <i>like</i> identarian associations in American English extrapolated to Irish English?	218
7.6.	Pragmatic functions of clause-medial <i>like</i> in CoFIrE and contrastive samples	222
7.6.1.	The sociopragmatics of focuser medial <i>like</i> in CoFIrE	226
7.7.	Concluding remarks	230
8.	'New' <i>over-the-top</i> intensifiers in the <i>Ross O'Carroll-Kelly</i> corpus	233
8.1.	The selection of New Intensifying <i>So</i> and Non-lexical intensifier <i>Totally</i>	233
8.2.	Stylizing North/South Dublin stereotypes in CoROCK	236
8.2.1.	Language and style in the <i>Ross O'Carroll-Kelly</i> series	239
8.3.	(<i>Over-the-top</i>) intensifiers	245
8.3.1.	New Intensifying <i>So</i> in the literature	249

8.3.2.	Non-lexical intensifying <i>Totally</i> in the literature	250
8.4.	New Intensifying <i>So</i> and Non-lexical <i>totally</i> in CoROCK: Methodology of analysis	255
8.5.	NISo: Distribution of occurrences in CoROCK-SO	257
8.5.1.	NISo: Colligational pattern preference in CoROCK-SO	259
8.5.2.	NISo: Marking prosody and <i>over-the-top</i> emphasis in text	264
8.5.3.	The pragmatic functionality of NISo in CoROCK-SO	270
8.6.	Non-lexical <i>totally</i> : distribution of occurrences in CoROCK-T	275
8.6.1.	NLT: Colligational Pattern Preference in CoROCK-T	277
8.6.2.	NLT: Marking prosody and <i>over-the-top</i> emphasis in CoROCK-T	281
8.6.3.	The pragmatic functionality of NLT in CoROCK-T	286
8.7.	Identity indexicality of NISo and NLT in CoROCK	291
8.8.	Concluding remarks	296
9.	CONCLUSIONS	300
9.1.	Review of thesis' aims and objectives	300
9.2.	Novel contributions of the thesis to the study of fictionalized IrE representations	302
9.3.	RQ 1: Orality reproduction in CoFIrE	303
9.4.	RQs 1 and 1a: Use, form, developments, and realism in CoFIrE	305
9.4.1.	RQs 1 and 1a: Studying use, form, and developments in CoROCK	308
9.5.	RQ 1 and 1a: Pragmatic multifunctionality of top 3 pragmatic items in CoFIrE	310
9.5.1.	RQs 1 and 1a: Pragmatic multifunctionality of New Intensifying <i>So</i> and Non-lexical intensifier <i>Totally</i> in CoROCK	312
9.6.	RQ 2 and 3: Identity and modern Irishness indexation in CoFIrE and CoROCK	313
9.7.	Limitations of this study	319
9.8.	Future directions	320
	BIBLIOGRAPHY	323
	APPENDIX 1	353
	APPENDIX 2	369
	APPENDIX 3	370
	APPENDIX 4	371
	APPENDIX 5	374

1. FICTIONALIZING LANGUAGE TO CONVEY NATURAL ORALITY AND IDENTITY

This chapter begins by providing an overview of the link existing between fictional (literary) dialects-dialogues and the representation of natural spokenness. The connection between language and identity representation will also be explored here, in particular as it pertains to the use of fictional (literary) dialects-dialogues as tools which enhance the characterization of books. The chapter will also outline this thesis' research questions, and will locate the study within the interdisciplinary frameworks it draws on, clarifying which part of those frameworks the thesis deals with and which parts it does not. In addition, the chapter will also survey what has already been done in the field of fictional Irish English (IrE) representation in literature, identifying a need for studies which investigate fictional IrE representation in a wide collection (or corpus) of fiction to investigate the type of modern Irishness that is being linguistically indexed in *contemporary* Irish fiction, paying particular attention to the use of pragmatic features. An overview of the thesis' chapters is offered at the end.

1.1. Fictional (literary) dialects as mirrors of natural orality

How can readers automatically decipher a character's regional ties or certain aspects of their socioeconomic status, for instance, as different from another character's simply by the language they use when the medium in which they are created prevents us from 'hearing' them? That is due to fictional (literary) dialogue/dialect and their conscious authorial use as stylistic tools which enhance the characterization of their fictional worlds.

Fictional (literary) representations of natural language have long been a subject of debate in the fields of literature and stylistics. Their validity has often been questioned by scholars who criticize their artificial nature as well as the absence of paralinguistic characteristics of orality, e.g. gestures, stress, pitch, etc., inherent to them. However, it is important to remember the fact that authors create these fictional renditions of discourse based on their observations and perceptions of the way natural discourse is used and functions. In addition, while most paralinguistic features are lost due to the written medium they are restricted to (Page, 1988, pp. 7-11; Amador-Moreno & Nunes, 2009, pp. 2-3; Amador-Moreno & McCafferty, 2011, p. 2), some other paralinguistic elements

may still be rendered in text. This can, for instance, be conducted through typographic renderings with capitalization, italicization, as well as via rendering of other stress and intonational features (as shall be discussed in chapter (8)). This is well illustrated through the capitalization of SO in ‘She is SO not going to the gym anymore, that’s for sure’ (Howard, 2005), which signals the extra phrasal stress the speaker (i.e. a fictional character) is placing on the intensifier.

Despite being manufactured by authors, fictional (literary) dialogue/dialect is still based on the writer’s observation of naturally-produced language and models of speech (Fowler, 1989; Simpson, 1997; or Schneider, 2002 *inter alia*). The author (re)creates on paper the type of language they are familiar with, be it their mother tongue, a regional variety, a sociolect, a specific register they may want to reproduce, etc. (see Ferguson, 1998, p. 3; or Amador-Moreno, 2010b, p. 90). To do so, writers include linguistic items that are characteristic of the type of language they are reproducing so that the readership is able to recognize it automatically (Hodson, 2014) and/or gauge their level of ‘authenticity’ *vis-a-vis* the original language or dialect (Amador-Moreno & McCafferty, 2011, p. 2). Given the fact that fictional (literary) dialogue/dialect is based on real language use, it functions as a great tool with which authors are able to flesh out their characterization (Amador-Moreno & Nunes, 2009, p. 4), as the features that are being reproduced may index aspects of the fictional speaker’s identity. Take, for instance, the following fragments:

(1.1.) ‘Oh my God,’ she goes, ‘that’s lollers. That is, like, *so*¹ lollers.’ (Howard, 2011)

(1.2.) ‘Hey, Ro, how the hell are you?’ He’s like, ‘Alreet, Rosser? [...] Ine grant, so I am.’ (Howard, 2013)

(1.3.) Mumbly Dave started to do a line with a girl from the city. Woo hoo, boy. I’ll be right for the Christmas, lads! This wan is mad for me! (Ryan, 2013)

The fragments above, gathered from various novels referenced above, illustrate the indexation of different identarian items. On the one hand, fragment 1.1. above is regionally neutral. Nevertheless, discourse pragmatic markers² such as *like*, and

¹ Emphasis in the original.

² The term *discourse pragmatic marker* refers to a broad variety of functional items such as *like*, *right*, *you know*, *in my opinion*, *kind of*, *well*, *yeah no*, *I think*, *so*, *that is to say*, *in consequence*, etc. However, there seems to be no consensus among academics on which term to use to refer to them, with various labels being applied over time. These include terms like ‘discourse connectives’ (Blakemore, 1992), ‘discourse particles’ (Schourup, 1985), ‘discourse markers’, ‘pragmatic markers’ (Schiffrin, 1987; Fraser, 1988), or ‘discourse-pragmatic features’ (Pichler, 2013) to name a few (the use, functions, and different names these

intensifiers *so* or *Oh my God* suggest that this may be a young female speaker (as will be argued in chapter (7)³). Example 1.2. above, which phonetically renders the accent of a Northside Dublin character, is much more regionally distinctive to the context of Ireland given its use of emphatic tag ‘*so I am*’, which scholars identify as “emblematic of Irish identity” (Timmis, 2015, p. 324; see also Harris, 1993; Simpson, 2004, p. 12; or Amador-Moreno, 2019b, pp. 104-105). Finally, 1.3. is completely indexical of Irish voice, given the inclusion of grammatical, lexical, and colloquial/slang terms distinctive to this variety. Notice the overuse of the definite article *the* in non-standard contexts such as with names of seasons or festive days like *the Christmas* (Bliss, 1984, p. 149; Filppula, 1999, p. 56; or Amador-Moreno, 2010b, pp. 32-33), as well as the lexical use of *lads* (boys) and pejorative *wan* in reference to a woman (Hickey, 2007, p. 363), or the Irish English⁴ (hereafter IrE) distinctive slang phrase *to do a line with somebody*, meaning to have “a regular romantic or sexual relationship” with another (Ayto, 2010, p. 211). These fragments, therefore, showcase how fictional (literary) dialogue/dialect (re)construction may create a great representation of language, especially for linguists, as long as one keeps in mind their artificiality (Ives, 1950, p. 138).

Decoding the valuable identity information that is indexed in the reproduced features will depend, on the one hand, on the level of ‘authenticity’ the fictional portrayal shows with regard to the real spoken dialogue/dialect being represented, while on the other, it will also rely upon how acquainted the audience is with it (Hodson, 2014, p. 200). If the reader is familiar with the real spoken language that is being fictionalized, then they will be able to decode what the use of certain items indexes (e.g. in fragment 1.3. above, *doing a line with someone* and *wan* are regionally distinctive to Ireland). However, if readers are unaware of their indexical value, then the authorial effort put into

items have been assigned over time are discussed at length in chapter (7)). This thesis follows Tagliamonte (2012) in using the more inclusive *discourse pragmatic marker* label.

³ See sections (7.1.) and (7.6.1.) for more on the popular age and gender (mis)associations linked to the use of *like*.

⁴ Studies exploring the variation of English spoken in Ireland often refer to this variety using different terms, all of which have different connotations (see Hickey, 2007, pp. 3-5; Amador-Moreno, 2010, p. 8 for detailed discussions on these terms). *Anglo-Irish*, for example, is reportedly an ethnically and religiously loaded term. Deriving from *Hibernia*, the name the Romans gave to Ireland, *Hiberno-English* is another label which may also be found in studies, mostly in reference to the fictional representation of the English spoken in Ireland, exploring in particular grammar and lexical features, rather than including also other fields, like pragmatics (Amador-Moreno, 2010, p. 8). Thus, while the thesis does include some references to *Hiberno-English* made by other authors or speakers in the form of direct quotation, the more neutral, more inclusive *Irish English* is preferred in this study.

characterizing the fictional cast linguistically will be wasted. The readership, therefore, plays a key role in the process of voice/dialect reproduction in fiction. Writers considering using fictional (literary) dialogue/dialect should, therefore, keep in mind the audience the book is aimed at so as to decide the type and amount of dialect they can (re)produce. This raises questions such as: is the book intended for a national or local readership who will easily identify the non-standard features and their (socio-cultural) identity indexical values? Or is the writer aiming to access a more international audience whose readers may have never been exposed to any of those features? If so, the (re)production of vernacular or dialectal items might require the addition of a glossary at the end of the book as, otherwise, the international audience will be unlikely to perceive any identity trait from the character's stylistic choices. In addition, an excessive use of these features may also alienate a non-local audience and could, potentially, affect the sale of the book in the global marketplace. Thus, to create successful characterizations⁵, the fictional (literary) dialogue/dialect needs to be purposely and carefully planned and crafted. This thesis, therefore, investigates the representation of spoken IrE in a corpus of contemporary IrE fiction, taking its fictional portrayal as a valid source of linguistic representation.

1.2. Research Questions

An inter-disciplinary, comparative study, this thesis endeavors to answer the following research questions pertaining to the fictional representation of IrE in contemporary fictionalized IrE literature:

1. Which are the most frequently (re)produced features of spoken IrE in contemporary IrE fiction?
 - a. How 'authentic' or valid is the fictional representation when contrasted against real, spoken IrE and British English⁶?

⁵ It is also important to mention the fact that in making conscious use of heavily indexical linguistic features, the author's construction of language may simultaneously reflect the state of contemporary culture (see Palma-Fahey, 2005, p. 2 for a discussion of this as reflected in the fictional language of soap operas).

⁶ For a detailed description of the spoken corpora chosen to contrast the fictional data against, see sections (3.8) through (3.8.2.).

2. What does the (re)production of these features in fiction index with regard to speaker identity (i.e. age, gender, geographic location, and social class)?
3. How is modern 'Irishness' indexed in contemporary IrE literature through the use of pragmatic items in fictional dialogue?

1.3. Locating the Study

This thesis is a wide-ranging, inter-disciplinary, comparative study that takes literary representations of oral IrE, or fictionalized IrE, as evidence for linguistic, contemporary usage. It understands fictionalized IrE as a tool through which modern Irishness is reproduced and (re)constructed in fiction by means of linguistic features. Given the dearth of studies that investigate the literary representation of contemporary IrE on a large scale (see section (2.3.1.) for an overview of these studies), this thesis endeavors to use a corpus (i.e. collection of electronic texts) of contemporary IrE fiction as baseline for this study. Despite the existence of other corpora of IrE fiction (outlined in section (3.1.)), at the time of writing there seems to be an absence of corpora of *contemporary* IrE fiction which fulfil the requirements needed for this project, all of which are explained in detail in sections (3.2.) and (3.3.).

To address the research questions outlined in section (1.2.), this thesis presents a comparative study of fictionalized representations of IrE using a novel corpus of *contemporary* IrE fiction which I compiled: the *Corpus of Contemporary Fictionalized Irish English* (hereafter CoFIrE). CoFIrE is a synchronic corpus comprising 1,123,601 words, which contains 16 works of fiction published in the Republic of Ireland by 8 Irish authors (see chapter (3) for an extensive explanation of the corpus construction criteria). In order to check the validity of the CoFIrE fictional representations against real language use, contrastive analyses were carried out against a one-million word corpus of spoken IrE, the *Limerick Corpus of Irish English*, and against the spoken component of the *BNC2014*; a corpus which documents British English (sections (3.8.) and (3.8.2.) provide detailed explanations regarding both of these corpora and the sampling techniques used in the thesis).

Although, as mentioned above, this is an interdisciplinary thesis (i.e. combining linguistics and literature) which comprises a variety of intersecting methodologies which will be explained below, it is important to lay out what this thesis is *not*. This thesis is *not* a dialectal or lexicographical study of IrE concerned with documenting features of this variety as portrayed in fiction. On the contrary, it takes fictionalized (literary) representations of IrE speech and applies a variety of intersecting frameworks to better explore the research questions concerning the analysis of the use, form, and pragmatic function of fictionalized IrE features present in the corpus, their identarian indexicality value, and their degree of authenticity when compared with natural linguistic uses. These methodological approaches range from Corpus Linguistics to Corpus Stylistics, Sociolinguistics and Pragmatics. However, it should also be noted that this is *not* a thesis that is prototypical to any one of those frameworks, but uses tools and methods of each depending on the subject of analysis. For example, it uses analytical Corpus Linguistics methods such as *word frequency lists* and *concordance lines* (explained in section (2.3.4.)) to investigate which are the most frequently produced items in CoFIrE, as well as to analyze the use of each feature in context through *concordance line* searches. However, *keyword analysis*, *n-grams*⁷ (see Biber et al., 1999, 2007, pp. 987-1024), *patterns*, *word plotting*, and the statistical data available through Corpus Linguistics methods (e.g. *type/token ratios*, *mean word length*, etc.) were not used. Corpus Linguistics software suites (i.e. *Wordsmith Tools*, (Scott, 2012), and *AntConc* (Anthony, 2018) are utilized in this thesis to facilitate the search and retrieval of pertinent data and linguistic features via searches for the specific tags each IrE item was assigned in CoFIrE (see (3.6.1.) for a detailed explanation of the original, manual annotation system designed and implemented in CoFIrE).

Pragmatic techniques are also employed in the manual coding of the corpus, particularly in the annotation of pragmatic items (e.g. *fuck*, *you know*, *Jaysus*, *like*, etc.) which were coded following a *form-to-function* approach (see (3.2.1.) for a description of the type of pragmatic features that were coded in the corpus). Taking their context of occurrence into consideration, pragmatic items in CoFIrE were annotated for a) function and/or b) for emotion conveyance. For example, the *fucking* in fragment 1.4. below was

⁷ Sequence of adjoining items (be they letters, words, tokens, etc.), such as set phrases. Other terms assigned to these items include *clusters*, *chunks*, *lexical bundles*, *lexical phrases*, or *multi-word units* (Greaves & Warren, 2010, p. 213)

annotated as *intensifier* (function), and as transmitting a *negative emotion; anger* (emotion conveyance).

(1.4.) I want my fuckin pension you little prick, Mickey roared and roared (Ryan, 2012)

It must be noted that not all pragmatic items in CoFIrE were annotated for emotion conveyance, only those which were found to be the most frequently produced features (which shall be examined at length in chapters (5) through (7)), and the case studies (chapter (8)) which were subsequently carried out in the *Ross O'Carroll-Kelly* corpus (which will be discussed below) which was also created in this study to address the need for a detailed examination of Paul Howard's books. This need arose after preliminary findings at annotation and analytical stage of CoFIrE highlighted Howard's novels as the ones which contributed the largest amount of IrE orality features to the corpus (for more on the creation of the *Ross O'Carroll-Kelly* corpus, see section (3.9.)).

Corpus Stylistics, which is explained at length in chapter (2), was selected as the main analytical framework/methodology due to its multifold usefulness. Blending the study of language in literature and in non-fiction texts (according to McIntyre & Walker's, 2019, p. 15 proposed definition) with the application of Corpus Linguistics computational techniques, corpus stylistic's more easily retrievable quantitative data lends empirical objectivity to what, in the past, may have relied on literary critics' intuitions. Corpus Stylistics, therefore, was used to better answer research question 1a (section (1.2.)) pertaining to the analysis of the use of fictionalized IrE in *contemporary* fiction, and the assessment of the level of realism inherent to its representation. Furthermore, in using Corpus Linguistics toolkits, Corpus Stylistics enables researchers to contrast the linguistic dialogue of a text or *corpus* against real-life uses in any given variety of speech, as is the case of this thesis, where the validity of the fictionalized rendition of IrE in CoFIrE is contrasted against the *Limerick Corpus of Irish English* and *BNC2014*. Finally, Corpus Stylistics was also central to answering research question 3 (see (1.2.)) regarding the potential indexation of speaker identarian values through fictionalized IrE, for, as Stockwell & Mahlberg (2015, p. 131) put it, the application of corpus methods and tools to stylistic analyses of fictional dialogue may also "help to identify textual features that contribute to the creation of a reader's sense of character".

Sociolinguistic factors also had to be taken into consideration to address research questions 2 and 3 regarding the encoding of speaker identities through fictional speech.

Thus, all the features in CoFIrE were manually annotated for speaker age, gender, social class, and geographic location. This annotation system (explained at length in sections (3.6.) and (3.6.1.)) was designed so that it would allow for a faster analysis of identity indexation through language use in the CoFIrE texts. The implementation of this particular type of sociolinguistic analysis was also carried out to address Barron & Schneider's (2005) call for a systematic analysis of the effects of sociolinguistic variables on (IrE) language use in a field they label *Variational Pragmatics*, which is at the junction between pragmatics and dialectology (this is further explained in section (1.4.1.) below). The employment of this sociolinguistic approach was very helpful in addressing and uncovering sociolinguistic variation within CoFIrE, but also in identifying items that are sociolinguistically distinctive to specific speakers/characters in the fiction world.

It must be clarified that while the thesis is concerned with language use and indexation depending on age, gender, class, and region, it does *not* address any other sociolinguistic area of research, such as *linguistic politeness, language planning and policy, attitudes to language, language contact and cross-cultural communication, conversation or critical discourse analysis*, or the *linguistic construction of gender*, among various other areas.

This methodologically complementary thesis, therefore, begins with a broader examination of the research questions which guide this study (outlined in section (1.2.)), identifying the most frequently reproduced items in CoFIrE. Each of these will, subsequently, be contrasted against representations of their real use in samples from the *Limerick Corpus of Irish English* and *BNC2014* so as to gauge the level of realism present in the fictional portrayal, which addresses research question 1a. Ultimately, and in order to answer research question 3 more fully, the thesis narrows its focus to a fine-grained analysis of specific, high-frequency, pragmatic elements (chapter (8)) in the works of one individual author: Paul Howard. As mentioned above, preliminary findings from CoFIrE indicated that Howard's two *Ross O'Carroll-Kelly* novels (see sections (3.9.), (8.2.) and (8.2.1.) for a detailed description of the series and its linguistic style) were responsible for contributing the largest amounts of features to the corpus. Howard's status as an outlier in the CoFIrE group of authors with regard to quantity of IrE orality representation, therefore, called for a more detailed study of his books. This led to the compilation of a separate, *Ross O'Carroll-Kelly* corpus (explained at length in section (3.9)), which allowed for a better analysis of 'new' linguistic developments that may echo modern

Ireland's youth, especially in the case of urban IrE and in urban Dublin English in particular (see Hickey & Amador-Moreno, 2020, p. 14, and Hickey, 2005a, 2016b, 2020).

Titled the *Ross O'Carroll-Kelly Corpus* (hereafter CoROCK), the corpus contains 1,567,256 words and comprises novels from the critically-acclaimed, satirical and humorous *Ross O'Carroll-Kelly* series (see sections (3.9.), (8.2.), and (8.2.1.) for a detailed description of the compilation of the corpus, the series and its linguistic style respectively). The same methods of analysis used in CoFIrE were applied to CoROCK where two pragmatic items which are not traditionally connected with IrE, but rather with American English, were investigated. These include New Intensifying *So* and Non-lexical intensifier *Totally* (both analyzed in chapter (8)), which are exemplified in fragment 1.5. below. While, as mentioned above, these intensifiers have not traditionally been associated with the IrE repertoire, their use in CoROCK merited attention since they featured prominently in the speech of Southside Dublin characters both in CoFIrE and CoROCK. Their representation in fiction could, therefore, be revealing as to a specific type of modern Irishness (i.e. Southside Dubliner) which is represented in the books, which would address research question 3. Fragment 1.5. illustrates their use well.

(1.5.) 'That is SO, like ... aaaggghhh!' and Sophie goes, 'I know. It's, like, totally ... duuuhhh!' (Howard, 2003)

These two intensifiers, which are often popularly associated with negative speaker stereotypes (see section (7.6.1.) for their popular (mis)association with the *Valley Girl* and *Surfer Boy/Jock Dude* personae in the USA) were analyzed for identity indexical value as perceived by Paul Howard. This thesis will contend that these American stereotypes which are connected to their use could be transported and extrapolated to the context of Ireland, and Dublin in particular, in the form of what I have come to label the *Posh Southside Girl/Socialite* and the *Rugby Jock/D4 head* personae (see also (7.6.1.) for more on the extrapolation of the US-bound stereotypes to the Irish context).

1.4. Rationale for this study

To be able to answer research question 3 regarding the identification of the type of modern Irishness that is being portrayed through fictional speech in *contemporary* IrE fiction (see (3.2.2.) for an explanation of what was taken as *contemporary* in this thesis), one needs

to be familiar with the type(s) of identity the Irish have, or rather, have been *perceived* and *portrayed* as having through time.

The term ‘identity’ is difficult to define as this is a complex and flexible concept which refers to an individual or group of people’s sense of self and group belonging. It is a cultural construct (Abercrombie et al., 2006, p. 190), which is neither finite nor permanent, but in constant change depending on people’s historical context, as well as on circumstances and/or major changes pertaining to their daily lives, their social, economic, political, or religious background, among other factors. Thus, identity can, and often is reconstructed (O’Donovan, 2009, p. 95). A way to do so is through language.

One can create, (re)construct, and/or *project* their identity through language, and this can all be achieved through linguistic choice(s) (i.e. different ways of expressing the same message) we make (Joseph, 2010; see also Eckert, 2000; Hickey & Amador-Moreno, 2020, p. 3). This is due to the fact that languages are ever-changing, living systems which are heavily influenced by the society and culture people are part of. As such, the use of different linguistic choices will, as was illustrated in section (1.1.) above, encode, or *index*, information about the speaker’s socioeconomic background, religion, or political affiliations, among other factors, or about the speech community the speaker belongs to. Thus, by consciously making use of certain linguistic choices, the speaker can project their own linguistic identity as being that of a member of the community they belong to, or can mislead the listener into believing they belong to said community, although that may not be the case (Hickey & Amador-Moreno, 2020, p. 4). For example, in their study of ‘Pittsburghese’, Johnstone et al. (2006) find that the monophthongization of /aʊ/ to /ɑ:/ in Pittsburgh indexes working-class, male voice. In the context of Ireland, Hickey (2018) identifies a recent lowering of short front vowels in Dublin English as being indexical to young “cosmopolitan” female Dubliners.

The portrayal of orality in fiction through fictionalized (literary) dialect-dialogues can, therefore, also be taken as “a cultural performance” (Terrazas-Calero, 2020, p. 252) of relevant information about the “speakers’ and readers’ identities, intentions, [or] beliefs” (Warner, 2014, pp. 362-364). As mentioned in section (1.1.), however, the successful rendering of the dialect or variety being reproduced in fiction will hinge on how close to real usage the fictionalized (literary) dialect-dialogue is, and on the audience/readers being familiar with the set of values encoded in its use. In this

case, readers must be familiar with IrE features in order to be able to understand, upon reading/‘hearing’ a character’s dialogue, that they are Irish.

Defining contemporary (linguistic) Irishness in fictional contexts, however, is a complicated task, as this concept has fluctuated throughout time and has often hinged on the use of the Irish or English languages (see Hickey & Amador-Moreno, 2020, p. 4 and various contributions to their volume for more on the complexities of Irish linguistic and cultural identity). For centuries, Irishness was linguistically rendered in literature through the caricature of the Irish Paddy or ‘Stage Irishman’. This was a colonial, dehumanizing stereotype promoted by England, especially in Elizabethan drama, which represented the Irish as a poverty-stricken nation of brutish, drunken, illiterate, dim-witted and/or ape-like people. This stock character had a number of associated, identifying features which Leerssen (1996, p. 97) finds in Tudor and Stuart plays. These range from personality traits like their irritability, wildness, unreliability, or naiveté, to their looks and fashion including “shaggy hair, narrow trousers, cloaks and darts” (as discussed in Duggan, 1937; see also Bartley, 1954, or Bliss, 1979 for more on this; and Leerssen, 1996, pp. 85-168; or Amador-Moreno, 2010b, pp. 89-100 for a detailed exploration of the evolution of this stereotype in literature). Nevertheless, their use of the Irish brogue, swearing, and other IrE features is, perhaps, the most recognizable Stage Irishman trait (Amador-Moreno 2010, p. 90), as will be explained below.

The Stage Irishman stereotype was often linguistically indexed through the character’s dialogue which would be full of characteristics the playwrights *perceived* as being very frequently used by the Irish. Whether that be the case or not was unimportant. What mattered was that the audience be “familiar with and could easily identify” (Amador-Moreno, 2010, p. 90) the linguistic traits with the stereotype being portrayed. Amador-Moreno (*ibid.*, pp. 89-97) offers a very detailed examination of the history and development of the Stage Irishman by both national and international playwrights, overviewing different academic sources, and listing some of its linguistic repertoire as including features like the Irish pronunciation of English consonants [t], [d], [s] often represented as fricatives *sh*. This is illustrated with examples from Ben Jonson’s *Irish Masque* such as ‘Chreeshes sake’ (Christ’s sake) (*ibid.*, p. 92). Hickey (2007, p. 8) also points out a lack of features which are “diagnostic” of the stereotype, noting that authors

emphasized mostly salient, rural, (often pronunciation) IrE items, such as [θ, ð] fortition⁸, or epenthesis in words like *farm* [farəm] (see Hickey, 2010 for an account of other pronunciation features linked with Stage Irish over time).

The Stage Irishman caricature was perpetuated in literature for centuries. As observed by Walshe (2012, p. 264), one need only look at 19th-century magazines like *Punch* in Britain to see the ape-like drawings which exported such a dehumanizing and degrading image of the Irish around the world. Take also, for instance, James Murdoch's representation of Irishness. Murdoch was a Scottish Orientalist scholar and journalist. In 1892 he published *From Australia and Japan*, a volume of short stories where he introduces the reader to Mick O'Donovan. The story this character features in recounts a trip by ship where Mick, the reader is told, is traveling in steerage. He is described as being funny and witty. He also swears very frequently and often acts foolishly due to his constant state of inebriation. In fact, the narrator recounts an episode where Mick is so drunk that he falls overboard into the ocean. The narrator also provides a subtle yet prejudiced commentary on Mick's moral character:

[his face] had the map of Ireland all over to prove it. And then his brogue, and the things it said,—especially the latter! I had met him in the Sydney Police-Court, —but that would be a digression, and besides, I don't want to give folks a low idea of his moral character to begin with, and thus prejudice them against him. (Murdoch, 1892, p. 38)

Mick's 'brogue' or accent is well illustrated in the next fragment:

'Now, ye omadhaun av the son av a dirty Cockatoo farmer, isn't that just what I'm goin' to demonstrate? You know them shpuds we gits for breakfast, and tiffin, and tay, and dinner, an' all?'

"We do," came in a chorus from all round the table, for we had plenty of opportunity to make their acquaintance. ("Shpuds," be it explained, is Hibernian for "spuds," which again is Anglo-Saxon vernacular for "potatoes,"—that is, at least in all civilized countries.) (*ibid.*, p. 44)

⁸ (Stage Irishman) *Fortition* implies a consonantal pronunciation change whereby fricatives [θ, ð] are produced as alveolar stops [t, d]. Thus, for example, the word *three* would become homophonous with *tree*, while *mother* may be pronounced/rendered as "mudder" (Howard, 2014).

Mick's dialogue firstly shows generally non-standard features which include pronunciation (e.g. *av* for *of*, reduction of final clusters: *an'* for *and*; *goin'* for *going*) and grammatical items (e.g. demonstrative *them* (i.e. *them shpuds*); overgeneralized –s marking (*we gits*)). However, it also showcases traces of Stage Irish features, which is the 'brogue' the narrator referred to above. These include grammatical items such as *ye*, the use of Irish terms like *omadhaun* (i.e. a foolish person), and the inclusion of pronunciation features such as T-lenition (*dirthy* for *dirty*), or the frication of *s* in *shpuds* (spuds). With regard to the latter, notice the narrator's negative view of it as a feature used in "uncivilized" countries.

The negative typecast of the Irish born in colonial times shifted with the passing of time and the 17th and 18th centuries saw attempts being made by authors who were either Irish, had connections with Ireland, or had lived in the country (Amador-Moreno 2010, p. 95-96) to dignify the dialect (Leerssen, 1996, pp. 85-168; Amador-Moreno, 2010b, pp. 96-97). Although the Stage Irishman typecast persisted well into the 19th and 20th centuries, it coexisted with a move by Irish writers such as Somerville and Ross, Lady Gregory, or Yeats, among many others (i.e. particularly those that participated in the Irish Literary Revival (ca 1890-1922)) to dignify IrE and use it in literature "artistically and as a means of seriously delineating Irish national character" (Garvin, 1977, p. 103). In doing so the dialect became connected with Irish nationalism (see Amador-Moreno, 2010, pp. 97-100 for more Irish linguistic identity portrayal in literature in this period). In addition, these two centuries also brought about a series of events which not only decimated the population of the country, as in the case of the Great Famine (1845-1849), but also affected its national identity.

The creation of the Irish Free State (1922-1937) in the south of the island following the Anglo-Irish Treaty (1921) and War of Independence (1919-1922) which split the island into two, and the subsequent declaration of the Republic of Ireland (1949), came at a cost. For years, Ireland was plagued with economic stagnation, high rates of unemployment, and mass emigration. So much so, that it gained the reputation of being one of the poorest countries in Western Europe (Alvarez, 2005). The old, colonial stereotype shifted then to accommodate the new circumstances the Irish found themselves in. Thus, Irishness became synonymous with poverty, parochialism, and illiteracy. Linguistic identity also changed. As Hickey (2020, p. 79) points out, there was an "endonormative re-orientation" whereby IrE switched from being a Southern British

English *affiliate*, which took this variety as a prestige model, to creating their own supraregional IrE variety (strongly influenced by changes led by educated South Dubliners) which accommodated their new sociocultural identity, and distanced them linguistically from their colonial past as well as from more locally-bound IrE forms.

The entrance of Ireland into the European Economic Community (later the European Union) in 1973 opened the doors to a wider market and, by the mid-1990s, the country found itself experiencing what some called an “economic miracle”, consequently becoming known as the *Celtic Tiger*. This was a period (roughly from the mid-1990s to the late 2000s) in which Ireland experienced rapid economic growth which was due, in part, to factors such as the “economic openness to global markets, low tax rates, and investment in education” (Dorgan, 2006). Ireland, thus, went from being one of the poorest countries in Western Europe to one of the richest worldwide (see, for example, Lucy et al.’s (2019) edited volume for a detailed overview of the events, personalities, trends, etc. leading to the Tiger and to its aftermath). However, the Celtic Tiger was much more than an economic boom, as it brought about an identity crisis in a country whose “national psyche [was] tied up in centuries of poverty” (Alvarez, 2005)

The Celtic Tiger wasn’t just an economic ideology. It was also a substitute identity [...which] expressed itself in a mad consumerism, in an arrogance towards the rest of the world, in a willful refusal of all ties of history and tradition. But there were other things wrapped up in it too—optimism, confidence, a new openness and ease, an absence of fear. (O’Toole, 2010, p. 5)

By 2008, however, the bursting of the property bubble and the collapse of the banks, sunk the country into a severe economic recession. As a consequence of such a brisk downturn of fortunes, enormous unemployment and emigration returned as some of the most important issues the country was facing, and its newfound, cosmopolitan identity took a hit. Contemporary IrE literature holds a mirror to these events with authors like Roddy Doyle or Paul Howard representing the hardships of the Irish working class in the case of the former, and the excesses of the Celtic Tiger years in the case of the latter (see O’Brien, 2015, and *fc.* for a monograph on the cultural significance of Howard’s *Ross O’Carroll-Kelly* series).

Nowadays, contemporary Irishness is hard to define as it seems to be trapped between the identity of the Irish Cosmopolitan and the Paddy (O’Donovan, 2009, p. 107),

while struggling to distance itself from the “repressive weight of [its] history” and the influence of globalization (Ni Eigheartaigh, 2009), which is very often synonymous with Americanization (Terrazas-Calero, 2020, p. 276) . However, if we want to investigate the concept of ‘modern’ Irishness, we need only look at the way IrE is used, in this case, in contemporary fiction. Indeed, Amador-Moreno (2010, p. 101) posits that the optimism and confidence that ran rampant across Ireland during the Celtic Tiger years was also reflected in literature, where certain authors “consider this [linguistic] variety as a key element of Irishness and an important component of Irish identity.” Writers such as Roddy Doyle, Paul Howard, Marina Carr, Nuala Ní Chonchúir, playwright Brian Friel, and poets like Paula Meehan or Sarah Clancy, to name but a few, can be counted amongst those authors that consciously exploit the use of IrE to enhance their characterization via the indexation of “shared socio-cultural references” (*ibid.*). For example, Paul Howard has been critically acclaimed for his acute rendition via representations of lexicogrammatical, pragmatic and phonetic features of the accent of affluent Southside Dubliners, which contrasts with the portrayal of working-class Northsiders found in his *Ross O’Carroll-Kelly* series (for more on Howard’s linguistic style, see section (8.2.1.)), and which is also present in Roddy Doyle’s acclaimed *The Barrytown Trilogy*. Another example would be Donal Ryan, whose debut novel, *The Spinning Heart*, and subsequent *The Thing about December*, were praised, not only for their portrayal of the consequences of the recession in rural and small town areas, but also for his linguistic renditions of the Limerick and County Tipperary dialects.

One of the reasons I chose to investigate the fictional portrayal of IrE in literature was the fact that while the use of Stage Irish language and the portrayal of the Paddy is well researched in academia from a literary perspective (see Duggan, 1937; Bartley, 1954; Bliss, 1979; Leerssen 1996, or Amador-Moreno, 2010, pp. 89-100, inter alia), its modern counterpart seems to have received considerable less attention. In addition, and according to Amador-Moreno (2010, p. 101), there seems to be a correlation between contemporary Irish authors and their portrayal of fictionalized IrE to encapsulate modern Irishness in their fiction world. Inspired by the work of scholars like Amador-Moreno or Walshe (along others who shall be discussed in section (2.2.1.)) who have examined the type of fictionalized (literary) dialect-dialogue being used nowadays in individual works of fiction and corpora of telecinematic discourse respectively, and given the lack of studies that analyze the portrayal of fictionalized IrE in a corpus of *contemporary* fiction, I

decided to examine, as outlined in (1.2.) and (1.3.), not only which linguistic items were more frequently used in a corpus of contemporary IrE fiction, but also to investigate whether the portrayal of Irishness through those elements varied in any way from the traditional, negative caricature of the Stage Irishman.

As discussed in (1.3.), Corpus Stylistics was chosen as the main framework methodology due to its blend of stylistic analysis and Corpus Linguistics, computerized analytic tools, which not only facilitated the annotation and faster retrieval of large amounts of data via the Corpus Linguistics tools, but also expedited the investigation of the type of identity values that the use of linguistic choices in literature suggests they may encode in real life.

1.4.1. Why focus on pragmatic IrE elements?

Before conducting any Corpus Stylistic investigation of identity indexation through language, it must be pointed out that linguistic identity can be conveyed through various levels of language (Hickey & Amador-Moreno, 2020, p.15). For example, on the phonetic level, Hickey (2016b) highlights the recent adoption of American English characteristics among young Irish speakers which serve to create new identarian profiles in the context of Ireland. Other examples can be found throughout IrE scholarship, which enjoys a very long and rich tradition of studies dating back to the late 18th century, with the focus of attention mostly being the examination of IrE lexis, phonology, morphology, and syntax. Indeed, Amador-Moreno et al. (2015, p. 1) state that lexical studies of IrE have a longer standing tradition tracing back to the middle of the 15th century, although subsequently more serious academic consideration was given to the fields aforementioned. According to Hickey's (2005b, p. 18) detailed account of the evolution of IrE scholarship, and despite earlier enquiries into IrE use, the investigation of this variety began to bloom in the 20th century, culminating in the publication of P.W. Joyce's (1910) monograph, *English as We Speak it in Ireland*. This was followed, in subsequent decades, by the publication of a number of studies aimed at a wider, 'lay' audience, on the subject of the use and particularities of the English spoken in Ireland. Christensen's (1996) *A First Glossary of Hiberno-English* also addressed the need for non-specialist publications that documented IrE lexicon which "is not always fully understandable to those who are

familiar only with Standard English” (*ibid.*, p. 7). The book offers a wordlist with accompanying quotations in context which is designed for “general readers and students outside of Ireland” who do not have the time or inclination to “embark on more detailed linguistic studies” before approaching a text written in IrE (*ibid.*). While there are other studies (see Hickey, 2005 for an overview of previous research into IrE), perhaps the culmination of IrE lexical analysis would be the more recent publication of *A Dictionary of Hiberno-English* (Dolan, 2006, 2020).

More popular works documenting IrE vocabulary published in the 21st century would include Kelly & Kelly’s *Overheard in Dublin* series. These are books that comprise compilations of humorous, every-day life anecdotes, quotes, and conversations overheard in the general area of Dublin, which were originally sent to and/or posted on the *overheardindublin.com*⁹ website by ordinary citizens. Despite their non-academic nature, the books effectively document language use in and around the capital phonologically, at times showing a high degree of linguistic awareness. The fragment below depicting a schoolgirl switching accents on the phone illustrates this perfectly, as the comedic value of this anecdote lies on the socio-pragmatic elements encoded in her switch from Northside to Southside Dublin¹⁰ accent:

‘Outside the school? Yeah, I’m on the Luas, I’ll be two minutes like. Yeah I’ll be there. I’m fookin comin alright?! Will ya hold on...Yeah, outside the school. Will ya fookin hold on....! I’m comin! Jaysus! Fookin givin’ me hassle.’

She gets off the phone to whoever it was and makes another call (in a blatant soutside [sic] accent):

‘Hello, yes, yes, I’m on my way. OK, excellent. Alroysh, heh, see you then.’

(Kelly & Kelly, 2007, p. 80)

Although these publications tend to be disregarded within academia due to their mostly casual perceptions and observations on language use from the point of view a non-

⁹ While the original website is now no longer in use, contributions may still be made to their accounts on *Twitter* (@OverheardDublin) and *Facebook* (*Overheard in Dublin*).

¹⁰ The Northside of Dublin City has traditionally been perceived as a working-class area where speakers use more locally-bound IrE features which often are indexical of *local* Dublin identity. Contrastively, speakers in the Southside, the more affluent area, try to dissociate themselves from the local identity by using less regionally-bound items. The use and connotation of using either accent are explained at length in sections (8.2.) and (8.2.1.).

academic, almost perceptual dialectologist audience, Amador-Moreno et al. (2015, p. 1) defend their value, suggesting more scholarly attention be paid to them as they “provide useful starting points for academic study”. Indeed, it seems they document well the use of pragmatic elements and their indexical value. Although IrE has a distinctive pragmatic profile (Barron, 2017) which functions as a strong index of Irish identity (Hickey & Amador-Moreno, 2020, p. 15), this field seems to have traditionally been a neglected area in Irish scholarship. Nevertheless, Amador-Moreno (*fc.*) notices a recent increase of interdisciplinary studies which look into the discourse-pragmatic profile of IrE and the contextual factors that influence it.

Perhaps responsible for inciting interest in this field is Barron & Schneider’s (2005, p. 11) edited volume, *The Pragmatics of Irish English*, which addressed the need for more research into pragmatics which addresses issues of language interaction in the “private, official and public spheres of Irish life” (*ibid.*, p. 3-11). In this ground-breaking volume, they emphasized the need to place “language in (inter)action” in the spotlight of IrE research, and shone a light on the need for a systematic investigation of the effects of sociolinguistic variables on language in use. This would, therefore, call for the creation of a sub-field they labeled *Variational Pragmatics*, which would stand at the junction “between pragmatics and modern dialectology” (*ibid.*, p. 12). Although there were some early, individual studies looking at IrE pragmatic/variational items which also used Corpus Linguistics methods, such as Kallen & Kirk’s (2001) analysis of the IrE verbal phrase (see O’Keeffe (*fc.*) for an overlook of other studies), the importance of Barron & Schneider’s (2005) volume lies on the fact that it brought about a crucial shift in IrE research which looked at it as a variety in its own right (rather than in comparison/contrast to British English), while Amador-Moreno (*fc.*) highlights the fact that it “addressed an important gap in the study of regional pragmatic variation”.

Since the publication of Barron & Schneider’s (2005) volume, research into IrE pragmatic phenomena in different domains seems to have flourished. For example, Clancy (2016) explores the pragmalinguistic repertoire of intimate discourse (i.e. vocatives, pragmatic markers, taboo language, and pronouns) in a small corpus of IrE family discourse: the *Limerick Corpus of Intimate Talk*. Murphy (2010) analyzes IrE pragmatic items in her *Corpus of Age and Gender Irish English* from a sociolinguistic, age and gender-focused perspective, while work on socio-pragmatic, perceptual dialectology (see Lonergan, 2016; or Lucek & Garnett, 2020) and/or pragma-stylistic

investigations of the creation and negotiation of diasporic Irish identity in a corpus of Irish emigrant letters (Amador-Moreno & Avila-Ledesma, 2020, Avila-Ledesma, 2019, Avila-Ledesma & Amador-Moreno, 2016) have also emerged in the field. Investigation of Irish identity issues from a socio-pragmatic perspective are also explored in many of the contributions to Hickey & Amador-Moreno's (2020) *Irish Identities* volume (see O'Keeffe (*fc.*) for a detailed overview of other contributions to the field of IrE discourse-pragmatic variation). Finally, Irish fiction has also been the source for research into the pragmatics of IrE with scholars using telecinematic corpora (Walshe, 2009, 2011, 2016, 2017), comic books (Walshe, 2012, 2013), and literature (Amador-Moreno, 2012a, 2015, 2016; Amador-Moreno & Terrazas-Calero, 2017, 2022; Terrazas-Calero, 2020) as valid sources for linguistic investigation (see (2.5) and (2.5.1.) for a more detailed overview of Corpus Stylistic studies that look at Irish fiction).

Despite this catalogue of pragmatic research into IrE, Amador-Moreno (*fc.*) and O'Keeffe (*fc.*) identify discourse pragmatic markers as the most prolific field of study. Take, for instance, all the contributions to Amador-Moreno et al.'s (2015) edited volume, *Pragmatic Markers in Irish English*. Building up on pragmatic issues raised in Barron & Schneider's (2005) volume, the editors indicate that this book was meant as a response to the neglect IrE pragmatics had endured within academia (Amador-Moreno et al. 2015, p. 3). The volume, whose leitmotif is highlighting the utility of discourse pragmatic markers as sources for linguistic investigation (Amador-Moreno, 2005), investigates them in terms of pragmatic and sociolinguistic variation in a wide range of informal, intimate, professional, and/or literary settings in IrE, including some cross-varietal comparisons with varieties of British English.

Given the fact that, as mentioned above, IrE has a distinctive pragmatic profile which is indexical of Irishness, this thesis endeavors to contribute to the traditionally understudied, yet currently proliferating area of IrE pragmatics. In doing so, it also addresses the need for more studies which observe IrE pragmatic use in large corpora, and caters to this by focusing research question 3 (see (1.2.)) on addressing what type of modern Irishness may be indexed through the use and reproduction of pragmatic items in contemporary IrE fiction.

1.5. Overview of chapters

This section offers a description of the layout of the thesis. As we have seen, Chapter (1) introduces the close connection that exists between fictional (literary) dialect-dialogue, particularly regarding issues of authenticity of representation and identity indexicality value. It locates the thesis as an interdisciplinary study which uses Corpus Linguistics, Corpus Stylistics, Pragmatics, and sociolinguistic analytic techniques to investigate the fictional representation of IrE in a corpus of contemporary IrE fiction. This chapter also outlines the research questions, which focus on establishing which are the most frequently reproduced IrE items in the corpus, as well as assessing their level of realism through comparisons with natural, spoken uses. Their value with regard to indexing modern Irishness will also be explored, while a closer look will be paid at discourse pragmatic items to examine the indexation of ‘new’ Irishness. Finally, the chapter establishes the thesis’ rationale with regard to the investigation of Irish identity as portrayed in literature, arguing that while there is plenty of academic research into past representation of Irish identity in literature, particularly in the form of the Stage Irishman, its modern counterpart, which may have been affected by crucial socioeconomic events in 21st-century Ireland, has not enjoyed such academic scrutiny. It argues that, given the distinctive pragmatic profile IrE has, which serves as a strong identity-indexing element, this neglect can be solved by exploring pragmatic items in a large corpus of Irish fiction.

Chapter (2) will survey the literature on the main frameworks which inform the thesis theoretically and methodologically, particularly Corpus Stylistics, and will pay close attention to those studies that are relevant to the examination of fictionalized IrE. Chapter (3) offers a detailed account of the design and building criteria in the creation of the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE), while also describing the sampling techniques used for the selection of authors and books included in the corpus. It will also describe the manner in which the texts had to be prepared for computerized analysis, and will provide a detailed explanation of the original, manual annotation system I designed to expedite corpus searches in CoFIrE (and subsequently in CoROCK) and the retrieval of pertinent data. Corpus data and metadata storage and management will also be explored at length. Finally, the chapter overviews the corpus analytic tools used in the thesis, and explains which corpora were used as comparative sources. Chapters (4), (5), (6), and (7) discuss findings drawn from the analysis of CoFIrE.

Chapter (4) outlines general findings from CoFIrE, investigating occurrences and frequency of production, and exploring production by social class. In addition, the chapter finds pragmatic elements to be the most prominent ones in the corpus, which centers the focus of research of subsequent chapters on investigating the top three most frequently reproduced pragmatic items. Thus, chapter (5) explores the use of non-standard quotative verb GO analyzing intralinguistic variables such as *tense*, *lexical form*, and *grammatical subject*, and investigating its pragmatic functionality and its social marking value within the corpus. These findings will be contrasted against the reference corpora to measure the realism of the fictional portrayal. In addition, this chapter will also examine complementarily the second most frequently reproduced, non-standard quotative, BE LIKE, with regard to their use, social, and pragmatic value in the books. Chapter (6) investigates the use of the second most frequently produced pragmatic item in CoFIrE: intensifying *fucking*. Beginning with an overview of the multifunctionality of taboo language in natural interaction as well as in fictional dialogue, the chapter outlines the functionality of intensifying *fucking* in the literature, surveying studies of its use in spoken and fictionalized IrE. It explores the use of intensifying *fucking* in CoFIrE with regards to colligation pattern preference, *tnesis*¹¹, pragmatic functionality, and social indexical value, all of which will be contrasted against the reference corpora. Finally, the chapter also offers a case study of the use of intensifying *fucking* by the homodiegetic narrator in the *Ross O'Carroll-Kelly* series, as its narrative use and pragmatic functionality was found to deviate from his use as a character within the story. Chapter (7) explores the third most recurrent pragmatic feature: discourse pragmatic marker *like*. The chapter begins with an overview of the use of *like* in the literature and will survey studies which have focused on investigating this feature in IrE. Then, it will focus on analyzing the use, form, pragmatic functions, and identity indexation value of *like*, contrasting the findings against the natural usage offered in the contrastive corpora samples.

Chapter (8), however, focuses on the case studies conducted on CoROCK. It begins with an examination of the linguistic style of the *Ross O'Carroll-Kelly* series and the stylization of North/South Dublin stereotypes in the books, and then introduces New Intensifying *So* and Non-lexical intensifier *Totally* as the features that will be examined at length. It overviews the literature on intensifiers and on *over-the-top* intensifiers in

¹¹ *Tmesis* refers to the insertion of *fucking* into non-standard contexts like single words or compounds, e.g. “abso-*fucking*-lutely”. For more on *tmesis*, see sections (6.2.) and (6.6.1.).

particular, as both of these items, and then surveys the literature on the use of New Intensifying *So* and Non-lexical intensifier *Totally* in other varieties. The chapter also explores their use, form, pragmatic multifunctionality, and sociodemographic indexical value in CoROCK in detail. Finally, all of the concluding remarks on findings drawn from the analyses of CoFIRE and CoROCK will be brought together in chapter (9), where the limitations of this study will also be outlined and future recommendations will also be offered.

2. CORPUS STYLISTICS AND THE STUDY OF FICTIONAL (LITERARY) DIALECTS

This chapter provides an overview of the literature that informs the frameworks and methodologies that are the foundation of this thesis, including Corpus Linguistics, Corpus Stylistics, Sociolinguistics, and Pragmatics. The chapter focuses on providing a detailed definition of what Corpus Stylistics is, its multifunctionality and advantages. The connection between Corpus Stylistics and Corpus Linguistics is also explored, detailing what *corpora* are, how they feature in Corpus Stylistics, overviewing also how to design and build a corpus so that it is as valuable for the Corpus Stylistics project as possible, and outlining Corpus Linguistic tools which can be used in Corpus Stylistics studies. Some of the most prominent studies which apply Corpus Stylistics to fictional speech representation in different varieties of English are surveyed, identifying key research themes. The same is done with studies which have applied Corpus Stylistics to the study of fictional IrE linguistic representation, also identifying several research trends in the context of IrE. The latter survey evidences certain gaps in the context of IrE research which this thesis addresses, such as the application of Corpus Stylistics to the study of frequency of production, authenticity of representation, or the indexicalization of contemporary Irishness in a larger corpus of *contemporary* fiction.

2.1. Introduction to Corpus Stylistics

In the last few decades, *Corpus Stylistics* (hereafter CS) has emerged as a new area of study in the field of text analysis. In essence, CS is a blend of traditional stylistic analyses complemented by the application of Corpus Linguistic techniques (see (2.2.) for a definition of what a *corpus* is, and (2.3.4.) for an overview of Corpus Linguistic techniques and analytic software). With its roots firmly planted in Stylistics, McIntyre & Walker (2019, p. 15) believe that the problematic definition of CS also extends to Stylistics, which is generally described as “the study of the language of literature”. In short, CS is concerned with the study of style in language (Leech & Short, 2007, p. 11) and the effect non-linguistic factors may have on it (Jeffries & McIntyre, 2010, p. 1). Most stylistic research (including its earliest studies) seemed focused on examining the

language of literature, with some scholars like Fischer-Starcke (2010) or Ho (2011) appearing to have carried this arguably ‘narrow’ definition of stylistics on to CS, describing it basically as the use of corpus techniques for the analysis of literary language. However, McIntyre (2015, p. 16) argues against this narrowness, maintaining that neither Stylistics nor CS are *only* concerned with literary language and that, while literature still is one of its main objects of study for many researchers, to think it only investigates literary language would be “unacceptably reductive”. Instead, McIntyre (*ibid.*) and McIntyre & Walker (2019, p. 15) propose that a definition of CS should not be restricted to literary language but inclusive of non-fictional texts as well. Examples of these types of investigations include Jeffries’s (2007, 2010) work on critical stylistic studies; Coupland’s (2007) sociolinguistic approach, Semino & Short’s (2004) ‘narrative’-focused CS analysis of speech, writing, and thought presentation in a corpus of prose fiction compared to non-fictional texts (e.g. newspaper news reports and (auto)biographies), or Chen et. al’s (2019) contrastive CS analysis of Clinton’s and Trumps’ campaign speeches during the 2016 USA presidential election, inter alia). Thus, and despite the general non-consensus, a basic definition of CS should mention the application of Corpus Linguistics techniques to the study of literary *and* non-literary written texts.

As a blended sub-discipline, CS merges computational corpus toolkits with qualitative (stylistic) analyses, utilizing quantitative data provided by Corpus Linguistic techniques to inject objectivity into stylistic investigations. One of the benefits of CS is the fact that making use of the (more easily retrievable) quantitative data allows for research and findings which, in the past, may have heavily relied upon literary criticism and the analyst’s own intuitions and observations. Furthermore, in applying corpus methods/tools to stylistic analyses, CS effectively conciliates stylisticians’ interest “in a detailed textual analysis and the concerns of data science and digital humanities approaches that look for trends across large amounts of data” (Mahlberg & Wiegand, 2018, p. 140). In addition, McIntyre & Walker (2019, p. 310) also argue in favor of CS being regarded, not as a “niche practice on the fringes of stylistics”, but as an integral part of stylistic analysis, which satisfies the demand for empirical evidence, allowing for transparent and systematic means of interrogating literary critical positions (Stockwell & Wynne, 2006). Carter (2014, p. 86) also provides a description of CS which effectively encompasses its wide-ranging functionality:

[...] for me corpus stylistics at its best illustrates the best of both stylistics and applied linguistics practice: it is evidenced in language use, it is retrievable in quantitative datasets, it does not hide from qualitative human assessment and evaluation, it offers rich possibilities for language learners at all levels and it expands the frontiers of applied linguistics and literary studies, even if some literary specialists and some applied linguists may be looking in other directions.

All of the properties aforementioned make CS a great tool to assess fictional (literary) dialects-dialogues. Despite its relative novelty when compared to more established fields, CS enables the researcher to determine the linguistic artistry of an author/authorial group, or to contrast the linguistic dialogue of a text or group of texts against real-life linguistic uses in any given variety of speech (Amador-Moreno & Terrazas-Calero, 2022, p. 519). Furthermore, and as Sinclair (1991, p. 100) points out, using computer-assisted methods to explore the language of literary texts is very useful, since “[t]he language looks rather different when you look at a lot of it at once”. Corpus Linguistics methods (such as those discussed in section (2.3.4.)) offer the researcher the possibility of analyzing large quantities of text for linguistic phenomena in a more objective and empirical manner, while also allowing for a faster retrieval of the pertinent data. In addition, the application of corpus methods and tools to stylistic analyses of fictional dialogue may also “help to identify textual features that contribute to the creation of a reader’s sense of character” (Stockwell & Mahlberg, 2015, p. 131). Of course the quantitative data must be qualitatively analyzed and interpreted by the investigator as well (Baker, 2006, p. 89). For all of these reasons, CS was selected as the main framework and methodology to investigate the research questions (see (1.2.)) of this thesis.

2.2. How to do Corpus Stylistics?: *Corpora* in Corpus Stylistics

One of the main elements necessary to conduct a CS study is a *corpus* or collection of electronic texts. While this may seem deceptively obvious, McIntyre & Walker (2019) point out that the data studied in CS might often fail to fall within the stricter parameters

of the traditional definition of what a *corpus* is for linguists like Sinclair (2005) or McEnery & Wilson (2001). According to them, corpora are collections of electronic texts which are representative of a language or variety. For that, the corpus has to be “maximally representative of a particular variety of a language” (*ibid.*, p. 73) so as to be considered “a source for linguistic research” (Sinclair, 2005, p. 16). In other words, *representativeness* (see (2.2.1.) for an in-depth explanation of this concept) must be a key criterion in the design of a corpus (detailed explanations of further corpus design and building criteria including *representativeness*, *sampling*, and *balance* are offered in sections (2.3.) through (2.3.3.)). In addition, Sinclair’s (*ibid.*) stringent criteria also list a number of databases which are *not* to be considered linguistic corpora. These include the Internet, archives, collections of citations and/or quotations, and individual texts. The justification for excluding the latter being that a single text does not fulfill the criterion of *representativeness* (see (2.3.1.) for more on this criterion) of a linguistic variety.

Nevertheless, CS studies often analyze one text, individual author, or group of authors within the same literary movement/genre, or examine the way a particular feature or set of linguistic items is utilized in one or more texts, among other research topics. As such, McIntyre & Walker (2019) do consider individual-text studies as Corpus Stylistics, with Mahlberg (2007) suggesting that the *corpus* part of CS evokes, not the dataset, but the various methodological practices adopted from Corpus Linguistics which are essential to CS. This thesis agrees with and is based on both Mahlberg’s (*ibid.*) and McIntyre & Walker’s (2019) views on single-text studies being acceptable *corpora* in CS along with linguistic corpora that contain larger amounts of text.

Having discussed the nature of CS, it is worth mentioning that its analyses may utilize different sorts of databases depending on the object of study. For instance, the researcher may want to investigate the use of a specific linguistic item (e.g. intensifying *fucking* which shall be explored in chapter (6)) in a given text. They may also want to examine the development of one or more items in a larger corpus that comprises the entire oeuvre of a writer (see, for example, chapter 8 on the investigation of New Intensifying *So* and Non-lexical intensifier *Totally* as used in the *Ross O’Carroll-Kelly Corpus*). Contrastively, the analyst could be interested, for instance, in comparing fictional linguistic uses against a corpus of naturally-occurring language of the variety they are interested in (e.g. the *Corpus of Contemporary American English* or the *British National*

Corpus). Thus, different types of corpora would be required for each of those hypotheticals.

It is also possible that the analyst's research questions call for the compilation of a new corpus, or specialized database (see sections (2.3.) through (2.3.3.) for more on corpus building criteria). Indeed, Gibbons & Whiteley (2018, p. 286) note that while CS can and does use different sorts of databases, there seems to be a preference for specialized corpora. These are substantially smaller in comparison to larger corpora such as the *British National Corpus* (100 million words), or some of American linguist Mark Davies' corpora, including the *Corpus of Contemporary American English* (+1 billion words), *iWeb* (14 billion words), the *NOW (News on the Web)* corpus (+6 billion), or his *Corpus of Global Web-based English* (1,9 billion), to name a few. Unlike these, specialized corpora are designed specifically for the project at hand. This gives the researcher an advantage over utilizing larger (more general) corpora which they, then, may need to spend countless hours sorting through in search for usable material. Despite often having smaller word-counts, Anthony (2013, p. 146) argues in favor of specialized corpora as valuable resources for linguistic examination, stating that the value of a database lies not in its size, but on "what kind of information we can extract from it".

Another advantage of specialized (smaller) corpora is their manageability (Vaughan & Clancy, 2013). Unlike larger databases, the corpus data and metadata can be more easily marked up and stored, enabling researchers to conduct more in-depth qualitative analyses. Furthermore, specialized corpora also benefit from being either automatically or manually annotated (see (3.6.) for a description of corpus annotation); a task which, despite being time-consuming if done manually, is made easier by the fact that the analyst is often both the corpus compiler and tagger. Wynne (2005, p. 225) argues in favor of annotation, suggesting that it is a more empirical procedure than choosing linguistic examples, while also making sure that the analyst fine-combs their categorization system to account for all cases. This procedure also enables the researcher to conduct fast analyses, allowing them to "extract statistics relating to frequency, distribution and co-occurrence of forms from the annotated text" (*ibid.*). Finally, Wynne (2005) also underlines the benefits of corpus annotation in the context of expanding or replicating the research in other literary/non-literary areas (see section (3.6.1.) for a detailed explanation of the design of the annotation system used in CoFlrE and its application to the corpus).

Drawing on the various advantages detailed in the literature that specialized corpora present, and given the current lack of a corpus of *contemporary* IrE fiction containing the restricting parameters set for this thesis (which shall be outlined in sections (3.2.) through (3.2.3.)), I decided to compile my own specialized, manually-annotated corpus to answer the thesis research questions (see (1.2.)) As mentioned in the previous chapter, the goal of creating the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE) is to investigate the portrayal of spoken IrE in *contemporary* IrE fiction. The compilation and manual annotation system designed for and applied to the corpus are detailed at length in chapter (3).

2.3. Criteria for designing and building corpora

Before creating a corpus, researchers need to be aware of a series of criteria which must be set and scrupulously followed so as to ensure the usefulness of the corpus for their project. These include corpus *representativeness*, *size*, *sampling*, and *balance*, and will be discussed in the following sections, along with how they applied to the design and building of CoFIrE.

2.3.1. Representativeness

Provided that, as mentioned in (2.1.), a *corpus* is a collection of electronic texts which is representative of a language or linguistic variety, the first criterion that must be taken into consideration during the design stage is *representativeness*. According to Biber (1993, p. 243) this refers to the “extent to which a sample includes the full range of variability in a population¹²”. However, Leech (1991, p. 27) contends that a corpus will only be representative if “findings based on its contents can be generalised to a larger hypothetical corpus”. In other words, if the findings can be generalized to the variety. This connects with issues such as ‘authenticity’ (especially in the case of literary/fiction corpora) because for a corpus to contain as ‘authentic’ a documentation of material as possible,

¹² Biber uses *population* to refer to the language or linguistic variety the corpus analyzes.

then it needs to capture the full linguistic variability of the language, variety or *population* (to use Biber's terminology) it studies.

According to McIntyre & Walker (2019, p. 65), the term *language variation* or *population* needs to be understood broadly, suggesting the adoption of a more sociolinguistic view or *variation* as language use in (social) context. As per Holmes & Wilson's (2017, p. 6) definition, this would encompass "different accents, [...] linguistic styles, [...] dialects, and even different languages" whose use varies according to different social circumstances. This would allow for the inclusion of, not only general corpora dealing with a variety of a language (e.g. *British National Corpus* for British English), but also for the inclusion of more specialized corpora, perhaps dealing with an author's oeuvre and style, or, as is the case of the corpus compiled for this thesis (see section (3.4.)), a corpus which studies the representation of IrE orality features (i.e. pragmatic elements) in *contemporary* Irish fiction. In all, McIntyre & Walker (2019, pp. 65-66) propose that *variety/population* in the context of a corpus be understood as "any grouping of language examples that are delimited by any number of non-linguistic parameters". Among other parameters, they mention *genre*, *author*, *geographical boundaries*, *mode/medium of communication* (spoken vs written), or *domain*. The linguistic and non-linguistic parameters set for CoFIrE are explained at length in sections (3.2.) through (3.2.2.).

To return to the definition of a *representative* corpus, the corpus must, therefore, represent a *variety/population* in its entirety. However, such a goal is quite unattainable, as it is very hard for a dataset to contain "every single example of a particular language or language variety" (*ibid.*, p. 66). This is why scholars mention the need for the corpus to be "maximally representative" (Sinclair, 2005, p. 73) or to reach a "maximal degree of representativeness" (McEnery & Wilson, 2001, p. 78) of the variety it studies in order to be considered a *reliable* "source for linguistic research" (Sinclair, 2005, p. 16). Given the practical impossibility of reaching absolute *representativeness* as well as the fact that it cannot be objectively ensured or measured (Tognini-Bonelli, 2001, p. 139), Leech (1991, p. 27) suggests that *representativeness* be regarded as an "act of faith" while McEnery & Hardie (2012, p. 10) refer to it, along with *balance* (which shall be discussed in section (2.3.3.)) and comparability, as an ideal which corpus builders "strive for but rarely, if ever, attain". Despite these challenges, McIntyre & Walker (2019, p. 66) maintain that *representativeness* (at least as much of it as possible) can only be achieved via good

knowledge of the *variety* being studied and meticulous corpus design. In addition, *representativeness* will also be dependent upon corpus *sampling* and *balance*, both of which will be discussed in the following sections.

2.3.2. Sampling

Sampling entails the selection and gathering of examples of the *variety* the corpus represents. However, before gathering the examples, the corpus builder must identify what is the *variety/population* the corpus will represent. In Biber's (1993, p. 244) words, "it is not possible to identify an adequate sampling frame or to evaluate the extent to which a particular sample represents a population until the population itself has been carefully defined". The term *sampling frame* refers to the full list of texts the samples will be drawn from. Sinclair (2005) lists a series of parameters, which Biber (1993, p. 243) suggests be non-linguistic, which could be used in this restriction. These include but are not limited to: *mode* (e.g. spoken, written, electronic), *text type* (e.g. book, ad, tweet...), *domain* (e.g. academic), *location* (English in New Zealand), or *text date* (e.g. New Zealand English in 19th century novels). Thus, the constraining parameters of the *variety/population* should be established before the collection of the frame. For example, *the study of pragmatic elements in English* might be a rather large and vague proposition. Thus, the researcher should delimit that broad research topic by creating more narrowing criteria. For instance, they might consider restricting by subject area (*taboo language*), mode (*written*), or location (*Ireland*). This would retrieve a narrower topic (i.e. the *study of taboo language in written IrE*), which, in turn, would simplify the collection of the sampling frame. Existing corpora in the field of IrE fiction (see (3.1.) for an overview) exhibit constraining criteria. Take, for example, Cesiri's (2012) *Corpus of Irish Fairy and Folk Tales*, which restricts its sampling frame by time (*19th century*), genre (*fairy and folk tales*), mode (*spoken*), geographical distribution (*rural Ireland*), and occupation (*peasantry*). Other IrE corpora follow different parameters. Walshe's (2017) telecinematic corpus, for example, is restricted by genre (*movies*), location (*Northern Ireland*), time (*the Troubles, 1960s-1998*), variety (*IrE features*), and mode (*spoken*

interactions). Sections (3.2.1.) and (3.2.2.) offer a detailed explanation of CoFIrE's *variety/population* and the non-linguistic parameters that were set at corpus design stage.

Having clearly defined characteristics that identify the *variety/population* the corpus should be representative of helps in creating an efficient sampling frame, which is the next step in corpus design. Bibliographies could be used to this end (McEnery & Wilson, 2001, p. 78; McIntyre & Walker, 2019, p. 69). Once the frame is created, the texts can be sampled, using different techniques (see Biber, 1993 for a discussion of *(non)-stratified sampling* and *random sampling*) which cater to the criteria and projects' research questions. However, none of these were used in the sampling of the CoFIrE frame whose creation and sampling techniques will be explained at length in sections (3.3.) and (3.3.1.).

Issues arising from the establishment of the frame are the number and size of the samples, or rather, whether full texts should be included or only fragments. The matter of determining the corpus size (i.e. *how many texts*) is hard to establish and will be dependent upon several factors, including *representativeness*, or the *variety* the corpus is representative of, the parameters set to restrict it, as well as on time constraints (Reppen, 2010, p. 31; McIntyre & Walker, 2019, p. 72). For example, a project wishing to study the pragmatic functions of *like* which had restricting parameters (e.g. spoken discourse at university level among 20-25 year-old students) may collect a small number of spoken texts per individual but wish to recruit as many participants as possible in order to capture the most number of functions. Similarly, a corpus designed to be representative of orality in a given variety of English, for example, will aim at gathering as much data as possible, resulting in large corpus sizes. This is the case of corpora such as the *Corpus of Contemporary American English*, which contains over 1 billion words from various genres including spoken, academic, newspapers, as well as telecinematic subtitles, blogs, etc., and spanning 1999-2019. In other cases, the size of the corpus is already pre-established, especially when the corpus represents finite varieties. For example, Quaglio (2009) created the *Friends* corpus, which comprises 604,767 words, and contains fan-made scripts from across the nine seasons this popular USA sitcom ran for. Similarly, Bednarek's (2011a) *Gilmore Girls* corpus also contains fan transcripts of all episodes (153) of this USA comedy drama. Finally, it must be noted that the size of the corpus will also be dependent upon the "time it will take to collect, computerise, annotate and, if required, tag and parse the corpus" (Clancy, 2010, p. 82).

The next issue to be considered is the length of the texts (i.e. the *variety*) to be collected. Given that texts (whether spoken or written) significantly vary in length, the inclusion of longer texts along with smaller ones may skew the corpus (*ibid.*, p. 85). This raises the issue of *how much* text should be incorporated. Biber (1993), for example, suggested that the use of extracts of approximately 1,000 words might be representative of a variety, whereas Sinclair (2005) disagrees, deeming fragments of a text as unrepresentative of the whole. Instead, he proposes that full texts be included when possible (see section (3.2.3.) for a detailed explanation of avoiding copyright issues when dealing with corpora of contemporary fiction), acknowledging the possibility that the difference in sample sizes may affect the overall balance of the corpus if one of the text were to be longer than the others. Similarly, McIntyre & Walker (2019, p. 73) also propose that the larger the amount of texts or extracts the corpus includes, the more representative the database will be of the *variety*. In addition, they state that if the researcher determines that a specific number of full texts will be representative of each parameter of the *variety*, then the number and size of the sample texts will be “determined at the design stage, but the size of the corpus [will] not [be] known until it is built” (*ibid.*, p. 75). The compilation of CoFIrE follows Sinclair’s proposition that full text be included, the number of which was pre-selected before the establishment of the frame due to time restraints (see (3.3.) and (3.3.1.) for a detailed account of the sourcing, sampling, and size of the CoFIrE texts).

2.3.3. Balance

The size and number of the sample texts is tightly connected to another major factor in corpus building: *balance*. Although vague in nature, *balance* would imply the inclusion of a wide range of genres (McEnery et al., 2006, p. 13) which are taken as representative of the *variety* the corpus is representative of. Thus, *representativeness* and *balance* are tied together since, as stated by Leech (2007, p. 134), the *balance* of a corpus “is an important aspect of what it means for a corpus to be representative”. In other words, *representativeness* will be dependent upon how balanced the corpus is. However, achieving *balance* is often difficult to accomplish. Similarly to *representativeness*, *balance* is also regarded as a desiderata in corpus building which all researchers seek “but rarely, if ever, attain” (McEnery & Hardie, 2012, p. 10) as there is also no measure for it,

which makes it difficult to demonstrate (Leech, 2007, p. 134). *Balance*, therefore, will be based on the builder's intuitions and judgment (McEnery & Hardie, 2012, p. 10; McEnery et al., 2006; Sinclair, 2005).

Balance is also interconnected with corpus *size* and *sampling*. McIntyre & Walker (2019, p. 73) determine that while the inclusion of the largest amount of (sample) texts or fragments will grant more *representativeness* to the corpus, the number and size of the samples brings *balance* to the fore since it is up to the builder's judgement to decide *how much* text will be included (Sinclair, 2005). They go on to describe a balanced corpus as that in which the sample texts assigned to the parameters of the *variety* do not necessarily need to be equal, but reflective of the frequency of the parameters in the variety, so "the relative popularity of some of the texts or genres might influence the number of texts collected" for each parameter (McIntyre & Walker, 2019, p. 73). For example, a corpus that investigates the representation of orality in 19th century fiction may include stage plays, for they are likely to contain more dialogue, which is where the more naturalistic forms of orality may be represented, rather than the narrative passages included in novels. Similarly, a corpus of youth language may include not only spoken conversations (face-to-face or online) but also electronic messages, as this group is more likely to text and communicate electronically than, for instance, elderly speakers. These decisions would, therefore, be examples of the corpus builder's judgments applied to the design and compilation of the corpus for the achievement of *balance*, which will be further assisted through *sampling* (as shall be explained in (2.3.2.)). In the case of CoFIrE, it was decided that novels and short stories would only be sourced if they contained large amounts of dialogue as opposed to large narrative passages, since there is a larger probability that IrE features be used in direct dialogue rather than in narration (a more detailed explanation of *balance* in relation to the design of CoFIrE is offered in section (3.4.)).

Finally, and according to Sinclair (2005, p. 4), it is important that the analyst "maintain comparison between the actual dimensions of the material and the original plan", but also that they document all the steps followed in the design and compilation of the corpus so as to ensure the suitability of the corpus and to leave room for future improvements.

2.3.4. Corpus analytical tools in Corpus Stylistics

When conducting Corpus Stylistics (CS) studies, the researcher may use the analytical toolkit available in Corpus Linguistic software, with tools such as frequency lists, concordance lines, keywords, inter alia. For example, *frequency lists* provide a list all words and characters in a corpus, and sort them, either alphabetically or by order of frequency, by identifying which word is more frequently used in the corpus. The list may also offer statistical information pertaining to (standardized) *type-token ratios* (Scott, 2010, p. 148). Given their use, *frequency lists* are often described as “a first approach to a corpus” (*ibid.*) which provides an initial look at the vocabulary or topics that are covered in the dataset. On the other hand, *concordance line* searches create a list of all the instances of a word or set of words as used *in context*, which allows the researcher to retrieve pertinent information about the feature’s different functions according to its context (Smith et al., 2008). As will be explained in the various methodological sections described in chapters (5) through (8) regarding the analysis of CoFIrE and the case studies carried out in CoROCK corpus, both *frequency lists* and *concordance line* searches were used in this thesis to determine 1) the most frequently produced, pragmatic items in CoFIrE, and 2) the various uses, forms and functions that the items under investigation had depending on their context of use.

Frequency lists, concordance lines, and other Corpus Linguistics tools, such as *collocation*, *keyword* analyses or *type-token ratios*, are made available to researchers through a number of specialized, text analytic software programs and suites. Examples include programs like *WordSmith Tools* (Scott, 2012), *AntConc* (Anthony, 2018), *Sketch Engine* (Kilgarriff et al., 2014), *Wmatrix*¹³ (Rayson, 2009) or TACT (Lancashire et al., 1996). Amador-Moreno & Terrazas-Calero (2022) list three tools worth employing which give the researcher the opportunity to perform certain tasks when exploring literary language through concordancing, key word analyses, inter alia. These are Hickey’s *Corpus Presenter* suite; *Corpus Linguistics in Context* (CLiC), and *WordWanderer*.

¹³ *Wmatrix* is a web-based, corpus analysis and comparative tool originally developed by Paul Rayson (Rayson, 2009) which incorporates the automatic semantic tagger UCREL and the *Constituent Likelihood Automatic Word-tagging System* (CLAWS) tagger (see Garside, 1987). *WMatrix* allows the researcher to conduct frequency list and concordance analyses.

Corpus Presenter was created by Hickey (2003) and includes 27 programs that allow the researcher to investigate their own corpus, as well as a corpus of generally dramatic IrE texts from the Middle Ages through the 20th century (see section (2.4.) below for a detailed explanation of this suite). A web part of the Dicken's Project started by the University of Nottingham and now run in collaboration with the University of Birmingham, CLiC is another program which contains 15 of Dicken's novels. It also includes (Mahlberg et al., 2016, p. 435) a reference corpus of 19th-century books and other subcorpora through which CS and cognitive poetics analyses may be conducted. Finally, *WordWanderer* (Dörk & Knight, 2015) is web app that allows expert and non-experts to visualize linguistic patterns in an uploaded text through corpus tools like concordance lines.

2.4. Corpus Stylistics Studies in Fiction

Since the number of Corpus Stylistics (CS) studies has recently increased, this and the following sections will focus on noting and providing a brief overview of some of the most significant research themes and publications in this field.

Having established how to create a new corpus and which tools to use to analyze it, this section will survey the research that has already been done in the field of CS and fictional (literary) language, particularly in the context of IrE. The research conducted in this particular area of study has led me to conclude that there are at least three distinct (often interconnected) research trends in the field (see (1)-(3) below). These align closely with those identified in Bednarek's (2017, pp. 144-150) pragma-stylistic investigation of the multifunctionality of televisual dialogue, which include characterization, authenticity/realism, and humor.

- 1) Language variation.
- 2) Testing the authenticity of the literary portrayal against natural uses.
- 3) Analyzing the use of literary dialogue for characterization purposes.

The introduction of computerized, electronic corpora and their methodology eases the investigation of language variation, be it in the form of diachronic (i.e. exploring linguistic evolution over time) or synchronic (i.e. investigating the language in a specific

time) studies. In the case of the former, the use of fictional (literary) dialect-dialogue provides the researcher with a written source(s) through which linguistic items and their evolution may be explored. For instance, the *Penn Parsed Corpus of Historical English* could be used for diachronic investigations, as it contains a collection of corpora that comprise British English texts from the mid-12th century through the early 20th century (for a more in-depth list of historical corpora, see Amador-Moreno, 2019b). Similarly, a more ‘specialized’ corpus with which diachronic analyses of IrE may be conducted is Hickey’s *Corpus Presenter* suite introduced in the previous section, which contains various written (mostly dramatic) IrE texts ranging from the Middle Ages through the 20th century. On the other hand, synchronic analyses may allow for comparative examinations of written and spoken sources. These types of investigations may be conducted on the various, 1-million word, *International Corpus of English* sub-corpora, which contain spoken/written data from several varieties of English. Other corpora may also be used for synchronic investigations of different varieties, like the *Limerick Corpus of Irish English*, which is explained at length in section (3.8.), or the *British National Corpus* and the more updated *British National Corpus 2014* (hereafter *BNC2014*) corpora (also explained in (3.8.)), among many other corpora (see Amador-Moreno & Terrazas-Calero (2022) for a more in-depth explanation and overview of different diachronic vs synchronic studies).

2.4.1. Corpus Stylistics Approaches

Whether diachronic or synchronic in nature, Adolphs (2006, pp. 65-66) identifies two types of approaches CS studies lend themselves to when investigating fictional (literary) dialects-dialogues: *intra-textual* and *inter-textual*. On the one hand, *intra-textual* studies involve the examination of one text or text collection with the aim of “reveal[ing] further information about the data” (*ibid.*, p.65) and its interpretation. Techniques that may be used in these studies include *concordance* and *collocation* analyses, *frequency lists*, *keyword* searches and *type-token ratio* calculations. On the other hand, the *inter-textual* approach involves comparing texts against other collections of electronic texts. Here, “individual lexical items and phrases in literary texts [can be compared] with those that occur in other, possibly non-literary, corpora” (*ibid.*, p. 66) to check for deviations and analyze their literary effect. In this case, reference or baseline

corpora may be used as contrastive source of evidence of linguistic use which “can inform the analysis of such [linguistic] items in a literary text or corpus” (*ibid.*, pp. 65-68). *Word* (and phrase) *list* comparisons, *concordance* and *collocation* analyses are useful techniques in this approach.

Intra-textual, CS studies using specialized corpora may vary from study to study. For instance, one could investigate only spoken interactions in the *BNC2014* or, as is the case of this particular thesis, decide to examine IrE orality representation in *contemporary* IrE fiction, as opposed to in 19th century prose. Examples of *intra-textual* studies using specialized corpora in the field of IrE non-fiction include those conducted on McCafferty & Amador-Moreno’s (2012b) *Corpus of Irish English Correspondence* (CORIECOR). CORIECOR is an almost 4-million-word corpus (which was inexistent up to this moment) comprising ca 6,500 letters by Irish emigrants sent to and from various countries from 1700s to 1900s, whose aim is enabling researchers to trace and study the “emergence and development of features of [IrE] and study syntactic, morphological, stylistic, regional and social variation” (Amador-Moreno 2019, p. x). Amador-Moreno’s & McCafferty’s studies offer diachronic, analyses of traditional IrE features and discourse pragmatic markers as reflected in the letters (McCafferty & Amador-Moreno, 2012c, 2012a, 2019; Amador-Moreno & McCafferty, 2015; Amador-Moreno, 2012b, 2019a), which also led to Amador-Moreno’s (2019b) publication of *Orality in written texts: Using historical corpora to investigate Irish English (1700-1900)*. This volume traces in detail the use and development of discourse pragmatic markers (*so, anyway/anyhow, like, and sure*), deictics (*this/that; here/there*), pronouns and embedded questions in CORIECOR, at times contrasting with other (historical) corpora like the *Corpus of Historical American English* (COHA), the *Corpus of Modern English Prose*, the *Corpus of Late Modern English Texts*, and the *Corpus of Modern Scottish Writing*, and with spoken contemporary IrE corpora such as *LCIE* and the *Limerick Corpus of Intimate Talk* (Clancy, 2016). These contrasts would be representative of *inter-textual*, historical linguistic case studies. CORIECOR is also used as a database for corpus pragma-stylistic, *intra-textual* examinations of the emigrants’ perceptions of life abroad as opposed to in their homeland. Words like *home/country* and related items (Amador-Moreno & Avila-Ledesma, 2020; Avila-Ledesma, 2019; Avila-Ledesma & Amador-Moreno, 2016) are examined with regard to their use in the creation and negotiation of diasporic Irish identity.

Other types of *intra-textual* studies may use CS to investigate language variation and characterization in one particular text. For example, Culpeper (2002, 2009) examined the speech of six characters in *Romeo and Juliet* and contrasted it against the speech of the rest of the cast. Culpeper's (2002) keyword analysis of style markers identifies words that are specific to each character, and which are also revealing of their motivation. For example, Romeo's top three keywords are *beauty*, *blessed*, and *love*, which validate our perception of him as "the lover of the play" (2002, p. 20). However, other keywords in his speech (i.e. *eyes*, *lips*, and *hands*) reflect his preoccupation with the physical, especially with describing his own body, which further expands his character, exposing his egocentric nature (*ibid.*). In contrast, Culpeper's subsequent (2009, p. 53) study finds that Juliet's keywords (i.e. *if*, *yet*, *or*, *would* and *be* (mostly subjunctive) highlight the anxious state of her character. Another example of an *intra-textual* analysis is Mahlberg & McIntyre's (2011) examination of keywords in *Casino Royale*, where they propose their classification into 'reader-centered' (i.e. normally evaluative and ambiguous) and 'thematic-centered' (i.e. concrete, world-building elements). Straddling the line between *intra-textual* and *inter-textual*, Mastropierro (2018) offers a specialized, monographic investigation of Joseph Conrad's *Heart of Darkness* and four of its Italian translations, combining CS and descriptive translation studies. In televisual dialogue, Bednarek (2012, p. 223) utilizes *concordance* and *keyword* analysis to investigate the linguistic construction of the character of Sheldon Cooper in CBS' sitcom *The Big Bang Theory* as a 'nerd/geek'. By examining occurrences of pronoun *I* and *I'm* in his dialogue and the information they disclose about him, Bednarek (*ibid.*, pp. 210-211) finds that the traits they identify are archetypal of the nerd stereotype: high intelligence, lack of social skills, physical unfitness, technological expertise, or his being a fan of the sci-fi *Star Trek*, among other features.

Other *intra-textual* CS studies may look at the oeuvre of a particular author, as in the case of Mahlberg (2013, p. 51), who examines Dicken's style in a corpus of 23 of his books, analyzing characterization techniques, arguing that clusters (i.e. repeated linguistic patterns) can "contribute new categories to the inventory of descriptive tools for literary stylistics" by functioning as devices that can be utilized as "[local] textual building blocks" (*ibid.*, pp. 26-73) which define and identify characters within his fictional world. Further examples of scholars investigating the works of individual writers are Moss' (2014) doctoral diachronic examination of Henry James' syntax in his early and late style,

or Archer et al's (2009) study and identification of keywords within key, automatically-identified semantic domains in Shakespeare's tragedies and comedies. To do this they use the historical semantic tagger tool of *UCREL Semantic Analysis System (USAS)*, a corpus tool currently part of the *Wmatrix* platform (as already defined in section (2.4.1.)), both originally developed by Paul Rayson (2009) for the automatic semantic annotation of present-day and historical English (spoken and written) texts, as well as texts produced in several other languages. Archer et al's (2009) study focuses on the domain of LOVE in three selected "love" tragedies and three "love" comedies, finding that, in contrast with the comedies, the domains of "intimate/sexual relationships" and "liking" are the most underused in the tragedies (*ibid.*, p. 143), which appear to represent "love" in a less ideal, darker manner (*ibid.*, p. 154).

Inter-textual CS studies are varied. For example, Montoro (2019) investigates the claim made by the general and critical perception that popular fiction (as opposed to 'highbrow' literature) is syntactically simple due to its under-use of subordinating conjunctions. She uses the *Lancaster Speech, Writing and Thought Presentation corpus*, her own *Chick Lit Corpus* (Montoro, 2012) and *Modern Vampire Corpus* (Montoro, 2015), and contrasts them with the written section of the *British National Corpus*. She believes this alleged phrasal simplicity, which her findings refute (Montoro, 2019, p. 66), may be caused by the under-use of (often less complex) nouns and prepositions introducing nominal post-modifiers that are evidenced in her corpora (*ibid.*, p. 75). Finally, Montoro (*ibid.*) argues that these two under-representations may be illustrative of the genre's willingness to follow only some stylistic trends of 20th-century English (i.e. less informative and colloquial language), which could perhaps contribute to the perception of the linguistic simplicity of popular fiction.

Another *inter-textual* CS study is Piazza's (2011) investigation of discursive deviance in the speech of serial killers in four horror/slasher movies by Italian director, Dario Argento. Her analysis uncovers a type of discourse which is self-absorbed and narcissistic. This is represented by the insistence on 1st person pronouns and verbs of perception and affection (e.g. *like* or *feel*) rather than of cognition. Their interactions (*ibid.*, pp. 90-99) tend to be one-sided, full of modality and with a tendency to unidirectionality, shifting to bi-directionality towards the end, after the killers have been apprehended or given in to the interrogators. She also develops a classification that assesses the level of informativeness in their discourse. Her analysis shows (*ibid.*, pp. 93-

103) that their discourse is highly deviant as it violates the pragmatic maxim of relevance by being irrelevant, unclear and purposely uninformative. This changes when their interactions become bi-directional, with the killers now being talked *about* by other characters or their discourse becoming over-informative (for the sake of the audience).

2.4.2. Approaches within *inter-textual*, synchronic CS analyses of fictional (literary) dialect-dialogue

Amador-Moreno & Terrazas-Calero (2022) argue that *inter-textual* CS research is an especially significant approach for synchronic investigations of fictional (literary) dialect-dialogue as it can be compared with naturally spoken language. Furthermore, Mahlberg (2007, p. 221) notes that the contrast between these fictional-spoken language and naturally-produced discourse corpora enables the researcher to analyze the representation of sociocultural contexts. Thus, and as mentioned in (2.4.), two of the major investigative topics within this angle of the approach can be identified as:

- 1) Testing the authenticity of the literary portrayal against natural uses.
- 2) Analyzing the use of literary dialogue for characterization purposes.

For instance, Mahlberg (2016, pp. 149-153) illustrates how the use of corpus methods in CS can help identify linguistic patterns that perform specific functions in narrative texts. In particular, she utilizes the *inter-textual* approach to investigate the use of ‘suspensions’ (i.e. interruptions of a character’s dialogue by the narrator) in a corpus of Dicken’s novels and a corpus of 29 other 19th century novels. Exploring 4-word clusters, she finds that the most frequent ones are those 1) describing the manner of speech, which include references to body language (e.g. *in a low voice*, or *laying his hand upon*, among others), and 2) those including temporal information (e.g. *after a short silence*). Functionally, she proposes that the first group can describe the character’s manner of speech, but can also,

when referencing their body language, contextualize or highlight in a specific context. Furthermore, they can also present the narrator's interpretation of the character's manner of speech and body language, especially when in cases such as "*as if he were*" or "*as if he had*". On the other hand, the temporal clusters are used to retrospectively report a pause in the character's dialogue, thus functioning more as a narrative device than as a tool to report about the time being narrated. Another example would be Ruano San Segundo's (2018) study of stage directions in *Waiting for Godot*, contrasting them against its French version for stylistic differences.

The connection between lexico-grammatical patterns and characterization is explored in Bednarek's (2011a, p. 55) study of televisual dialogue in the TV drama-comedy *Gilmore Girls*. Her findings are contrasted against a corpus of 10 other TV shows and several reference corpora of unscripted spoken and written language, including the *American National Corpus* (spoken and written parts), the *Longman Spoken American Corpus*, and the *Cambridge International Corpus*. Using CS techniques, she investigates "expressivity" or "expressive character identity", that is, identity-constructed aspects that are drawn from evaluative, emotional, and ideological meaning which can be relayed either verbally or non-verbally. In her study, Bednarek notes that some expressive items in the dialogues like endearments (e.g. *honey*), interjections (e.g. *wow*) and evaluative adjectives (e.g. *weird*), among others can be character distinctive, while also being illustrative of the way groups share expressive resources.

Since the number of both *intra-textual* and *inter-textual* CS studies is large, this section provides a brief overview of the significant publication growth that has taken place on this field in the last decade. A relatively short overview of the field is found in Semino's (2011, pp. 547-548) description of what she terms "corpus-based or computer-aided stylistics" in *The Routledge Handbook of Applied Linguistics*. A few years later, the growth of the field among academics is made evident by the devotion of entire handbook chapters to it. That is the case of Mahlberg's (2014, 2016) CS-focused contributions to the same handbook and to *The Bloomsbury Companion to Stylistics*. Arguably, a large quantity of the work conducted on this field from 2010 onward has been done, as exemplified above, in the shape of individual papers or as (hand)book chapters (see, for example, Mahlberg's references above, Amador-Moreno & Terrazas-Calero (2022; 2017), McIntyre & Walker (2022), Montoro (2019); or Bednarek (2011b, 2017), Piazza (2011), or Walshe (2011, 2017) for telecinematic discourse). However, there is a

growing number of monographs devoted to demonstrating the approach and its practical application. Examples include Fischer-Starcke's (2010) exploration of Jane Austen's novels, Ho's (2011) investigation of John Fowles' *The Magus*, Mahlberg's (2013) monographic study of Dicken's language, while Mastropierro's (2018) book explores the connection between its practice and translation studies. In Hoover et al.'s (2014, p. 5) edited volume, all contributions employ CS or what they label 'digital literary' techniques, and despite its focus on the pragmatics of fiction, many of the contributions in Locher & Jucker's (2017) volume demonstrate use of CS techniques. Finally, in 2019, McIntyre & Walker published a monograph devoted to Corpus Stylistics. While these are all significant contributions to the field, this thesis is concerned with what has been done regarding CS in the field of IrE fiction, as surveyed below.

The number of CS studies concerning IrE fiction is much smaller. Despite the recent increase in the number of investigations which incorporate CS, most tend to take a Corpus Linguistics or Applied Linguistics approach. As will be explained in detail in the following sections (see (2.5.) through (2.5.1.)), of those that do use CS, most focus on examining non-standard language, (individual) discourse pragmatic markers, and/or on investigating how they may index specific Irishness. After surveying current CS studies in the context of Ireland which utilize CS (see (2.5.1.)), this thesis identifies three main search trends and sees the need for studies which address the research questions (outlined in (1.2.)) which are key to this thesis utilizing a large corpus of fiction. The thesis, therefore, also endeavors to contribute to this need by compiling and using a corpus of *contemporary* IrE fiction as valid evidence of real present-day IrE use, and by taking a CS approach in conjunction with Corpus Linguistics techniques to investigate RQ1: which are the most frequently (re)produced linguistic items in the corpus, and RQ1(a) how authentic the (re)production is when contrasted against real usage. In using a corpus of contemporary IrE fiction, the thesis will also address RQ2 pertaining to what the use of the commonly reproduced features may index with regard to speaker identity, at least as perceived by the authors, and will try to answer RQ3 by exploring how modern 'Irishness' may be indexed in contemporary Irish literature through the use of pragmatic IrE items in fictional dialogue.

2.5. Studying Irish English in Fiction

The analysis of the portrayal of IrE orality in fiction has enjoyed a long tradition in the field of Irish research. However, the number of studies that take a CS approach is rather small when compared to other varieties, and those that do incorporate it tend to mostly fall under the Corpus Linguistics or Applied Linguistics umbrella as mentioned in the previous section. Thus, we find studies that explore the construction and representation of the Stage Irishman (described in section (1.4.)) through the linguistic rendition of IrE in literature (see Duggan, 1937; Bartley, 1954; Bliss, 1970; or Amador-Moreno, 2010b, pp. 89-100; and Leerssen, 1996 for a detailed examination of the evolution of this typecast throughout literature). Others investigate the use of IrE by individual authors like John Millington Synge (Kiberd, 1979/1993), James Joyce (Dolan, 1991, 1990), William Carleton (McCafferty, 2005, 2008, 2009), Patrick McGill (Amador-Moreno, 2006), or Marina Carr (Lynch, 2006). More recent work has been conducted on analyzing the conscious use of IrE to (re)construct identity in Howard's *Ross O'Carroll-Kelly* series (see Amador-Moreno, 2005, 2012a, 2015, 2016; Amador-Moreno & Terrazas-Calero, 2017, 2022; and Terrazas-Calero, 2020). Further studies look into the general portrayal of IrE (see, for example, Taniguchi, 1972; Todd, 1989; Amador-Moreno, 2010c; or Connell, 2014), oral narratives (Cesiri, 2012), telecinematic portrayals (Walshe, 2009, 2011, 2016, 2017), and comic book fiction (Walshe, 2012, 2013) but most works take a Corpus Linguistics or Applied Linguistics approach.

For example, with regard to comic books, Walshe (2012, 2013) examines the representation of Irishness through features American writers *perceive* to be salient in IrE in the Marvel and DC¹⁴ comic book universes. On the one hand, he studies (2012, pp. 285-7) a corpus of 150 Marvel comic books, identifying the features the books portray as being indexical of IrE, yet notices a salient use of stereotypical Stage Irish features (e.g. *begorrah, musha, top of the mornin' to ye*, etc.; see section (1.4.) for a definition of the Stage Irishman stereotype and other features comprised in its linguistic characterization), and confusion with Scottish English items (e.g. negation '*cannae, dinnae*, etc.', or lexis '*laird, lassie, bonnie*, among others). Contrastively, his examination (2013, pp. 114-119)

¹⁴ DC stands for *Detective Comics, Inc.*; the original name of this comic book company.

of the DC corpus (also including 150 comic books) evidences fewer occurrences of Stage Irish and Scottish English confusion, indicating the use of actual IrE lexis (e.g. *aye, lad*, etc.; oaths '*faith, by my troth*'; or expletives *bloody, bleedin'*, among other items). Nevertheless, the DC data also highlights the use of grammatical features (e.g. non-standard verb forms or lack of subject-verb agreement) which are either not exclusive to IrE or simply not IrE items.

Walshe (2011, p. 128) also examines the language used in the popular Irish sitcom, *Father Ted*, to determine whether it provides an 'authentic' representation of this variety. He finds that, contrary to what would be expected, stereotypical Stage Irish features such as *begorrah* or *top of the morning*, among others, are absent from the show. Instead, he discovers a number of grammatical, lexical, and discourse features which, despite not being as regionally recognizable by the British audience the show was tailored for, manage to transmit a valid representation of Irishness and Irish discourse on screen (*ibid.*, p. 132). Walshe contends that while the use of these items serves to lend a layer of authenticity to the show, they can also "be subverted for humorous purposes" (*ibid.*, p. 144). His study finds that common IrE features which are used across all seasons, albeit rarely, include fronting¹⁵, negative imperatives in the progressive form (e.g. Daddy tells me don't be looking at him (Ryan, 2012)), irregular verb forms, or the use of second-person plural pronouns like *ye, yiz*, and *youse* (Walshe, 2011, p. 131). The most recurrent grammatical items, he notes, are salient features of this variety like embedded questions (i.e. use of non-standard word order in *if/whether* indirect questions like 'I wonder *is he dead*'), the use of first-person *will* for *shall* in all contexts, unbound reflexive pronouns—in particular *himself* and *herself* in cases such as "I went out with *himself* last night" (*ibid.*, pp. 132-33). Walshe (*ibid.*, p. 136) also notices that while discourse pragmatic markers often associated with IrE (e.g. *arraah, what, like, here, or lookit*) are slightly present, the more frequently-used ones are those which are perhaps not as regionally distinctive, yet are still recognizable features of Irish discourse, like *sure*, which he describes as the most typical and highly versatile discourse pragmatic marker in IrE (*ibid.*, p. 137). Religious expressions like *Holy mother of God, Lord God Almighty!, Oh God!, or By God!* (*ibid.*) are also abundantly used in a pragmatic way to convey "shock, excitement, surprise,

¹⁵ *Fronting*: focusing device which gives prominence to specific parts of a sentence/utterance by moving those to front position. See, for example, the following fragments taken from CoFIrE: "Trouble? [...] Good as gold, she was, the same girl" (Keegan, 2010) or "Six classes there are, all and every one of them in the schoolhouse" (Ní Dhuibhne, 2003).

impatience, anger, or just as fillers” (*ibid.*, p. 138; see also Stenström, 2006, p. 6). Unlike the grammatical and pragmatic items, which infused realism to the series, Walshe (2011, p. 139) proposes that the lexical features create humor due to their frequency and context of use. He cites taboo words as a prime example, given their status as the most frequent lexical item in the show, and a distinctive feature of Irish speech; a fact which is also corroborated in this thesis (see chapter (6) for a detailed exploration of intensifying *fucking*). *Feck*, a salient IrE milder form of FUCK, appears to be quite a frequent taboo word, along with variants *fecker* and *feckin’* (*ibid.*, p. 140). Other recurrent lexical items which are highly representative of Irish discourse in the show (*ibid.*, p. 144) are the use of *lad(s)*, lexicalized *eejit* (i.e. idiot), *your man/one*, or adjective *grand*, among various others.

Despite the evidently robust tradition of research into the representation of IrE orality in fiction briefly outlined above, the number of studies that use CS as their main methodology is considerably smaller, and, as mentioned above, they often use Corpus Linguistics or Applied Linguistics as their main frameworks. However, CS investigations have begun to flourish in the context of fictionalized IrE portrayals in recent years. The following section will provide an overview of those studies, all of which incorporate both *intra/inter-textual* approaches, identifying also various research trends that exist in this field.

2.5.1. CS Investigations of Fictional IrE: Research Trends & Existing Studies

Studies on the fictional representation of IrE appear to mirror the three research trends already outlined in section (2.4.). Thus, the main topics (outlined on Table 2.1. below) could be divided into: (1) analyzing the reproduction of IrE items in fiction, (2) exploring the correlation that may exist between fictional vs spoken IrE with respect to authenticity (often pertaining to use, pragmatylistic functions, and identity indexing and/or styling through fictional IrE discourse), and (3) examining the fictional reproduction of these IrE items with regard to their characterization/identity indexing value.

Topics	Subtopics
1. Reproduction of IrE items in fiction.	
2. Authenticity of the portrayal	a) Style of an author(s)
3. Characterization/identity indexical value.	b) Specific items in a work(s) of fiction c) Identity indexical value of items in (specialized/large) corpora

Table 2.1. Broad research topics and subtopics in CS studies of fictional IrE orality in fiction.

Topics (2) and (3) in Table 2.1. above may also include subtopics, such as (a) examining the style of an author(s). An example would be Amador-Moreno’s (2006) study of linguistic style in Patrick McGill’s early novels. Subtopic (b) would involve the analysis of specific items in a work(s) of fiction. Amador-Moreno & O’Keeffe’s (2018) study, for instance, would fall into this category as it explores the use of the *Be + after+ V-ing* structure looking at its various meanings, pragmatic uses, but also checking for issues of ‘authenticity’ by contrasting its use in a fiction corpus against a corpus of real spoken IrE. They use historical and contemporary literary texts from Hickey’s *Corpus of Irish English*, which, as part of his *Corpus Presenter* suite, comprises a small collection of (mostly dramatic) Irish Medieval through 20th-century texts (see (3.1.) for more on Hickey’s corpus). In particular, Amador-Moreno & O’Keeffe (2018) analyze 5 historical works from 1800 to present and compare them with 49 contemporary works from 20 Irish writers from 1951-2007. Their data is contrasted against the *Limerick Corpus of Irish English* using *concordance line* searches which are qualitatively analyzed. They find this structure is “not simply a feature of fictional discourse, used to caricature the Irish, but a characteristic element of the English used in Ireland throughout time” (*ibid.*, p. 62). They also discover that aside from traditional functions (e.g. indicating immediacy), its use may also dramatize the narration of events by incorporating the speaker’s attitude. Finally, subtopic c) involves investigating the use of linguistic features in a (specialized or larger) corpus with regard to the indexing of identity (as is the case of this thesis), although I find that, in most cases, (b) and (c) tend to intertwine and become the most predominant subtopic.

The blend of subtopics (b) and (c) outlined in Table 2.1. above often raises questions of ‘authenticity’, bringing on the discussion of the role the audience/reader has

in the perception of the identity that is being weaved into the fictionalized portrayal of IrE. Amador-Moreno (2010) and Hodson (2014) make the case that the perception of the ‘authentic portrayal’ of IrE, in this case, by an audience depends on whether the text is intended for an *inter-cultural* context (i.e. an international audience that is only able to identify the Irishness of the characters through their fictional use of *general* linguistic features), or *intra-cultural* context, which “assumes shared content and knowledge of the variety that both speaker and hearer are able to understand” (Amador-Moreno, 2010, p. 106). Amador-Moreno (*ibid.*) also proposes the existence of a type of trend within contemporary fictional representations of IrE that responds to the latter. This is reflected in the variety of studies that have been carried out in this field, encompassing areas such as literature, comic books (described in section (2.5.)), (radio) advertisement, telecinematic scripts, as well as comedic performances and YouTube videos; all part of fictional representations of this variety which have been conducted using Corpus Linguistics and some parts of CS. Although not exhaustive, Table 2.2. outlines some of these (mostly Corpus Linguistics-based) studies dealing with authenticity, identity indexation issues (i.e. subtopics (b) and (c)), or a blend of both.

AREA	STUDIES	FOCUS OF STUDY
	Amador-Moreno (2012a, 2015, 2016)	Non-standard language and pragmatic markers as stylistic hallmarks in Paul Howard’s <i>The Curious Case of the Dog in the Nightdress</i> .
Literature	Amador-Moreno & Terrazas-Calero (2017, 2022)	Form, pragmatics, and identity value of New Intensifying SO in a corpus of three <i>Ross O’Carroll-Kelly</i> novels.
	Terrazas-Calero (2020)	Form, pragmatics, and identity value of non-standard quotatives and pragmatic markers (<i>like, fucking</i>) in the <i>Corpus of Contemporary Fictionalized Irish English</i> .
Comic books	Walshe (2012, 2013)	Portrayal of Irish characters through perceived salient speech items in two corpora of 150 Marvel universe and 150 DC comic books respectively (see section (2.5.) for a description).

(Radio) advertisement	O’Sullivan (2013, 2015, 2018, 2020)	Stylization of social, hybrid Irish identities through accent variation in her <i>Irish Radio Advertisement Corpus</i> .
Comedic & online performances	Vaughan & Moriarty (2018)	Indexation of Limerick ‘inner-city’ identity through IrE in a corpus of <i>The Rubberbandit</i> ’s comedic performances on YouTube.
	Murphy & Palma-Fahey (2018)	Construction of the “Irish Mammy” identity in a specialized corpus of TV sitcom, <i>Mrs Brown’s Boys</i> .
Telecinematic scripts	Walshe (2011)	Authenticity of rendition of IrE language use in <i>Father Ted</i> scripts (see section (2.5.) for a description).
	Walshe (2009, 2016, 2017)	Portrayal, identification of key items, and level of authenticity of IrE rendition in cinematic Irish corpora.

Table 2.2. Summary of (mostly CL-based) studies of fictional IrE portrayal dealing with subtopics (b) and (c).

In literature, the majority of investigations that have included CS elements have been devoted to examining non-standard language and discourse pragmatic marker use, investigating their indexation of specific Irishness. For example, Amador-Moreno (2012, 2015, 2016) analyzes the language used in Paul Howard’s (2005) *The Curious Incident of the Dog in the Nightdress*. Part of the critically acclaimed, annually best-selling Ross O’Carroll-Kelly series, Amador-Moreno looks at features like non-standard quotatives (e.g. *go, be like, be there*) and discourse pragmatic markers *like* and *sure*, positing that the phonetically rendered portrayal of IrE evidenced in the novels is one of Howard’s stylistic hallmarks. This would serve the purpose of stylizing two types of Dubliner identities: the upper-class, D4-accented, Southside Dubliner (embodied in the character of Ross) as opposed to the Northside Dublin speaker of more locally-restricted IrE features (see sections (8.2.) and (8.2.1.) for a detailed explanation of the series and its distinctive linguistic style(s)). Amador-Moreno & Terrazas-Calero (2017; 2022) further investigate this stylization in a corpus of three Ross O’Carroll-Kelly novels, focusing on the use of New Intensifying *So* (e.g. ‘Ross, you are *so*¹⁶ a snob’ (Howard, 2013)). This is a non-standard intensifier (which shall be studied at length in chapter (8)) which has never

¹⁶ Emphasis in the original.

been traditionally connected to IrE, yet is popularly perceived as having originated in the United States, where it is indexical of the (often negative) identity stereotype of the California ‘Valley Girl’. This personae is representative of young females from San Fernando Valley, California, who are often perceived as vain and dizzy (see (7.5.1.) for more on this stereotype and its linguistic construction). Using *concordance* and *collocation* searches, as well as pragmatic functional and emotive marking, and taking a sociolinguistic approach, Amador-Moreno & Terrazas-Calero (2017; 2022) find that this particular feature is only ever present among the Southside Dublin cast in Howard’s books, with Northsiders preferring to use traditionally IrE distinctive tag-*so*¹⁷ instead.

With regard to (radio) advertisement, O’Sullivan (2013, 2015, 2018, 2020) studies the stylization of social, hybrid Irish identities through the use of accent variation in her *Irish Radio Advertisement Corpus* (IRAC). IRAC is a corpus of radio advertisements containing 200 ads broadcast on RTÉ Radio 1, which is subdivided into 5 subcorpora according to the time they aired (i.e. 1977, 1987, 1997, 2007, and 2017). In her 2020 (p. 241) publication, she provides a diachronic analysis of the corpus, discovering the substitution of the prestige Standard Southern British English accent (i.e. ‘Received Pronunciation’) in the earlier decades in favor of Advanced Dublin English, also known as *non-local* Dublin English (in Hickey’s (2005) terminology) or ‘D4 accent’ (see (8.2.1.) for a detailed explanation of this accent and its sociocultural associations in Dublin). Since the Advanced Dublin English accent contains elements from IrE, American English, and British English, O’Sullivan (2020, p. 241) posits that its increased use in the ads as an audience-designed style reconstructs Irish identity as “a more hybrid, cosmopolitan one, rather than one based on the colonizing variety”.

Despite drawing less on language representation and CS, Vaughan & Moriarty’s (2018) examination of the indexation of Mid-western Irish identity in comedic performances also merits attention. Concentrating on the representation of linguistic ideologies and sociocultural realities, they investigate the use of Limerick¹⁸ ‘inner city’ English with a focus on the representation of linguistic ideologies and sociocultural realities in a corpus of performances by Limerick-city, comedy hip-hop duo, *The Rubberbandits*. The corpus includes (*ibid.*, p. 22) their performances (e.g. songs,

¹⁷ Example of tag-*so*: ‘You probley shoultn’t have any mower, Rosser. You’re fooked, so you are!’ (Howard, 2014, p. 39).

¹⁸ Limerick City is the capital of County Limerick (in the Mid-West of Ireland).

interviews, tweets, sketches, etc.), meta-performances (newspaper and online articles as well as miscellanea about them), and audience/media commentaries (YouTube comments, comments on their Facebook page). Using *frequency/keyword* lists and *concordance* analysis of the performance corpus, and contrasting them with the *Limerick Corpus of Irish English*, they find linguistic items (e.g. *yokes*, *hash*, and *joint*) that allude to the drug culture that is linked with the concept of the *knacker*¹⁹ and with Limerick inner-city (*ibid.*, p. 27). Furthermore, their study reveals the reproduction of the Limerick English accent in words like *daycent* (decent) or *beiting* (beating), as well as terms of address and more drug references. They argue that the validity of their representation of local voice is proven by the audience's reiteration of these words and catchphrases, including the item *yurt* (*ibid.*, pp. 28-30), and that their parodying stylization of the *knacker* not only connects inner-city voice/stereotype, but also disrupts social schemata.

With regard to telecinematic discourse, Murphy & Palma-Fahey (2018) explore the construct of the Irish "Mammy" identity through the forms *mammy*, *mother*, and *Agnes* as used in a specialized, 40,000-word corpus of the TV sitcom *Mrs Brown's Boys*. They use *keyword* analysis to determine the social and cultural concepts surrounding the "Mammy" stereotype, which revealed five themes: *love*, *control and interference*, *sex*, *sexual orientation* and *culture* (*ibid.*, p. 17). These themes are both constructed and deconstructed by means of utilizing traits that were traditionally associated with the stereotype of the Irish "Mammy", while also breaking away from those expectations, thus acknowledging yet shattering "the assumptions and traits [...] central to characterizing the stereotype" (*ibid.*). For example, with regard to the theme of *love*, they find that Agnes, the protagonist, shows fierce love for her children yet also emotional repression, which is in line with this particular stereotype (*ibid.*, pp. 17-19). However, the stereotype is also subverted by portraying her as confrontational when her children are negatively commented upon by others. Murphy & Palma-Fahey posit that language also functions as another subversive element, especially when presented in the form of Agnes' extensive use of "sarcasm, inappropriateness, rudeness" and taboo language (*ibid.*, p. 20). This type of linguistic performance, they believe, indicates her social status as a resident of Dublin's

¹⁹ *Knacker* or *Skanger*: derogatory IrE slang terms for lower-class "young, uncouth youth" (Dolan, 2006, p. 199) from a high criminal rate area often recognized for their outfit (e.g. casual sports), and poor education level. See section (8.2.) for a more detailed definition.

inner city, an area traditionally known for its high level of unemployment, poverty, crime, but also for its community closeness.

Drawing closer to CS, further work on telecinematic discourse includes Walshe's (2009, 2016, 2017) exploration of the portrayal of IrE in telecinematic Irish corpora. His (2017) study offers a comparative analysis of the use and distribution of IrE features in a new corpus (in the sense of a collection of non-annotated texts) of Northern Irish movies, set during the Troubles (1960s-1998) and their aftermath. This corpus provides a more complete view of the cinematic representation of IrE, which complements his (2009) corpus of 50 movies produced between 1939-2007 in the Republic of Ireland, spanning a period setting between the late 19th century and present-day Ireland. The 2009 corpus was used to investigate a variety of Southern IrE features diachronically and synchronically. In his study of Northern IrE, Walshe (2017, p. 286) transcribes only sentences that contain recognizable IrE items, and contrasts his findings with the spoken component of ICE-Ireland (for a definition of this corpus, see section (3.9.)). He finds that his corpus reveals the preference for some lexical items which appear to be regionally bound to the South while being absent from or being used significantly less in the northern movies (e.g. *your man/your wan*, *yoke*, and *give out (to someone)*). Contrastively, *to catch oneself on*, which is described in the literature as an Ulster expression, only appears in the Northern corpus (*ibid*). Regional-preference is also evident in the use of intensifiers like *wild* and *rightly*, which occur only in the north, while *fierce* and *altogether* are used only in (in the case of the former) or occur much more often in the southern movies (2017, pp. 290-291). Southern-bound discourse pragmatic markers such as *come here* (i.e. appealing for attention) or *lookit* only appear in the southern movies, while *how's about you?* and vocatives *big man* and *big lad* occur only in the northern films, yet are absent from the reference corpora (*ibid*, 291-293).

Walshe (2017, p. 293) proposes telecinematic discourse, not only as a source representative of real regional variation distribution, but also of its use in context. An example is the discourse pragmatic item *sure*, which is slightly more prominent in the southern movies. In terms of context, his results contrast with Kallen's (2013, p. 198) whose findings indicated clause-initial *sure* was the most frequent context in both regions, followed by phrasal, and tag *sure*, with clause-final *sure* being the least frequent in both north and south. Walshe's (2017, p. 295) telecinematic corpora, however, reveal tag *sure*

as being more frequent in the southern movies, which contrasts with Kallen's (*ibid.*) previous findings.

Perhaps one of the most compelling parts of Walshe's studies is his argument in favor of using telecinematic discourse (in particular in the context of IrE) as evidence for real-life, language use, but also his call for it to be used to make way for future inquiry into linguistic research. Drawing on Walshe (2012), Palma-Fahey (2015, p. 350) also defends televisual discourse (particularly in soap operas) and its authenticity, maintaining that fictional dialogues are meant to be spoken in media genres such as films, TV serials, and soap operas. While I do agree with that statement, I do not believe that applies *only* to mediatized/telecinematic discourse. My contention is that the same can be said of 'prototypical' fictional (literary) dialect-dialogues. The difference between fictional *televisual* and fictional (*literary*) being that, with the latter, the audience has to *read* the dialogue to be able to *hear* it. In essence, the reader has to be almost an actor that receives the written directorial/authorial instructions on how to (mentally) 'perform' the fictional dialogue to be able to correctly interpret the message encoded in the literary representation (of IrE in the case of this thesis). That being said, I take the aforementioned investigations as a basis on which to expand.

2.6. Concluding Remarks

This chapter has surveyed the literature that informs the framework and methodological basis of this thesis, as well as it has outlined certain gaps in the context of IrE research that need to be explored.

On the one hand, the chapter introduces this thesis as an inter-disciplinary study of fictional representations of IrE in a corpus of *contemporary* Irish fiction, which also explores what type of modern Irish identity they index in it. While it uses an eclectic array of methodologies to investigate this subject, ranging from Corpus Linguistics, to Pragmatics, and Sociolinguistics, the chapter presents Corpus Stylistics (CS) as this thesis' main framework. A detailed explanation of what CS is, its advantages, and different uses was also provided, along with a survey of some of the most prominent studies conducted on fictional speech representation in other varieties and in the context of Ireland. In addition, it also explains the link that exists between CS and Corpus

Linguistics, describing what corpora are, how corpora feature in CS, how to design and create one so that it is as valuable for one's (CS) project as possible, and which Corpus Linguistics tools can be utilized in CS studies. The review has also shown that while smaller, specialist (as opposed to general) corpora have been developed for the study of different linguistic features and/or varieties of English (including IrE) in CS studies, the most prominent of which are surveyed here, none of the existing studies serve the focus of this thesis. Indeed, in the survey of literature regarding CS studies in the context of fictionalized IrE, where this chapter has also outlined various research trends that seem prevalent in the field, it is clear that although much work has been conducted on general aspects of IrE, and/or individual authorial style, as represented in literature/fiction, most studies did not take a CS approach, nor did they examine frequency of production, authenticity of representation, or the indexicalization of contemporary Irish identity (particularly through pragmatic items) in a larger corpus of *contemporary* fiction.

Thus, given the lack of CS research addressing the issues that are key to the research questions (see (1.2.)), this thesis will investigate whether contemporary fictionalized IrE is a valid representation of real-life linguistic use in present-day Ireland. To aid this process, the thesis endeavors to cover the three main research trends that are generally studied in CS studies and which have been outlined in this literature review, namely (1) linguistic variation, (2) authenticity and speaker identity indexing, and (3) the use of fictional reproductions for characterization purposes, all of which align with the research questions (hereafter RQs, see (1.2.)). Thus, (1) linguistic variation will be explored by investigating how fictionalized Irish English is being used in contemporary Irish fiction, and examining which features are the most prominently represented in the corpus texts (RQ1), their form and pragmatic functions. To further examine empirical evidence as to whether contemporary fictionalized IrE provides an (2) authentic representation of real-life usage and speaker-identity indexation, the findings will be contrasted against a reference corpus of spoken IrE, namely the *Limerick Corpus of Irish English*, and against the spoken component of the *BNC2014* (RQ 1.a.), which may offer evidence of any potential linguistic developments that may have taken place or may be currently in progress. Finally, the thesis will also (3) explore the connection that exists between the fictional use of these features for characterization purposes, as I will investigate what type of speaker identity indexing/stylizing value (RQs 2 and 3) they

might encode with regard to the sociolinguistic factors of age, gender, geographical location, and social class.

To do this, and as mentioned throughout this chapter, Corpus Stylistic and standard Corpus Linguistics analytic methods (e.g. *word frequency* lists, *concordance* and *collocation* analysis) will be utilized on the specially curated CoFIrE corpus. As Chapter 3 will explain in detail, much consideration is given to the annotation of the corpus and the process of data analysis whereby results generated from the use of CL tools are given more qualitative analytic treatment, including pragmatic and emotion classification (e.g. does the use of the feature transmit anger, insults, or compliments, etc.?).

3. METHODOLOGY

This chapter will provide a detailed explanation of the linguistic and non-linguistic criteria that were set in the process of designing and building the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE), as well as it will detail how the corpus data and metadata were stored. The *sampling frame* and techniques used in CoFIrE will be explained in detail, along with some of the limitations of using *contemporary* literature. The chapter will also survey the authors and texts that were included in the corpus and the reason for their inclusion. An overview of the different types of existing annotation systems will be provided, but the chapter will focus on explaining at length the original and manual annotation system that was implemented in CoFIrE, as well as the process of text preparation the texts had to undergo before annotation. A description of the spoken English corpora (i.e. the *Limerick Corpus of Irish English* and *BNC2014*) which have been selected to serve as comparison for the findings of the fiction corpus will also be provided. Finally, one of the authors in the fiction corpus, Paul Howard (author of the *Ross O'Carroll-Kelly* series whose style is better described in sections (8.2.) through (8.2.1.)) was found to be an outlier with regards to spoken IrE representation in CoFIrE at annotation and analytical stage. This led to the creation of a separate corpus which was used to investigate fictionalized IrE as a case study on one author. Thus, section (3.9.) provides a detailed explanation of the design and compilation of the *Ross O'Carroll-Kelly Corpus*.

3.1. Rationale for creating the *Corpus of Contemporary Fictionalized Irish English*

In order to examine representation of IrE and its use in the portrayal of modern Irishness in fiction, the thesis needed a corpus of *contemporary* (i.e. 21st century) Irish fiction to answer the research questions outlined in section (1.2.) (section (3.2.2.) offers a detailed definition of what was taken as *contemporary* in this thesis). As explained in section (2.2.), in order for a corpus to be as useful as possible for the research questions, a series of restricting criteria must be established at corpus design stage. The first step, therefore, was the search for an existing corpus which may fulfill those parameters.

While in the past academics seemed to focus on the study of individual, often lexical or phonological items (as discussed in (1.4.1.)), recently some scholars have undertaken a more systematic examination of IrE usage by compiling specialized fiction corpora that lend themselves to diachronic analyses. For example, Hickey's (2003) *Corpus Presenter* suite (already introduced in (2.4.)) contains the *Corpus of Irish English*, which consists of a small collection of Irish Medieval through 20th-century texts, most of which are dramatic in nature, but which allow for stylistic and diachronic studies (Amador-Moreno & Terrazas-Calero, 2022). Another potentially useful database could be Cesiri's (2012) *Corpus of Irish Fairy and Folk Tales*, which includes transcriptions of oral stories by 19th-century, Irish peasant storytellers. Despite the spokenness of the data, Cesiri's corpus does contain IrE fiction, "rescu[ing] a type of material that is very close to orality and that had not been studied exhaustively before" (Amador-Moreno & Terrazas-Calero, 2017, p. 256). The compilation of Connell's (2014) *Corpus of Hiberno-English Literary Dialects* (CHELD) also contributes another database for the study of fictional IrE. In this case, CHELD incorporates 50 plays by 17 Irish authors produced in the 20th century by the Abbey Theater; the country's national theater which was key in the revival of artistic culture in Ireland. While excluding plays produced by other theaters at the time, CHELD allows for a detailed, diachronic analysis of 20-century IrE, as the corpus is annotated including phonology and syntax, as well as for metadata such as decade of production, geographic area of setting, and author gender, or character age/gender.

Despite the usefulness of all three corpora for the study of IrE across different time periods, none of them contained material that aligned with the restricting parameters I set at design stage for my corpus to be as useful for the thesis research questions as possible, and which are explained at length in sections (3.2.) through (3.2.2.) below. Therefore, a new corpus had to be built: the *Corpus of Fictionalized Irish English* (hereafter CoFIrE).

3.2. Designing the *Corpus of Fictionalized Irish English*

While analyzing a ready-made corpus can be a complicated task due to accessibility, copyright, privacy and distribution issues (Lee, 2010, p. 108), designing and compiling a

synchronic corpus of contemporary texts from scratch is an arduous, complicated, and time-consuming task. In addition, and as explained in (2.2.), it forces the corpus-builder to create specific criteria which need to be meticulously followed so that the data is as useful as possible for the needs of their investigation. Thus, the design of the *Corpus of Fictionalized Irish English* (CoFIrE) had to be carefully planned in advance. The following sections discuss in detail the criteria which were followed in the design of CoFIrE (outlined in (3.2.1)-(3.2.2.)), including the *sampling frame* and sampling technique (see (3.2.3.) and (3.2.4.)), the actual corpus size, texts, and balance (discussed in (3.4.)), the selection of authors and books (which shall be explained in (3.4.) through (3.4.8.)), the process of preparing the texts for analysis (see (3.5.)), the original annotation system created and implemented in CoFIrE (explained in detail in (3.6.) through (3.6.1.)), how the corpus metadata was stored and managed (see (3.7.)), and the corpora of spoken English used to compare and contrast the fiction data against (sections (3.8.) through (3.8.2.)).

3.2.1. CoFIrE design criteria: Establishing the ‘variety’

Before beginning the process of building the corpus, the researcher must first establish *what* the *variety* the corpus will be representative of is. Since *Corpus of Fictionalized Irish English* (CoFIrE) was created as a specialized database with the purpose of catering to the thesis RQs (see (1.2.)) which aim at investigating the reproduction of IrE orality in Irish fiction, and how its use may index modern Irishness, the first thing to be delimited and established at design stage was the *variety* the corpus would be representative of. In this case, the *variety* was closely connected to another non-linguistic parameter also set at design stage (which are outlined in Table 3.1. in section (3.2.2.) below), which is *location of publication*. Thus, the latter parameter was restricted to the Republic of Ireland. It was, therefore, decided that the *variety* the corpus would investigate would be Southern Irish English, corresponding with the variety of English spoken in the Republic of Ireland. The exclusion of Northern Irish English (NIE) was based on the fact that while both Northern and Southern IrE share Irish substratum influence, they are often examined separately in Dialect Studies due to the influence NIE received from Scots and Scottish English. Indeed, their contact led to the development of Ulster Scots in the northern

province, and to NIE having features which differentiate it from Southern IrE. These include different prosodic and intonational patterns, or phonologically shared traits with Scots and Scottish English like the lowering of short front vowels or fronting of /u/ to [-u], among others (see Hickey, 2007, p. 15 for more on Northern vs Southern IrE phonological features, and Corrigan (2010) for a monographic volume on Northern IrE)²⁰.

Following on from this, the features coded in CoFIrE include grammatical, lexical, and pragmatic items which were identified in different sources as idiosyncratic to and/or as non-standard Southern IrE (see Table 4.1. in section (4.1.) for an overview of the amount of grammar, lexis, and pragmatic items coded in the corpus and the number of occurrences per category). These sources include grammar manuals (i.e. Filppula, 1999; Hickey, 2007; Kallen, 2013), dictionaries (Dolan, 2006), and other academic works (Walshe, 2009; Amador-Moreno, 2010b) which provided an overview of this variety of English. Online etymological searches on websites like O’Byrne’s (2006) *Dublin Slang Dictionary and Phrase-Book* or *Irishslang.info* (2016) were also used specifically for the purpose of consulting certain lexical and slang terms (e.g. ‘to be *up the duff*’²¹ is Irish slang for *pregnant*). The process of coding features was also iterative since some linguistic features stood out during the reading of the texts as part of the text selection process (see (3.2.3.) and (3.2.4.) for a detailed overview of this process). These features were, subsequently, checked in the sources mentioned above to determine whether they were typical IrE features.

Examples of the features coded in the books include grammar items like the ‘after’ perfect (*‘I’m after calling him’* for Standard English Present Perfect) or order inversion in embedded *if/whether* questions (e.g. *‘I asked was he coming back early’*). Lexical features like suffix *-een* for diminutives, *spud* (potato), *shebeen* (illicit bar without a license), *auld* (old), common Irish phrases/words such as *mas dhea* (as if; yeah right), as well as quasilexicalized phrases like *your man* inter alia were also included, along with pragmatic elements such as *Jaysus*, *bleeding*, *sure*, among many others (see Appendix 1 for a full list of all annotated items in CoFIrE). Phonological features were, however, excluded as their investigation would have required a separate study. In including features

²⁰ Scholars interested in the study of the representation of NIE in fiction may wish to create a corpus of contemporary fictionalized NIE which replicates and is complementary of CoFIrE, thus providing a broader look at the fictionalized representation of both varieties of English in the island of Ireland.

²¹ *Up the duff* and variants (e.g. *Up the Damien*) were coded because they were identified as Irish slang on sources such as *Irishslang.info*.

identified in academic sources as idiosyncratic to and/or non-standard in IrE, the corpus attempts to ensure the maximum representativeness of the *variety* it compiles.

During the reading and annotating stage of the texts it became obvious that the books included many non-standard, pragmatic items which appeared in addition to traditionally IrE pragmatic features. Examples of these types of non-necessarily traditionally IrE items include FUCK (and all its variants), discourse pragmatic marker *like*, or non-standard quotatives (e.g. *be like*), among others which are not necessarily and/or popularly associated with IrE but with other varieties or are present across varieties. While the pragmatic items that were distinctive to IrE were indeed annotated, it was decided that CoFIrE would also code the non-standard, traditionally non-distinctive IrE pragmatic items. This was done to a) address the need for more variational IrE pragmatic studies, especially observing IrE pragmatic use in large corpora, which was outlined in (1.4.1.), b) to answer RQ3 by gaining a better understanding of the type of *modern* Irishness which may be undergoing indexation through the use of universal items used across varieties, and to c) explore the potential adoption and/or adaptation of globalized features into the Irish context (see, for example, the analyses of *New Intensifying So* and intensifier *Totally*, in the CoROCK corpus offered chapter (8)). Finally, by including these pragmatic items which have not traditionally been linked to IrE, the thesis would be able to audit whether the CoFIrE authors utilize more traditionally, regional-distinctive items (e.g. *after* perfect, *begorrag*, *arrah*, etc.), which are often associated with the negative stereotype of the Irish Paddy (described in (1.4.)), or whether their linguistic portrayal includes more globalized items, thus conveying modern Irishness by distancing from the more traditional and stereotypical identity.

3.2.2. CoFIrE design criteria: Other parameters

Once the *variety* had been properly identified, I followed previous researchers' advice (e.g. Sinclair, 2005; Biber, 1993, p. 243; McIntyre & Walker, 2019, pp. 65-66) and created a series of non-linguistic parameters (outlined in Table 3.1. below) to aid in further identifying and delimiting the *variety* under study, as well as in the subsequent development of the *sampling frame*, which will be discussed in sections (3.2.3.) and (3.2.4.).

CoFIrE Restricting Parameters	
Variety	<i>IrE features</i>
Location	<i>Republic of Ireland</i>
Author nationality	<i>Irish</i>
Mode	<i>Written</i>
Genre	<i>Fiction</i>
Text type	<i>Novel/short story</i>
Date	<i>Contemporary</i>

Table 3.1. Restrictive parameters in CoFIrE.

As illustrated in Table 3.1. above, these parameters established the mode (*written*), genre (*fiction*), text type (*novels & short story collections*), location of publication (*Republic of Ireland*) and date (*contemporary*) of the texts to be included in the CoFIrE. Before explaining these parameters in more detail, it is important to describe what was taken as *contemporary* fiction here. This project draws on the definition of the word *contemporary* outlined in the *Oxford Dictionary*, which refers to something “belonging to the same time” or “belonging to the present time” (*Oxford Advanced Learner’s Dictionary*, 2010, p. 325). Thus, the thesis takes *contemporary* fiction to be literature published during the 21st century.

Having delimited the time span, the next issues were genre and text type. In this case, it was decided that the scope of *fiction* genres would include *novels* and *short story collections*. Given that they tend to be rich in dialogue, and often have large word counts, the inclusion of these genres offered the potential for greater numbers of IrE dialectal or non-standard features of orality. Plays, however, were excluded because their word counts tend to be smaller.

Word count was in direct connection with *how much sample text* would be included in the corpus. In that regard, I followed Sinclair’s (2005) recommendation to include full texts, rather than fragments. In addition, and drawing on McIntyre &

Walker's (2019, p. 73) idea that the larger the amount of text (or *variety*) included in the corpus (i.e. in this case, the larger the amount of IrE dialectal/non-standard features the texts contained), the more representative the corpus would be, it was decided that in order for a book to be included, it needed to contain as many features of the *variety* (i.e. Southern IrE) as possible (see section (3.2.1.) above for examples of annotated features). In other words, it needed to have large amounts of IrE features, preferably in the form of (inter)character dialogue, as dialogue represent orality in written texts.

It is worth noting that features which appeared in narrative passages were only included if the narrator was a) homodiegetic (i.e. both narrator and active character in the story), b) autodiegetic (i.e. narrator and protagonist of the story), as in the case of the *Ross O'Carroll-Kelly* series (see (3.9.), (8.2.), and (8.2.1.) for detailed descriptions of this series), and/or b) Irish. Thus, if the text was 'rich' in IrE dialect, then it could be added to CoFIrE. An example of a dialect-rich exchange is offered below. This conversation takes place in a shared taxi ride between a working-class male in his late 20s (i.e. Alan) and a girl the narrator describes as a "painfully fashionable south county darling" from Blackrock; a suburb in the affluent Southside Dublin area.

"Alan. Hi. Oh, thanks for the lift by the way."
"Not at all, sure it's bloody freezin' out there. Couldn't leave ye standing there all night. At a party were ye? Office do, yeah?"
"Not so much an office thing. Clients bringing us out for dinner."
(Ruane, 2003, p. 10)

Notice how the higher status speaker (i.e. the girl) evidences no dialectal traces. Instead, it is Alan, the lower status character, who uses all the regionally distinctive items (underlined), some of which are identified in academic sources as characteristic of IrE (i.e. *ye*, discourse pragmatic marker *sure*), while others are not necessarily only Irish though commonly found in this variety (*fronting*²²), or are universal pragmatic elements (i.e. *bloody*).

²² *Fronting*, also known as *topicalization* (Filppula, 1999, pp. 260-270; Hickey, 2007, p. 267), is a device with which the element that is to be emphasized in the sentence is moved to the front of the main clause. For example, "Six classes there are, all and every one of them in the schoolhouse" (Ní Dhuibhne, 2003).

Novels and short stories were also selected in order to answer RQ3 (see (1.2.) for all the RQs) regarding the indexation of modern Irishness through speech. To do this, the texts needed to provide as much sociolinguistic information (i.e. age, gender, social class, and location) about the ‘fictional demographics’ (i.e. speakers/characters) as possible. This type of metadata is, however, often missing from plays, which is also why they were excluded, yet it is usually commonly a) found in narrative passages through descriptions and/or (inter)character explanations, or can b) be inferred by the reader from the information provided in the text (e.g. information regarding their age, occupation, their outfits, their location, etc.) in the case of novels and short stories. For example, the homodiegetic narrator in fragment (a) below, a construction foreman, describes an exchange with an unemployment office clerk who informs him that his previous employer scammed him out of unemployment benefits. Notice how in the same passage, the narrator explicitly tells the audience the background (i.e. unemployed, working class) of other characters who will also be speakers in the book. Similarly, the narrator in (b) introduces readers to the next speaker whom he sarcastically describes as an entitled, affluent Southside Dublin young man.

- (a) Did you never look for a P60 from your employer? A what, now? You're some fool, she said with her eyes [...] I think she started to feel sorry for me then. But when she looked at the line of goms²³ behind me - Seanie Shaper, innocent Timmy, fat Rory Slattery and the rest of the boys, all clutching their dirty payslips - she started to feel more sorry for herself. (Ryan, 2012)
- (b) At this stage, the south county UCD²⁴ frat boy superiority complex sets in, and he gives me a snide my-daddy's-a-partner-in-KPMG kind of look. (Ruane, 2003)

It must be pointed out at this stage that there were times when not all factors concerning a speaker's age, gender, location, or social class could be gathered or inferred from the text (see (3.7.) for a detailed explanation of the social class distribution system followed in the annotation of CoFIrE). In those cases, whichever factor(s) was missing was left blank in the database which was created to store the corpus metadata (this shall be discussed at length in section (3.7.)), but the given linguistic feature was still annotated if

²³ *Gom*: IrE slang for ‘fool, idiot’.

²⁴ *University College Dublin (UCD)* is located in the affluent area of Southside Dublin. It is Ireland's largest and one of its most prestigious universities.

the character was Irish. In cases where non-Irish characters or characters whose place of origin was not the Republic of Ireland produced features, these were not annotated (more on the process of annotation in (3.6.1.)).

Finally, the establishment of these restraining parameters was not only so that the corpus could better cater to the thesis' RQs, but also because, at the time of compilation and time of writing, there exist no corpora of *contemporary* IrE fiction which follow these parameters. The novelty of compiling such a corpus however, posed several challenges, ranging from establishing the *sampling frame*, sourcing the texts, to dealing with copyright issues. These are all explained in the following sections.

3.2.3. Developing CoFIrE's sampling frame and sampling technique and limitations of CoFIrE

The size of CoFIrE (detailed in section (3.3.) along with the authors and books) was predetermined at design stage. To arrive at the optimum corpus size in terms of designing, sourcing, sampling, analyzing, coding, storing metadata, etc., and in order for the corpus to be as representative a specialized corpus (i.e. contain as much of the *variety*) as possible, it was decided that it would aim to contain approximately 1 million words.

Once size was established, the compilation of the corpus posed a challenge: creating the *sampling frame* (i.e. list of texts to be included) and *sourcing texts* which followed the restraining parameters explained in (3.2.1.) and (3.2.2.). Unlike in previous centuries where literature was produced over the span of 100 years, at the time CoFIrE was compiled (i.e. 2016), we were only almost two decades into the 21st century. This means that the amount of *contemporary* fiction using large numbers of IrE features produced in the Republic of Ireland from the 2000s onward is likely to be proportionally smaller than that of the same type of books produced in previous centuries (see (3.3.) for measures taken to compensate for the dearth of *contemporary* fiction that aligned with CoFIrE's parameters).

The process of creating the *sampling frame* was complex, mostly due to the fact that bibliographies, literary manuals and/or 'lists' of contemporary writers whose books are rich in IrE dialect / non-standard features were scarce. Thus, a wide variety of

academic (listed in Table 3.2. below) and non-academic sources were consulted, which, as explained below, ensured that the corpus contained a balanced and wide catalogue of IrE orality-rich books.

Author/Year	Title
Hunt	
Mahoney (1998)	<i>Contemporary Irish Literature: Transforming Tradition</i>
Harte & Parker (2000)	<i>Contemporary Irish Fiction: Themes, Tropes, Theories</i>
Cahill (2011)	<i>Irish Literature in the Celtic Tiger Years 1990-2008: Gender, Bodies, Memory</i>
Mianowski (2017)	<i>Post Celtic Tiger Landscapes in Irish Fiction</i>

Table 3.2. Academic sources consulted to create CoFIrE’s *sampling frame*.

Another helpful source I consulted was *Ricorso*. Created and compiled by Dr. Bruce Stewart, Reader Emeritus in English Literature at the University of Ulster, *Ricorso* is a website that catalogues Irish writing and authors. It contains biographical records, primary works, as well as extracts and commentaries compiled from a variety of sources, including “book notices, reviewing organs, and academic journals” (Stewart, 2011).

While helpful in providing explanations of literary movements and/or detailing the authors’ careers, these academic sources very rarely, if ever, mentioned the writers’ style or their representation of orality. Thus, ‘non-academic’ sources had to be consulted so as to develop the *sampling frame*. These included online reader and critics’ reviewing sites like *The Irish Times Book Club*²⁵ (2022), as well as literary magazines such as *Electric Literature* (2016) or *The Millions* (2016), both of which are online magazines which offer essays, reviews, and coverage on literature. It was also decided that readers’ reviews may be another valuable source of information especially since they are the receivers and decoders of the authors’ conscious reproduction of orality via the use of IrE dialectal/non-standard items. Thus, (non-academic) reviews which were posted to reading

²⁵ Accessed at: <https://www.irishtimes.com/culture/books/the-book-club>

recommendation sites heavily informed the *sampling frame*. In particular, I looked at reviews which were posted on *Goodreads* (2016).

Similar to other social media, *Goodreads* is a free social network for bibliophiles, currently owned by Amazon, but originally created and launched by Otis Chandler and Elizabeth Chandler in 2006. The site offers members the chance to discuss literature in different forums, create book lists, and participate in reading challenges, as well as it provides customized recommendations based on the member's 'previously read' shelf. What was more useful for this project was the fact that members can rate and review books, sometimes offering future readers very detailed descriptions of the books, their content, and their linguistic styles. Examples 3.1. through 3.3. below are fragments from reviews which led to the inclusion of Barry's (2011), Ryan's (2012), and Howard's (2013) books into CoFIrE:

- (3.1.) "City of Bohane is not an easy read and requires work of the reader. There's [...] dense dialect and colloquialisms [...] Hell, Barry even makes up words throughout, and delightfully so. For me, it is the strange, twisted and beautiful language that makes this novel so compelling." (Rohan, 2011)
- (3.2.) "All together "The Spinning Heart" has hundreds of Hiberno-English expressions; some pages have one or two in almost every sentence. I think that for Ryan, this is just a way of being true to the way people speak." (Elkins, 2015)
- (3.3.) "I can totally utterly recommend this book ... if you have spent at least a good couple of the boom 'n' bust²⁶ years in Ireland. The accents, the jokes, the people's traits are a pitch-perfect satire of nowadays Ireland" (Dunne, 2014)

While the academic and non-academic sources aforementioned were helpful when developing the *sampling frame* in the sense that they led to certain authors and individual books, it was still difficult to find an actual list of books which featured the prominence of IrE orality in text. The process of sampling the texts was of help in this regard.

Selecting the texts for inclusion into the corpus involved an iterative procedure. While informed by the academic and non-academic sources mentioned above, it was necessary for the books to contain as many IrE dialectal/non-standard features as possible

²⁶ Reference to the Celtic Tiger years of economic prosperity (see section (1.4.) for a description of this period).

if CoFIrE was to be balanced and representative of the *variety* under examination. To ensure this, the books were initially thoroughly skim-read for features after their selection. Thus, if the book being sampled contained very few features of the *variety*, then it would be excluded. This was the case with David Park's (2012) *The Light of Amsterdam*. At a total of 384 pages, the amount of features of the *variety* (outlined in (3.2.1.)) which CoFIrE was to represent that were found after reading the book was minimal. As illustrated in the fragment below, notice how there is only one feature (i.e. the underlined contracted auxiliary) in the exchange. Thus, the books from the *sampling frame* which, after being skim-read, proved to contain abundant features were included in the corpus.

'I never mix business and pleasure,' Marty answered.
'You should try it some time,' Lisa said before asking if she could smoke.
'No you can't—you know you can't.'
'And I'm in a hurry because I've four kids about to wake up looking for their
Coco Pops and me to get them out to school,' Pat said. 'and after that there's a
day's work to be done.' (Parks, 2012, p. 30)

Another challenge in the compilation of CoFIrE was determining *how many texts* it would include given the considerable dearth of academic sources that listed authors/books which followed the restrictive parameters already mentioned in (3.2.2.). Seeing as orality reproduction is a consciously-made, authorial choice, further sources were consulted, including critical reviews which overtly mentioned voice reproduction (such as 3.1. through 3.3. mentioned above), as well as interviews where the authors themselves described their own style and creative process as concerned with the rendition of IrE orality. The use of these sources led to the decision to narrow down to a list of 8 authors (all of whom are described at length in sections (3.4.) through (3.4.8) along with why they were included in the corpus). In addition, the number of authors determined *how many texts* the corpus would include, as it was decided that CoFIrE would contain 2 books by each author (see Table 3.3. in section (3.3.) below for an outline of all CoFIrE authors and books).

A practical limitation on the *sampling frame* was the availability of texts. While books published in previous centuries are often copyright-free and easily available for download, *contemporary* (i.e. 21st century) literature is copyright-protected. Thus, accessing sample electronic or preview copies to identify whether they contained large

numbers of IrE features so they could, therefore, be added to the corpus, was slightly challenging (see sections (3.2.) and (3.2.2.) for more on the criteria texts should fulfil to be included in the corpus). Another explanation for how challenging gaining access to sample copies concerned how relatively recently the books which were selected for sampling had been published at the time of compilation. Another reason was the ‘niche’ nature of the *variety* the corpus had to be representative of (which is detailed in section (3.2.1.)). When compared to major literary markets, such as the North American or British ones, where most books are readily available on all major (online) booksellers, and most include a sample view before purchase, the Irish fiction market, especially regarding fiction containing large amounts of dialect-vernacular features, is rather small and restricted. In fact, several of the books which were included in CoFIrE lacked electronic copies (e.g. Ruane’s²⁷ or Ní Dhuibhne’s) or were only available for purchase in specialized or second-hand bookshops. To solve this, some of the authors were contacted for access to electronic copies when the books in the *sampling frame* were unavailable for purchase or preview.

Contacting the authors also served another purpose: *avoiding copyright infringement*. It is important to bear this issue in mind, especially when investigating and reproducing fragments from contemporary books. The authors included in CoFIrE and/or their literary agents were contacted via email to prevent the corpus from copyright infringing. The emails (a template of which has been added to Appendix 2 informed the authors/agents about this project. They were asked 1) whether they gave me permission to use their books in future publications, and, in some cases, whether 2) they could provide me with a copy of a book of theirs which was part of the *sampling frame* but was unavailable for preview or purchase at the time of compilation. As mentioned above, this would aid in the sampling of the book, but it would also prevent me from having to scan a physical copy to make it corpus-software readable. It must be noted that that there were cases where the writers could not be reached (i.e. Ruane and Doyle), and others who could not provide copies of their books. However, all of the writers²⁸ who replied were very generous and all expressed their support and acceptance in having their books be a part

²⁷ I would like to thank my dear friend, John O’Grady, without whose generous gifting of his own copy of Ruane’s (2003) book I would not have been able to sample and later analyze it.

²⁸ I would like to thank O’Connor, Ryan, Howard and Howard’s Penguin House agent, Ms Patricia Deevy, for having very generously sent electronic copies of their books. My gratefulness also extends to Keegan, Ryan, and Ní Dhuibhne for their acceptance of the inclusion of their books in the corpus, along with their best wishes for this thesis.

of CoFIrE. Both the authors and books which were ultimately included in CoFIrE are outlined in the next section while sections 3.4. through 3.4.8. detail why the authors were chosen along with their books.

3.3. The *Corpus of Fictionalized Irish English*

CoFIrE is a specialized, synchronic corpus of *contemporary* Irish fiction, containing 1,123,601 words. As illustrated in Table 3.3. below, it comprises 16 works fiction, including 13 novels and 3 short-story collections (colored blue), all published in the Republic of Ireland between 1993-2014. The books were written by 8 authors, five of whom are male and three, female.

Author	Book	Year	Book Code	Novel (N) or Short Story collection (SC)	Approx. Word Count
Roddy Doyle	<i>Paddy Clarke ha ha ha</i>	1993	RDPC	N	82,503
	<i>The Woman Who Walked into Doors</i>	1996	RDWD	N	65,733
Éilís Ní Dhuibhne	<i>The Dancers Dancing</i>	1999	ENDD	N	70,073
	<i>Midwife to the Fairies</i>	2003	ENMW	SC	64,133
Donal Ruane	<i>Tales in a Rearview Mirror</i>	2003	DRTRM	SC/N	73,396
	<i>I'm Irish Get Me Out of Here!</i>	2004	DRII	N	70,016
Claire Keegan	<i>Walk the Blue Fields</i>	2007	CKWBF	SC	43,553
	<i>Foster</i>	2010	CKF	SC	14,045
Nuala Ní	<i>You</i>	2010	NCY	N	63,109

Kevin Barry	<i>The Closet of Savage Mementos</i>	2014	NCSM	N	65,772
	<i>City of Bohane</i>	2011	KBCB	N	69,572
	<i>Dark Lies the Island</i>	2012	KBDLI	SC/N	50,149
Donal Ryan	<i>The Spinning Heart</i>	2012	DRYS	N	49,366
	<i>The Thing About December</i>	2013	DRYTAD	N	70,073
Paul Howard	<i>Downturn Abbey</i>	2013	PHDA	N	140,075
	<i>Keeping up with the Kalashnikovs</i>	2014	PHKU	N	132,033
Total					1,123,601

Table 3.3. CoFIrE authors, texts, book codes, and approximate word count/book.

It is important to mention that while two of the books (i.e. Ruane's *Tales in a Rearview Mirror* and Ryan's *The Spinning Heart*) contain short stories, all of them are intertwined into one overarching story. In the case of the former, the narrator is a taxi driver recounting his daily adventures. In the latter, 21 different characters from the same town tell their own story which, in the end, is one interconnected tale. This differs from the 3 other short story collections in CoFIrE (colored blue in Table 3.3. above), all of which contain stories by different characters which are thematically disconnected from one another. The thematic interconnection in Ryan's and Ruane's books led to their being catalogued as novels in CoFIrE. Furthermore, Table 3.3. evidences the inclusion of Doyle's and Ní Dhuibhne's books which were published in the 1990s and could, therefore, be perceived as not fulfilling the parameter of *contemporariness* (i.e. published in the 21st century) set for this study. Their inclusion is justified by the limited amount of *contemporary* Irish fiction which followed the parameters set for CoFIrE existent at the time of compilation, already mentioned in section (3.2.2)). To compensate, it was decided that only three books would be incorporated from the last decade of the 20th century. A detailed explanation of the CoFIrE authors, the books selected for inclusion, and the reason why they were selected is offered in sections (3.4.) through (3.4.8.).

Table 3.3. above also includes information regarding the year the texts were published, the book code every text was assigned (see (3.6.1.) for more on the creation of the codes and their use as tags in the CoFIrE annotation system), and their approximate word count. As already pointed out in section (2.3.3.), scholars believe *balance* to be a hard-to-achieve desiderata given that it is based on the corpus builder’s judgment (McEnery & Hardie, 2012, p. 10; McEnery et al., 2006; Sinclair, 2005). Thus, it was decided that the CoFIrE texts had to *ideally* be relatively similar in word count (see Table 3.3.). However, there are some exceptions. The small number of texts which followed the CoFIrE parameters limited the spectrum from which to select texts, which led to texts from the *sample frame* being included (regardless of their word count) if they contained a minimum of 100 IrE features (see chapter (4) for a detailed analysis of the production of features per book) and aligned with the restricting parameters (outlined in (3.2.1.) and (3.2.2.)). Notice how the 3 short story collections average at 40,577 words and are, therefore, smaller in word count than novels. The 13 novels logically have larger word counts, ranging from 49,000 through 140,000 words, and averaging at 77,066 words.

Text Type	Average Word Count
3 collections	40, 577
13 novels	77, 066
11 novels (without Howard)	66, 342

Table 3.4. Word count averages/book in CoFIrE.

A caveat is noted in relation to the *balance* of the corpus in terms of the amount of *variety* (i.e. Southern IrE dialectal/non-standard items) in each text. Provided that, as mentioned in other sections, the use of dialect/non-standard elements (and their quantity) is a conscious authorial choice which varies from one writer to another, achieving *balance* was practically an impossibility because not all authors used the exact same amount of features, as will be illustrated below with regard to Paul Howard. As much *balance* as possible was sought by following Sinclair’s (2005) recommendation to include full texts, and why only books with large amounts of features were selected in the process of developing the *sampling frame* (see chapter (4) for a discussion of the books’ word

counts, their feature contribution, and overall number of annotated features in CoFIrE). It is important to mention that Paul Howard's *Ross O'Carroll-Kelly* novels, which chronicle modern Irish culture and identity, and are critically acclaimed for their acute portrayal of Dublin orality, were comparatively larger in word count than the other CoFIrE books (for more on Howard's linguistic style see sections (8.2.) and (8.2.1.)). However, and as shown in Table 3.4. above, there is only a 10,724 word difference between the corpus novels with and without Howard's books, so it was believed that this slight difference would not affect the overall word count balance of the corpus.

At corpus annotation and analytical stage, however, it was found that Howard's texts contained an abundance of pragmatic elements which surpassed that of other books (as shall be further discussed in sections (4.1.) and (4.2.)). This would somewhat skew the data regarding pragmatic features in CoFIrE; a risk which other researchers (Clancy, 2010, p. 85) warn a corpus may run when some sample texts are longer than others. It is, perhaps, the conversational nature of the *Ross O'Carroll-Kelly* books, which are presented as an oral account by the protagonist to the author of his daily adventures (see (8.2.) and (8.2.1.) for a detailed description of the series and its style), which accounts for the large amounts of pragmatic items in the novels, which Amador-Moreno & Terrazas-Calero (2017, p. 257) take as being Howard's stylistic hallmark. To address this imbalance in feature number contribution, the total amount of occurrences of features produced by both Ross O'Carroll-Kelly and all other characters in CoFIrE is normalized in analytical chapters (5) through (7). The outlier nature of Howard with regard to feature production in CoFIrE, however, merited attention on its own. Thus, it was believed that the compilation of the first ever *Ross O'Carroll-Kelly Corpus* would complement the findings from the broader analysis carried out through CoFIrE, and would also aid in answering the RQs (section (1.2.)) of this study. The compilation of the *Ross O'Carroll-Kelly Corpus* will be explained in detail in section (3.9.), while a detailed description of the series and its stylization of North vs Southside Dublin (linguistic) stereotypes shall be provided in sections (8.2.) and (8.2.1.). The books included in CoFIrE and their authors will first be detailed.

3.4. CoFIrE authors and books

As previously mentioned in (3.2.3.), various academic and non-academic sources were consulted, along with reviews and author interviews, to create a *sample frame* of Irish authors whose linguistic styles were concerned with the (re)production of IrE orality. Eight authors were selected, and 16 works of fiction were included into CoFIrE, all of which span three decades. The books are not only representative of IrE dialectal and non-standard features, but also include themes and historical events in Ireland (pre, mid, and post Celtic Tiger), which may have influenced the way Irish people use (and reproduce) language to convey their evolving identity. The following sections indicate why these authors/books were included in the *sampling frame* and in the corpus after sampling.

3.4.1. Kevin Barry

Award-winning, Limerick-born, short story writer, novelist, and playwright, Kevin Barry, was a clear candidate for inclusion with *Dark Lies the Island* (2012), which contains a “dialect-heavy prose [that] produce[s] a startlingly unique voice” (Krempels, 2012). The second book of his to be added for its highly detailed mirroring of IrE orality was *City of Bohane* (2011), which was critically described as “an extravagant experiment in language, [which is] rich in Irish English slang and vernacular” (Carey, 2018). Indeed, in an interview on his creative process and use of language, Barry recognized the important role voice reproduction plays in his books, especially in *City of Bohane*:

[Bohane] is an invented place but its language is sprung directly from working class speech in the cities I grew up in, Limerick and Cork. Those kinds of voices have never really shown up before in Irish literature. (Lee, 2013)

Notice how the author himself makes reference to the type of class-influenced language (i.e. working class) that infuses his novel. Barry’s detailed reproduction of Limerick and Cork speech led to the inclusion of his two books to CoFIrE.

3.4.2. Claire Keegan

Tradition and modernity are blended in along with dialect in Keegan's short story collection *Walk the Blue Fields* (2007), and in her novella, *Foster* (2010), both part of the corpus. An award-winning, short-story writer from County Wicklow, Keegan's style is distinctive for its blend of modern, globalized Ireland with elements from its pagan past (Lupino, 2014, p. 9). In her review of Keegan's award-winning collection, fellow writer Éilís Ní Dhuibhne (see (3.5.3.) for a description of Ní Dhuibhne's style and books added to CoFIrE) praised how the local rural dialect is reproduced in a way that "expresses its colour, richness and wit [while being] mercifully unpatronising and never overplayed" (Ní Dhuibhne, 2007), which is why *Walk the Blue Fields* was included in CoFIrE.

Although published originally as a short-story in *The New Yorker*, *Foster* (2010) was also highly praised for its linguistic and dialectically-rural richness which "is deeply rooted in [the characters'] background in terms of age, education, and geography" (Lupino 2014, p. 8). The short-story, which was declared Best Book of Year by *The New Yorker*, was later expanded and published in book format, and was also included in the syllabus for the Leaving Certificate Comparative Literature Examination in the Republic of Ireland.

3.4.3. Éilís Ní Dhuibhne

Modern Irish short stories, or short fiction, are a genre where plenty of shifts are occurring. Contemporary Irish short fiction seems to be characterized by its critical attitude towards tradition, with writers "reclaiming and revisioning rather than rejecting [it]" (Peach, 2004, pp. 11-12). Thus, short fiction Irish writers can be found experimenting with language and with complex, non-linear narratives to "tackle current ideas and topics (gender as performance, retrieving women's history, immigration)" (Ingman, 2009, p. 255), blending contemporary life with oral tradition in the form of reimagined folk tales and myths.

The modern short fiction Irish writer selected for inclusion into CoFIrE was Éilís Ní Dhuibhne with her short story collection, *Midwife to the Fairies* (2003), and the novel, *The Dancers Dancing* (2000). This renowned Dublin-born author is very versatile, shifting easily from novel, to play, to children and young adult literature, among other genres, and writing in both Irish and English. Her fiction often “explore[s] the divergences and continuities between tradition and modernity in Irish society” (Fogarty, 2003, p. xi). Such blend is also represented in the use of dialect in *Midwife*, as shown in (c) below. Notice the use (my emphasis) of the definite article, habitual *do be*, and discourse pragmatic markers *sure* and clause-final *like*.

- (c) After the day’s work and getting the bit of dinner ready for myself and Joe, sure I’m barely ready to sit down when it’s on. It’s not as relaxing like. I don’t know, I do be all het up somehow on Fridays on account of it being such a busy day at the hospital and all (Ní Dhuibhne, 2003, p. 22).

The connection between language and identity is also featured in her semi-autobiographical novel, *The Dancers Dancing* (2000), which is set both in Dublin and in a Donegal Gaeltacht²⁹ in 1972. Excerpt (d) illustrates the use of 2nd-person singular *youse* (my emphasis) and the prejudiced view of a Dublin girl with regard to the use of ‘local’ dialect.

- (d) ‘What did youse bring?’ Sandra says youse. It damns her for all eternity as far as Orla is concerned. Orla has a special linguistic mission in life [...] to stamp out every trace of local English dialect from her surroundings [...]. Words like youse cannot be tolerated. Orla has her work cut out correcting the terrible English of her mother and her brother. (Ní Dhuibhne, 2000, p. 12)

The link between language use and Irishness evidenced in her books is the reason why *Midwife* and *The Dancers Dancing* were both included in CoFIrE.

²⁹ *Gaeltacht*: an Irish-speaking region.

3.4.4. Nuala Ní Chonchúir

Voice reproduction also seems to be a prominent element in the novels written by Nuala Ní Chonchúir's included in CoFIrE, namely *You* (2010b) and *The Closet of Savage Mementos* (2014).

A multi-award winner, Ní Chonchúir is a well-established poet and fiction writer. Her debut novel, *You*, describes the hardships of a dysfunctional Irish family in the context of 1980s, suburban Dublin. Despite the fact that the plot might seem traditional for an Irish novel, the book gathered critical acclaim due to its second-person-singular narration style. When asked about the attention to detail illustrated in the reproduction of different Irish accents in her novels, Ní Chonchúir described herself as a “fan of stylistic writing” (Ní Chonchúir, 2010a) stating:

- (e) I am in love with language, always have been. I love the way people talk, the expressions they use and invent. I'm from Dublin and the book [*You*] is written in a Dublin vernacular that I am very familiar with (*ibid.*).

She also acknowledges in (f) the connection that exists between language and identity portrayal in real life and how that may be used for characterization purposes in fiction:

- (f) I always know who a character is and where they are from [*in terms of their voice*]. Place is so important to Irish people, it has to be important for a character. I love accents and Irish people's accents differ from parish to parish, which is intriguing, so I always listen out for variations in dialogue (Ní Chonchúir, 2013).

The reproduction of Irish themes and voice is also a core stylistic feature in *The Closet of Savage Mementos*, which is an emotional depiction of identity, death, grief, and motherhood, the latter of which seems to be one of the most popular imageries in IrE literature. However, her portrayal of this theme in this book is subversive as it presents it in the form of un-maternal motherhood.

- (g) In all the years I have been imagining this reunion I have never heard his voice in anything but a Dublin accent. It seems so stupid now, but there you have it; I can be a very stupid woman at times (Ní Chonchúir, 2014, p. 193)

The Closet of Savage Mementos, which is divided into two parts, one set in the Scottish Highlands and the other in Dublin, spans two decades and shows a profound awareness of the link joining voice and identity. This is shown in (g) above, where a woman is about to reunite with the son she gave up for adoption who lives in Scotland.

3.4.5. Roddy Doyle

Investigating the portrayal of modern Irishness through orality in fiction led to the inclusion of Dublin born, award-winner, Roddy Doyle. Dublin city clearly exerts a great influence over his fiction which has received critical acclaim for his unadorned and “flawless recreation of working-class Dublin speech” (Mahoney, 1998, p. 247). When asked about his portrayal of the dialect as shown in (h), Doyle attributed its success to his love of Dublin English:

- (h) I love the rhythm and the bullet of it. Even though I’ve grown up with it, I still actively listen to it. [...] Now as a writer, if you have got something as strong as that: why wouldn’t you use that language to create stories from? (O’Malley, 2013)

The novels written by Doyle included in CoFIrE are *Paddy Clarke ha ha ha* (1993) and *The Woman Who Walked Into Doors* (1996). The former, which won the 1993 Booker Prize for fiction, is a dialect-heavy depiction of the effects of alcohol abuse, unemployment, and poverty narrated from a ten-year-old boy’s perspective whereas the latter, also rich in linguistic features, presents an unvarnished depiction of spousal abuse.

3.4.6. Donal Ruane

While having retired from writing after publishing three books, the inclusion of Ruane's *Tales in a Rearview Mirror* (2003) and *I'm Irish: Get Me out of Here* (2004) infuses CoFIrE with a more comical and orality-reproducing note. *Tales in a Rearview Mirror* is presented as an amalgam of twenty five "true" stories Ruane claims to have gathered when working as a taxi driver in Dublin. Despite its more serious nature, as it describes the situation of Ireland pre and post economic crash, *I'm Irish: Get Me out of Here* is also filled with IrE linguistic items. Indeed both books very proficiently present the reader with a wide spectrum of modern-day Dublin life, heightened through phonetic reproductions and the use of dialect, which "catch the nuances and patois of the north and southsiders perfectly" (McKeon, 2003). Furthermore, the books also show an acute awareness of the importance of voice reproduction for the enrichment of the cast. An example would be one of the short stories where the narrator is surprised by the fact that a female customer is clearly "from the country [...but] there's no sign of an accent" (Ruane, 2003, p. 219). Accent awareness and (re)presentations is also heavy in the next fragment³⁰:

- (i) "Howiya goin' horse? Are ye well? I'd say yer quare busy tonight boy, ha? Tis fuckin' manic out there like!" said the lad in the front. "Savage, boy, fuck-
ing savage! Jesus, would ye get us home te fuck outa here?" echoed the lad in the back. "Where are ye off to, lads?" I asked. "Dun Laoghaire and then on te Dalkey, good man". Dun Laoghaire and Dalkey, if you don't mind. Fair play to the boys, from the back arse of nowhere - outside of Dublin - and they land themselves in the prime residential areas when they come up. (Ruane, 2003, p. 187)

The previous exchange illustrates Ruane's dialect-heavy and slang-filled portrayal of IrE voice between the narrator and two customers he refers to earlier in the story as "pure country" (i.e. lower class) men from outside of Dublin. Notice the judgmental comment made by the narrator on the fact that these customers live in an affluent area of the city, the inference being that their accent, which seems to be representative of lower social status (at least to the narrator), does not correspond with their current address. While, at

³⁰ The underlined words are the features which were annotated in CoFIrE.

the time of writing, none of his books have been investigated by academics with regard to language representation, I believe they should be used for such purposes as they evidence a high level of IrE spokenness which is shown in the example above.

3.4.7. Donal Ryan

Since the Celtic Tiger crash and the subsequent economic recession it brought about can be described as one of the main events to have undeniably affected modern-day Ireland in the 21st century³¹ (see (1.4.) for an overview), it was to be expected that it would have a clear influence over contemporary fiction. Thus, I decided to include books which were concerned with its depiction. This would also enable the examination of the type of linguistic features used in those books, and how (if at all) they may reflect the ‘modern’ Irish identity (corresponding with RQs 2 and 3 outlined in (1.2.)), as opposed to books which were published before the Tiger, or books that are not concerned with its portrayal.

Multi-award winning, Limerick-based, Donal Ryan, was selected because his first two novels, *The Spinning Heart* (2012) and *The Thing about December* (2013), set him up as a “chronicler of contemporary Irish life” (Boland, 2013) who records the flip side of the metaphorical Tiger coin and the effects of the economic downturn it brought about. *The Spinning Heart*, which garnered Ryan multiple awards, is written as a polyphony of 21 well-differentiated, dialect-heavy voices, which present a harrowing portrayal of the effects the recession had on a small, middle-class, rural, Irish community. Written in multiple viewpoints, his debut novel has often been praised for its truthful depiction of accent. Ryan himself has frequently attributed his ear for the accent to his former position as a labor inspector. He also acknowledges in fragment (j) below the influence orality has in his fiction and the importance of voice reproduction for the portrayal of identity:

- (j) I’m writing in the slang and grit of my own language, which is true because we’re all immersed in the [sic] model of growing up. You know, unless you grew up in a house where people speak with received pronunciation [sic] which is unlikely unless you’re a British aristocrat I think. You know, you will hear the demotic or the language of common people [...] as we all are all

³¹ At the time of compilation (i.e. 2016), COVID-19 had not broken out yet, which is why I refer to the Celtic Tiger as one of the main events in 21st-century Ireland. Seeing as how the pandemic has also had devastating effects world-wide, it would be interesting to study the fiction that is produced after COVID is eradicated and the economy recovers so as to examine the portrayal of post-pandemic Irishness.

around you all the time. And that's where people kind of reveal themselves, when they're speaking as the person that they are (Cleary, 2018, p. 31)

The orality he so masterfully reflects and purposely uses in his books is illustrated in (k) below, which evidences IrE features (my emphasis) such as habitual *do*, fronting, vernacular lexical items (i.e. *yoke*), or lexicalized pronunciations (i.e. *auld* for *old*).

(k) I do see that boy of the Mahons nearly every day [...] He's beautiful, that boy, tall and fair-haired, like his mother. His auld father is a horrible yoke (Ryan, 2012, p. 31).

The second novel by Ryan included in CoFIrE, *The Thing about December*, is set at the height of the Tiger in the same rural village featured in the previous book, to which it is undeniably connected both thematically and stylistically, as the portrayal of voice also heavily imbues it.

3.4.8. Paul Howard

Not many names in contemporary Irish literature are as synonymous with the Celtic Tiger (explained in (1.4.)) as *Ross O'Carroll-Kelly* (RO'CK). RO'CK is a fictional character as well as the pseudonym under which the former sports journalist, playwright, and fiction author, Paul Howard, writes his best-selling, eponymous comedy series. The RO'CK saga began as a sports column in *The Sunday Times* in 1998 which documented the (legal and illegal) events Howard witnessed when he was asked to write a weekly piece on school rugby teams. The column's popularity soon skyrocketed and was turned into full length novels. At present, the series, which enjoys almost cult-like status (see O'Brien (*fc.*) for a monograph on the cultural significance of this series), comprises 21 full-length novels, 4 plays, a spoken word album, a spoof travel guide, a book of mock interviews, and a weekly column and podcast in *The Irish Times*.

The RO'CK books humorously satirize the wealthy and snobbish Southside Dublin society that arose during the Celtic Tiger period, which is embodied in the character of former school rugby jock, Ross, who documents his many (mis)adventures experienced pre, mid, and post Tiger in the first-person. The value of these books as

chronicles of modern Irishness is often critically expressed, with scholars describing the long-running series as “[a] prominent early example of the popular culture spawned by the “New Ireland” of the Celtic Tiger years [which is a] potentially significant site for exploring the subject of Ireland and the contemporary” (Kelly, 2017, p. 50). Thus, this series represents a valuable source of data with which to explore RQ3 (see (1.2.)) pertaining to the indexation of modern Irishness through language. One of the aspects of modern Irishness the series revolves around is the geographic, socioeconomic, and linguistic divide which exists between the traditional working-class Northside and the more affluent Southside Dublin areas, both of which are embodied in the cast of characters, and in Ross in particular.

Aside from its great value as a chronicler of modern Irishness, the annually best-selling RO’CK series was also selected for inclusion in CoFIRE, and subsequently became a separate corpus (see (3.9.) for more on the compilation of CoROCK, and sections (8.2.) and (8.2.1.) for a detailed description of the stylization of North and South Dublin identities and of the linguistic style of the series), was the fact that it has repeatedly been critically acclaimed for its portrayal of the Dublin accent. Narrated in the first person by Ross, the books are presented to the readers as a story “as told to Paul Howard” (for a detailed description of the linguistic style of the RO’CK series, see (8.2.1.)). Through Ross and his many (mis)adventures, Howard satirizes the affluent society that arose in Southside Dublin during the Tiger years, which contrasts heavily with the working class society of North Dublin (see (8.2.) for more on the North/Southside Dublin stereotyping in the books). This contrast, which revolves around the traditional geographic and economic North-South division, is also linguistically rendered in the books (for more on the linguistic style of the series, see (8.2.1.)). The high level of orality and accent representation, coupled with their representativeness of modern Irish identity, led to the incorporation of *Downturn Abbey* (2013) and *Keeping up with the Kalashnikovs* (2014) into CoFIRE. Excerpt (1) below showcases the prominence of orality encapsulated in the series through the use of non-standard quotatives (e.g. *go, be there*), taboo words and expletives (*fooked, focking*³², and *Jesus*), tag-*so* (*so you are*), and the phonetic contrast between of Ronan’s Northside and Ross’ Southside Dublin accents.

³² *Fooked* is a phonetic representation of the Northside pronunciation of the word *fucked*, while *focking* reproduces the Southside Dublin pronunciation of *fucking* (see (6.5.4.) for more on orthographic renditions of geographic location through FUCK variants).

- (1) [Ronan] goes, ‘We want to provide for ear thaughter – gib her all the things we nebber had groan up.’³³
I just shake my head. I focking love this kid.
‘Jesus, Ro, when you talk like that,’ I go, ‘you make me want to face up to my own responsibilities.’
He’s there, ‘You probley shouldn’t have any mower, Rosser. You’re fooked, so you are!’ (Howard, 2014, p. 39).

Despite the opulence of the ‘good years’ of the Celtic Tiger, which is so well reflected in the RO’CK novels always from the perspective of the affluent class, it all ended with the economic downturn which sank the country into a recession. This led to a return to high unemployment and emigration rates of previous years from which Ireland has been struggling to recover, and which is also represented in later books in this series, as well as in Donal Ryan’s novels.

3.5. CoFIrE Analytic Process: Text preparation

Once all the books were sampled, the next stage was preparing the source texts (whether electronic or physical) to be readable using corpus software. The electronic books were easy to prepare. Cover pages, tables of contents, acknowledgments, page numbers, and any other type of data that was not related to the actual fiction story of the book was deleted. Afterward, the text (usually in .docx format) would be converted into plain format (.txt), which is the formatting style required by the corpus suites (i.e. *Wordsmith Tools version 6* (Scott, 2012), and *AntConc* (Anthony, 2018)) used in this project.

Converting the physical copies of the books into corpus-readable files was more time consuming. Firstly, the books were scanned using the text scanning and optical character recognition software Abbyy©. After their digitization, the texts had to be manually edited due to the fact that the typesetting and layout of the text would often be in disarray. They were, then, cleaned off all formatting (e.g. indentations, page numbers, title and cover pages, or illustrations, among others) and saved as (.docx) files. Afterwards, the documents were converted into (.txt) format so they could be analyzed using corpus software. Since (.txt) files do not contain any type of formatting, any italics

³³ Non-phonetic equivalent: ‘We want to provide for our daughter—give her all the things we never had growing up.’ His second line would be: ‘You probably shouldn’t have any more, Rosser. You’re fucked, so you are!’

or bold letters the authors had originally purposely used but which had been deleted in the process of deformatting were added to the (.docx) files. These would be used for reference in qualitative analyses of prosodic rendition. Finally, copies of both the (.txt) files, which were the ones used in the analyses, as well as of the (.docx) files were saved.

3.6. Data Annotation: Designing the CoFIrE Annotation System

Seeing as how the conversion of source texts into corpus-friendly files strips the texts of all contextual information which may have been contained in them, *annotating* the raw text becomes a “key part of the process of annotating language resources” like corpora (Adolphs & Knight, 2009, p. 47). *Annotating*, thus, refers to the process, manual or automatic, of adding “interpretative, linguistic information to a corpus” (Leech, 1997, p. 2), marking “features in texts that are not immediately observable when the raw text is seen with the naked eye” (Anthony, 2013, p. 148). The corpus can, then, be annotated, or coded, for parts of speech, lexical items, or pragmadiscursive features, among many others items, depending on the researcher’s interest. For example, in the case of CoFIrE, the items to be *annotated* included all those features which were IrE-distinctive and/or non-standard grammatical, lexical and pragmatic items (e.g. *Jaysus, I’m after arriving home, auld*, etc.), as well as some pragmatic items which are not necessarily exclusive to IrE, such as *fuck* or *like*, among others (see (3.2)-(3.2.2.) for more on design and building criteria for CoFIrE, and (3.6.1.) for a detailed description of the annotation system designed for it).

While some scholars argue against *annotation*, claiming that the use of tags compromises the integrity of the raw text and hampers the study of the original linguistic patterns (Sinclair, 2004, p. 191), others advocate its use. Indeed, *annotation* is widely recognized as “a crucial contribution to the benefit a corpus brings” (Leech 1997, 2), which amplifies the usefulness of the corpus, making it function as a “repository of linguistic information” (McEnery & Wilson, 2001, p. 32). In addition, it is a very valuable tool, for while the querying of raw corpora naturally tends to retrieve irrelevant or useless linguistic items (Smith et al., 2008, p. 164), *annotation* enables the researcher to retrieve pertinent information in a faster, detailed, and automated manner. When a corpus is annotated, its pertinent data is, therefore, more readily available, as it can be

systematically accessed and retrieved faster than if done manually. This can be done using, for example, tag searches. Take, for instance, the word *crack*, which functions as noun and verb, in examples 3.4. through 3.6. below:

(3.4.) There is a *crack* [N] on the window

(3.5.) Don't throw a rock at the window! It will *crack* [V]!

(3.6.) [At a dinner party] Hey, what's the *crack* [N][SL] like? Did I miss anything?

Imagine a researcher is interested in examining the way nominal *crack* functions in a corpus. To avoid retrieving non-nominal occurrences, they could sort out functions in context by part-of-speech-*annotating* the corpus with tags like [N] (i.e. noun), or [V] (i.e. verb). Thus, a search for tag [N] using corpus analytic software will instantaneously retrieve all nominal occurrences of *crack* (i.e. 3.4. and 3.6. above). If the researcher were interested in pragma-discursive elements, then an extra tag such as [SL] may be added to 3.6., for that particular case is an IrE slang term (also spelt *craic*) which means 'gossip', 'fun', 'entertainment', 'how are things?', etc. Adding pragma-discursive *annotation* to the hypothetical corpus would allow the researcher to retrieve specific pragmatic data much faster, preventing them from having to sift through irrelevant occurrences. Although this is seemingly a simple example, it illustrates some of the uses of *annotation* well.

When annotating, the investigator must also decide the type of corpus *annotation* they want to implement, as it can be either automatic, automatic but manually corrected, or fully manual (McEnery & Hardie, 2012). However, it is important to bear in mind that none of these three processes are error-free. While existent automatic and computer-assisted *annotating* systems are useful when it comes to coding a corpus quickly, they seem to focus on *annotating* a text semantically or code for parts of speech. This is the case of *UCREL Semantic Analysis System* (USAS, Rayson, 2009), which was designed for the semantic analysis of texts, as well as word-class taggers like the *Constituent Likelihood Automatic Word-tagging System* (CLAWS), which was used to tag the *British National Corpus* (Garside, 1987; Leech et al., 1994), or Atwell et al's (1994) *Automatic Mapping Among Lexico-Grammatical Annotation Models* (AMALGAM), which

annotates correspondences between word-class tagsets and parsing schemes, among others (Baker et al., 2006, p. 18).

To make CoFIrE as useful a corpus as possible, it was decided at design stage that the corpus would be *annotated* so that it would allow for in-depth quantitative and qualitative analyses by means of tag searches through corpus software. Thus, the texts needed to be marked for the relevant linguistic features mentioned above prior to their electronic *annotation* (see (3.2)-(3.2.2.) for more on the *variety* CoFIrE is representative of). Thus, the texts were all read twice. During the first reading, all features of the variety were manually marked up (i.e. highlighted) to hasten the subsequent electronic *annotation* of the corpus (see (3.6.1) for a detailed description of the annotation system designed for CoFIrE). The second reading ensured that none of the features had been accidentally overlooked. Once the reading process was finished, the texts were ready to be *annotated*.

With regard to the type of *annotation* used in this project, it must be noted that all the available automated taggers were deemed to be too part-of-speech and word-class-oriented, which was not useful for this thesis. Since there was no other automatic tool available that catered to the thesis' RQs (see (1.2.)) and corpus design criteria (outlined in (3.2.) through (3.2.2.)), especially because the corpus would eventually be mostly tagged for discourse-pragmatic items (see chapter (4) for a discussion on the abundant numbers of pragmatic elements), I decided to create a *manual annotation* system (described in detail in section (3.6.1.)) which was specific to the purpose of this project. However, designing and implementing a brand new, manual annotation system is a major undertaking for various reasons. Firstly, and similar to designing a corpus, its creation implies the establishment of restricting parameters which will ensure the usefulness of the system when it comes to enriching the corpus and facilitating linguistic analyses of it.

The first parameter set for CoFIrE's annotation system was the *variety* it needed to annotate. It is well known that the usefulness of using corpora lies in the fact that they allow the researcher to investigate various linguistic topics, ranging from phonology, to pragmatics, or syntax, among many other fields. However, it was decided at design stage that this project would only investigate the use of IrE-distinctive and/or non-standard grammatical, lexical, and pragmatic features which were identified as such in grammar manuals and dictionaries, as well as some (mostly pragmatic) items not necessarily exclusive to IrE in order to examine what type of modern Irish identity their use indexes

in the texts (see (3.2.1.) for more on the *variety* coded for in CoFIrE and the reference manuals and (1.2.) for the thesis' RQs). Thus, the annotation system was only applied to those features which are part of the *variety* CoFIrE is representative of, rather than to the entirety of the raw text (i.e. Standard English). For example, the only annotated word in fragment³⁴ (m) below would be *Jaysus*, as that is an expletive interjection typical of IrE.

- (m) 'Jaysus, keep talking like that and you'll fit right in. But please don't come home with some kind of County America accent.' (Ní Chonchúir, 2014, p. 23)

Another criterion was that the annotation system had to code items at macro and micro levels. For example, discourse pragmatic markers such as *like* were annotated (i.e. *macro level*), but the annotation system also marks whether they occur, as in the case of *like*, in clause-initial, medial, or final position (i.e. *micro level*).



Figure 3.1. Diagram of the *top-bottom* approach used to annotate CoFIrE.

Two approaches were also adopted to annotate the corpus: *bottom-up* and *top-bottom*. In the *bottom-up* approach, linguistic items are coded as they are found during the meticulous, manual marking of the texts which took place prior to annotation and which was outlined above. However, a *top-bottom* approach (visually illustrated in Fig. 3.1.) implies a more focused search for individual features, which are retrieved through

³⁴ My emphasis

concordance enquiries, after which relevant data about the functions of the feature in context are marked (Smith et al., 2008, p. 164). Both approaches were used for the annotation of CoFIrE and the CoROCK corpora (see (3.9.) for a description of the latter). In fact, the two case studies conducted on CoROCK (which are detailed in chapter (8)) were analyzed using a *top-bottom* approach whereby the individual features (i.e. New Intensifying *So* and Non-lexical intensifier *Totally*,) were searched for using *concordance line* searches which allowed for the exclusion of non-intensifier uses.

To answer RQ3 (see (1.2.)) regarding the type of modern Irishness being indexed through the recurrence of certain linguistic features, it was also necessary for the system to mark up as much information about the speakers (i.e. characters) as possible. This would expedite subsequent quantitative and qualitative sociolinguistic analyses of the corpus data by means of tag searches with corpus software, as will be explained in the next section.

3.6.1. CoFIrE's Annotation System

The annotation system used in CoFIrE comprises three main tags which include 1) an opening tag, 2) a general linguistic category tag, and 3) a tag which indicates the specific item that is being coded. Table 3.5. provides a visual summary of the system.

EXAMPLE			
She says she's always, <PHDASR> like <PDM><ML></PHDASR>, wanted a music room?			
Tag type	Annotated Data	Tag	Tag Metadata
Opening/Closing tag	Author, Book, Speaker codes	<PHDASR>	Paul Howard; <i>Downturn Abbey</i> ; Sorcha Reported
General linguistic category	Item category	<GR> <VOC>	Grammar; Vocabulary; Pragmatics

Specific item	Annotated item	<P>	Clause medial <i>like</i>
		<ML>	

Table 3.5. Summary of annotation system developed for and implemented in CoFIrE.

The *opening tag* precedes the item that is being annotated and contains essential authorial, textual, and speaker information. To create this tag, all books were assigned a book code which includes the initial letters of the author’s full name, as well as the initials of the book title (see Table 3.3. in section (3.3.) or Figure 3.3. in section (3.7.) below for a list of the book codes). Furthermore, all the speakers’ (i.e. characters) names were reduced to initials and added to the *opening tag*, which doubles as their name tag in the corpus metadata. An *N* was added to the character code if the feature was produced by a narrator, and an *R* was added when the voice of a speaker was being reported by another character. Thus, the *opening tag* in Table 3.5. above, <PHDASR>, corresponds to “Paul Howard; *Downturn Abbey*; Sorcha Reported”. The *opening tag* is mirrored in an identical *closing tag* at the end which signals the completion of the annotation of a particular feature. This tag is especially useful when annotating items consecutively, for it enables an easier quantitative study of the corpus features by separating them from one another, while also helping in quantitative analyses of individual speakers’ language use. In addition, the *opening tag* also allows the researcher to retrieve information pertaining to the speaker’s identity much faster for, as shall be explained in (3.7.), information regarding their age, gender, social class, and geographical location was stored separately from the text.

The second tag is a *general linguistic category tag*, which is applied immediately after the item that is being coded. For this, each feature was distributed into three main linguistic categories, which are represented by the tags <GR> (*Grammar*), <VOC> (*Vocabulary*), and <P> (*Pragmatics*). Since pragmatic characteristics included different types of items (e.g. taboo words like *feck* or *shite*, slang terms such as *gob*, discourse pragmatic markers like *sure*, etc.), category <P> was further subdivided into 5 subtags so as to aid in subsequent quantitative and qualitative analyses. Table 3.6. below lists them, their meaning, and provides an example of each. Thus, the example in Table 3.5. above tags the feature as <PDM>. In other words, “Pragmatics; Discourse Pragmatic Marker”.

Sub-tag	Meaning	Example
PQ	Non-standard quotatives ³⁵	<i>Be like, be there</i>
PT	Taboo language and expletives	<i>Jaysus, bloody</i>
PSL	Slang ³⁶	<i>Gaff</i> (house), <i>Gas</i> (funny)
PDM	Discourse pragmatic markers	<i>You know, arrah</i>
PBO	Boosters/intensifiers	<i>Pure, fierce, solid</i>

Table 3.6. Pragmatic sub-tags applied to CoFIrE texts.

Finally, the third tag marks the *specific item* being annotated. To do this, all features from across the different categories were assigned a tag which consists of the abbreviated name of the item (see Appendix 1 for a list of all feature tags). For example: <G><HD> (grammar, habitual *do*), <VO><GRA> (vocabulary, *grand*), <PQ><GO> (pragmatics, quotative *go*). Thus, the specific tag in Table 3.5., <PDM><ML>, would signal the fact that the feature being annotated is <ML> (clause-medial *like*).

Thanks to the design criteria of the CoFIrE annotation system, its implementation was critical in hastening the process of analyzing the corpus. In fact, had the system not been implemented, it is likely that the study would have required a team of researchers to analyze CoFIrE. Its extreme usefulness resides in the fact that its uses are manifold. On the one hand, it provides detailed contextual information about a specific feature or about the features in a linguistic category. On the other, it allows the researcher to conduct several types of specialized, fine-grained, corpus searches. For example, a search by *opening tag* could 1) shed light on an author's style, 2) retrieve quantitative data regarding dialectal features in one or more books, and/or could even 3) provide insight into a specific speaker or group of speakers' idiolect, among other uses. The researcher can also gather information about which type of grammatical items, for example, are more

³⁵ Quotatives were taken as pragmatic elements due to the fact that they functioned as markers within the text and not as grammatical units. Furthermore, the quotatives that were annotated in CoFIrE were those which deviated from standard ones, such as *say*, *reply*, or *answer*, among others, which were not coded. See chapter (5) for a detailed description and analysis of the non-standard quotative repertoire annotated in the corpus.

³⁶ Taboo and slang words were not included in the vocabulary category because they were annotated for their pragmatic effect. However, a future revision of the annotating system will endeavor to double code all <PT> and <PSL> features as being lexical items as well.

frequently produced in the corpus if they search by *linguistic category* tag. In addition, if the object of study is a particular item, then one can refine the search by searching for its *specific item* tag. For instance, if the researcher is interested in studying the use of discourse pragmatic marker *like* in different clausal positions in CoFIrE, they can tag-search for <IL>, <ML>, or <FL> (corresponding with clause *initial*, *medial*, and *final* respectively) using corpus software. This will automatically retrieve, as illustrated in Fig. 3.2. below, a *concordance line* list of all of the occurrences of the feature in context.

N	Concordance
1	<IL></DRIIN> , where <DRIIN> like <PDM><ML></DRIIN> , is home exactly, <DRIIN> like
2	with? 'Em ... well, yeah <DRIPW> like <PDM><ML></DRIPW> .. we did, but ...' But nothing.
3	</DRTRMGE> , she's <DRTRMGE> like <PDM><ML></DRTRMGE> , <DRTRMGE> totally
4	stopping. We are <DRTRMMEG> like <PDM><ML></DRTRMMEG> , so dead," shrieked the
5	met him, Amy. He's... <DRTRMBE> like<PDM><ML></DRTRMBE> , well when he was..."
6	And then you're cut off, <DRYSBR> like <PDM><ML></DRYSBR> , all of a sudden? That's what
7	I lucky to have a job? Ya, <DRYSH> like <PDM><ML></DRYSH> , I'm really lucky. <Seanie> I
8	us up, though. God, <DRYSK> like <PDM><ML></DRYSK> , I have to pick my steps
9	something about that, <DRYSK> like <PDM><ML></DRYSK> , it's scientifically proven that
10	kid and I'd keep him for <DRYSLL> like <PDM><ML></DRYSLL> , a night or some <DRYSLL>

Figure 3.2. Example of a search for tag <ML> across CoFIrE using *Wordsmith Tools* (Scott, 2012).

Despite its undeniable usefulness, it is important to point out that creating a manual annotation system from scratch is a time-consuming task, especially when the process relies solely upon one researcher rather than on a team. Indeed, the annotation of CoFIrE was done manually by the present researcher, and it involved iteration and some external validation from my supervisors in order to ensure its maximum practicality. Thus, amendments had to be made. For example, the *opening tags* initially did not include speaker code names. However, upon consultation, my supervisors advised me to include speaker tags to ensure the expediency of searches regarding speaker identity indexation. Other amendments included the correction of tags which, during the corpus data analytical stage, were found a) not to correspond with the features they coded, or b) contained a typological error (e.g. missing a letter). In the latter case, the feature did not appear in tag searches using corpus tools, which ultimately would skew numerical counts. The current version of the system, with all the amendments and corrections, has proven to be so useful that I decided to implement it on the case studies conducted on the CoROCK corpus (chapter (8)), albeit with small adaptations that catered to those studies.

3.7. Processing and managing the CoFIrE Metadata

The usefulness of a corpus does not only lie in the linguistic data from its texts, which may or may not be enriched with annotation, but also in its metadata, which is literally ‘data about data’. The corpus metadata is the information the researcher can gather and mark about the content of the corpus, its quality, speakers, and many other characteristics of the texts that are comprised in it (Adolphs, 2006, p. 25). If managed and stored properly, the metadata becomes an essential tool which allows for a better and more accessible analysis and interpretation of the corpus. With regard to metadata storage, Sinclair (2005, p. 6) recommends that it “be stored separately from the plain text and merged when required in applications”. While metadata could be stored in the form of ‘headers’ at the beginning of each text signaling, for instance, a speaker turn-taking or text genre, storing it separately from the corpus provides easy, unencumbered access to it, thus facilitating comparative and qualitative analyses.

In the case of specialized corpora, managing and analyzing a manually annotated written corpus of slightly over a million words such as CoFIrE can be challenging if the metadata is not processed correctly. In this case, the CoFIrE metadata consists of:

- Source text information (i.e. author full name, book code, and year of publication)
- List of the annotated grammar, lexical, and pragmatic items
- Sociolinguistic information about the speakers (i.e. age, gender, class, and geographic location)

In addition, it was decided at design stage that, following Sinclair’s (2005) recommendation, the CoFIrE metadata would be stored separately from the source texts in a *Microsoft Access* relational database comprising three interconnected tables, labeled *Books*, *Features*, and *Characters*. *Table Books* (illustrated in Fig. 3.3.) lists all the book

codes³⁷, which are part of the *opening tag* in the annotating system (see (3.6.1.) for a description of the system), which were created to identify the authors and their books. As already mentioned in the previous section, the book codes consist of the initial letters of the authors' names and last names, and of the titles of the books.

Idbook	bcode	Title	Author	PubYear
1	DRYS	The Spinning Heart	Donal Ryan	2012
2	NCCSM	The Closet of Savage Mementos	Nuala Ni Chonchiúir	2014
3	CKF	Foster	Claire Keegan	2010
4	CKWBF	Walk the Blue Fields	Claire Keegan	2007
5	KBDLI	Dark Lies the Island	Kevin Barry	2012
6	NCY	You	Nuala Ni Chonchiúir	2010
7	DRYTAD	The Thing About December	Donal Ryan	2013
8	RDPC	Paddy Clarke ha ha ha	Roddy Doyle	1993
9	RDWD	The Woman Who Walked into Doors	Roddy Doyle	1996
10	PHKU	Keeping up with the Kalashnikovs	Paul Howard	2014
11	DRII	I'm Irish, Get Me Out of Here!	Donal Ruane	2004
12	ENDD	The Dancers Dancing	Éilís Ni Dhuibhne	1999
13	ENMW	Midwife to the Fairies	Éilís Ni Dhuibhne	2003
14	PHDA	Downturn Abbey	Paul Howard	2013
15	KBCB	City of Bohane	Kevin Barry	2011
16	DRTRM	Tales in a Rearview Mirror	Donal Ruane	2003

Figure 3.3. Screen capture of *Table Books* in CoFIrE MS Access database.

Table Features (visually represented in Fig. 3.4.) catalogues all the annotated items (see Appendix 1 for a full list) and distributes each into the linguistic categories they belong to (i.e. grammar, vocabulary, and pragmatics, including its subcategories).

Idfeature	feature	category
1	Embedded questions	Grammar
2	Non-standard verbal concord with inflectional -s (plural subject + -s ver	Grammar
3	After perfect	Grammar
4	Reduction of vb form: preterite for past participle	Grammar
5	Relativization 0 subject	Grammar
6	Focusing devices: fronting	Grammar
7	Focusing devices: clefting	Grammar
8	Subordinating "and"	Grammar
9	Demonstrative "them"	Grammar
10	Non-standard verbal concord (singular sub + inflected vb)	Grammar
11	Def. art. "the" + uncountable nouns	Grammar
12	Def. art. "the" + occupations	Grammar
13	Def. art. "the" + languages	Grammar
14	Quasi lexicalized phrase possessive pronoun + man "Your man"	Vocabulary
15	Medial object perfect	Grammar

Figure 3.4. First 15 items³⁸ on *Table Features* in CoFIrE MS Access database.

³⁷ The reason the book codes are not listed alphabetically or grouped by author is that books were analyzed in random order.

³⁸ These are only numerical references assigned to each feature by input order and do not correspond to their frequency of production in the corpus.

Cataloguing all features and their distribution into the categories was extremely helpful, especially when a) retrieving information about individual features b) examining which category is more prominent in the corpus, or, among other things, c) investigating which items are more frequently used in each category, all of which relate to RQ1 (see (1.2.)). The latter is particularly useful when studying the indexation of identity in CoFIrE (corresponding with RQ2) as it allows for an investigation into whether contemporary writers (at least the ones comprised in the corpus) continue to use grammatical and lexical features commonly linked with the Stage Irishman linguistic stereotype, or rather, whether they use ‘new’, less locally-bound features which move away from the old stereotype, indexing a more modern type of identity (which addresses RQ3).

Finally, *Table Characters* contains the most crucial data concerning identity indexation (i.e. RQs 2 and 3). It stores essential sociolinguistic information pertaining to each speaker’s identity which was retrieved from what was indicated or inferred from the texts, as well as metadata.

Id	chara code	social backgrc	refidfeature	num	refidbcod	fem c	male	geographic loc	age approx	fem writer	male writer
1504	RDPCP	MIDDLE	6	15	8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	DUBLIN	PRETEEN 10?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
1505	RDPCP	MIDDLE	2	7	8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	DUBLIN	PRETEEN 10?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
1506	RDPCP	MIDDLE	18	2	8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	DUBLIN	PRETEEN 10?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
1507	RDPCP	MIDDLE	3	1	8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	DUBLIN	PRETEEN 10?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
1508	RDPCP	MIDDLE	31	18	8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	DUBLIN	PRETEEN 10?	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figure 3.5. Screen capture of *Table Characters* in CoFIrE MS Access database.

As illustrated in Fig. 3.5., the data is distributed into columns which document the a) speakers’ nametags (see MS Access database, *Table Characters* for all information pertaining to the characters gathered from the CoFIrE texts), which were used in the annotation system as the *opening tag*. As explained in (3.6.1.), nametags were assigned to each individual character and consisted of their book code and the initial letters of the character’s name. For example, the tag <RDPCP> indicates that the character belongs to *Roddy Doyle’s Paddy Clarke ha ha ha* novel, and the <P> was assigned to the character of *Paddy*. The table also includes b) metadata regarding book code (i.e. 8 above is the random, numerical reference assigned to that book), author gender, as well as each

speaker's feature (i.e. 'refidfeature 6', for example, is the numerical reference the database assigned to 'fronting') and production of occurrences (i.e. how many times the speaker produces the feature), the latter of which aids in quantitative analyses. Furthermore, this table also documents sociolinguistic information about the speakers, including their c) (approximate) age cohort, when explicitly mentioned in the text or gathered from the book, gender, geographical location, and socioeconomic status.

Distributing the speakers into different socioeconomic strata for sociolinguistic analysis was challenging to do in CoFIrE because the social class of the fictional characters is normally not explicitly mentioned. Thus, a series of criteria had to be developed for the distribution of speakers into social ranks. The criteria were informed by all the information pertaining to their status which was readily available in text whether in dialogue or narrative form. Thus, and following Nevalainen & Ramoulin-Brunberg (2017, p. 133) social stratification criteria developed for historical linguistic analysis, the criteria used to distribute the CoFIrE speakers into ranks include:

- Level of education: e.g. Is the speaker literate/illiterate?; do they have third-level education?; did they drop out of school?, etc.
- Occupation: e.g. Are they an unemployed construction worker? Or a white-collar lawyer? Are they sex workers or daycare owners?, etc.
- Economic status: e.g. are they described either in narrative passages or by other characters as being wealthy? Or is the character clearly in economic dire straits? Are they described as being financially unstable to the point where they must put their children into temporary foster care?, etc.

The speaker's economic status was informed by their occupation as well as by other contextual information available in the texts. For example, is the speaker a street criminal, a blue-collar worker, or a wealthy character? Are they described by others as wearing outfits that align with certain social subcultures/groups? For instance, speakers described as *skangers* or dressing as such in the texts would be classified as lower class. *Skanger* is a derogatory slang IrE term for a lower-class "young, uncouth youth" (Dolan, 2006, p. 199) from a high criminal rate area (usually in Northside Dublin) often recognized for their outfit (e.g. casual sports clothes like tracksuits, (often up-tipped) caps, tennis shoes, fake gold jewellery for men), and poor education level (for more on the 'fashion'

stylization of Northside vs Southside Dubliner stereotypes in the *Ross O'Carroll-Kelly* series see (8.2)).

Based on Nevalainen & Ramoulin-Brunberg's (2017) social rank distribution criteria, I used my own 3-tiered system to investigate the social stratification of the corpus (Table 3.7.) which distributes speakers into *upper*, *middle*, and *lower* class. However, the *middle class* was further fine-grained into *lower-middle*, and *upper-middle class*, to account for the different levels of education and different occupations the characters are described as having in the books.

STRATA	OCCUPATIONS
Upper class	<p>Affluent and/or highly educated individuals; affluent Southside Dubliners; high income professionals including: lawyers, popular radio hosts, musical composers, famous (erotica) authors, and famous poet.</p>
Lower class	<p>Generally low class such as uneducated and/or having very little income; uneducated small farmers; criminals, former gang leaders, gang members, thieves, 'fresh out of juvvie' individuals; (low-rent) bordello madams, pimps, sex workers.</p>
Middle class	<p>Generally middle-class individuals; builders, foremen, brick layers; shop clerks, cashiers; auctioneers; priests; funeral directors; chefs; secretaries; police officers; taxi drivers; bakers; farmers; hairdressers; surveyors; terrorists with middle-class background; unemployed former owners of building companies; (small) magazine editors; some Northside Dubliners; (small) store owners; daycare owners; owner of brick-laying company.</p>

Low-Middle	<p>Generally low-middle income; (small) farmers; fishermen; fishmonger; butcher;</p> <p>truck drivers; itinerant sellers; fortune tellers; some terrorists; unemployed individuals; innkeeper owner in criminal area; individuals described 'pure country'³⁹; some Northside Dubliners described as scammers or 'slappers'⁴⁰; drug dealer boss.</p>
Upper-Middle	<p>Generally have higher income positions; doctors; nurses; accountants; bull-semen distributors; building-company owners; secretaries at law firms; hotel managers; primary and secondary school teachers; but also big farmers, affluent gang leaders and their wives; and (not necessarily very famous authors poets).</p>

Table 3.7. Illustration of the social stratification system and occupations in CoFIrE.

Those characters whose social status was difficult to determine were all assigned to an 'unknown' category. This also includes all the narrative voices which clearly did not belong to any of the characters in the books. Finally, the stratification system applied in CoFIrE was subsequently supplemented at analysis stage (as shall be discussed in chapters (5)-(7)) with the demographic classification system applied in the *LCIE* and *BNC2014* samples used for contrastive purposes in this thesis. This sociodemographic cross-examination is explained in detail in section (3.8.2.).

In all, the benefit of storing all the CoFIrE metadata in the interconnected database is that crucial sociolinguistic information is readily available. Indeed by clicking on a given feature from *Table Features*, a drop-down menu folds out and provides information gathered from *Table Characters*, which documents *how many times* the item is produced in the texts, *who* produces it (i.e. name tag), and also retrieves any identitarian information about the speaker that is available in the database. Thus, the database ensures that sociolinguistic analyses are conducted faster and in a more detailed manner.

³⁹ Derogatory: from a rural background

⁴⁰ IrE derogatory slang: woman who has many casual sexual encounters.

3.8. Corpora used to compare and contrast the CoFIrE fictional data

Since the representativeness of a (fiction) corpus will depend on whether findings based on its data can be “generalised to a larger hypothetical corpus” (Lee, 1991, p. 27), the analyses conducted on CoFIrE were contrasted against the *Limerick Corpus of Irish English* (Farr et al., 2004) and the spoken component of the *British National Corpus 2014* (Love et al., 2017).

Initially, the CoFIrE data was to be compared and contrasted against two corpora that were representative of *contemporary* spoken IrE, namely the Irish component of the *International Corpus of English* (ICE: Ireland) and *SPICE-Ireland Corpus*, which comprises the spoken component of *ICE: Ireland*, annotating discourse, pragmatic, and prosodic items (Kallen & Kirk, 2012, p. 7). *ICE: Ireland* is part of the larger *International Corpus of English* (ICE) project, which endeavors to represent standard, spoken and written World Englishes. At slightly over 1 million words (i.e. 1,079,775), the compilation of *ICE: Ireland* began in the 1990s and spanned a period of 14 years. Despite being of invaluable use to IrE scholars, due to its inclusion of formal and informal Southern and Northern IrE in a wide variety of discourse situations (e.g. face-to-face interaction, to legal presentations, examination essays, scripted speeches, student essays, administrative prose, inter alia, see Kallen and Kirk (2008, p. 9) for a full list), the majority of *ICE: Ireland* texts tend to be more formal than the more conversational CoFIrE ones. Thus, it was thought that the high formality of the *ICE: Ireland* domains would likely prevent it from including the most frequent items (i.e. pragmatic features) annotated in CoFIrE. Pilot searches for elements such as discourse pragmatic items like *yeah*, *no* would later confirm that initial theory. Thus, despite its invaluable contribution to the field of IrE studies, both *ICE: Ireland* and *SPICE* were discarded as baseline corpora.

The only other existing corpus of *contemporary* spoken IrE to date is the *Limerick Corpus of Irish English* (LCIE), which became the first corpus against which the CoFIrE data was contrasted to check the representational validity of the fiction findings (corresponding with RQs 1, 1a, and 2 outlined in section (1.2.)). *LCIE* is a one-million word corpus of naturally-produced IrE, collaboratively compiled in the Republic of Ireland between Mary Immaculate College and the University of Limerick, under the co-supervision of Dr. Anne O’Keeffe and Dr. Fiona Farr (Farr et al. 2004). *LCIE* comprises

a variety of conversational contexts, ranging from *intimate* (e.g. a family chat) to *transactional* (e.g. customer chatting with a taxi driver); *professional* (e.g. academics discussing a conference paper); *socializing* (e.g. chat between friends); and *pedagogic*, such as a lecture (for more information, see Farr et al., 2004). The reason *LCIE* was chosen as a basis for comparison and contrast was the highly conversational nature of most of its texts. However, the use of *LCIE* as a source of comparison for CoFIrE's data has its limitations, as its compilation took place in 2002/2003, while the fiction texts span up to 2014. It must also be noted that *LCIE* is currently being updated. The presently ongoing update began in 2017 and is being undertaken under the direction of Dr Joan O'Sullivan at Mary Immaculate College. *LCIE 2.0.*, as its preliminary name stands at the time of writing, contains the same texts domains as *LCIE*, but unlike its predecessor, *LCIE 2.0.* is pragmatically annotated. However, its small size (63,857 words including annotation at the time of writing), was the reason it was not added as second comparative corpus, although it will definitely be of great use in future investigations.

Given the absence of any other comparable corpus of *contemporary* spoken IrE, and *LCIE*'s time-span limitations, it was decided that the findings based on CoFIrE data would also be cross-variationally compared with a corpus of contemporary standard British English; a variety of English which is geographically close and shares cultural ties with the Republic of Ireland and with IrE. Contrasting against a standard British English dataset would also allow this project to disambiguate any Irish linguistic uses from those utilized in the neighboring variety. Upon consultation with my supervisors and external reviewers⁴¹, the spoken component of the *British National Corpus 2014* (Love et al., 2017) was chosen for such purposes. The reason the original *British National Corpus* was not selected is the fact that it was compiled in the 1990s, with compilation finishing in 1994, which means that the corpus data does not align with the *contemporary* parameter in CoFIrE (see (3.2.2.) for a detailed description of what was taken to be *contemporary*). Compiled between 2012 and 2016 by Lancaster University and Cambridge University Press, the spoken component of *BNC2014* comprises 11,5 million words which take place among friends and relatives in a wide variety of real-life and informal contexts (Love et al., 2017). The use of *BNC2014* allows for the 1) comparison of the findings based on CoFIrE data with a corpus of spoken English which is more up-to-date and aligns better

⁴¹ I would like to thank Dr Brian Clancy (Mary Immaculate College) for recommending the use of *BNC2014* as a second corpus given the lack of comparable corpora of spoken IrE.

with the *contemporary* parameter followed in CoFIrE (linking with RQ 1 and 1a). In addition, contrasting against another variety of English might 2) shed some light into potential linguistic developments that may have or are currently taking place in IrE and are being documented in the corpus (corresponding with RQs 1a and 2 outlined in (1.2.)), and could also be helpful in 3) investigating whether the annotated non-standard pragmatic items in CoFIrE coincide or differ from their use in a neighboring variety (which connects with RQs 2 and 3).

3.8.1. Method of comparison/contrast with LCIE and BNC2014

In order to contrast the CoFIrE findings against *LCIE* and *BNC2014*, and given the large size of both datasets, down samples of 100 randomized occurrences per corpora were taken for each individual item analyzed in the thesis. It must be noted that not all *LCIE* text files were used in the creation of its sample, for some did not contain speaker information, which would hinder the comparison of CoFIrE sociolinguistic data against the *LCIE* sample necessary to answer RQ 2 regarding the type of speaker identity features index. Thus, only those files which were accompanied by demographic information (i.e. 90) were utilized in this thesis. Since *LCIE* is not annotated nor searched via tag, *Wordsmith Tool's* (Scott, 2012) concordancer was used to recover individual items. In the case of *BNC2014*, the corpus was searched using its concordancing tool and tag searches. For example, the tag *<be/V like>* was used to analyze non-standard quotative BE LIKE in sections (5.8.1.) and (5.8.2.). However, this would retrieve all instances where the verb *to be* was accompanied by *like* (e.g. *see what it's like, it used to be like the, or you were like we're just laying down the marble*), which would be manually sorted afterward.

Once all sample occurrences had been gathered from both corpora, they were exported onto an MS Excel file and randomized so as to ensure as much representativeness in such small samples as possible. The first 100 randomized items were, then, compiled as the down samples for each corpus. After their gathering, they were manually sorted (i.e. for example, all cases where *be like* did not function as a quotative verb were classified as non-quotative uses, and, therefore, not counted). The remaining occurrences (i.e. the ones that *did* function as the item under investigation)

were, then, manually classified for 1) form and function. Where necessary, they were classified for clausal position, tense, and subject. In addition, they were also coded for 2) speaker information (i.e. age, gender, geographical location and social class) as this would facilitate subsequent enquiries into the indexicality value of the features being examined (see the methodology sections in chapters (5)-(7) for more on how each individual feature was contrasted against the *LCIE* and *BNC2014* samples, and how the samples were annotated to fit the requirements of each investigation).

3.8.2. Sociolinguistic cross examination

The demographically rich nature of *LCIE* and *BNC2014* enabled the sociolinguistic cross examination of the CoFIrE data against their samples, which would address RQ 2 pertaining to the indexation of speaker identity through the fictional rendition of IrE features. Both corpora contain crucial information regarding speaker age, gender, geographical location, place of birth (or country of origin in the case of *BNC2014*), and location at the time of recording. The length of time living in the area is also available in *BNC2014*.

Furthermore, the speakers' socioeconomic status was also available in both datasets, with *LCIE* providing their level of education, and *BNC2014* coding their social class. *BNC2014* is extremely helpful when investigating social class as it follows the demographic, occupation-based classification system created by the National Readership Survey's social grade classification (National Readership Survey, 2016) illustrated in Table 3.8., and distributes speakers accordingly.

Code	Description
A	Higher managerial, administrative and professional
B	Intermediate managerial, administrative and professional
C1	Supervisory, clerical and junior managerial, administrative and professional
C2	Skilled manual workers
D	Semi-skilled and unskilled manual workers

E State pensioners, casual and lowest grade workers, unemployed with state benefits
only

Table 3.8. National Readership Survey's social grade occupation-based classification in the UK.

This is further refined, as the *BNC2014* speakers are also distributed according to the nine major grade classes included in the *National Statistics Socio-Economic Classification* provided by the British Office for National Statistics (see *The British National Corpus 2014: User Manual and Reference Guide (Version 1.1)*, 2018, pp. 27-28 for more on NS-SEC classes). However, the only sociodemographic classification used in the down samples was their NRSSG classification due to time-restraints.

3.9. The Ross O'Carroll-Kelly Corpus and subcorpora

As already mentioned in section (3.3.), it became evident during the annotation and analytical process of the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE) that the two *Ross O'Carroll-Kelly* novels, written by Paul Howard, included in CoFIrE contained a larger number of IrE features (particularly pragmatic items) than the other books (see sections (4.1.) and (4.2.) for a detailed account of the production of occurrences per book in CoFIrE). Thus, it was decided that the compilation of the first *Ross O'Carroll-Kelly Corpus* (hereafter CoROCK) would be beneficial to answer the research questions (section (1.2.)) of this study through a case study of this 'outlier' author in terms of spoken language features.

In all, CoROCK contains 1,567,256 words and includes 15 novels⁴², all of which span a period of 18 years (2000-2018). Table 3.9. below outlines the novels and plays included in the corpus, along with their date of publication, book code, and approximate word count. Notice how *Downturn Abbey* and *Keeping up with the Kalashnikovs* (colored blue), both of which are part of the CoFIrE corpus, were also included here.

⁴² I would like to thank Paul Howard for his generosity in sending me electronic copies of the novels and the scripts to the unpublished RO'CK stage plays, the latter of which will be used in future publications.

Ross O'Carroll-Kelly Corpus			
Year of publication	Title	Book code	(Approx.) word count
2000	<i>The Miseducation Years</i>	TMY	75,745
2001	<i>The Teenage Dirtbag Years</i>	TDY	71,071
2003	<i>The Orange Mocha-Chip Frappuccino Years</i>	OMCFY	58,914
2005	<i>PS. I Scored the Bridesmaids</i>	PS	82,615
2005	<i>The Curious Incident of the Dog in the Nightdress</i>	CIDN	89,861
2007	<i>Should Have Got off at Sydney Parade</i>	SHGSP	84,884
2008	<i>This Champagne Mojito is the Last Thing I Own</i>	TCMLTIO	102,201
2009	<i>Rhino What You Did Last Summer</i>	RHINO	114,799
2011	<i>NAMA Mia!</i>	NAMA	145,495
2013	<i>Downturn Abbey</i>	PHDA	133,873
2014	<i>Keeping up with the Kalashnikovs</i>	PHKU	126,347
2015	<i>Seedless in Seattle</i>	SS	122,663
2016	<i>A Game of Thrown-ins</i>	GOTI	118,376
2017	<i>Operation Trumpsformation</i>	OT	118,589
2018	<i>Dancing with the Tsars</i>	DWT	121,823
Total word count			1,567,256

Table 3.9. CoROCK books, year of publication, book codes, and approximate word counts

The compilation of CoROCK was conducted in two stages linked to two investigations which involved separate coding: the use of 1) New Intensifying *So* and 2) Non-lexical intensifier *Totally* subsequently. As chapter 8 details, these items arose as salient and worthy of indepth case studies in preliminary analyses of CoROCK. The corpus, therefore, can be divided into two sub-corpora (i.e. CoROCK-SO and CoROCK-T) depending on the books used in each case study. On the one hand, CoROCK-SO

comprises 8 novels which were compiled to investigate the use of New Intensifying *So* across a span of 13 years (i.e. from 2005 through 2018). Having identified NISo as an *over-the-top* intensifier collocating in non-standard contexts (see (8.3.) for a detailed explanation of *over-the-top* intensification), a frequency list conducted on CoROCK-SO identified *Totally* as the next *over-the-top* intensifier, ranking comparatively lower on the list (position 286 with 363 occurrences) than SO (position 50). Given the lower frequency of *Totally* in CoROCK-SO, it was decided that this sub-corpus would be expanded so as to take a diachronic look at the use of *Totally*. The second sub-corpus configured for this investigation, CoROCK-T, therefore, comprises 11 books (6 of which are part of CoROCK-SO), published from 2000 through 2018 and spanning 18 years (see Appendix 3 for a detailed listing of the books used in each subcorpus).

With regard to text preparation, the CoROCK texts were stripped of all editing (e.g. illustrations, chapter titles, editorial information, acknowledgments, etc.) following the same process implemented in CoFIRE and outlined in section (3.5.). Furthermore, and unlike CoFIRE, the CoROCK books were *not fully* annotated for grammar, lexical, and pragmatic items. Instead, the texts used in each case study were manually annotated for the specific items under investigation, and subsequently classified for pragmatic functions and emotion connotation following the classifications developed and implemented in CoFIRE (see chapters (5)-(7)). The methodology of analysis used in each case study is explained in detail in section (8.4.).

3.10. Concluding remarks

This chapter has described in detail the rationale for the building of CoFIRE. It has explained in detail the variety CoFIRE needed to be annotated along with the restricting parameters set for the inclusion of books. The chapter has also outlined the sampling frame and sampling techniques used in the compilation of the corpus, highlighting accessibility and copyright as limitations of this study. In addition, the chapter has described the authors whose works have been curated, while it has offered a detail view of the original and manual annotation system designed and implemented on CoFIRE. The MS Access database which was created to contain the corpus metadata was also described at length here. As the chapter explains, in the building of the main corpus for this study,

it became clear that Paul Howard's representation of spoken IrE features as characterization tools for his novels stood out from all of the other CoFIrE authors. Thus, I saw a clear need for a case study of this author's works through the lens of high frequency features. This led to the compilation of the first corpus of Paul Howard's *Ross O'Carroll-Kelly* books, i.e. CoROCK, which were sub-divided into two datasets for coding and analysis of the high frequency intensifiers which will be analyzed in chapter 8. We now move to the analysis phase of the study. Thus, chapter (4) focuses on analyzing general findings with regards to IrE orality representation in CoFIrE, while chapters (5) through (7) will investigate the top three most frequently reproduced pragmatic items in CoFIrE.

4. OVERVIEW OF GENERAL COFIRE FINDINGS

While CoFIRE contains a total of 1,123,601 words, not all of these were annotated. As mentioned in chapter (3), the annotated *variety* marked those grammatical, lexical, and pragmatic items which were identified as IrE distinctive/non-standard, and certain non-standard (pragmatic) items that are not necessarily exclusive to this variety of English. This chapter provides a detailed account of the number of annotated items and item occurrences, looking also at how these features were distributed across the texts in CoFIRE (sections (4.1.) and (4.2.)). To answer RQ 1 pertaining to which are the most frequently reproduced items of IrE orality in CoFIRE, the chapter also examines the most salient category (i.e. grammar, lexis, pragmatics) in the corpus, and outlines the three most frequently used pragmatic items (sections (4.2.) and (4.3.)), all of which will be analyzed in detail in chapters (5) through (7). Finally, section (4.4.) summarizes the findings pertaining to feature production by social class in the corpus so as to answer to RQs 2 and 3 regarding the type of speaker identity CoFIRE may index in the texts. Concluding remarks are offered in (4.5.).

4.1. Summary of features and frequency of occurrence in CoFIRE

The first step taken to answer RQ 1 regarding which are the most frequently reproduced items of IrE orality in CoFIRE was to explore how many annotated items the corpus contained. In all, CoFIRE has a total of 202 annotated items. It must be noted that in many cases, various individual items were grouped under the same ‘item category’. For example, discourse pragmatic markers such as *you know* are annotated separately depending on whether they appear in clause initial (i.e. <IYK>), medial (<MYK>), or final (<FYK>) positions. Individually, these are annotated as 3 separate items, but could be grouped as 1 feature (i.e. YOU KNOW). Similarly, in the grammar category, the overuse of definite article *the* was annotated differently if it collocated with uncountable nouns, occupations, languages (e.g. “I was well able for *the English* and geography and history” (Ryan, 2013)), quantifiers, etc. In all, there are 9 different cases of overuse of definite article *the* in the corpus, yet one could argue that it is 1 feature. Similar cases may

be found in the vocabulary and pragmatic categories, where variants of the same word or *lemma* (e.g. lemma FUCK, variants *fucking*, *fucker*, *fucks*, etc.) were annotated separately. Table 4.1. summarizes the overall amount of individual annotated items in CoFIrE as explained above along with the number of *occurrences* (i.e. the amount of times the item is produced in the corpus) of each item in the corpus. Table 4.1. also illustrates the amount of items there are when grouped, as discussed above, all of which are outlined with their tags and examples from the books in Appendix 1. For the purpose of this thesis, however, I will take individual items to be representative of the features annotated in CoFIrE.

ANNOTATED ITEMS IN COFIRE			
	Individual Items	Grouped	Occurrences
Grammar	53	28	3,386
Lexis	56	42	3,326
Pragmatics	93	53	12,857
<i>Total</i>	<i>202</i>	<i>123</i>	<i>19,569</i>

Table 4.1. Annotated items and frequency of production in CoFIrE

The quantitative analysis of the annotated items in terms of frequency of occurrence (i.e. amount of times each feature occurs in the corpus) illustrated in Table 4.1. above indicates that CoFIrE contains a total of 19,569 occurrences of annotated items, which represents 1.74% of the corpus word count. While the percentage might seem low, such amount of linguistic items is quite large given the size of this specialized corpus.

The analysis of annotated feature/occurrence per book illustrated in Fig. 4.1. reveals that while all texts contribute more or less annotated items to the corpus, a number of books are responsible for contributing larger amounts of items to CoFIrE.

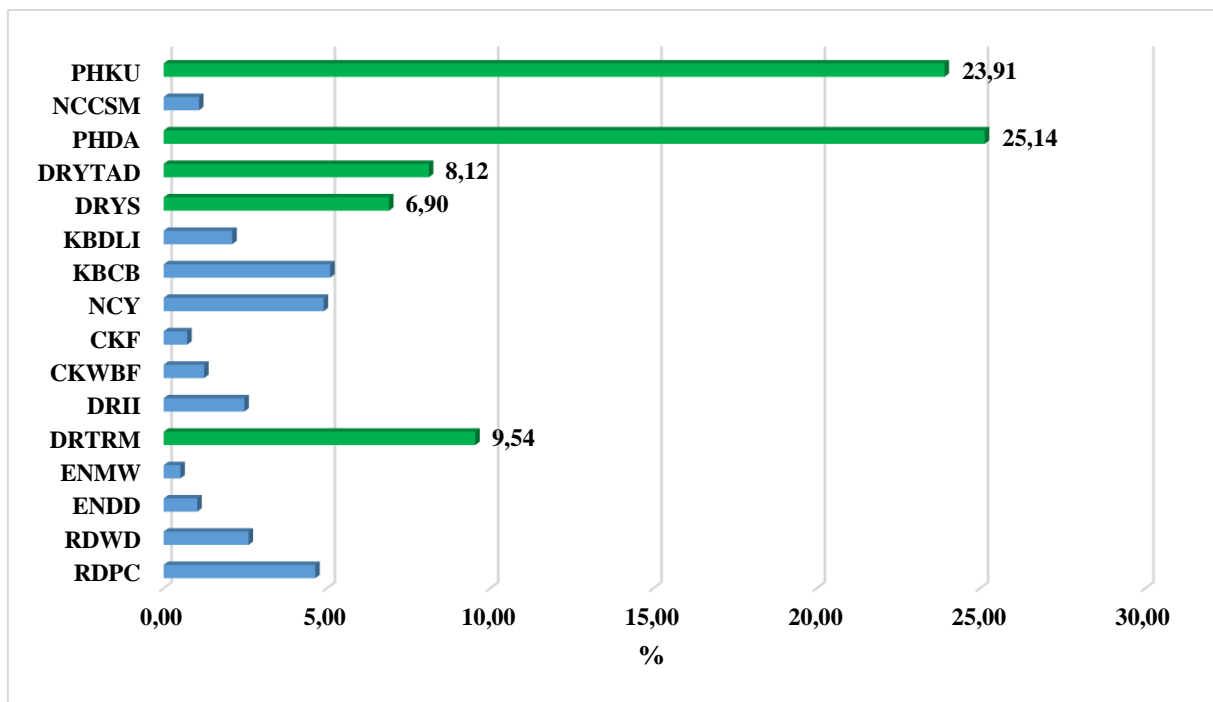


Figure 4.1. Percentage of annotated items per text in CoFIrE.

As would be expected, those books (colored green above) are the ones that contain the most amount of dialogue, with some being highly praised for their depiction of orality. In addition, all are written as first-person narrations, with some claiming to be ‘faithful’ reproductions of actual conversations (i.e. Paul Howard’s novels and Ruane’s *Tales in a Rearview Mirror*). It is perhaps their claim to orality which led to the authors’ making a conscious stylistic choice to use IrE distinctive and non-standard items, as well as other pragmatic elements which could be considered ‘universal’ to most or all varieties of English (see Appendix 1 for a full list). Orality reproduction, therefore, is not just a tool to develop characterization but also a goal which influences the linguistic style of the book, and buttresses the based-on-true-events nature of these texts.

4.2. Pragmatic prominence in CoFIrE

Before continuing with the analysis, it must be noted that, as already mentioned in section (3.3.) and as was illustrated in Fig. 4.1. in the previous section, Howard’s two *Ross O’Carroll-Kelly* novels are responsible for producing a substantially elevated amount of

annotated items, which would see the need for the compilation of the CoROCK corpus (explained in section (3.9.)). In addition, this also led to the undertaking of a more a more detailed quantitative examination of the pragmatic elements in CoFIrE which, as illustrated in Table 4.1. (section (4.1.)), are the most prominent category.

The analysis of feature/occurrence per book revealed that Howard's *Ross O'Carroll-Kelly* books contribute 71.9% (i.e. 9,219 occurrences) of the overall amount of pragmatic items in CoFIrE. Nevertheless, such abundance is to be expected, particularly if we take into consideration the oral nature of these books, all of which, as shall be mentioned in (8.2.1.), are prefaced with "as told to Paul Howard", and are first-person narrations by the main character, Ross, recounting his (mis)adventures (see (8.2.1.) for more on linguistic style in this series). It is, however, clear that Howard functions as the outlier in the group of CoFIrE authors which is due, perhaps, not only to the longer length of his novels when compared to the rest of books (Sinclair, 2005), but also to the series genre and premise. Nevertheless, for the purpose of this study Howard's input was examined in conjunction with that of the other writers, while his elevated reproduction of pragmatic items led to the compilation of the CoROCK corpus (explained in (3.9.)) on which two individual case studies were conducted (as discussed in chapter (8)). Despite the elevated pragmatic depiction in the *Ross O'Carroll-Kelly* novels, pragmatic items continue to be the most represented features in CoFIrE. In fact, if Howard's pragmatic production were to be discounted altogether, the corpus would still have a total of 3,638 pragmatic annotated features, which continues to place this category in the lead for largest amount of items and occurrences in the corpus.

The abundance of pragmatic features/occurrences in CoFIrE, therefore, is important for various reasons. On the one hand, this would answer RQs 1 and 3 respectively (see (1.2.)), highlighting the reproduction of pragmatic linguistic features as the most significantly rendered characteristics of IrE orality in the CoFIrE texts. On the other, the prominence of pragmatic items could also suggest a higher awareness of these items by the CoFIrE authors than was common in older literary renditions of this variety. This prominence could further answer RQ 3 (see (1.2.) pertaining to how modern Irishness may be linguistically indexed in the CoFIrE texts, as it could be the result of a shift in perception, with authors using pragmatic elements as a way to distance themselves from the rendition of IrE found in older books. These often centered on reproducing primarily grammatical and lexical traits, all of which had been perpetuated for centuries via the

Paddy and Stage Irishman stereotypes (described in (1.4.)). The abundance of pragmatic items in the corpus may also represent a stylistic shift whereby authors also utilize pragmatic elements to render *orality* more faithfully. This is certainly the case with Howard's books. From scribbling down what he overhears people (including friends and family) use in conversation which he later assigns to his characters, to italicizing features for stress, placing interrogation marks at the end of sentences to mark uprising intonation, and reading aloud while writing to ensure a faithful rendition (my transcript of Howard, 2019), it is clear that the reproduction of orality is a key element not only in the *Ross O'Carroll-Kelly* series but in Howard's writing style (see (8.2.1..) for more on linguistic style in this series).

I believe that by 'prioritizing' pragmatic elements over grammatical and lexical ones which may be perceived as connected with Stage Irish repertoire (see Amador-Moreno (2010) for an overview of the features connected to the stereotype and section (1.4.) for an example from Murdoch's (1892) short stories), the CoFIrE authors may use these features, as will be shown throughout this thesis, as a) means for documentation of linguistic changes and developments. Furthermore, and given the distinctive nature of the IrE pragmatic repertoire (Barron, 2017), which is key to the indexation of Irishness (Hickey & Amador-Moreno, 2020, p. 15), these writers may also use pragmatic elements as b) markers through which a more modern Irishness may be indexed in contemporary IrE fiction (i.e. RQ 3). Thus, the rest of the thesis will concentrate on examining pragmatic elements.

Another reason for this focus on the pragmatic items of CoFIrE is the fact that, as mentioned in (1.4.1.), despite the recent flourishing of variational (IrE) pragmatics, this field has been largely neglected in academia (Barron & Schneider, 2005, p. 11; Hickey, 2005a; Amador-Moreno et al., 2015, p. 3). However, pragmatic phenomena like the ones annotated in CoFIrE (see Appendix 1 for a list of the annotated items) can be of invaluable use to linguists investigating language change (Amador-Moreno et al., 2015, p. 3). My belief is that these items can be great tools through which we can look "[at] the study of dialect in literature", as pointed out by Amador-Moreno (2005, p. 74) in reference to discourse pragmatic markers, but which can also apply to the investigation of the indexation of modern identity through language in literature/dialogue (i.e. RQ 3). Thus, and so as to answer this thesis' RQs (see (1.2.)) the rest of the thesis and findings chapters

will concentrate on investigating the three most frequently produced pragmatic elements in CoFIrE.

4.3. Top pragmatic items in CoFIrE

The quantitative analysis of CoFIrE with regard to RQ1 pertaining to which are the most reproduced items of (pragmatic) IrE orality in the texts and illustrated in Fig. 4.2 indicates that the most prominent pragmatic items in the corpus are: 1) non-standard quotative verbs, particularly quotative GO (3,174 occurrences), 2) taboo word FUCK in its lemma form, with intensifying *fucking* being the most recurrent variant (1,457 occurrences), and 3) discourse pragmatic marker *like*.

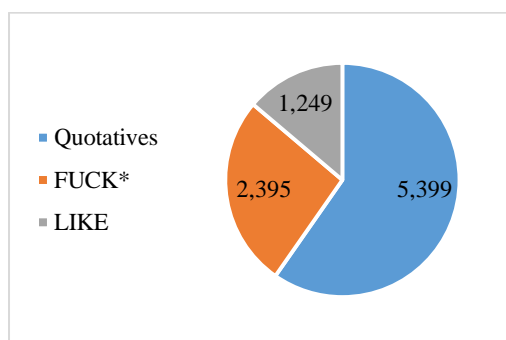


Figure 4.2. Most recurrent pragmatic items/occurrences in CoFIrE.

The following findings chapters (i.e. (5) through (7)) examine these three pragmatic items at length, providing quantitative and qualitatively analyses that investigate in detail their use, form, and pragmatic functionality as portrayed in the texts. To address RQ 2 regarding their indentarian indexation value, these items will also be inspected in terms of the potential influence the macro-social factors of age, gender, geographical location, and social class may have exerted over their use. This part of the analysis will also shed light on the type of modern Irishness these features may index (i.e. RQ 3), at least according to the authors' portrayals in fiction. In addition, all findings will be contrasted against *LCIE* and *BNC2014*, which will aid in answering RQ 1a, assessing 1) the level of realism of the fictionalized rendition of these features by examining whether they are

being used similarly in real spoken IrE and contemporary British English. This contrast will also 2) reveal any potential linguistic trends or developments that are perhaps being recorded in contemporary Irish fiction which speakers may be unaware of. In carrying out these contrastive analyses, the thesis will also be contributing to the field of Variational Pragmatics (Barron & Schneider 2008), as it analyzes pragmatic variation across regional (i.e. in the case of this project Irish and British English) and social varieties.

4.4. Annotated item production by social class in CoFIrE

The shift from traditionally negative stereotypes to the representation of ‘new’ Irishness in fiction which may be indicated by the prominence of pragmatic elements pointed out in section (4.1.) is also mirrored in the frequency of item occurrence production by social class evidenced in Table (4.2.) below.

SOCIAL CLASS ITEM DISTRIBUTION IN COFIRE		
	Number of Characters	Freq. of Occurrence
UPPER	69	9,384
MIDDLE	<i>Low-Mid</i>	2,495
	<i>Middle</i>	3,563
	<i>Upper-Mid</i>	1,192
	<i>Total Middle-class production</i>	7,250
LOW		648

UNKNOWN	35
NARRATORS	2,252

Table 4.2. Number of characters and frequency of item occurrence by social class in CoFIrE.

The quantitative analysis of the frequency of occurrence of annotated items by socioeconomic rank (Table 4.2.) retrieved several findings which are significant when answering RQ 2 regarding their identarian indexical value. On the one hand, at first glance it appears that *upper-class* characters produce the largest amount of item occurrences. However, a more detailed, qualitative investigation into this class revealed that one character (i.e. *Ross O'Carroll-Kelly*) is responsible for producing 92.5% of upper-class occurrences in CoFIrE. Table 4.3. below provides a breakdown of the distribution of *upper-class* annotated item occurrences. So as not to skew the results given *Ross O'Carroll-Kelly's* nature as an outlier within this group, his input was normalized to occurrences per 100 words and subsequently added to the remaining *upper-class* input, leaving this particular rank with an overall production of 794,51 occurrences. Thus, the *middle class* would stand as the major 'item occurrence-producer' rank in CoFIrE (see Table 4.2.), followed by the *upper-class*, and *low class*. Finally, the remaining 2,287 occurrences were produced, as mentioned in section (3.8.), by characters whose social background could not be determined and/or were narrators who were not part of the story's cast.

UPPER CLASS OCCURRENCES			
		Raw	Normalized (PHW)
Overall total	9,384		
	<i>Ross O'Carroll-Kelly</i>	8,682	92.51
	Occurrences produced by other characters	702	7.48
Upper-class production with	794.51		

**RO'CK's
normalized
(PHW) input**

Table 4.3. Annotated item-occurrence raw distribution and normalized per hundred words (PHW) within the *upper-class* in CoFIrE.

The analysis also shows that the *middle class* is over-represented in terms of number of characters when compared to the other ranks, comprising a wide array of occupations (outlined in Table 3.7., section (3.7.)), amongst which many characters are former builders and/or unemployed construction workers. This, coupled with its status as the highest occurrence-producing class, may also be indicative of the fact that the corpus may document a distinct shift in terms of Irish identity representation in fiction (Terrazas-Calero, 2020, pp. 259-260). Moving away from the traditional, negative typecast of the Irish Paddy (explained in (1.4.)), the identity the corpus seemingly registers regarding social class (i.e. RQs 2 and 3) appears to be that of post-Celtic-Tiger-recession Ireland. This seems to be further enhanced and expanded by the documentation of the involvement of contemporary Ireland in a globalized world which appears to be reflected in and indexed through the use of the most frequently produced pragmatic features in CoFIrE.

4.5. Concluding remarks

This chapter has overviewed the number of annotated items (both individual and grouped) identified in CoFIrE as well as their frequency of occurrence, answering RQ 1 by identifying pragmatic items as the most frequently represented category in CoFIrE. While Paul Howard's *Ross O'Carroll-Kelly* books have been found to be outliers with regard to their substantially larger contribution of pragmatic item-occurrences to the corpus, due to which *Ross O'Carroll-Kelly's* production is normalized hereafter so as not to skew results, the production of pragmatic items in the rest of CoFIrE texts is still much more salient than their representation of grammatical and lexical annotated items. The chapter theorizes that the elevated portrayal of pragmatic features in CoFIrE may be representative of a shift in the rendition of IrE in literature (RQ 3), which traditionally

centered on the use of grammatical and lexical items most of which were perpetuated through the negative stereotype of the Stage Irishman (see (1.4.)). This shift may also be due to contemporary authors' wish to stylistically render modern Irish orality in a more faithful manner, documenting potential linguistic developments speakers may be unconscious of. Due to the distinctive nature of the IrE pragmatic repertoire pointed out by Barron (2017), the shift in the linguistic representation of IrE evidenced in CoFIrE could also be a tool through which the corpus authors may encode modern Irishness in literature (i.e. RQ 3). This chapter has also discovered that, with regard to social class/item occurrence and potential modern Irishness indexation (i.e. RQs 2 and 3), the *middle class* is the social rank which is the most represented in CoFIrE, with a majority of characters who are builders and/or unemployed construction workers, among other occupations. The fact that this is also the rank that contributes the most amount of item occurrences to the corpus could further buttress the theory that the linguistic rendering of IrE orality in CoFIrE may be representing a shift in the portrayal of Irishness with regards to social class, perhaps indexing the effects of post-Celtic-Tiger-recession Ireland.

Given the distinctiveness of the IrE pragmatic repertoire with regards to how Irishness is indexed, and the salience of pragmatic rendition in CoFIrE, the following chapters will focus on exploring at length the top three most frequently reproduced pragmatic items in the corpus, namely the use of non-standard quotative verbs and GO in particular (chapter (5)), the prominence of intensifying *Fucking* (chapter (6)), and the use of discourse pragmatic marker *Like* (chapter (7)).

5. NON-STANDARD QUOTATIVE GO

This chapter concentrates on exploring (RQ 1) the most frequently produced pragmatic item in CoFIrE: non-standard quotative GO. The chapter, therefore, begins with an overview of non-standard quoting devices (5.1.), identifying the different non-standard quotative devices in the CoFIrE repertoire (5.2.), and surveying the literature on the use of GO in different varieties of English (5.3.) and in IrE (5.3.1.). The chapter, then, outlines the methodology (5.4.) used in the analysis of the use of GO in the corpus, and then explores in detail the influence the intralinguistic variables of *tense*, *lexical form*, and *grammatical subject* (sections (5.5.) and (5.5.1.) respectively) may exert on the use of this quotative as represented in the corpus. Its pragmatic functionality is also explored in (5.6.), outlining the subfunctions GO may have in IrE as represented in the corpus texts (section (5.6.1.)), while the social markings of this verb are also studied in (5.7.) to address RQs 2 and 3. The validity of the fictional rendition (RQ 1a) is tested in sections (5.8.) through (5.8.4.) which provide contrastive analyses against samples of the *Limerick Corpus of Irish English* and *BNC2014*. The second, most frequently reproduced quotative, i.e. BE LIKE, will also be explored in sections (5.9.) through (5.9.5.) with regards to its use in CoFIrE as contrasted against two samples from *LCIE* and *BNC2014*. Finally, concluding remarks are offered in (5.10.).

5.1. Non-standard quoting devices

People talk most of all about what others talk about—they transmit, recall, weigh and pass judgement on other people's words, opinions, assertions, information [...] Were we to eavesdrop on snatches of raw dialogue in the street, in a crowd, in lines, in a foyer and so forth, we would hear how often the words 'he says', 'people say', 'he said...' are repeated. (Bakhtin, 1981, p. 338)

The above quotation perfectly illustrates how important the “transmissions and interpretations of other people’s words” (Bakhtin, 1981, p. 338) are in everyday interactions. Quotation is, thus, an essential element in (spoken and written) narrations. One of the ways in which quotation can be introduced into narration is by means of quotative verbs.

Quotatives, or *verba dicendi*, are verbs and phrases speakers use to introduce or represent the voices/thoughts of others and/or their own, whether it be in the form of direct or indirect speech. Standard examples of quotatives include *say*, *think*, *ask*, *demand*, or *explain*, among many others. The use of quotation devices is highly frequent in natural interaction, particularly in (oral and written) action-oriented narratives (Barbieri, 2005, p. 231; Tagliamonte & Hudson, 1999) which, consequently, become dramatized. Take, for instance, examples⁴³ below from CoFIrE:

- (5.1.) She **said** a man could not know himself and live alone. (CKWBF)
- (5.2.) Mercy, the Gant **thought**, there's no shutting the kid up. (KBCB)
- (5.3.) 'Where is she?' I **shouted** at Cormac, racing from room to room (NCCSM)
- (5.4.) She's on the screen **going**, 'Fionn told moy to run. He told moy to go aaht the beck door and just, loyke, run as fast as oy could? [...]' (PHKU).

The narrators in example 5.1. above report speaker utterance in indirect speech using the standard quotatives *say*, and applying the internal changes regarding subject, tense, and adverbial variation required by indirect speech form. Examples 5.2. through 5.4., however, relay the utterance in direct form in three different ways. For example, 5.2. reports thoughts directly using the standard quotative *thought*, while the first-person narrator in 5.3. simply transmits to the reader her utterance and how she produced it (i.e. she shouted it). The direct speech quote in 5.4. varies in that the reporter relays to the listener (i.e. the author) a conversation he had with a New Zealander. The direct quotation is introduced by a non-standard quotative GO which prefaces his animated reproduction of the New Zealander's utterance, which almost reads like a 'performed' version of her accented voice.

In the field of Critical Discourse Analysis, Fairclough (2010, p. 76) describes texts (both spoken and written) as elements which "connect events that are removed from each other in time and space". This is particularly true of narratives, where the narration takes place at a different time and place as that of the action it reports. As a result, the report of events can never fully recreate in speech or writing the reported utterance verbatim or otherwise (Romaine & Lange, 1991, p. 230; Blackwell et al., 2015, p. 2). Instead, they seem to reimagine "what it would be like to hear, see, or feel what the original speaker

⁴³My emphasis.

did” (Wade & Clark, 1993, p. 818). In that sense, Tannen (1986) posits that, rather than reported speech, these utterances are ‘constructed dialogues’ which are recreated by the reporter in a narrative. Moreover, when the quotation is done in direct speech, the narration can acquire a level of dramatic presentation that speakers find more attractive and engaging (Schiffrin, 1981; Tannen, 2007b).

One of the ways in which constructed dialogue can be deployed in direct speech is by ‘ventriloquizing’ (Tannen, 2007a, p. 52). This concept refers to the way a constructed utterance is framed by the reporting voice as if it were the dialogue/thoughts of another *in their own voice* (Tannen, 2007b, pp. 21-22), as illustrated in 5.4. above. In doing so, the content becomes a reframed and recontextualized utterance set in the current dialogue (*ibid.*, p. 17). When ventriloquizing, the reporting speaker enacts and animates the other’s dialogue by means of assuming their “pitch, amplitude, intonational contours, voice quality, pronoun choice, and other linguistic markers of point of view” (*ibid.*, p. 22), which also allows them to embellish their own “evaluation of it and to create a recognizable scene as well as captivating rhythm” (*ibid.*, p. 9), which breathes life and theatricality into the narration itself. In ventriloquizing, the ‘performing’ speaker, therefore, adopts the deictic position of the reportee (Romaine & Lange, 1991, p. 229), whether it be a person or animals, mimetic sounds, and other non-speaking reportees, thus effectively switching the conversational footing from authoring the story to bringing another’s voice to life through the tale (Winter, 2002, p. 14; Tannen, 2007b, p. 17).

Ventriloquized reporting, therefore, differs from regular reporting in that it represents another’s voice, thoughts, or sounds in direct speech using a representation of *their own voice*, thus functioning as an almost performance-like reporting technique. As prefacers to dialogue, quotatives can be considered to be one of the most recurrent ventriloquistic reporting tools used in narration, particularly in literature, with which the speaker (i.e. the narrator or characters) prepares the reader to be immersed into the (re)constructed dialogue of the scene or story at hand.

The English quotative repertoire, however, is highly complex and has expanded over time to introduce a wider variety of quotative forms, as shown in example 5.5. below. Functionally, their use can go beyond that of merely reframing prior utterances, for they are dependent on internal linguistic variables (like grammatical subject or tense), but also on external variables, like speaker age and gender (Barbieri, 2005, p. 223), which indicates that they could be great indexes for speaker identity.

- (5.5.) “**I’m there**, ‘**Anto, how the hell are you?**’ **He’s like**, ‘Howiya?’ I say the exact same thing to Tina and her old pair and **they’re like**, ‘Howiya?’ as well. Anto just stares Kennet out of it and **goes**, ‘Look at dat fooken scoombag.’” (PHDA)

Fragment 5.5. is a perfect example of ventriloquized constructed dialogue, where the reporting speaker introduces his own voice (*I’m there*), but switches quotatives (*he’s like*; *they’re like*; *he goes*) to reconstruct the voice of other characters whose dialogue he animatedly reproduces by taking on their accents. This gives the narration a level of theatricality that would be absent if the narrator were to restrict himself to using traditional reporting verbs and ‘un-ventriloquized’ direct speech. The following section explores the non-standard quotative repertoire annotated in CoFIrE, focusing on overviewing the literature on the use of the most frequently reproduced quotative verb (GO) in IrE.

5.2. CoFIrE’s Non-Standard Quotative Repertoire

The non-standard quotative coding system in CoFIrE consists of 7 verbs in total: GO, BE LIKE, BE THERE, BE THERE +*ing*, BE ALL, GIVE, and (PRONOUN) BE HERE. As mentioned in chapter (4), they constitute the most frequent pragmatic item in the corpus, which is to be expected, seeing as how verbs of quotation are such a significant part of narration, and given the highly oral nature of the CoFIrE texts. The quantitative analysis of the CoFIrE quotative system (see Table 5.1. below, and section (5.4.) for a description of the methodology of analysis) indicates an uneven distribution of use among the verbs, showing a clear dominance of GO, followed closely by BE LIKE, and BE THERE, with GIVE, BE THERE+*-ing*, GIVE, BE ALL, and (PRONOUN) BE HERE being less frequently produced.

Quotative	Occurrences
GO	3,174
BE LIKE	1,124
BE THERE	1,089
GIVE	7
BE THERE <i>+ing</i>	2
BE ALL	2
(PRONOUN) BE HERE	1

Table 5.1. Raw non-standard quotative occurrence count in CoFIrE.

It should be pointed out that while GO and BE LIKE appear in 6 out of all 16 texts, Table 5.2. below illustrates how the production of non-standard quotatives in CoFIrE is overly represented (perhaps to call the reader’s attention to them) in the *Ross O’Carroll-Kelly* (RO’CK) novels. Notice how aside from GO and BE LIKE, the remaining quotative verbs are only produced in Howard’s novels. The prominence of these quotatives devices in the RO’CK series was already noted by Amador-Moreno (2015, p. 381; 2016), who finds quotatives to be a paramount feature for the imitation of natural interaction in Howard’s style. Thus, this chapter will focus on analyzing GO in detail, also offering an analysis of the second most frequently reproduced quotative, i.e. BE LIKE. Since BE THERE and the remaining, more infrequent, quotatives appear only in Howard’s books, they will *not* be analyzed in this thesis but will be considered in future publications.

PRODUCTION PERCENTAGE PER AUTHOR					
Quotatives	Authors				
	Howard	Doyle	Ní Chonchúir	Ryan	Ruane
GO	99.8%	0.06%	0.06%	0.06%	0
BE LIKE	99.7%	0	0	0.17%	0.08%

Table 5.2. Percentage of GO and BE LIKE production per CoFIrE author.

Despite their salience in the RO'CK books, the appearance of these quotatives in other books, however, is still significant, for it may evidence the exceptionally oral style of the CoFIrE authors, and their conscious use of these items as framing devices to animate their narration as much as possible. Furthermore, if the narration (in the corpus texts) is understood as a conversation being had between a fictional speaker (i.e. the characters) and a hearer (the readers), then the ventriloquized, almost enacted, rendition of another's voice could be considered a strategy through which authors build "interpersonal involvement [and] create a sense of identification" (Amador-Moreno, 2016, p. 304). Thus, authors effectively produce a sense of closeness and intimacy between the interactants involved in the fictional conversation, which may extend to the readers as well.

The fact that the spoken dialogue examined in this project was purposely constructed by the authors (as the creators of the stories) must also be taken into consideration. However, within those stories, utterances and dialogue are reported, and further (re)constructed and (re)imagined by the characters themselves, which makes quotatives and ventriloquized dialogues all the more recurrent. This is particularly true in the RO'CK novels, which, from the outset, warn the reader that their content is being presented to them "as told to Paul Howard" (as shall be explained in section (8.2.1.)). In other words, they are the (re)constructed oral narratives of a fictional character—Ross O'Carroll-Kelly—as he relays them to the author. In turn, Howard relates them to the reader as if transcribed from a recording, thus heightening the sense of intimacy between author-character-reader (see (8.2.1.) for more on the linguistic style of this series), which also becomes strengthened through the use of other linguistic items like intensifying *fucking*, which will be explored in chapter (6).

The next section provides a brief overview of the use of quotation devices in fiction and how, when carried out through "verbalized" constructed dialogue, they can breathe life into the narration by conferring vividness and dramatism to it. Given the uneven distribution of quotative occurrences within the CoFIrE quotative system, and the dominance of GO, the following section will focus on studying its use in detail. A review of the literature on the use of this verb if offered in (5.3.) and a survey of studies examining its use in IrE is included in (5.3.1.).

5.3. GO in the literature

The study of quotative GO appears to have often taken a back seat to the investigation of its competitor: BE LIKE. While Butters (1980) seems to have been the first to notice the use of what he labeled ‘narrative go’ in American English, brief empirical studies have flourished since then looking at its use in this variety (see Butters, 1980; Blyth et al., 1990; Romaine & Lange, 1991; Ferrara & Bell, 1995; Cuckor-Avila, 2002; Buchstaller, 2002, *inter alia*). Others have documented its surge in varieties like Canadian English (Tagliamonte & Hudson, 1999; Tagliamonte & D’Arcy, 2004), British (Macaulay, 2001; Stenström et al., 2002; Rühlemann, 2008), Australian (Winter, 2002), or even New Zealand English (Baird, 2001; D’Arcy, 2012).

Although seemingly in use longer than BE LIKE (Blyth et al., 1990; Vandelanotte & Davidse, 2009, p. 779; D’Arcy, 2012, p. 353), GO is still a common feature in informal narrations and conversations (Butters, 1980, p. 305; Tagliamonte & Hudson, 1999; Stenström et al., 2002; Tagliamonte & D’Arcy, 2004), despite what perceptions by non-specialist language observers may have suggested in early attitudinal research into the use of GO. For instance, in Blyth et al.’s (1990, pp. 223-24) attitudinal survey in the USA, most respondents claimed that GO and BE LIKE were both “stigmatized, ungrammatical, and indicative of casual speech”, with most speakers perceiving GO as a marker of “uneducated, lower-class males” (*ibid.*). In contrast, Macaulay’s (2001, pp. 11-14) study of quotatives in Glasgow found that GO was favored by middle-class, adolescent boys but more frequent among working-class adult women. Despite the negative response this quotative (and other non-standard verbs) originally received, previous empirical studies into its use in other varieties indicate that GO is one of the most common quotatives to introduce direct reported speech along with *say* and *be like*, (Tannen, 1986, p. 315; see Blyth et al. 1990 for a discussion on alteration between the three in American English), although the latter appears to have taken over as the go-to marker for constructed dialogue as of late, at least in the USA (Macaulay, 2001, p. 6).

The use of quotatives overall, and of GO in particular, has been found to be subject to a range of factors, including internal and extra-linguistic variables. For instance, there seems to be consensus among scholars on the fact that this verb is affected and constrained by external variables pertaining to the speakers’ age and gender, as some investigators

have noticed that it seems to be more prominent among younger (often male) speakers (see Tagliamonte & Hudson, 1999; Macaulay, 2001; D’Arcy, 2012, or Baird, 2002, inter alia). Contrastively, others like Macaulay (2001) have noted preference among females. Female preference was also remarked upon by Stenström et al.’s (2002, pp. 126-7) study of the *Bergen Corpus of London Teenage Language* (COLT), where they found GO was favored by lower-class, adolescent females. This is mirrored in Rühlemann’s (2007, pp. 135-136) study of *I goes* in the *British National Corpus*, in which this quotative also marks lower-class females in their teens and 20s.

In terms of intra-linguistic variables, scholars have taken into consideration the effect of *grammatical subject*, *tense-modality*, and the range of *discourse-pragmatic functions* this quotative frame serves. Thus, studies find that, with regard to *grammatical subject*, GO tends to be used to report *others*, particularly third-person subjects (see Blyth et al., 1990; Baird, 2001, p. 14; or Winters, 2002, p. 13 for Australian English). *Tense/modality* also seems to be a pivotal element in the use of this quotative. GO is strongly associated with the *Historical Present Tense*, which entails the use of Present Simple morphology with Past reference, across all varieties (see Blyth et al., 1990, p. 216; or Schiffrin, 1981, among others). An example from CoFIrE would be “she wearing a little denim mini, which I specifically told her not to wear in work. Oh, it’s so *hot*, she *goes*⁴⁴. Can I not just wear it *this* week?” (Ryan, 2012, italicized emphasis in the original). Notice how *goes* (i.e. *Historical Present Tense*) is reporting an utterance which was made in the past and is embedded within a Past-reference descriptive narration. More examples are offered in section (5.5.), which analyzes at length the influence the variable of *tense* may exert over the use of GO in CoFIrE.

Regarding the potential influence of tense over the use of GO, Baird (2001) investigated the use of quotative devices in a corpus of 25 taped, student interviews in New Zealand during 1995-1996 and 2000. Her findings indicate that GO, which was the most frequent marker in the earlier corpus but had lost its position to BE LIKE by 2000, was most frequently used in conjunction with *Historical Present Tense*, although it could also appear in Past Tense. Similarly, D’Arcy’s (2012, p. 363) diachronic study of quotatives in New Zealand also highlights the dominance of *Historical Present Tense* during what is revealed to be the peak of GO use (i.e. 1998-2001). However, in later years

⁴⁴ My emphasis.

(i.e. 2002-2006) the quotative appeared to favor *Present Simple* and other tense (except for *Simple Past*) and modality aspects. *Tense-modality* variation is also present in Winter's (2002, p. 10) study of quotatives in informal interviews with Australian adolescents. Her study also underscores the dominance of *Historical Present Tense*, while interestingly highlighting the occurrence of GO in *Historical Present Continuous* and *Simple Past Tense*. Finally, Engel & Ritz (2000, p. 136) also note its narrative use with *Historical Present Perfect* in Australia. While the dominance of *Historical Present Tense* is not surprising given the well documented preference for quotatives to appear in *Present Tense* for dramatic purposes (Mustanoja, 1960, pp. 488-506; Poplack & Tagliamonte, 2001, p. 138), the prominence of *Historical Present Tense* with GO, and other non-standard quotatives such as BE LIKE, is significant, highlighting its usefulness as a dramatizing tool which infuses more vividness into the reported content (Carter & McCarthy, 2006, p. 822).

With regard to *discourse-pragmatic functions*, GO can introduce a broader range of constructed dialogue than other traditional, 'neutral' verbs like *say* (Butters 1980, p. 305). For instance, it can present a dramatic *reimagination of constructed dialogue* in the form of direct speech (Schourup, 1982; Blyth et al., 1990, p. 220). It can also *reconstruct inner monologue/thoughts*, as well as function as a *mimetic device* through which another's voice can be expressively animated by accent changes, or intonation and pitch fluctuations, among others (D'Arcy, 2012, p. 348; Buchstaller, 2008), including *mimicking of non-lexical sounds and bodily gestures* (Butters 1980; Romaine & Lange 1991, p. 240; Winter, 2002, p. 14). In fact, GO seems to be associated with "highly [vocal and bodily] demonstrative quotations" (Blackwell et al., 2015, p. 6). As such, this verb functions as a device through which the narration can be infused with dramatism (Blyth et al., 1990, p. 222).

This level of theatricality and vividness is further enhanced by its use in *Historical Present Tense*, as explained above. The deployment of the quotation in *Historical Present Tense*, which is reportedly a very common feature in English narrative (Wolfson, 1979; Schiffrin, 1981), particularly when introducing direct quotations (*ibid.*), is used for dramatic purposes, as it provides readers with a sense of immediacy (Mathis & Yule, 1994) and contiguity that other tenses which make the utterance perfective, completed, and finished lack. Finally, the tense/aspect alternation with GO switching from *Historical Present Tense* to *Historical Present Progressive* (e.g. "The cross-tatto [sic] lad was [...]

ignoring the woman in the tight jeans and for a finish she just stood there **going You bastard.**” (Ryan, 2012)), to *Past Tense* (e.g. “She said God the cold or God the heat, depending on what the weather was like. When she walked across the kitchen **she went Tea tea tea tea tea**” (Doyle, 1993)), or even to the *Present Progressive* (e.g. “He loses it with me then. He storts **going, ‘Ross, don’t go in there! Do not go in there!’**” (Howard, 2013)), can serve to provide even greater dramatization (Carter & McCarthy, 2006, pp. 822-823). This is because it can move the narration and the readers from past to present, as if they were part of the story themselves. Further examples of different tenses are offered in section (5.5.).

5.3.1. GO in IrE

The number of studies conducted on the use of this item in IrE is comparatively smaller than those carried out in other varieties of English. The first to examine this quotative in the context of IrE was Höhn (2012), who performed a comparative analysis of quotatives in *ICE: Jamaica* and *ICE: Ireland* spanning the years 1990-1994 and 2002-05. Her findings reveal *say* as the most frequent quotative, despite BE LIKE and GO also being used. The latter, which was more prominent than BE LIKE, was almost exclusively restricted to informal interactions and private dialogues (2012, pp. 272-274). In her investigation of quotative markers in Irish fiction, Amador-Moreno (2012a, 2015, 2016) looks at BE THERE, GO, BE LIKE, and SAY in Paul Howard’s *The Curious Incident of the Dog in the Nightdress* (2005), finding GO to feature as the most prominent non-standard quotation device in the book.

From a sociolinguistic perspective, Höhn (2012, pp. 283-84) found GO to be slightly more prominent among females, although she observed a trend toward male preference over time. This shift is also found in Amador-Moreno’s study (2016, p. 313), where GO is favored by male speakers in fiction. Furthermore, Höhn (2012, pp. 280-81) noted that while often framing direct speech more frequently than inner monologue, this divide seemed to decrease over time, with GO appearing to have begun to also introduce “male words/thoughts/actions” (Amador-Moreno, 2016, p. 312) in the context of the novel. Finally, and as observed in other varieties, GO is predominantly used in the *Present Tense* (Amador-Moreno, 2015, p. 383), and also appears to be less frequently used for

self-representation (e.g. “‘Okay,’ I hear myself go ‘I’ll be the one who says it to him.’” (Howard, 2014)), favoring the report of others’ voices (e.g. “You tell him that the postman lost your letter to Gwen and he goes, ‘Oh, really? That’s nice’.” (Ní Chonchúir, 2010)), and showing a clear preference for third-person subject pronouns (Höhn, 2012, 278; Amador-Moreno, 2015, p. 385; 2016, p. 311) in the context of IrE. For a detailed analysis of the variables of *grammatical subject* and *pragmatic functionality* and *subfunctions* with examples, see sections (5.5.1.), (5.6.), and (5.6.1.) respectively.

The following sections explore the use of GO in CoFIrE in detail. Section (5.4.) will outline the methodology of analysis used here, while sections (5.5.), (5.5.1.), (5.6.), and (5.6.1.) will explore the extent to which the internal variables of *tense*, *lexical form* and *grammatical subject*, *pragmatic function*, and *pragmatic subfunction* might have influenced its use in the corpus. A detailed insight into the potential *indexical value* (RQs 2 and 3) of this quotative in terms of the external variables of age, gender, location and social class will be offered in section (5.7.). Finally, the findings on its fictional representation will be contrasted (RQ 1a) against *LCIE* and *BNC2014* in sections (5.8.) through (5.8.5.) to test for authenticity of representation.

5.4. Methodology of analysis

The methodology used for the quantitative and qualitative analyses of GO in CoFIrE should be described before outlining any findings since, as mentioned in chapter (3), each analytical chapter (from chapters (5) through (8)) utilizes a slightly different method. In this case, the analysis of the CoFIrE non-standard quotative repertoire was carried out by means of a search for tag <PQ> (ie. *Pragmatics: Quotatives*) using *Wordsmith Tools*’ (Scott, 2012) concordancer, which retrieved the occurrences of all the annotated non-standard quotatives. The quantitative study of this output highlights, as mentioned in in Table 5.1. (section (5.2.)), the dominance of GO as the most frequent verb of quotation in CoFIrE, totaling 3,174 occurrences, which led to the investigation focusing on this particular verb.

In order to examine the potential constrains morphosyntactic and extra linguistic factors may have exerted over GO in the corpus (see (5.2.)), all of its concordance lines

were extracted onto a Microsoft Excel file where they were manually coded for the intralinguistic variables of *lexical form*, *grammatical subject*, *tense*⁴⁵, and *pragmatic function* and *subfunctions*. Furthermore, they were also manually classified for extra-linguistic factors including speaker *age*, *gender*, *geographical location*, and *social class*. This process was also followed in the examination of BE LIKE which is offered in sections (5.9.) through (5.9.5).

5.5. Intralinguistic variables of GO in CoFIrE: *Tense*

As already mentioned in (4.2.), the *Ross O'Carroll-Kelly* (RO'CK) novels are responsible for producing a large amount of the pragmatic items and occurrences in CoFIrE. That also holds for the quotative system and GO. Indeed, the great majority of quotative occurrences (see Table 5.2. in section (5.2.)) occur in these two books, which are presented as 'transcribed' recollections from Ross as he relays them (allegedly verbatim) to the author in the *Present Tense*. As such, this series epitomizes the use of the narrative *Historical Present Tense*. Despite the high frequency of this quotative in the RO'CK books, it must be noted that the quantitative and qualitative analyses of the 6 occurrences produced in the other CoFIrE books also occur in *Historical Present Tense*, although other tenses are also present (see Figure 5.1. below).

The preliminary quantitative analysis of tense(s) indicated the prominence of *Historical Present Tense* (3,113 occurrences), which mirrors previous findings by Höhn (2012) and Amador-Moreno (2015; 2016) in the context of IrE fiction. Nevertheless, their studies do not specify whether it is being used in *Present Simple* or in *Historical Present Tense*, as noted in other varieties (see (5.2.)). Thus, a more thorough examination was undertaken to determine the time reference GO takes in CoFIrE.

⁴⁵ Following Winter (2002), this thesis takes the progressive *aspect* as included within the label of *tense*.

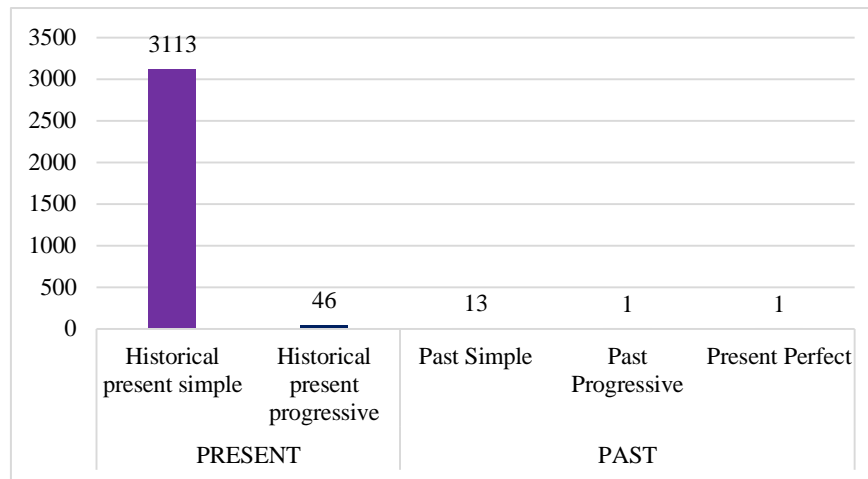


Figure 5.1. Raw occurrences of GO per *tense/aspect* in CoFIrE.

Figure 5.1. above highlights *Historical Present Tense* (in cases like 5.6. below) as the preferred tense in cases such as 5.6., which corroborates previous findings from other varieties (see section (5.3.) for an overview of the literature on GO). In addition, this is indicative of the fact that GO is used as a dramatizing tool in the corpus, which also aligns with previous research (Höhn, 2012; Amador-Moreno, 2015, p. 385). By impersonating the speaker’s own and/or others’ voices, authors are able to inject a level of theatricality into their stories that may not have been possible using any other narrative tenses.

- (5.6.) Honor **goes**, ‘Oh my God, you don’t even know that you’re getting this baby. **Hashtag desperado!**’ (PHDA)
- (5.7.) The cross-tattooed lad was [...] ignoring the woman in the tight jeans and for a finish she just stood there **going You bastard**. (DRYS)
- (5.8.) We were getting up from our knees and Uncle Eddie **went**,—**Grand, grand**. (RDPC)
- (5.9.) ‘He fucking pegged it,’ Oisinn goes. ‘We were shouting at him. **We were going**, “**Fred! Open the fucking door!**” but he just pegged it.’ (PHKU)
- (5.10.) She’s given you an ultimatum, hasn’t she? **She’s** turned around to you and **gone**, ‘**It’s him or it’s me**’. (PHDA)

Although *Historical Present Tense* is clearly the most prominent tense, Fig. 5.1. above also underscores a variety of past and present tenses, as well as progressive/perfect aspects GO may take, some of which have not been documented in other varieties yet. Thus, we find that the second most recurrent form is the (*Historical*) *Present Progressive*,

which is also a well-documented narrative form in English (Schiffrin, 1981). Just like *Historical Present Tense* (e.g. 5.6. above), the use of the progressive in cases like 5.7. appears to infuse the narration with a sense of immediacy and vividness that brings the reader into the story. It also focuses their attention on the progression of the scene and, in 5.7. in particular, on paralinguistic elements of the framed utterance like the physical stance of the speaker in this case.

Unlike in New Zealand (D'Arcy, 2012, p. 363), the use of GO in CoFIrE can also take, although to a lesser extent, *Simple Past Tense* in cases like 5.8. above, where the speaker reenacts his uncle's double reassurance. While almost insubstantial, the use of the *Past Progressive* (in example 5.9.) and *Present Perfect* (see 5.10. above) is worth noting, as existing literature has not documented such cases in IrE or other varieties (with the exception of Winters (2002), and Engel & Ritz (2000) who briefly mention its use in Australia). Thus, I believe the fictional representation of use in CoFIrE presents a valid rendition of real-life uses, while also documenting developments with regard to *tense/aspect variation* which may be reflective of the ongoing process of *pragmaticalization* this feature may be experiencing in IrE. Erman & Kotsinas (1993, p. 79) describe *pragmaticalization* as the process whereby a lexical item in a given context develops "directly into a discourse marker without an intermediate stage of grammaticalization", the latter of which is described by Hopper & Traugott (2003, p. xv) as the change "whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions" (*ibid.*). The potential *pragmaticalization*⁴⁶ of GO in CoFIrE could, therefore, be indicative of the adaptation of GO in the IrE repertoire. More diachronic examinations should, however, be undertaken in order to test this theory.

5.5.1. Lexical form and grammatical subject of GO in CoFIrE

As evidenced in the qualitative analysis of *tense/aspect*, GO appears in various lexical forms (e.g. i.e. *go*, *going*, *went*, and *has gone*), all of which are reflective of time

⁴⁶ For more on the process of *grammaticalization* and *pragmaticalization* and their different features, especially pertaining the *grammaticalization* and *pragmaticalization* of discourse pragmatic markers, see section (7.1.)

reference. A quantitative analysis (Table 5.3.) reveals that *goes* is the favored form, which corroborates previous findings drawn from British English (Biber et al., 1999, 2007, p. 1119; Stenström et al., 2002, p. 118; Rühlemann, 2008) and Irish English (Amador-Moreno, 2015).

	GO	GOES	GOING	WENT	GONE
<i>Raw occurrences</i>	732	2,287	141	13	1
<i>Production Percentages</i>	23%	72%	4.44%	0.40%	0.03%

Table 5.3. Distribution of GO by *lexical form* in CoFIrE.

The investigation focuses now on the use of *goes*. In terms of *grammatical person*, a search for *collocation* patterns containing *goes* as the node word was conducted using *Wordsmith Tools* (Scott, 2012) to gauge which was its most frequent *grammatical subject collocation pattern*. The term *collocation* was coined by Firth (1957) who proposed that the meaning of a word cannot be gathered by studying it in isolation but by looking at the words that accompany it or co-occur with it. Thus *collocation* is a phenomenon by which words or phrases tend to co-occur in context. However, scholars and different corpus analytical software utilize different terminology to refer to the *collocation* patterns that may be found in a corpus. For example, Harris (2006) labels collocations of two or more words as *n-grams*, while these may be referred to in Corpus Linguistics as *multi-word units*, *chunks*, *lexical bundles*, or *clusters* (McEnery & Hardie, 2012, p. 123). The investigation of clusters, which Scott (2012) defines as “words which are found repeatedly together [... and] represent a tighter relationship than collocates”, is significant since the repetition of patterns is one of the traits of constructing fictional dialogue in novels (Biber et al. 2007). The use of *clusters*, or *lexical bundles*⁴⁷, also serves an interpersonal function, both in natural interaction and fictional discourse. According to Carter & McCarthy (2006, p. 835), *clusters* “reflect the interpersonal meanings (meanings which build and consolidate personal and social relations) created between speakers and listeners (writers and readers)”. It must be noted that *collocations* may also occur with grammatical items or categories which do not necessarily carry semantic meaning. These

⁴⁷ Very frequent, multi-word collocations.

are labeled *colligations* (*ibid.*, p. 130). A *colligate*, therefore, would be a grammatical item/category that collocates very frequently with the word being studied. A frequent *colligate* in the English language would be definite article *the*. Figure 5.2. below illustrates the most frequent colligates occurring with *goes* in CoFIrE.

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	THE	THE	THE	ME	SHE	GOES	YOU	YOU	THE	TO	TO
2	TO	TO	TO	THE	HE		IT	THE	LIKE	THE	THE
3	IT	IN	MY	OLD	SORCHA		WHAT	TO	YOU	YOU	YOU
4	YOU	OF	HER	IT	AND		AND	RE	TO	AND	AND
5	OF	ON	OF	THEN	CHRISTIAN		TO	MY	THERE	IT	IT
6	IN	IT	AND	ROSS	OISINN		HE	AND	GOD	OF	OF
7	LIKE	LIKE	IN	AND	HONOR		OH	DON	AND	THAT	THAT
8	HER	YOU	WITH	YOU	JP		WE	IS	IT	IN	LIKE
9	HE	AND	YOU	HER	FIONN		ROSS	IT	IS	HE	SHE
10	THAT	HER	IT	OKAY	SUDDENLY		SHE	WAS	NOT	HER	IN
11	ME	OH	IS	GOD	DEAR		THAT	VE	HE	IS	WE
12	AND	THAT	THAT	WELL	JULIUS		THE	HE	JUST	SHE	ME
13	WHAT	IS	AT	WAY	SHADDEN		ON	CAN	GOING	ROSS	IS
14	OUT	WITH	HIS	THAT	KENNET		IS	WHAT	HAVE	GOING	HE
15	IS	FOR	FOCKING	HONOR	MAN		WELL	ARE	HER	WE	FOR

Figure 5.2. GOES (pro)nominal colligates in CoFIrE.

The quantitative and qualitative analysis of the most frequent *colligation patterns* occurring with *goes* one place to the left retrieved a variety of pronominal forms and personal nouns (illustrated in Fig. 5.2. above). The findings show that *goes* is exclusively used to frame and recreate the voice of third-person subjects. It is necessary to note its absence in non-standard subject-verb copula (e.g. *I goes*) which has, however, been encountered in other varieties (see Strenström, 2002, and Rühlemann, 2007; 2008 for detailed discussions on the use of *I goes* in *The Bergen Corpus of London Teenage Language* (COLT) and the *British National Corpus* respectively, and section (5.3.) for more on the use of GO in the literature). An examination of the most frequent third-person pronominal forms colligating with *goes* was undertaken to determine the most recurrent subject pronoun patterning. The data (Table 5.4.) reveal that while often signaling male voice, *goes* is more frequently used to reframe the voice of female subjects.

	He	She	It	They
<i>Raw occurrences</i>	1,003	1,282	1	1
<i>Percentage of occurrence</i>	43.8%	56%	0.04%	0.04%

Table 5.4. Raw and percentage distribution of subject pronoun + *goes* colligation in CoFIrE.

Although to a minimal extent, the occurrence of *goes* with an impersonal third-person singular subject to introduce *reported writing*, as shown in 5.11., must be noted. Here the narrator uses the quotative to frame his reconstruction of the content of a book that he quotes to his audience. Finally, the occurrence of *goes*+third-person plural, indefinite *everyone* (female) subject in 5.12. is also present in the corpus.

(5.11.) [...] there's, like, a paragraph that **goes**, 'The night had been a rollercoaster ride, not a cheap, three-dollar trip either, but a double-double dipper' (PHDA)

(5.12.) Everyone laughs, then **goes**, 'Aaawww!' thinking it's so funny but at the same time – oh my God – *so* a cute thing to say. (PHDA)

It is worth taking a broader look at the effect the intralinguistic variables of *lexical form* and *grammatical subject* may have on the use of GO (in all lexical forms) in IrE as represented in CoFIrE and visually illustrated in Fig. 5.3. below.

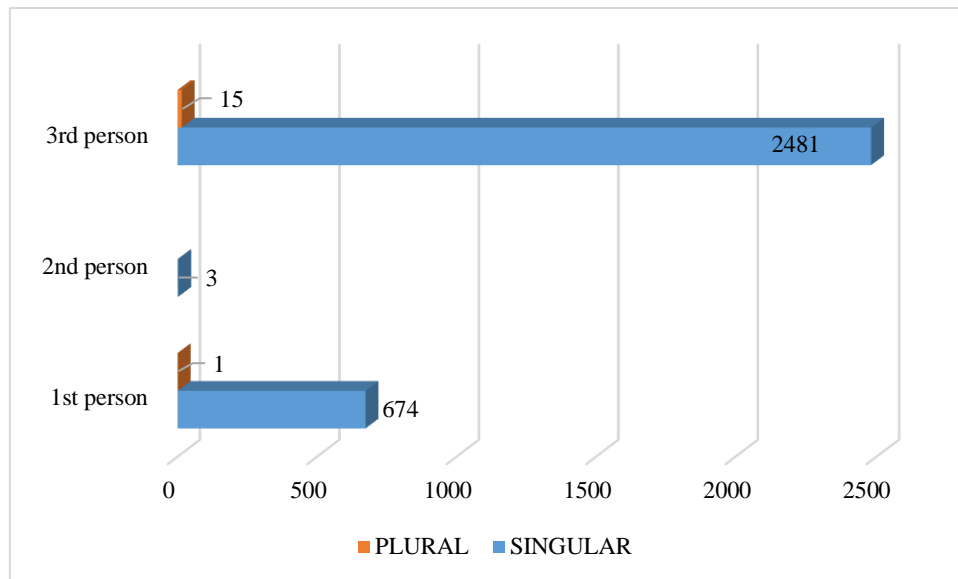


Figure 5.3. Raw number of occurrences of *pronominal spectrum+GO* colligational patterns in CoFIrE.

In terms of *grammatical person*, the quantitative analysis of subject pronoun colligations displays a pronominal spectrum (Fig. 5.3.) which highlights the sheer dominance of 3rd-person singular pronouns, particularly in cases like: “Oh, it’s so *hot*, she *goes*. Can I not just wear it this week?” (DRYS). This signals a clear tendency for the mimetic recreation of others’ voices which corroborates findings in the literature from IrE (see Höhn 2012 and Amador-Moreno, 2015, 2016) and other varieties, where GO also shows propensity to coalescing with 3rd-person subjects (see Blyth et al., 1990, and Barbieri, 2005 for American English; Winters, 2002, p. 13 for Australian English; and Baird, 2001, p. 14 and D’Arcy, 2012, p. 363 for New Zealand English).

In addition, the qualitative and quantitative analyses of these 3rd-person subjects for gender suggest that GO encodes a mimetic reenactment of a female speaker (i.e. 1,397 occurrences) in CoFIrE which is much more prominently rendered than its male counterpart (1,083 occurrences), with only one case of an *it* subject in fragment 5.11. above.

- (5.13.) I go, ‘Honor, what’s wrong?’ (PHKU)
 (5.14.) We were going, ‘Fred! Open the fucking door!’ (PHKU)
 (5.15.) I don’t think you’d just stand there [...] and go, ‘Oh my God, random! When were you first with each other – as in, like, with with?’ (PHDA)

The remaining pronominal subjects also showcase other collocations, such as 1st-person singular subjects (see 5.13.), some of which have not as of yet been documented in other varieties. For example, GO may also frame non-standard subjects, including 1st-person plurals (5.14.), and 2nd-person singular pronouns (5.15.). The presence of these previously undocumented grammatical pronouns is significant because it may hint at a further layer of adaptation and *grammaticalization/pragmaticalization* in IrE which is being recorded in the corpus.

5.6. Pragmatic Functions of GO in CoFIrE

Like in other varieties (sections (5.3.) and (5.3.1.)), the use of GO in CoFIrE serves the overarching pragmatic function of introducing constructed dialogue in a summarized, at times theatrical manner which resembles a verbatim report. However, qualitative analyses find that these recreated, direct utterances can be carried out in two distinct ways in the corpus as illustrated in the fragments below.

- (5.16.) I go, 'What do you mean? To keep doing what we're doing.' (PHKU)
(5.17.) 'Roth,' little Ross Junior suddenly goes, 'woulth you like a thandwidge?'
(PHKU)

On the one hand, there are instances where the direct dialogue is presented to the reader in a simple, verbatim-like manner, which limits itself to summarizing the content of the utterance in a realistic-‘sounding’ fashion. For example, the narrator in 5.16 above simply reports the content of his own utterance. On the other hand, there are cases where the content of the quotation is reproduced mimetically (D’Arcy, 2012, p. 348; Buchstaller, 2008). Thus, the narrator functions as a ventriloquist (Tannen, 2007b, pp. 21-22), introducing additional linguistic elements that adorn the constructed dialogue. In most occurrences, this involves the reproduction of the reportees’ accent and other idiosyncratic linguistic features, like the speech impediment reproduced in 5.17. above. Notice how the narrator not only reconstructs the content of the utterance—an offer of a sandwich—but also reproduces Ross Jr.’s lisp.

The study of the pragmatic uses of GO in CoFIrE to recreate direct dialogue in this manner, therefore, corroborates previous findings (Butters, 1980, p. 305; Schourup, 1982; Blyth et al., 1990, p. 220), as it clearly serves dramatic purposes, as well as a range of other functions. On the one hand, by reframing the content in others' voices and the speaker's own, they adopt the deictic position of the reportee (Romaine & Lange, 1991, p. 229). In addition, in ventriloquizing their accents or speech conditions, as in 5.17. above, the speaker (and the author) brings to life another character (Winters, 2002, p. 14; Tanner, 2007, p. 17) who may or may not be part of that particular scene. This further embellishes and animates the dialogue, while it may also introduce the speaker's "evaluation of [the content of the utterance]" (Tannen, 2007b, p. 9). For example, the narrator in 5.17. reconstructs the offer, but he also mimicks the reportee's lisp. This allows him to enliven the story, while also conveying his animosity and disdain for the child's lisp, which is further stressed earlier in that scene (i.e. 'Hi, Roth!' he goes. He's still got that focking lisp. *I thawt I thaw a puddy cat* and blahdy focking blah. (PHKU)).

In addition, the use of GO to frame life-like reproductions of content also functions as an authorial device in CoFIrE which allows authors to provide the readers with a sense of immediacy (Mathis & Yule, 1994), which is also reinforced by the use of the *Historical Present Tense* (Mustanoja, 1960, pp. 488-506; Poplack & Tagliamonte, 2001, p. 138). This might also create a sense of intimacy between narrator and reader which, coupled with the element of immediacy, could weave a make-belief illusion whereby readers are made to feel as though they are an active member in the story.

5.6.1. Pragmatic Subfunctions of GO in CoFIrE

The quantitative and qualitative analyses of the type of constructed dialogue introduced by this quotative in CoFIrE indicate a wide range of pragmatic subfunctions. As illustrated in Fig. 5.4., its pragmatic functions can, therefore, be divided into two major categories: 1) *representing others' dialogue* and 2) *representing the self*.

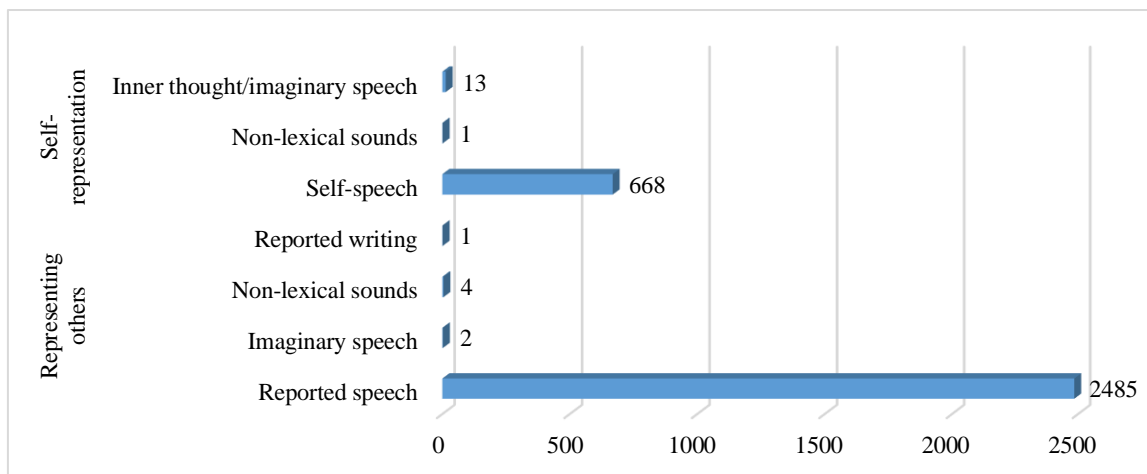


Figure 5.4. Discourse-pragmatic range of GO in CoFIrE.

The findings illustrated in Fig. 5.4. above indicate that while *self-representation* is a significant function, GO is more frequently used in the corpus to *represent others*, which mirrors previous findings on the use of this quotative in the context of Ireland (Höhn, 2012; Amador-Moreno, 2015, 2016), thus suggesting a highly realistic fictional rendition of this quotative in the fiction corpus (i.e. RQ 1a).

Besides offering mimetic recreations of constructed dialogue in direct speech, the qualitative analysis of GO's functional range also exposes an array of extra subfunctions, outlined in Fig. 5.4., some of which have already been documented in the literature (as overviewed in (5.3.) and (5.3.1.)). GO, therefore, may *reconstruct inner thoughts* as in 5.18. below or *imaginary speech*⁴⁸ (see 5.19.) which was never uttered neither by the narrator nor by the characters being reported.

- (5.18.) 'I'm about to go, 'Yeah? And how did that work out for the two of you?' when I suddenly notice that Sorcha's old dear is upset about something. (PHKU)
- (5.19.) 'He probably looked at your text this morning and went, "Oh my God, who is this randomer suddenly texting me?" He probably doesn't even know who you are.' (PHDA)

⁴⁸ I use the term *imaginary dialogue* in reference to occurrences where GO frames dialogue that was never uttered. Instances of *GO+imaginary dialogue* include phrases such as "[subject] would go", "[subject] want her to go", which further highlight the hypothetical nature of the dialogue.

It should be pointed out that both the quotative and content of the utterance appear to be mitigated in the previous fragments. On the one hand, the GO in (5.18.) is used in conjunction with the relative future periphrasis *be about to*, which indicates speaker's *intended* dialogue, although he never delivers the utterance. Thus, this was classified in the corpus as *inner thoughts*. On the other hand, 5.19. exemplifies the reconstruction of *imagined dialogue*. Prior to this fragment, the speaker's (Honor) father enthusiastically relayed to the family a complimentary message he received from a rugby player he admires. However, Honor, who is described as a hellish teenager, prefers to snub his happiness by producing this derisively sneering comment, which is made more so by the introduction of her own *imagined* interpretation of what she thinks the player must have said. Notice how this *imagined dialogue* is deceptively mitigated by the premodifying, downtoner adverb *probably*, which infuses even more sarcasm into the overarching comment. Thus, I believe the use of GO to express *inner thought* and *imagined dialogue* in the corpus, particularly in cases where the quotative is mitigated, could potentially function as a narrative device through which the narrator/author negotiates the fictitious/hypothetical/imaginary nature of the constructed dialogue with the characters and ultimately with the readers.

Another function of GO in CoFlrE, which also corroborates the literature (see, for example, Blyth et al., 1990; Buchstaller & D'Arcy, 2009; Romaine & Lange, 1991; Winters, 2002; Tagliamonte & D'Arcy, 2007; or D'Arcy, 2012, among others), and further underscores the realistic portrayal of use in the texts (RQ 1a), is that of *reconstructing non-lexicalized sounds* in cases such as, “I go, ‘Wow!’” (PHDA), or “She all of a sudden jumps and goes, ‘Owww!’” (PHKU). Finally, it must be noted that, while speech/thought/sounds are clearly the dominant reconstructed content in the corpus, GO can also *frame reconstructed text/writings*. Although proportionally less frequent than any of the other subfunctions outlined in Fig. 5.4. above, GO is found to be used as a frame which reimagines the content of a book in fragment 5.20.

- (5.20.) I stort opening all of the books and reading other lines that are morked. In one – *Members Only* by Anna Pryce – there's, like , a paragraph that goes, ‘The night had been a rollercoaster ride, not a cheap, three-dollar trip either, but a double-double dipper, full of neckbreaking shunts, loop the loops and breathless highs and lows.’ (PHDA)

The fragment above is clearly a reconstructed quotation from the book at issue, since it is virtually impossible for an individual (especially for *Ross O'Carroll-Kelly*) to recollect an entire paragraph and quote it *verbatim* in a time subsequent to their reading it. However, the use of GO in *Historical Present Tense* does not appear to convey the same sense of immediacy and theatricality as in other occurrences in the corpus (see examples in section (5.5.)). On the contrary, it would rather seem like the narrator uses it to make the reader believe the content it frames is an exact—non-dramatized—reproduction of the quotation, thus perhaps serving as a marker of (content) *veracity*. This cannot be conclusively proven since there is only one such occurrence in CoFIrE. Contrastive analyses against real-life uses of GO in such contexts, therefore, should be conducted in the future. However, and despite its virtual inexistence, this function, which is seemingly absent from other varieties, opens up the possibility of the potential pragmatic expansion of this quotative verb in Ireland, as well as the extent to which it may have become *grammaticalized* as recorded in the fiction corpus.

5.7. The Social Markings of GO in CoFIrE

The connection between the use of GO and certain language external (social) parameters, particularly speaker age and gender, has already been established (Barbieri, 2005, p. 223) and explained in detail in sections (5.3.) and (5.3.1.). To investigate the social markings of GO in CoFIrE, and answer RQs 2 and 3 (see (1.2.)), all occurrences were analyzed for indexical value regarding the social parameters annotated in the corpus (i.e. include *age*, *gender*, *geographical location* and *social class*). However, a disclaimer must be issued at this stage. It must be noted that no conclusive inferences can be made as to whether these factors play a role in the use of GO in Ireland (according to the authors' portrayal) due to the reduced amount of available occurrences (only 6) which were produced by characters other than *Ross O'Carroll-Kelly*. In total, 3 out of the 6 occurrences were uttered by narrators whose input in terms of *age*, *social class*, and *geographical location* was labeled “unknown” in Fig. 5.5.

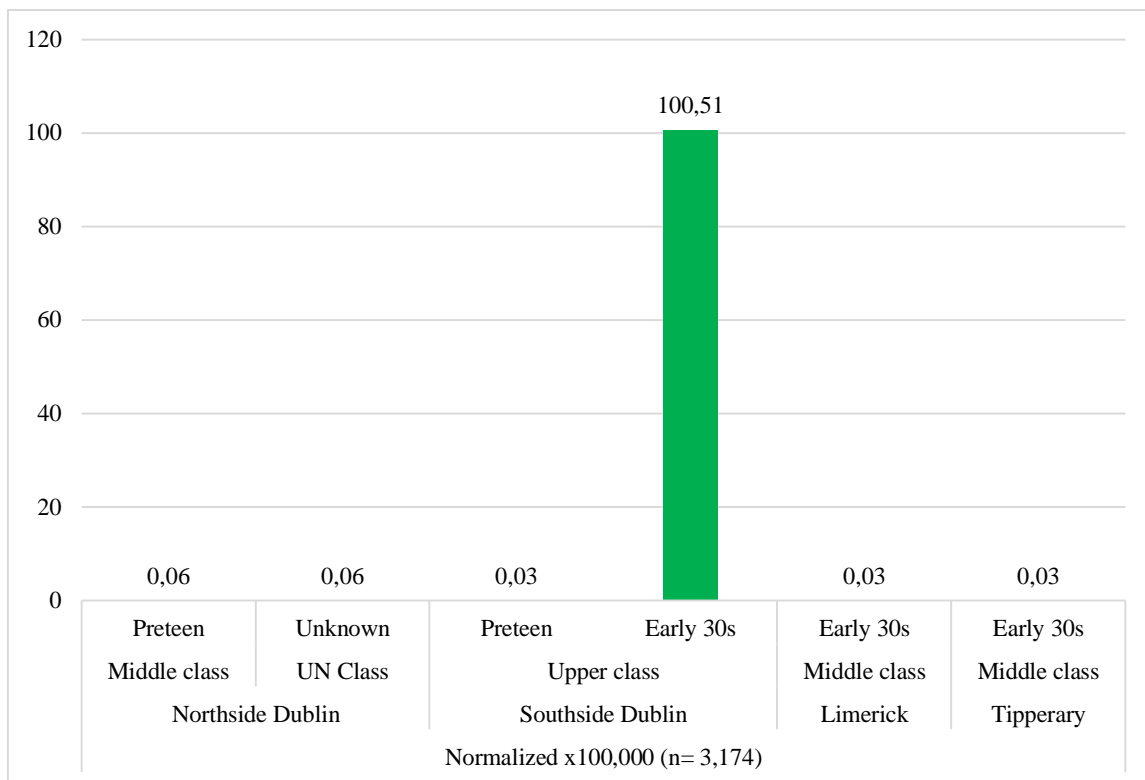


Figure 5.5. GO distribution of occurrence for *age*, *class*, and *location* in CoFIrE normalized by 100,000 words.

The examination of the 6 occurrences produced by characters other than *Ross O'Carroll-Kelly* (RO'CK) seems to suggest less identity indexing than when produced by Southside Dublin speakers of fictionalized Dublin English, especially by RO'CK, whose vast production was normalized so that the results would not be skewed. In terms of age, if RO'CK's input is excluded, then it appears that GO is equally used in the corpus by speakers in their (pre)teens and in their early 30s. Gender-wise (Table 5.5.), the findings indicate that GO is slightly more prevalent among male characters.

Gender	Raw Occurrences	Normalized x100 words
Unknown	3	0.09
Female	2	0.06
Male	3,169	99.8

	Raw	Normalized x100 words
<i>Ross O'Carroll-Kelly</i>	3,166	99.93
Occurrences produced by other male characters	3	0.09

Table 5.5. Gender distribution of GO in CoFIrE raw and normalized by x100 words.

With regard to social class (as illustrated in Fig. 5.5. above), and although at first it would appear that GO is more prominent among the *upper class*, the great majority of those normalized occurrences (i.e. 100,48) were produced by *Ross O'Carroll-Kelly*. If we exclude his production altogether, only 1 other occurrence is produced by an *upper-class* Southside Dublin character. Conclusions regarding class-association cannot be drawn, however, given the fact that the remaining occurrences are equally spread between the middle class and speakers of unknown class ranks. While generalizations must be avoided due to the minimal amount of occurrences produced outside of fictionalized Dublin English, especially due to the over-representation of this item by *Ross O'Carroll-Kelly*, it is possible to hypothesize that the use of this quotative in other regions of the Republic—as represented in the corpus—might also index the voice of a young, *middle-class*, male speaker in his 30s, which would align with the literature.

Although it is not within the scope of this thesis to study the use of this verb in fiction that records different dialects of IrE, these tentative findings warrant a future investigation into such books. This would allow us to get a broader spectrum of use, as portrayed in fiction, which, in turn, would provide an opportunity to determine the indexical value of this verb in regions other than Southside Dublin. Further research into the use of GO and its social markings will be conducted in section (5.8.5) through a contrastive analysis of its production in *LCIE* and *BNC2014*. In addition, the following sections will test the validity of the fictional portrayal offered in CoFIrE by contrasting it against samples of real-life usage drawn from *LCIE* and *BNC2014*.

5.8. Contrasting the CoFIrE data against the reference corpora

As stated in chapter (3), the corpora chosen for comparison were *LCIE* and *BNC2014*. However, and due to their large sizes, the comparison were conducted by means of random samples of 100 occurrences (more on the randomized samples in (3.8.1.)). It is important to mention that four samples of 100 concordance lines were retrieved per main lexical form (i.e. *go*, *goes*, *going*, *went*). This totals an overall sample of 1,600 concordance lines per reference corpora. Once retrieved, all lines were manually analyzed to sort out non-quotative uses. The process of sorting identified 44 and 36 occurrences of quotative *GO* in *LCIE* and *BNC2014* respectively. However, it is worth noting that 34 out of the 44 *LCIE* occurrences were produced in one single conversation by the same speaker, leading me to consider the possibility that this might be an idiolectal feature. The remaining 10 occurrences were uttered by different speakers. All occurrences were, subsequently, examined to evaluate the potential impact morphosyntactic and extra-linguistic factors may have had on the use of *GO* in the contrastive corpora, and to (i.e. RQ 1a) check the level of realism inherent to the fictional portrayal in CoFIrE when contrasted with real-life usage.

5.8.1. *Tense/Aspect* contrast between CoFIrE and *LCIE/BNC2014* samples

As already discussed in (5.5.), the study of the influence of the variable of *tense*⁴⁹ in CoFIrE highlighted the preference for narrative *Historical Present Tense*, while also signaling a spectrum of other *tense/aspect* co-occurrence which is very much replicated in the reference corpora, as evidenced in Figure 5.6. below.

⁴⁹ As previously stated in this chapter, the progressive aspect is included within the concept of *tense* here.

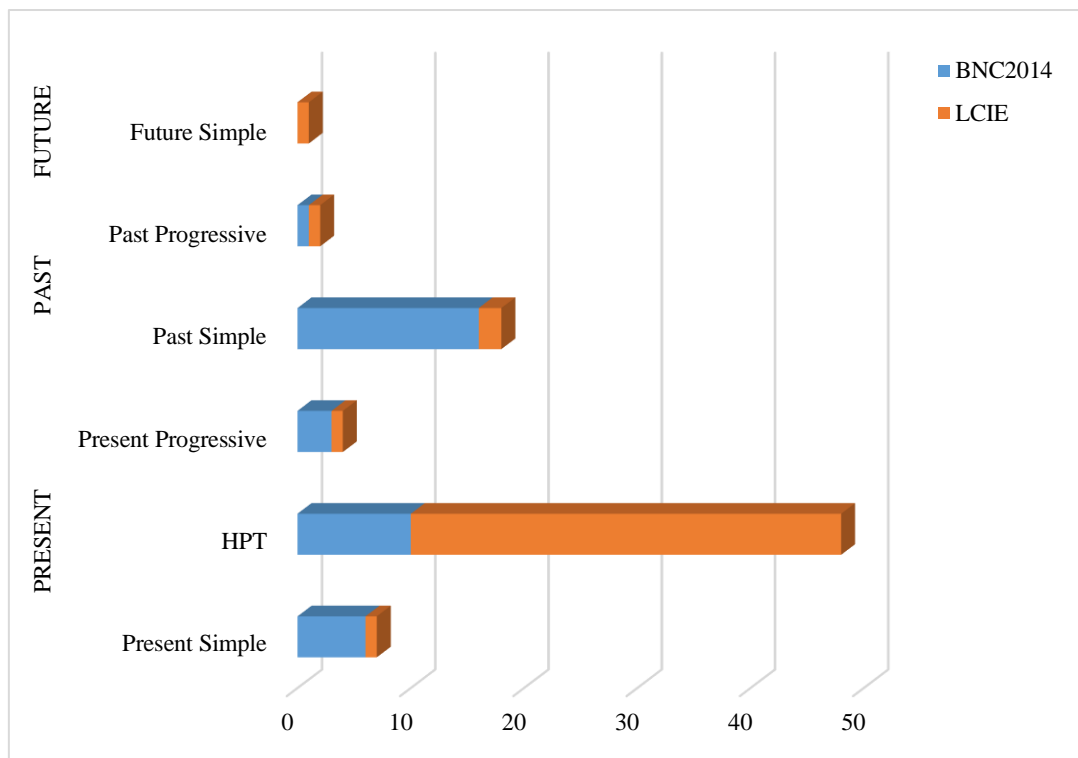


Figure 5.6. Raw distribution of GO by *tense* across the 100-item random *LCIE* and *BNC2014* samples.

As illustrated in Fig. 5.6., the findings from the contrastive samples indicate that GO is generally produced in the *Present Tense*, particularly in *Historical Present Tense*, which validate the level of realism present in the fiction corpus (see Fig. 5.1. in section (5.5.)), thus addressing RQ 1a. However, it should be pointed out that the prominence of *Historical Present Tense* in IrE (as portrayed in CoFIrE and *LCIE*) does not seem to be as clear in the British data, where *Historical Present Tense* and *Past Simple Tense* occur in equal measure (i.e. 16 occurrences respectively). This could, perhaps, underscore a more even *tense* preference for this quotative in British English, which contrasts heavily with the use of this verb in the context of Ireland. The only other noticeable finding in terms of *tense/aspect* is the occurrence of GO in *Future Simple* in *LCIE* in the fragment below:

S1⁵⁰: remember I was supposed to go for an interview at the Hibernian and the statistician the two of them came up at the I time or something and mama said to ring like.

⁵⁰ *LCIE* speaker code names have been changed.

S2: Oh yeah did Mike? No?

S1: Dan goes, “oh yeah you apply for them now next year and they’ll go oh that the fellow now who rang last year who didn’t want it” <\$G> laughter </\$E>
(LCIE)

Notice how in the fragment above, *Future Simple* GO (*they’ll go*) is enclosed in another quotative frame in *Historical Present Tense* (i.e. *Dan goes*). Here, Speaker 1 constructs Dan’s dialogue—which is introduced by *goes* in *Historical Present Tense* — in what can be assumed to be a dramatic imitation of Dan, seeing as the *LCIE* transcriber placed the content of the utterance in inverted commas. Despite being extremely rare in narratology, the use of this *Future Simple* is speculative in nature, functioning as a way of predicting or hypothesizing the story (see Fludernik, 2012, pp. 90-92 for a detailed discussion on the use of the future in narrative). Notice how Dan’s reconstructed voice introduces the second narrative GO in *Future Simple*, which frames the recreated content of what he speculated the interviewers at Hibernian might say. In spite of the absence of inverted commas, this hypothetical content appears to have been rendered humorously, perhaps adopting intonation and paralinguistic elements attributed to the interviewers, which is further reinforced by the speaker’s annotated laughter at the end of their turn. Finally, the occurrence of GO in *Future Simple*—a form that is absent from *CoFIrE* and *BNC2014*—in IrE might delineate another pragmatic function of this quotative, which is that of speculating as to the content of a future utterance in the story in a dramatic way (see (5.8.3.) for more on the pragmatic uses of GO in the contrastive corpora samples).

5.8.2. Lexical form and grammatical subject in CoFIrE and LCIE/BNC2014 samples

The contrastive analysis of preferred *lexical form* (Table 5.6.) also appears to support the findings from *CoFIrE* (see Table 5.3. in section (5.5.1.)), illustrating the salience of *goes* as the most significantly dominant form in the IrE sample, which highlights the level of authenticity (RQ 1a) present in the fictional representation of language use in the books.

Form	CoFIrE (%)	LCIE	BNC2014
GO	23.06	3	1
GOES	72	35	7
GOING	4.45	3	12
WENT	0.40	2	16

Table 5.6. Distribution of GO by *lexical form* across CoFIrE (by %) and raw 100-item randomized LCIE/BNC2014 samples

Contrastively, the British sample shows a preference for *went* and *going*, although it cannot be conclusively determined whether this is truly indicative of lexical differences across both varieties due to the small size of the samples.

Another authorizing finding is the clear preference of *goes* for third-person singular subjects in both corpora evidenced in Fig. 5.7. which had already been highlighted in the fiction corpus (as discussed in (5.5.1.)).

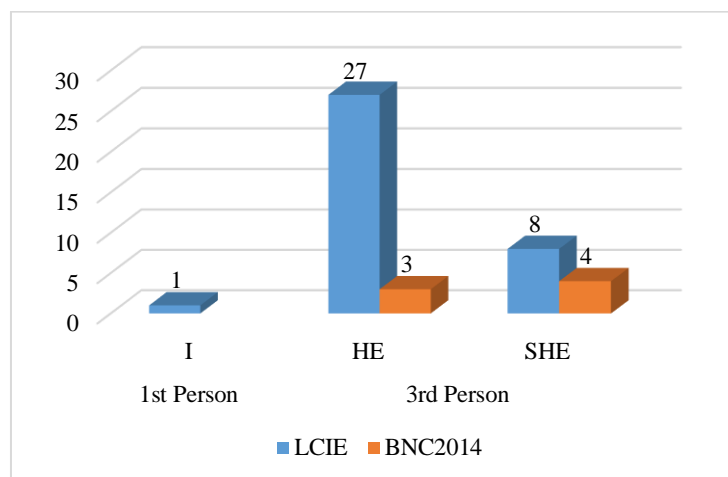


Figure 5.7. Raw distribution by *subject pronoun+GOES* colligational pattern in randomized LCIE/BNC2014 samples.

Furthermore, the examination of the most recurrent third-person subject pronouns of the reconstructed utterances this quotative frames revealed that, in the case of IrE, *goes* clearly favors colligation with masculine pronouns. This differs from its use in CoFIrE, where *goes* introduces female voices more frequently (as pointed out in (5.5.1.)). Perhaps this difference is due to the time the texts in either corpora were produced. For example, the length of time the *LCIE* occurrences of *goes* were produced in ranges from 2000-2003, while the time period covered by the CoFIrE texts spans up to the year 2014. In that regard, the dissimilarity in pronominal preference could perhaps be due to a morphosyntactic development this particular quotative form may have experienced or was in the process of undergoing in the early 2000s, and which is fully documented in the more contemporary CoFIrE texts. This development might be supported by the findings from the more contemporary British sample⁵¹ which, despite showing a quite balanced gender-pronoun distribution, still appears to indicate a slightly higher preference for female-pronoun colligation (i.e. 4 female to 3 male occurrences). However, the size of the samples prevents us from establishing the existence or lack thereof of a gender-pronoun bias in this particular variety of English.

It is also interesting to note the occurrence of non-standard subject-verb agreement in the *Present Tense* paradigm (*I goes*) for self-representative purposes, as illustrated by “And I goes “how the fuck did you know that I was talking to Marie?””(LCIE). The inclusion of theatricality and ventriloquism is also evident in this speaker’s self-report, as the corpus transcriber has placed the content of the constructed dialogue in inverted commas. Despite the presence of this one single occurrence of *I goes* in *LCIE*—absent from CoFIrE and *BNC2014*—it cannot be conclusively determined whether this was a grammatical mistake on the part of the speaker as they also produced two more occurrences of *goes* with third-person subject pronouns. However, Rühlemann (2008, p. 159) notes that the use of this form is “not infrequent in conversation”. In the same study, Rühlemann explored the sociolinguistic variables affecting the production of *I goes* in the *British National Corpus*, and found that it was overwhelmingly used by female teenagers or females in their 20s (*ibid.*, pp. 35-136), which aligns with the findings from *LCIE*, as this occurrence was produced by a female in the same age cohort. This would further suggest a highly realistic fictional portrayal (RQ 1a) of use offered in the CoFIrE texts.

⁵¹ The occurrences of *goes* in the *BNC2014* sample were produced in conversations recorded between the years 2014 and 2016.

Since the quantitative examination of the variable of *grammatical subject pronoun* across all lexical forms of GO in CoFIrE signaled that it was prominently used the mimetic recreator of female voice (as argued in section (5.5.1.)), it was decided that the contrastive samples would also be analyzed for grammatical subject pronoun colligation across all forms of GO. The study of the pronominal spectrum across all lexical forms in both datasets (Fig. 5.8.) lends credibility to the results gathered from CoFIrE as they also display a clear tendency for the reproduction of third-person pronouns.

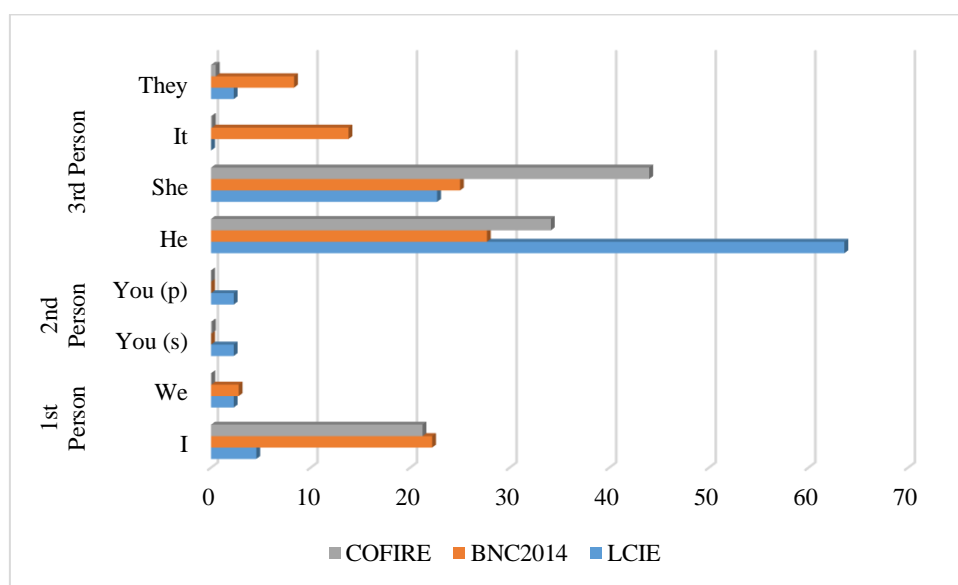


Figure 5.8. Distribution across *grammatical subject+GO* patterns in CoFIrE and LCIE/BNC2014 samples with data normalized to 100.

The analysis of the pronominal spectrum, however, indicates a general preference for colligation with the masculine pronoun *he* in both varieties, which contrasts heavily with its use in CoFIrE, where GO is more prominently used to recreate female voice (as discussed in (5.5.1.)). However, and as mentioned above, this pronominal-preference discrepancy in IrE could be due to a potential morphosyntactic development which might have been in its infancy during the compilation of LCIE but could have progressed and peaked afterward, as the more contemporary CoFIrE data appears to document. In the case of the British sample, and while the representation of male or female voice is rather balanced (at 10 and 9 occurrences respectively), this sample suggests that GO is also more recurrently used to frame male voice. However, more research needs to be conducted into

the influence of this particular variable in British English, as the size of the sample prevents us from drawing conclusions as to whether GO frames male voice more frequently female voice.

A closer look at the pronominal spectrum underlines colligation patterns which were also present in CoFIrE but undocumented in other varieties (see Fig.5.8. and section (5.5.1.)). These include the use of GO to *report collective voice* (i.e. *we*), and the *recreation of the voice of the addressed audience* (at the time of speech within the story) with *you*, at least in the context of IrE. While third-person singular is clearly the preferred subject pattern, it is worth noting the colligation of GO with the third-person plural *they*, but, more importantly, the appearance of impersonal third-person *it*⁵² in the British sample. In this case, the 4 occurrences of *it* present in the *BNC2014* sample represent inanimate or non-verbal subjects, among which we find animals (see 5.21.), inanimate subjects (5.22.), and text messages.

(5.21.) I would ever see with [sic] be a grizzly bear tearing its way into my tent rather than a black bear solemnly sloping off into the woods **going oh he was sweet** (*BNC2014*)

(5.22.) I think some of the science things are a bit too um like babyish or like basic? [...] Like pointing to an arm **going This is an arm like**. (*BNC2014*)

Notice, however, the absence of this pronoun from *LCIE* and yet its documentation in CoFIrE, where it is used to *reconstruct written content*. I believe its occurrence in the fiction dataset not only attests to the representative validity of the corpus (RQ 1a), evidencing morphosyntactic uses present in other varieties of English, but it could also record the *pragmaticalization* of this quotative, which may have undergone a development—in this case, its co-occurrence with inanimate third-person subjects—which is prevalent in modern-day Irish and British English.

⁵² In this case, the 4 occurrences of impersonal *it* appeared in full, nominal phrase form (i.e. *a grizzly bear, this is [...] like pointing to an arm going ...; a cute baby animal; and a text message ...*). However, for simplification purposes, they were represented as *it* in the figure.

5.8.3. Pragmatic Functions across CoFIrE and the contrastive samples

The quantitative and qualitative examination of the range of discourse-pragmatic functions that GO serves in CoFIrE and the comparative corpora (illustrated in Fig. 5.9. below) highlights that this quotative is primarily used to mimetically *report other's speech* in a manner that lends drama to the utterance.

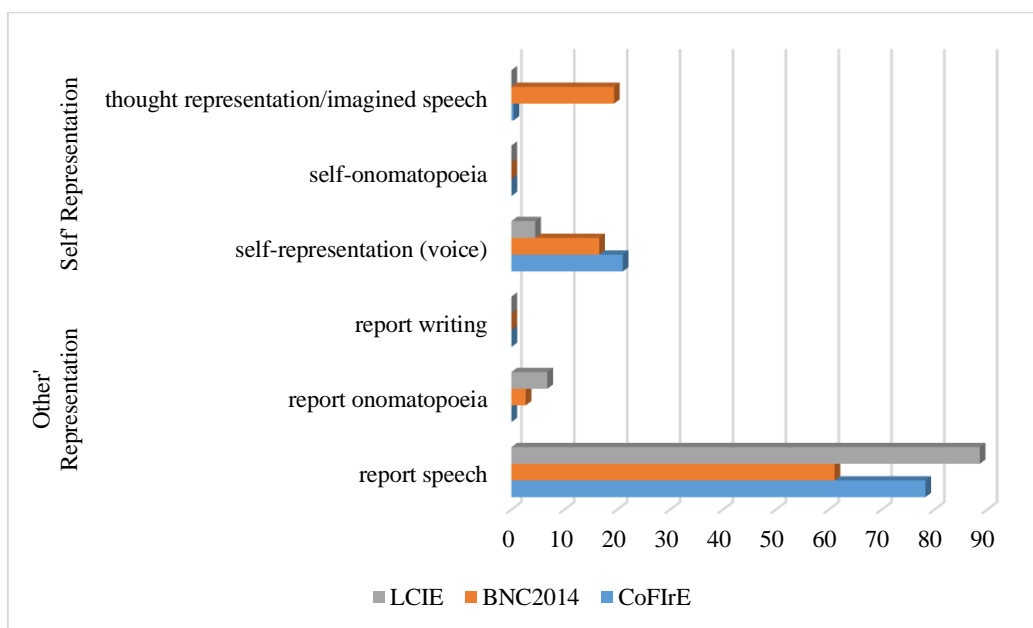


Figure 5.9. Distribution of GO per function with CoFIrE and LCIE/BNC2014 data normalized to x100 words.

While *reporting others* is significantly the most frequent function across all corpora, it is important to examine the other functions in the pragmatic spectrum that were uncovered by the qualitative analysis of all the corpora. For example, the pragmatic range in LCIE includes speakers representing others not only to *report speech* but also to *recreate non-lexicalized onomatopoeic sounds* like 5.23., while self-representation is limited to *reconstructing the speaker's own voice* 5.24. Notice, for instance, how the performance-like recreation of the non-lexicalized sound of the flare gun in 5.23. is annotated by the transcriber of the corpus with a descriptive metadata tag, while the speaker's reimagined self-utterance in 5.24. is transcribed in inverted commas.

- (5.23.) No you go into the office you open up the top drawer get out the flare gun and go <\$E> sound imitating the firing of a flare gun. (LCIE)
- (5.24.) I goes “how the fuck did you know that I was talking to Marie?” (LCIE)

Comparatively, Fig. 5.9. indicates that the *LCIE* pragmatic range is smaller than CoFIrE’s. The CoFIrE data shows GO can be used to frame a series of ‘new’ pragmatic contexts, most of which are documented in the *BNC2014* sample, but missing from *LCIE*. Thus, we find that in British English quotative GO can report *thought/imagined speech* like in the fragment below:

S1⁵³: it's quite harsh I mean Disney's like famous for pushing people's buttons like that
 S2: mm
 S1: like right straightaway a parent dies
 S2: cos they did exactly the same in The Lion King didn't they?
 S1: uhu
 S2: and it's the same thi- you know
 S1: the same exact
 S2: they kill off the parent and just this whole having a cute baby animal going mummy?
 S1: or daddy?
 S2: d- dad and er
 S1: why won't you wake up?

(*BNC2014*)

It is important to mention that 3 of the 4 occurrences of *thought/imagined speech* in the British sample are used to frame the voice of non-verbal subjects, namely ‘speaking’ animals and body parts. The fragment above is a good example. In this case, the speakers discuss Disney’s dramatically controversial (at least for these speakers) decision to kill Simba’s parents in *The Lion King*. At one point, S2 decides to emphasize how emotional that decision was for the audience by recreating and ventriloquizing Baby Simba’s *recreated speech* when finding his deceased father (i.e. colored in green). What is interesting is the fact that both speakers become involved in the dramatized reconstruction of the lion’s dialogue. Notice also how both speakers convey the lion’s discourse with an interrogative tone, which infuses more melodrama into the dialogue by conveying his disbelief and devastation at finding his father unresponsive and dead.

⁵³ The BNC2014 speaker code names have been altered.

The British sample also includes instances where GO introduces *reconstructed written content* like 5.25., where the speaker relates to a friend a (rather unsatisfactory) text message she received as reply to a card.

(5.25.) so I sent her an Easter card [...]and just got a text message back *going this isn't my address anymore* (.) um (.) *I've just been on this crazy hen night* (.) whatever (.) blah blah. so I said oh okay (.) cool (.) blah blah (.) er what's your address then? she wrote back and she *went okay well enjoy your Sunday*. (BNC2014)

There are two occurrences of quotative GO (*going, went*) introducing recreated written content in 5.25. above. However, it is obvious that while the speaker is relaying the *general* content of the written message, this is not a verbatim account, as evidenced by her use of *whatever* and *blah blah*. Although absent from *LCIE*, the use of GO to *reconstruct written content* is also present in CoFIrE, a fact which may be indicative of a later pragmatic development in IrE. However, its presence in the *BNC2014* sample and absence from *LCIE* might also be due to the British corpus being much larger in general than the Irish corpus (see (3.8.) for a description of the corpora sizes), which, therefore, offers the possibility that it may have captured more types of casual conversations. Future research should be conducted to test the validity of this theory. Finally, the British sample also shows cases of *reimagined reimagined (inner)thought* as in 5.26. below, which is a function that was also documented in the fiction dataset but is missing from *LCIE*.

(5.26.) the eighth week I was just sitting there *going well everything I have to do now I have until May to do so em* (BNC2014)

The comparative, cross-variational analysis of the pragmatic functions of GO in CoFIrE has, therefore, revealed a wider variety of uses than those present in *LCIE*, but most of which align with the *BNC2014* sample. I believe this could underscore the authentic portrayal of this quotative in casual conversation that the CoFIrE authors present in contemporary Irish fiction, while also indicating that the corpus attests to the *pragmaticalization* of this quotative in the context of contemporary Ireland. Such process may have been in progress during and after the compilation of *LCIE*, as the fiction data

seems to suggest a functional extension into ‘new’ pragmatic contexts that are also present in other varieties.

5.8.4. Social markings of GO across CoFIrE and contrastive samples

The analysis of the social markings of GO in CoFIrE (already discussed in (5.7.)) indicates that it is used in the texts to index the distinct voice of an *upper-class*, male speaker from Southside Dublin who is in his 30s. However, that is mostly due to its over-representation by the character of *Ross O’Carroll-Kelly* (RO’CK) in Howard’s books. If RO’CK’s input is disregarded, then GO may index the identity of a *middle-class* male in his 30s. To address RQs 1a and 2, this section examines whether the CoFIrE findings present a valid sociolinguistic representation of the use of GO in spoken Irish and British English with regard to age, gender, and social class.

	GENDER			AGE				
	Male	Female	Unknown	Teens (14-19)	20-30	40-50	60+	Unknown
<i>LCIE</i>	22	21	1	-	40	1	1	2
<i>BNC2014</i>	10	26	-	7	19	9	1	-

Table 5.7. Raw age and gender distribution of GO across *LCIE/ BNC2014* samples.

The indexation of gender was quantitatively analyzed in both samples. At first, the findings (Table 5.7.) from *LCIE* seem to validate the fiction data, as GO would appear to index masculine voice in IrE. However, the gender disparity is minimal, so generalizations cannot be made as to gender encoding. Interestingly, however, the cross-variational analysis revealed that GO is seemingly more recurrent among female speakers in British English. While at this stage it cannot be determined whether this marks cross-variational gender distinction between IrE/British English given the restricted amount of British data in the sample, it is still important to note this potential discrepancy. Furthermore, the

quantitative study of the effect of age also seems to validate the fictional portrayal in CoFIrE, with GO being more frequent among speakers in their 20s-30s in both varieties.

Before analyzing quantitatively the influence of social class, all *LCIE* occurrences were manually coded using the demographic classification system followed in the *BNC2014* which is outlined in detail in section (3.8.2.). It is also worth noting that given the broad nature of these categories, the distribution of *LCIE* speakers was conducted according to their occupation, with most GO-users being students (and classified under category E).

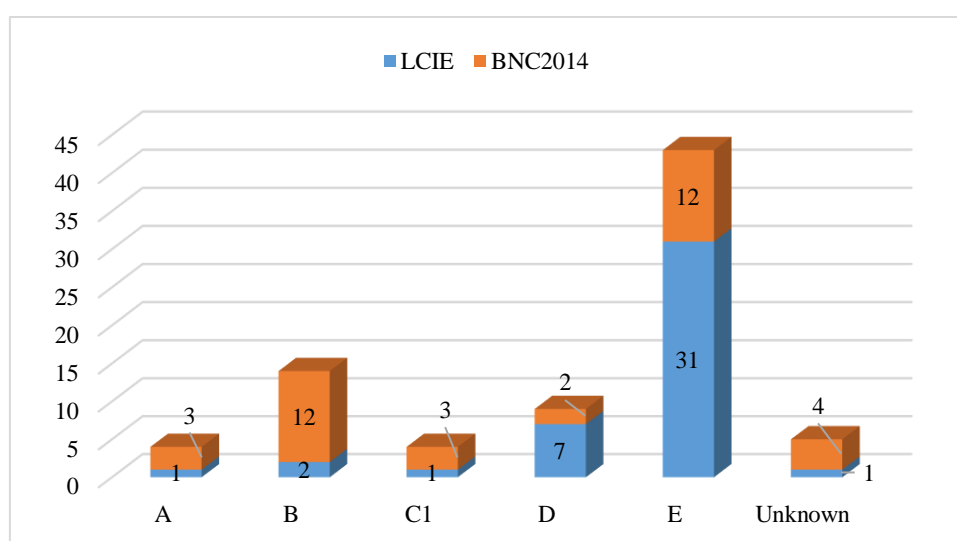


Figure 5.10. Raw distribution of GO by social class in 100-item *LCIE/BNC2014* samples.

Thus, the quantitative examination of the social class of the users of this quotative in IrE and British English (Fig. 5.10.) indicates that, in the case of *LCIE*, it is mostly used by students (category E) and people from a (low) *middle-class* background (category D). This shows some contrast with the CoFIrE findings, which indicated prominence among *upper-class* speakers (in the case of fictionalized Dublin English), and *middle-class* speakers from outside South Dublin; a fact which perhaps records a spread to other social ranks within Irish society in modern times (i.e. RQs 2 and 3). The correlation between the use of GO and the *middle-class* also seems to be prevalent in the case of British English, where we find an even distribution of use within the middle-class between young students (category E) and category-B speakers. This corresponds with findings from earlier

research, which also established a favoring among *middle-class* adolescent boys and working-class adult women in the case of Glasgow (Macaulay, 2001, pp. 11-14). However, these findings contrast with Stenström et al's study (2002, pp. 126-7) which highlighted a prevalence among *lower-class* adolescent females in the *Bergen Corpus of London Teenage Language*, also mirrored in Rühlemann's (2007) study of *I goes* in the *British National Corpus*, who found it was most recurrently produced by *lower-class* female adolescents and females in their 20s. Again, the findings from the British sample might indicate a recent shift of social indexicality in the use of GO in the context of British English, signaling upward social mobility to the middle class.

In all, it could be said that the fictional texts in CoFIrE offer a realistic representation of GO (RQ 1a), which appears to mirror its use in real, spoken Irish and British English. Nevertheless, it is necessary to reiterate the fact that the CoFIrE data offers a portrayal of real use as perceived by the authors, and that the findings from the contrastive, cross-variational analyses cannot be taken as conclusive evidence of class-preference due to the size of the samples examined. Therefore, I would like to address the need for a more detailed investigation of the use of this quotative in both varieties which allows researchers to examine contemporary usage with more updated demographic data.

5.9. The remaining CoFIrE quotative repertoire: BE LIKE

We now turn to the second, most frequently reproduced non-standard quotative: BE LIKE. The use of quotative BE LIKE in cases like 5.27. below is often commonly (mis)associated with a very specific linguistic stereotype within American English: the *Valley Girl*. This is a persona that typecasts young, middle-class, females from the San Fernando Valley in California (see Tagliamonte, 2016; D'Arcy, 2017 for further discussion), and which spread through pop culture with songs like Frank Zappa's 'Valley Girl' (1982), or in cinema through the homonymous Atlantic Pictures (1983) movie and Paramount's (1995) *Clueless* (see section (7.6.1.) for more on the cultural nuances and linguistic construction of this stereotype).

(5.27.) [...] so I said yeah that's totally cool, and he was like, yeah cool well I'll be in Renard's after the gig. (Ruane, 2003)

While BE LIKE is generally perceived to have originated in the United States (Romaine & Lange, 1991, p. 248), where it might have likely appeared in the 1960s and 1970s (D'Arcy, 2017), its use in American English was first reported in the 1980s by Tannen (1986). Since then, however, it (or at least awareness of it) seems to have spread outward to the rest of the English-speaking world. Thus, it is documented in African American (see Cuckor-Avila, 2002; and Kohn & Askin Franz, 2009 for Hispanics and African Americans in North Carolina), Canadian (Tagliamonte & Hudson, 1999; D'Arcy, 2004; Tagliamonte & D'Arcy, 2004, 2007), English English (Buchstaller, 2006a, 2006b, 2011), Scottish (Macaulay, 2001), New Zealand (Baird, 2001; Terraschke, 2010; Buchstaller & D'Arcy, 2009; D'Arcy, 2010, 2012), Australian (Winter, 2002; Rodríguez Louro, 2013), and Irish English (Höhn, 2012; Nestor et al., 2012; Schweinberger, 2015; Amador-Moreno, 2015, 2016; Terrazas-Calero, 2020).

Sociolinguistic analyses to test its 'Valley Girl' association, examining the correlation between BE LIKE and speaker age, region and social class, have been conducted and seem to coincide in connection with younger speakers (see Buchstaller & D'Arcy, 2009; Baird, 2001; Winter, 2002; and Rodríguez Louro, 2013, *inter alia*), often indexing teenage voice across varieties (Andersen, 2001; D'Arcy, 2017; King, 2010; Tagliamonte, 2016). Gender-wise, this quotative appears to be cross-variationally unstable, with some scholars noting its preference among female speakers (e.g. Blyth et al. 1990 for American English, Tagliamonte & D'Arcy, 2004, p. 506 for Canadian English, or Rodríguez Louro, 2013 for Australian English), while others remark upon its saliency among men (e.g. Buchstaller & D'Arcy, 2009, p. 309 in English English). In addition, Ferrara & Bell (1995) noticed that gender differences neutralized as the quotative expanded, and yet other studies document more discernible gender-distinctions leaning toward females as it becomes *grammaticalized* (see Tagliamonte & Hudson, 1999, and Tagliamonte & D'Arcy, 2004; 2007).

BE LIKE has also received scholarly attention in IrE, where it seems to index distinct identities, encoding the voice of young female speakers in the 20-30 age cohort (Schweinberger, 2015, p. 131). Contrastively, Amador-Moreno's (2015) investigation of its use in Howard's *The Curious Case of the Dog in the Nightdress* evidences an overwhelming use by young (20-30s), upper-class, male characters from Southside Dublin.

5.9.1. The social markings of BE LIKE in CoFIrE

The data from CoFIrE as regards the use of BE LIKE initially mirrors Amador-Moreno's (2015) findings, answering RQ 2 by indicating prominence of by Southside Dublin, upper class, male characters. This is, however, to be expected seeing as how her investigation uses a novel that is an earlier installment of the *Ross O'Carroll-Kelly* series. Upon close inspection, the quantitative analysis of those occurrences revealed that the majority (i.e. 1,120) were uttered by one single character: *Ross O'Carroll-Kelly* (RO'CK). The remaining 4 occurrences were also produced by *upper-class*, Southside Dublin, *female* characters in their 20s-30s.

It should be noted that while BE LIKE in CoFIrE is almost exclusively present in Howard's RO'CK novels (i.e. 1,121 occurrences), 3 of the 4 remaining occurrences were produced in Ryan's (2012) and Ruane's (2003) books, where the females are also young South Dubliners. Given RO'CK's overproduction, however, it cannot be conclusively determined whether or not BE LIKE indexes a feature of male speech in IrE. However, the correlation between this quotative and various female characters from Southside Dublin *across other books* in CoFIrE leads me to consider the fact that BE LIKE may be developing into (RQs 2 and 2) a clear marker of female identity in Dublin, indexing linguistic trendiness, cosmopolitanism, and status.

5.9.2. Social markings of BE LIKE in LCIE and BNC2014 samples

Two 100-line samples were used for the contrastive analyses against the reference corpora to check how realistic the fictional CoFIrE portrayal was (RQ 1a). The quantitative analysis of the *BNC2014* sample revealed 71 occurrences of the quotative. With regard to *LCIE*, and since this corpus is not annotated, the occurrences of BE LIKE were retrieved searching for the word LIKE using *Wordsmith Tools*' (Scott, 2012) Wordlist function. Its concordancer tool was subsequently used to retrieve all occurrences of LIKE (i.e. 2,803), which were manually examined for cases of the non-standard quotative. Only 71 occurrences of quotative LIKE were found in the *LCIE* texts examined in this project (see (3.8.1.) for more on which *LCIE* texts were examined). A manual, qualitative analysis (illustrated in Table 5.8. below) sorted BE LIKE variants, indicating a total of 58 occurrences of non-standard quotative BE LIKE out of the original 71, with the remaining being variant forms. For the purpose of this section, only the 58 BE LIKE occurrences

will be analyzed from here onward. When contrasted against the frequency of GO in *LCIE*, the occurrence of BE LIKE may be indicative of the on-the-rise status of this particular verb at the time of compilation.

Variant Form	Occurrences	Normalized to x100 words	Examples
<i>Be like</i>	58	81.6	Mary <i>is like</i> "I can't hear you I can't hear you."
<i>Go like</i>	2	2.81	Am a friend of mine did that she <i>went like</i> oh it was all loving and whatever but am but last summer
<i>Go on like</i>	1	1.4	<i>he kept goin' going on like</i> "oh every girl for the last six years I've been with has done the dirt on me"
<i>Say like</i>	9	12.6	Triona rang him and eat him and he <i>said like</i> "I'm really sorry but Charlie done it to my house"
<i>Start off like</i>	1	1.4	she's <i>start off web [sic] like</i> "I think it's a fucking disgrace like"

Table 5.8. Distribution of BE LIKE by form and occurrences in *LCIE* sample.

The correlation between BE LIKE and *female* voice and age found in CoFIrE (outlined in section (5.9.1.)) is also evident in the samples. The findings from their contrastive analyses with regard to gender (illustrated in Fig. 5.11. below) indicate an overwhelming use by females in both varieties which mirrors previous findings from American (Blyth et al., 1990), Canadian (Tagliamonte & D'Arcy, 2004, p. 506), or Australian English (Rodríguez Louro, 2013). This further highlights the level of authenticity and realism (RQ 1a) that is offered in the fiction dataset with regard to BE LIKE reproduction in the texts.

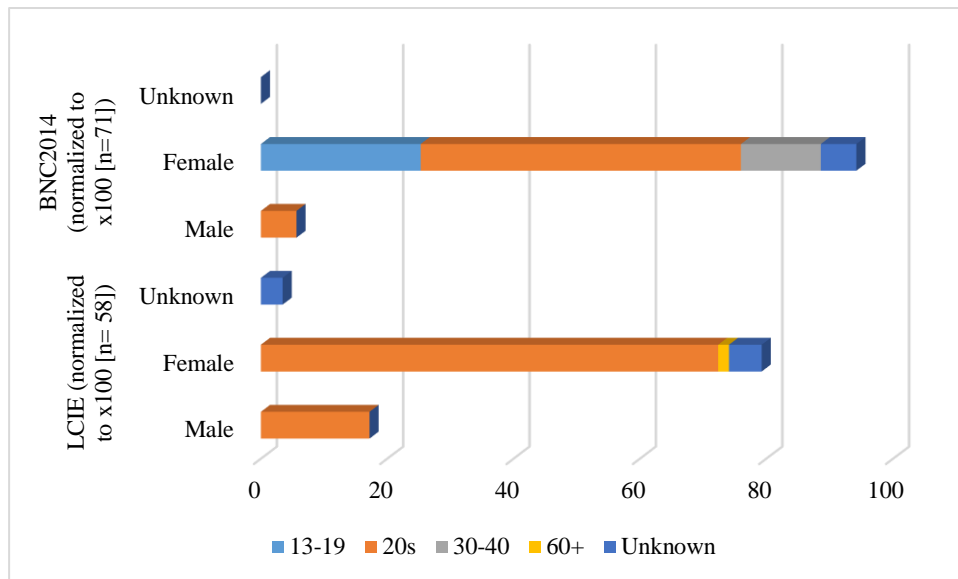


Figure 5.11. Age and gender distribution of BE LIKE in *LCIE/BNC2014* samples normalized to x100.

Age-wise, BE LIKE is clearly favored by young speakers, which aligns with the literature. However, and unlike in other varieties where it is perceived as being common among teenagers (see (5.8.2.)), the use of this quotative in the Irish and British English samples appears to be more prominent among speakers who are in their 20s. This is also in line with the use of discourse pragmatic marker LIKE which Schweinberger investigates in the context of IrE and Southern British English, whereby, by analyzing *ICE: Ireland* and the *Bergen Corpus of London Teenage Language*, he finds a strong connection between LIKE and females in their 20s. Chapter (7) provides more detailed investigations and analyses of discourse pragmatic marker LIKE in CoFIrE and the contrastive samples.

5.9.3. BE LIKE *Tense* in CoFIrE and contrastive samples

The examination of the effect of tense in CoFIrE (Table 5.9.) echoes the use of this quotative in other varieties, for it is overwhelmingly produced in *Historical Present Tense* for dramatic purposes, although it may also appear in *Past Simple*.

	CoFIrE	LCIE	BNC2014	Example
Historical Present Tense	99,2	22,5	16,9	See look at them <i>they're just like</i> "oh Christ I got to go scratch my hole when I go home" (LCIE)
Past Simple	0,8	65,5	80,3	<i>I was like</i> there's no way I 'm having some blonde probably bimbo in my bedroom pissing me off. (BNC2014)
Modal Form		12	1,4	Mam thought <i>he'd be like</i> "Oh we've got the world altogether lads"
Other forms			1,4	<i>I want to be like</i> oh I 'm so proud of you

Table 5.9. Contrastive tense distribution of BE LIKE normalized to x100 across CoFIrE and LCIE/NC2014 randomized samples.

In contrast, the analysis of the contrastive corpora samples from *LCIE* and the *BNC2014* indicates that the favored tense is *Past Simple*. The preference for *Past Simple* in *LCIE* could underscore a tense shift towards *Historical Present Tense* which may have taken place after the compilation of the corpus, and which is documented in CoFIrE. Furthermore, the tendency for *Past Simple*. With regard to the British sample, the analysis illustrates a wider variety of tenses, as well as the appearance of BE LIKE in the modal WOULD form and *want to be like* pattern, none of which is evidenced neither in the fiction nor in the IrE corpora.

5.9.4. Grammatical person +BE LIKE across CoFIrE and contrastive samples

The qualitative and quantitative analyses of the variable of *grammatical person* (Figs. 5.12. and 5.13. below) in CoFIrE indicate it is generally used to present a mimetic, more dramatized, recreation of dialogue. Indeed, BE LIKE seems to be more recurrently used to reproduce others, particularly the voice of third-person *female* subjects.

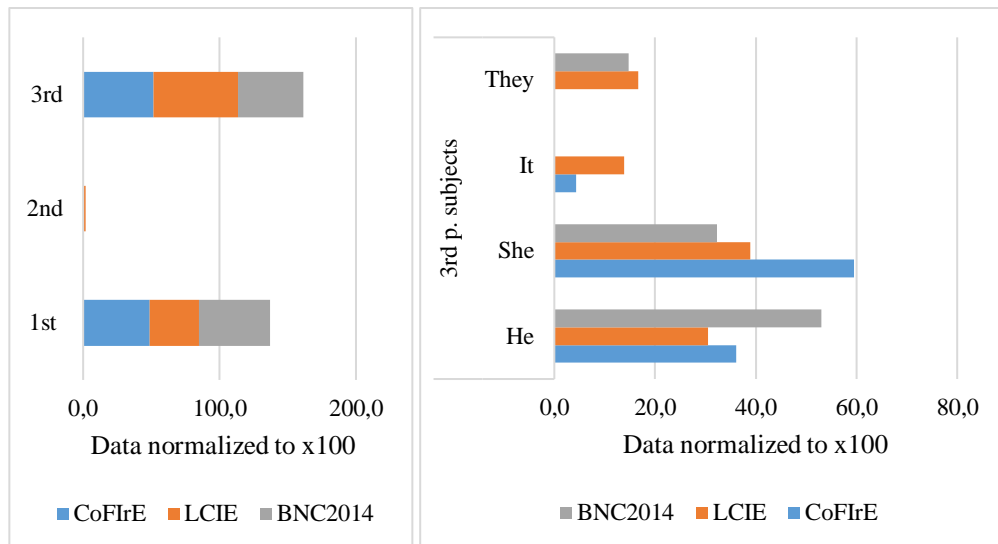


Figure 5.12. Grammatical person distribution⁵⁴. Fig. 5.13. 3rd-person subject pronouns⁵⁵.

By recurrently reproducing the voice of third-person *female* subjects, the use of BE LIKE and its discourse-pragmatic range in CoFIrE aligns with its functional evolution in American English (see (5.9.1.)), also mirroring Amador-Moreno's (2015) study, which documents the same expansion in Howard's early rendition of fictionalized Dublin English. The CoFIrE data, however, contrast with Höhn's (2012) analysis of BE LIKE in *ICE: Ireland*, where no expansion from first to third-person pronouns was found.

The functional expansion appears to also be present in the contrastive samples (Fig.5.13.). For example, *LCIE* documents the prominence of 'other' representation, particularly the representation of *female* voice, which corroborates the findings from CoFIrE. The occurrence of BE LIKE+neuter *it* and *they* subject pronouns found in *LCIE* must be noted, since the latter is absent from the fiction corpus. Contrastively, the results from the analysis of the *BNC2014* sample indicate less of a distinction in terms of *grammatical person*, with BE LIKE seemingly being more commonly used for self-representation. However, the difference between self and 'other' voice representation is minimal (52.11 to 47.9 normalized occurrences respectively), which prevents us from establishing whether this particular quotative is more recurrently used as a self-representative tool in this variety.

⁵⁴ Distribution of BE LIKE by grammatical person across CoFIrE and LCIE/BNC2014 randomized samples with data normalized by 100 words.

⁵⁵ Distribution of BE LIKE by 3rd-person subject pronouns across CoFIrE and LCIE/BNC2014 randomized samples with data normalized by 100 words.

5.9.5. Pragmatic uses of BE LIKE in CoFIrE and in contrastive samples

While *self-representation* is a particularly significant function, with some studies finding that BE LIKE favors *first-person mimetic inner-thought representation* (Tagliamonte & D’Arcy, 2007; Buchstaller & D’Arcy, 2009; D’Arcy, 2012; Durham et al., 2012; Tagliamonte et al., 2016; D’Arcy, 2017, p. 130), pragmatic expansion from *self-representation* to the *recreation of others* (D’Arcy, 2017, p. 21), including *emails and text messages* (Tagliamonte, 2016, pp. 71-76) has also been documented. As established in the previous section, BE LIKE is mostly used for the *representation of others* in CoFIrE. However, the qualitative analysis of its discourse-pragmatic uses in the fiction corpus (Fig. 5.14.) reveals a range of functions in ‘new’ contexts which merit attention.

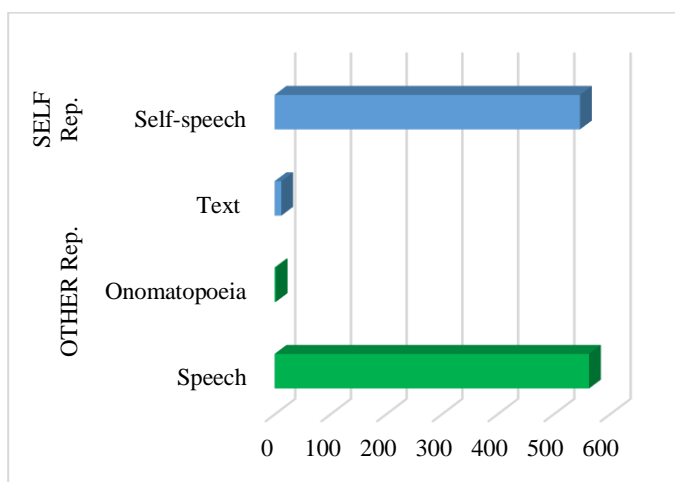


Figure 5.14. Raw functional distribution of BE LIKE across CoFIrE.

As observed in Fig. 5.14., the CoFIrE speakers use BE LIKE more prominently to *reproduce others’ speech* in a dramatic and theatrical form. However, its use can also frame *onomatopoeic reproductions of non-lexicalized sounds*, which, in the case of 5.28. below, not only lends stage-like value to the utterance, but also seems to substantiate the speaker’s conjecture of what ‘the boy’ wants to do to him.

(5.28.) One of the boys who *has*⁵⁶ a gun points at me and makes this noise with his mouth. *It’s like, ‘Tat- tat- tat . . . Tat- tat- tat . . . Tat- tat- tat . . .’* and I can tell he’s imagining what it would be like to shoot me. (PHKU)

⁵⁶ Emphasis in the original.

- (5.29.) My phone beeps in my hand. It's a text message from Sorcha's sister [...].
It's like, 'U wundt believ wot oisinn just told me hed like 2 do 2 me!' (PHKU)

The expansion of BE LIKE to 'new' pragmatic contexts aligning with its use in Canadian English (Tagliamonte, 2016, pp. 71-76) is worth noting, as there is a considerable number of occurrences where it *reproduces written content* in the form of 'recalled', word-for-word reproductions of paragraphs, instant messages, and tweets. Notice how the speaker in 5.29. frames the *reconstructed text message* as if it were a verbatim account of it. In addition, by conveying the content in a form that is natural to the written medium—abbreviations—the speaker further increases the level of faux realism to his report, which is also emphasized by the reproduction of the reportée's accent (*wot* for 'what'). This pragmatic function, along with the *portrayal of onomatopoeic non-lexicalized sounds*, is also present in *LCIE* (2 occurrences in total), which substantiates the fiction corpus' findings. This also leads us to consider that, at least in the context of Ireland, this 'new' function might have begun during the compilation of *LCIE*, and may have continued its expansion in later years, as documented in the novels. The *reconstruction of text* is absent from the British sample⁵⁷, however, assumptions as to whether this is a function that BE LIKE serves in this variety should be avoided due to the size of the sample. It is also possible that its absence be representative of cross-variational, pragmatic differences, so further examination needs to be conducted on this issue.

In addition, the qualitative exploration of the *LCIE* and *BNC2014* samples with regard to the remaining 'new' functions retrieved cases where BE LIKE frames others' *imaginary/hypothetical voice* (see 5.30.) and *thought* (in 5.31.), the latter of which was also documented in Amador-Moreno's (2015, p. 382) study of this quotative in Howard's (2005) rendition of Dublin English.

- (5.30.) <S1⁵⁸> I was like but if somebody reads that like your boss then they'd
be like <S2>: oh what's she got planned? (*BNC2014*)

- (5.31.) Mam thought he'd be like "Oh we've got the world altogether lads"
(*LCIE*)

⁵⁷ It is important to mention the fact that, in this case, we are examining only the pragmatic functions BE LIKE serves when reconstructing the *voice of others* and not when used for *self-representative* purposes.

⁵⁸ Speaker tag names from *BNC2014* have been assigned a random tag so as to further anonymize their identity.

Contrastively, while the great majority of occurrences of BE LIKE in CoFIrE—with the exception of 3—are produced in the two *Ross O'Carroll-Kelly* (RO'CK) novels, no occurrences where this quotative was used to *recreate thought* were found. Seeing as how the two RO'CK books in CoFIrE were published around 10 years after the novel analyzed by Amador-Moreno (2015) and the compilation of LCIE, this discrepancy in pragmatic use could highlight a longitudinal development in Dublin English (at least as perceived by Howard) which is documented within the RO'CK series itself. It may also underscore the *pragmaticalization* of this quotative which functions slightly differently than in British English.

Finally, I believe that the existence of the 'new' contexts in IrE, as well as the seeming absence of previous pragmatic functions from the fiction corpus, which are prevalent in the contrastive samples, lends validity (RQ 1a) to the fictional representation of BE LIKE use in the novels, while also diachronically documenting the potential *pragmaticalization* of this quotative verb in Dublin English.

5.10. Concluding remarks

The detailed study of the non-standard quotative repertoire in CoFIrE has illustrated the dominance of GO (RQ 1), which seems to function as a distinct index of identity in fictionalized Dublin English, marking the voice of *upper-class*, male speakers in their early 30s (RQ2). As such, its use establishes a clear geographical and socioeconomic linguistic distinction between its speakers, and those who use less prestigious, more locally-bound items (i.e. Northside Dubliners) and who may belong to a different social class (RQs 2 and 3).

The patterning of GO in fictionalized IrE mirrors the literature for the most part, although offering some divergent results when contrasted with *LCIE* and *BNC2014*. For example, the dominance of *Historical Present Tense* in CoFIrE and *LCIE* is less marked in the British sample, where *Historical Present Tense* and *Past Simple* are equal in occurrence count. The analysis of *grammatical subject patterns* between the two IrE corpora also appeared to foreground a potential morphosyntactic development documented in the fiction corpus, whereby GO seems to have gone from voicing male voice (in *LCIE*) to female speech (in CoFIrE).

This study has also highlighted the acquisition of ‘new’ pragmatic contexts, most of which are absent from *LCIE* but present in the British sample. Amongst these are the use of GO to *report imagined speech* and to *reconstruct written content*. I believe this functional extension into ‘new’ pragmatic contexts present in other varieties (i.e. British English in this case) may indicate that CoFIrE could attest to the potential *pragmaticalization* of this quotative in IrE; a process which may have been in progress during and after the compilation of *LCIE*. As such, this extension would also evidence the level of authenticity (RQ 1a) inherent in the portrayal of quotative GO in contemporary Irish fiction, and in Howard’s rendition in particular.

The analysis of the BE LIKE, the second most reproduced quotative in the books, has highlighted the adoption of this more globalized quotative form (RQs 1 and 3) which is strongly perceived as and associated with the American English linguistic repertoire. Such an adoption might highlight the process of *supraregionalization* taking place within IrE (Hickey 2005, p. 351) and Dublin English more specifically, according to Howard’s representation (RQ 3).

Howard’s fine-grained attestation of potential innovative quotative use (including the use of BE LIKE and the remaining non-standard quotatives which are only present in his novels) is praiseworthy as it coincides with their real-life use in other varieties while introducing contrastive and ‘new’ developments diverging and mirroring their occurrence in IrE and British English. It also underlines the outstanding level of orality present in his rendition of contemporary Dublin English. Furthermore, it demonstrates a dedicated authorial commitment to representing natural spokenness in Dublin (and IrE by extension), chronicling linguistic uses and developments that bring a uniquely high level of realism to his storytelling.

Finally, it is important to note that, as mentioned in previous sections, the use of these innovative, ‘ventriloquizing’ devices over other traditional ones also allows the authors (and Howard in particular) to include the reader into the narration, thus strengthening the fantasy bond of familiarity that exists between author-narrator-reader. This bond is further reinforced by means of the other two most recurrent linguistic items in CoFIrE, namely intensifying *fucking*, which will be explained at length in chapter (6), and discourse pragmatic marker *like*, which is explored in detail in chapter (7).

6. TABOO LANGUAGE: *FUCKING*

This chapter explores the second most prominent pragmatic item in CoFIrE: intensifying *fucking*. Firstly, the chapter provides an overview of studies that have investigated the multifunctionality of taboo language in natural interaction (6.1.) and in fictional dialogue (6.1.1.), with sections (6.2.) through (6.2.1.) focusing on overviewing studies which center around the use of intensifying *fucking* in the literature and in spoken and written IrE respectively. Section (6.3.) details the methodology of analysis, while (6.4.) overviews the use of intensifying *fucking* in CoFIrE. The sociolinguistics of intensifying *fucking* will be explored in sections (6.5.) through (6.5.4.), contrasting against randomized samples from *LCIE* and *BNC2014*, and colligational pattern preference will also be studied in detail in section (6.6.), with section (6.6.1.) focusing on exploring cases of *tnesis*. The pragmatic functionality of this intensifier will be outlined in (6.7.), and section (6.8.) offers a case study of the use of intensifying *fucking* in the narrative voice of *Ross O'Carroll-Kelly*. Finally, concluding remarks will be offered in section (6.9.).

6.1. The pragmatic functionality of taboo language

Many researchers have identified taboo language⁵⁹ as one of the most common features in spoken interaction (see e.g. Jay, 2000; Murphy, 2009; 2010 among many others) whose pragmatic value is multifaceted. As such, taboo language becomes one of the items to look out for when gauging and examining the realistic representation of orality in a corpus of, in this case, contemporary IrE fiction.

Commonly believed to be a borrowing from Polynesian *tabu*, the English word *taboo* refers to a series of behaviors that are socially and culturally sanctioned as they may cause harm, discomfort, or offense to others or to the individual (Allan & Burrige, 2006, p. 27). Stapleton (2010, p. 289) identifies *swearing*, another term for taboo language, as “a linguistic practice based on taboo, or that which is forbidden”. These words can appear in a wide array of forms (see Jay, 2009, p. 154), ranging from discriminatory language

⁵⁹ Despite the various terms that exist to refer to this particular type of feature, encompassing words like *profanity*, *swearwords*, *'dirty' words*, *cursing*, or *offensive language*, among various others, I have chosen to use *taboo language/words* as an interchangeable, umbrella term throughout this thesis.

and sex-related words (*fucker, shagging, bugger*), to religious expressions (*Jesus!, God dammit!*), including references to socially sanctioned topics like genitalia (*cunt, dickhead*), bodily parts and effluvia (*bloody, shit, crap*), animal names (*cow, bitch*), or pejorative words to describe people (*queer, retard*), among others. The practice of using taboo language, therefore, becomes a way through which these taboos can be risked, broken, or circumvented.

Pragmatically, these terms appear to be multifunctional, for as Allen & Burridge (2006, p. 2) point out, “[they can] be used as a shield against malign fate and the disapprobation of fellow human beings; [...] as a weapon against enemies and as a release valve when we are angry, frustrated or hurt”. This quote illustrates well one of the main pragmatic uses of taboo language which this chapter is concerned with: the conveyance of strong emotive meaning. Depending on the context (Stapleton, 2010, p. 294) and manner in which they are used, they can transmit a strong level of emotionality on the part of the speaker (Andersson & Trudgill, 1990, p. 195) which other items are unable to project. Thus, they function as an emotional outlet which allows speakers to express a wide range of strong feelings (Jay, 2000, pp. 51-52; Jay & Janschewitz, 2008, p. 269; Potts, 2007), both personal and interpersonally, ranging from positive to neutral or negative emotions (Jay, 2009, p. 155). Indeed, swearing is often linked with the expression of negative emotions, and with *anger* in particular (Stapleton, 2010, p. 294). Other emotions include, *frustration*, or *surprise* (Murphy, 2010, pp. 168-169), while speakers may also indicate contempt for the interlocutor, or might use them for their shock value (de Clerk, 1991, p. 157). Furthermore, Jay (2000, p. 52) argues that the expression of strong (speaker) emotions through taboo language also affects the interlocutor since they “understand that the feelings [conveyed by the speaker] are intense because the language accompanying them is”. This is better illustrated in my own examples (see (a) – (c)) below:

- a) I hate horror movies
- b) I *fucking* hate horror movies
- c) I hate *fucking* horror movies

While these sentences arguably convey the same message (i.e. the speaker does not like horror movies), there is a noticeable difference in terms of the level of emotional intensity

applied in each. For example, the speaker in (a) presents a neutral statement of their dislike, whereas the positioning of *fucking* as a premodifying verbal intensifier in (b) adds a layer of emotiveness which emphasizes their hatred, almost expressing strong aversion. Finally, the use of premodifying nominal *fucking* in (c) would place the emphasized aversion on the genre itself. While deceptively simplistic, I believe this example illustrates well the way in which taboo language, and in this case, *fucking* , can function as an intensifier to transmit strong emotions on the part of the speaker.

The multifunctional value of taboo language is further emphasized when we consider that it has been described as a ‘social creation’ (McEnery & Xiao, 2004) which enables us to communicate our emotional and psychological state of mind, for “[w]hen we hear a person cursing, we hear emotionality, hostility, aggression, anxiety, and religiosity” (Jay, 2000, p. 107). Taboo language can, therefore, also be an effective stylistic marker of identity (Stapleton, 2010, p. 298), indexing age, gender, class, and even geographical location, among other factors. For instance, society (and early investigations) tended to associate the use of cursing with men. Lakoff’s (1973) controversial study, *Language and Woman’s Place* , for example, listed the avoidance of strong taboo language as a main feature of what she labeled “women’s language”. This could arguably be attributed to the popular belief that women were/are expected to “be correct, discreet, quiet and polite in their behavior” (Chambers & Trudgill, 1998, p. 85), including in their linguistic practices, which are also expected to be proper and ‘ladylike’. According to Lakoff (1973, p. 50), this would explain why “stronger expletives are reserved for men, and the weaker ones for women”, especially euphemisms (Coates, 2004, p. 14). However, despite the belief that men curse more than women (Jay 2000, p. 166), scholars like de Clerk (1992, p. 288) call for a re-examination of this theory as it “do[es] not match commonly held perceptions”, while Coates (2004, p. 98) suggests that the “stereotypes of the tough talking male and the pure, never swearing female are false”.

Taboo language also appears to be linked to other social factors. Research (see Jay, 1992; 2000, p. 163; Stenström et al., 2002) finds it to be an age-graded feature, appearing more prominently in the speech of younger speakers, especially adolescents, and decreasing over time. de Klerk (1991, 2005) argues that taboo language can also function as marker of group identity and solidarity, often differentiating between male and females and indexing group membership and/or strengthening group affinity (Ljung, 1986, pp. 14-15), especially in adolescent subcultures (Stenström et al., 2002, p. 77;

Murphy, 2010, p. 169). As regards social class, Jay (2000, p. 158) maintains that while cursing is used across all social strata, “[a]nxiety about cursing is a middle class problem” (Jay, 2000, p. 159), which Fussell (1983, p. 151) attributed to their “fearing offending others”, at least in the context of the USA. Hughes (2006, p. 80) also notes the middle-class avoidance of swearing, which appears to be more prominent among the upper and lower class, at least in British society. Furthermore, Hughes (1992, p. 301) challenges the perceived connection between taboo language and male speech by noticing that lower-class (English) women frequently used strong expletives, which middle-class men avoided, to maintain social cohesion.

Despite the frequent appearance of taboo language in natural conversation (Jay, 2000; Murphy, 2009), scholars like Murphy (2010, p. 167) remark upon the scarcity of academic attention these words have received. Jay (2000, pp. 9-10) attributes this lack of scholarly interest to the forbidden nature of the topic itself and observes the marginalization of (offensive) emotional speech within linguistics and psycholinguistics. He proposes a Neuro-Psycho-Social (NPS) theory of cursing which understands taboo language as the result of the integration of “neurological control, psychological restraints, and socio-cultural restrictions” (*ibid.*, p. 19). Its aim is explaining why cursing occurs and why speakers select the taboo words they use. NPS advocates in favor of the integration of cursing into the definition of language “because language has to represent speakers’ knowledge of pragmatics, politeness, figurative language, vulgarity, insults, sex talk, humor, verbal abuse, and anger.” (*ibid.*, p. 11). Finally, Jay (*ibid.*) also argues that treating taboo language as a by-product of language only leads to marginalizing it further. Thus, and following Murphy (2010), this thesis embraces Jay’s NPS theory, and applies it to the detailed examination of the second most frequent pragmatic item in CoFIrE, intensifying *fucking*, which is hereafter understood as a multifunctional word.

6.1.1. The multifunctionality of taboo language in fictional dialogue

As discussed throughout the thesis, the essence of fictional (literary) dialect-dialogue lies in the portrayal of natural orality for realistic or ‘authenticity’ value. Given the general high frequency status of taboo language in natural (English) interaction corroborated in the literature, its reproduction in fictional (literary) dialect-dialogue serves as a way to

lend realism to the interaction within the story world, carefully stylizing it as a casually informal and spontaneous (at least it was created to appear unplanned) exchange between characters (Price, 2015, p. 58; Bednarek, 2019, p. 44). In their investigation of FUCK in the *British National Corpus*, for example, McEnery & Xiao (2004, p. 247) make the point that its high frequency in imaginative (mostly fiction) texts might be due to the genre's prominence of speech–representation in the form of dialogues. The use of taboo language can also function as a tool to create humor, as exemplified in Walshe's (2011, p. 139) study of the fictional dialogue in Irish TV series, *Father Ted*, where taboo words create humor through incongruence/unexpectedness. In this case, the humor behind the swearing lay in the fact that the audience did not expect the characters to swear due to their identities (i.e. priests and elderly people).

The construction and display of (character) identity through taboo language use in fictional (TV) dialogue is also explored by Bednarek (2019, p. 45), who argues that it can construct speaker identity and even inter-character relationships (Kozloff, 2000, pp. 43-46), which is supported by various other studies (Bednarek, 2015; Price, 2015; Queen, 2015). Bednarek (2019, p. 46) exemplifies this by pointing out how *bloody* is only used in her *Sydney Corpus of Television Dialogue*, comprising 66 contemporary US TV series, by non-American characters. The audience, therefore, automatically connects its use with speakers of other varieties. In addition, she argues that the display of character emotion taboo words carry can function as a way of affecting/manipulating the audience's emotional and evaluative responses (Bednarek, 2019, p. 47; Kozloff, 2000, pp. 49-51). I believe that Bednarek's argument can be extrapolated to fictional (literary) dialect-dialogue, where taboo language can also function as a tool through which the speakers in the fictional story world, and the author in reality, can manipulate or guide the readers' emotional reactions to characters, their inter-relationships, and/or events in the books.

6.2. Intensifying *Fucking*

In order to better understand the nature of the second most frequently produced pragmatic item in CoFIrE, intensifying *fucking*, it is necessary to clarify what *intensifiers* are (this definition will be further expanded in section (8.3.)). Bolinger (1972, p. 17) describes intensifiers (e.g. *so, really, very, etc.*) as “adverbs that maximise or boost meaning”,

which often “denote a place on a scale of intensity, either upward or downward” (Greenbaum, 1996, p. 151). Although Quirk et al. (1985, p. 429) categorize intensifying adjectives into *emphasizers*, *downtoners*, and *amplifiers* (the latter of which corresponds to Bolinger’s label), the term ‘intensifier’ is preferred in this thesis.

Intensifiers, therefore, strengthen the emotional value of a content word or a full sentence. Their use in general is often associated with colloquial speech, and given the fact that speakers use it to convey and boost emotive content, they are considered part of emotional language. As such, they have traditionally been connected with female speech (see, for example, Lakoff, 1973 for British English; Ito & Tagliamonte, 2003 for American English; Tagliamonte & Roberts, 2005; and Tagliamonte, 2008 for Canadian English; Saarenpää, 2016, and D’Arcy, 2015 for Canadian/New Zealand English). While the variety of intensifiers that exist in the English language is large, Carter & McCarthy (2006, p. 227) note that taboo words frequently function as intensifying adverbs/adjectives, and remark upon the prominence of intensifying *fucking* in cases such as, “you are a *fucking* idiot”, among others. For more on the use of intensifiers as sociodemographic indexes and their contexts of use and functionality see sections (8.3.) through (8.3.2.).

According to the *Oxford English Dictionary*, the origin of intensifying *fucking* in cases like the example above appears to be rather recent, as it was attested in *Suppressed Book about Slavery!* (Carleton, 1864, p. 211), where a Reverend described to his congregation the punishment that a particularly cruel slaveholder gave one of his slaves: “Hush, you ***** b–h, will you take the name of the Lord in vain on the Sabbath day?” (asterisks in original). However, its insertion into non-standard contexts like single words or compounds in cases like “abso-*fucking*-lutely” is much more modern, and was first recorded in *Notes & Queries: A Medium of Intercommunication for Literary Men, General, Readers, etc.* (1921, p. 415), and originally noted by Scheidlower (2009, p. 172), in a note by a former English soldier who documented the English army slang used by World War I soldiers:

“[It is] absolutely impregnated with one word which [...] was used adjectivally to qualify almost every noun[...] Words were split up to admit it: "absolutely" became "**abso----lutely**," and *Armentières* became "Armen---teers." [my colored emphasis].

There does not seem to be agreement within academia, or even in dictionaries, regarding whether to refer to this phenomenon as *tmesis* or *expletive infixation* (McMillan, 1980;

McCarthy, 1982; Bauer, 1983, pp. 89-91; Zwicky & Pullum, 1987; Plag et al., 2009). The difference between them is hard to describe, and they are sometimes used interchangeably to refer to the same process, which the *Oxford English Dictionary* describes as involving "the separation of the elements of a compound word by the interposition of another word or words" (Simpson & Weiner, 2021, p. 162). In this thesis, that would be intensifying *fucking* in examples like the fragment above and (d) and (e) below.

d) Every.*Fucking*.Thing (PHKU)

e) Blahdy *focking* blah (PHKU)

Nevertheless, intensifying *fucking* can also be inserted in a compound (i.e. (d)), often being placed before the primary stressed syllable of the second component in the collocation (Mackenzie, 2019, p. 73), or it can also separate collocations (e) which would normally be uninterruptible (McMillan, 1980, p. 167). Thus, Hegedűs (2013, pp. 164-165) suggests that the term 'expletive infixation', which involves the insertion of "offensive intensifiers" which are "free morphemes or bound nonce-morphemes [whose] expressive colloquial effect [...] cannot be regarded as a part of plain morphology" (Zwicky & Pullum, 1987, p. 7), not be used as it misconceptualizes the concept of 'infixation' (i.e. bound morphemes are affixed for morphological purposes and do not lead to register downstep).

Seeing as how intensifying *fucking* is a full word and not an infix, I follow Mackenzie's (2019) study, classifying all occurrences of this particular collocation as *tmesis* (see section (6.6.1.) for the analysis of *tmesis* in CoFlrE and the contrastive *LCIE* and *BNC2014* samples). However, I have separated the occurrences where it follows the traditional definition more closely (i.e. splitting one word into two units) and classified them as *standard tmesis*. The other occurrences appearing in contexts where the intensifier divides compounds and/or collocates which would normally be inseparable were labeled *non-standard tmesis*.

6.2.1. Intensifying *fucking* in the literature

While largely under-examined, the study of intensifying *fucking* frequently appears to be covered in passing, and it is often contingent upon its status as a variant of lemma FUCK. For example, in British English, Stenström et al. (2002) investigate teen talk in the *Bergen Corpus of London Teenage Language* (COLT), where FUCK (including all its inflected forms) is found to be the most frequently used taboo word among teenagers. Their comparative study of adult speech in the *British National Corpus* (BNC), however, identified this lemma as the third most recurrent taboo word, with *bloody* being the preferred one (*ibid.* , pp. 80-84). Their analysis also indicates that *fucking* is the most used form in COLT, with male speakers using it more often than females. McEnery & Xiao (2004, pp. 258-259) looked at the pattern distribution of FUCK in the BNC across the spoken and written registers. Their study found *fucking* to be more significantly used for emphatic purpose, which can also appear as an ‘infix’ separating one word as well as two parts of one name (e.g. Jesus *fucking* Christ). In their examination of the spoken register, *fucking* is found to be the most popular variant across all age groups, ranking especially high in the 15-24 and 25-34 cohorts, and across all educational levels (*ibid.* , pp. 241-246). More interestingly, the study of FUCK in the written register (*ibid.* , pp. 247-252) indicated a higher salience of this particular lemma in imaginative texts (of which most are fiction), with *fucking* seemingly being preferred by male writers (as opposed to female authors who favored *fuck*) writing for an (adult) mixed audience.

To date, however, I find that Mackenzie (2019) offers the most in-depth analysis of the syntactic patterns of intensifying *fucking* . In his study, he uses items found by others such as McEnery & Xiao (2004) or Sheidlower (2009), along with web sources such as the blog *Strong Language* , or online dictionaries like the *Urban Dictionary* or *Green’s Dictionary of Slang* . He also adopts a Functional Discourse Grammar approach to the identification of the syntactic distribution of expletives *fuck* , *fucking* , *fucking well* , and *the fuck* (including *tmesis*), which he substantiates by studying their discourse/grammatical acceptability according to his own “self-consultation as a native speaker” and to the sources mentioned above (Mackenzie, 2019, p. 62). Mackenzie proposes that these expletives are not lexical, but grammatical, and that the “bleaching out of their literal meaning, [the syntactic...] loss of positional flexibility” and

morphological form fixation evidences their *grammaticalization*⁶⁰ (*ibid.*, p. 76). In other words, they have become semantically empty but syntactically functional. With regards to intensifying *fucking*, he points out its status as a syntactic optional (although not meaningless) marker, as it triggers the speaker's emphasized emotive meaning (*ibid.*, pp. 77-78). Thus, this thesis follows Mackenzie's classification (see Appendix 4 for a summarized version of his syntactic distribution with examples from CoFIrE) and applies it to the use of intensifying *fucking* in CoFIrE, and in the contrastive analyses against the *LCIE* and *BNC2014* samples.

6.2.2. Intensifying *fucking* in spoken and fictionalized IrE

Despite the well-known Irish propensity for swearing (Bartley, 1954), the number of studies covering taboo language is significantly small (Asián & McCullough, 1998; or Farr & Murphy, 2009 for a discussion of profanity and religious expressions in IrE; Murphy, 2009; 2010 for a sociolinguistic investigation of taboo language in an age and gender differentiated corpus of spoken IrE conversations; and Walshe, 2009, pp. 150-152 for swear words in a corpus of Irish films). The number of studies looking at intensifying *fucking* in spoken IrE is even smaller. Clancy's (2016) investigation of intimate talk in the *Limerick Corpus of Intimate Talk* (i.e. a sub-corpus of *LCIE*), highlighted taboo language as a highly salient feature of this type of discourse, where *fucking* functions as a bonding device. Murphy (2009) studied the form and pragmatic functions of FUCK (and its variants) in her *Corpus of Age and Gender differentiated Irish English* (CAG-IE), comprising 90,000 words of casual conversations. CAG-IE contains two subcorpora, namely the *Male Adult Corpus* and *Female Adult Corpus*, which are further divided into three age cohorts (i.e. 20s, 40s, and 70s/80s). She found FUCK to be a highly frequent item, especially in the voice of males in their 20s (*ibid.*, pp. 90-94). Her analysis also revealed that *fucking* was the most commonly used variant, functioning predominantly as an amplifier (*ibid.*, p. 92). She noticed it was more salient among males in their 20s, who used it as their most common intensifier (*ibid.*, pp. 92-100). These findings were mirrored

⁶⁰ For a brief definition of the process of *grammaticalization* see section (5.5.). Section (7.1.) offers a more detailed description of the process of *grammaticalization* and its different features, especially pertaining the *grammaticalization* of discourse pragmatic markers.

in her examination of taboo language in the *Female Adult Corpus* (Murphy, 2010, p. 180), where *fucking* was also the most popular amplifier among female speakers in the 20s age cohort. Finally, Murphy (2009, pp. 96-102) noticed that, while conveying emotive meaning, intensifying *fucking* could also communicate attitude, especially in the form of positive or negative evaluations, finding that both genders seemed to transmit almost exclusively negative evaluations, especially about others.

The number of studies that examine the use and pragmatic functions of intensifying *fucking* at length in fictional dialogue in general, and in fictional IrE in particular, is considerably low. Aside from Terrazas-Calero's (2020) publication of preliminary findings on taboo language in CoFIrE, no other publications that discuss the use of intensifying *fucking* in fictional IrE dialogue in depth have been found at the time of writing. Thus, this chapter will utilize quantitative and qualitative analyses to provide a detailed picture of the use of intensifying *fucking*, its colligational preferences, pragmatic functions, and sociolinguistic indexical value as represented in CoFIrE, contrasting its findings with samples from *LCIE* and *BCN2014*.

6.3. Method of analysis

Before analyzing the use of intensifying *fucking* in CoFIrE, we must clarify which forms were taken as occurrences of this intensifier in this thesis. For example, *fecking*, which is the intensifying variant of *feck*, a well-known IrE euphemistic taboo word, was excluded. While Dolan (2006, p. 91) catalogues lemma FECK as a euphemism of FUCK, *feck* is perceived as being less taboo than its counterpart (Walshe, 2009, p. 149). Following Murphy (2009, p. 92), I do not regard *fecking* as a phonological variation of *fucking*, but as an item of its own. This, coupled with its less vulgar nature, is the reason why *fecking* was excluded. The forms which were included as cases of intensifying *fucking* comprise: *fucking*, *focking*, *fooking*, *fuckin'*, *fooken*, and *fucken*, all of which are orthographic representations of the various pronunciations of intensifying *fucking* in CoFIrE (for an analysis of how these may be indexical of regional identity in the corpus, see section (6.5.4.)). Concordance lines were then used to identify all occurrences of intensifying *fucking* in CoFIrE (and in the contrastive samples). Cases in which *fucking* maintained its literal sense (i.e. having sexual intercourse) were automatically discarded.

1	BOOKCODE	LINE	COLLIGATES	FUNCTIONS(SUB)FUNCTIONS	AGE	GENDER	CLASS	LOCATION	VARIANT
2	CKWBF.txt	ack into the rented trousers: 'A <CKWBFWBFD> fuc N		neg SARCASM	20-30	m	mid	x	fucken
3	CKWBF.txt	You mad bollocks! he roars. 'Could y<PREMOD N		neg ANGER	20-30	m	mid	x	fucken
4	DRII.txt	ped for a punter just ahead of us. '<DRIIP2> fuckin N		neg INSULT	30s	m	mid	dublin	fucking
5	DRII.txt	dem ...' he slurs, as we pass by. '<DRIIP5> Fuckin' N		neg INSULT	20-40	m	low-mid	dublin	fuckin'
6	DRII.txt	></DRIIP2> , but no, he gets up an' <DRIIP2> fuc!PREMOD V		neg DISLIKE	20-40	m	low-mid	dublin	fuckin'
7	DRII.txt	e response it would provoke. 'Ged a <N		neg ANGER	20-40	m	low-mid	dublin	fuckin'
8	DRII.txt	<DRIIP5> yew <GR><YE></DRIIP5 tmesis (collocation)		neg ANGER	20-40	m	low-mid	dublin	fuckin'
9	DRII.txt	IP5> me <GR><ME></DRIIP5> door by a <DRIIP5 N		neg INSULT	20-40	m	low-mid	dublin	fuckin'
10	DRII.txt	RIIP5> ye <GR><YE></DRIIP5> St<N		neg ANGER	20-40	m	low-mid	dublin	fuckin'
11	DRII.txt	ye <GR><YE></DRIIP5> stewpid priPREMOD V		neg ANGER	20-40	m	low-mid	dublin	fuckin'
12	DRII.txt	riggers <GR><THEGEIMP></DRIIP5> Big <DRIIP5 N		neutral NEUTRAL	20-40	m	low-mid	dublin	fuckin'
13	DRTRM.txt	en..." "I don't believe dis! I don't <DR PREMOD V		neg ANGER	x	f	low-mid	dublin (n)	fuckin'
14	DRTRM.txt	fanny! It's jus' nod on, it really <DRTIPREMOD V (aux)		neg ANGER	30s	f	mid	dublin (n)	fuckin'
15	DRTRM.txt	use the common parlance, she was 'a <DRTRMN> fi N		neg ANGER	x	m	low-mid	dublin (n)	fuckin'
16	DRTRM.txt	X></DRTRMST> <GR><YE></DRTFN		positive SEXUAL ADMIRATION	32	m	upper-mid	dublin (s)	fuckin'
17	DRTRM.txt	DRTRMBI> busy tonight boy, ha? 'Tis <DRTRMBI> fadj (pred)		neg ANGER	x	m	low-mid	dublin (n)	fuckin'
18	DRTRM.txt	ove it when I'm good and ready to <[PREMOD V		neutral NEUTRAL	20-30	m	low-mid	dublin (s)	fuckin'
19	DRTRM.txt			neg ANGER	middle-aged	m	mid	x	fuckin'

Figure 6.1. Sample image of the Excel file illustrating the classification of intensifying *fucking* in CoFIrE.

All occurrences of intensifying *fucking* were extracted onto an Excel file (illustrated in Fig. 6.1. below), where they were subsequently analyzed and classified for speaker information (i.e. age, gender, class, and geographic location) so as to investigate its indexical value regarding speaker identity as presented in the books (i.e. RQ 2 in section (1.2.)). Next, all occurrences were classified for colligational patterns following Mackenzie's (2019) classification (see Appendix 4) to account for differences or similarities. My supervisors were consulted during the classification stage for clarification in cases which lent themselves to confusion, such as examples 6.1. and 6.2. below⁶¹. For example, while the colligate in 6.1. was easy to discern, i.e. *bury* (verb), the colligate in 6.2. proved more challenging. In this particular case, I wondered whether to class *what* as an interrogative pronoun heading a phrase. Upon consultation, I was advised of its grammatical complexity, with interrogative pronoun *what* being the head of an idiomatic phrase (*what's his face*) which substitutes a proper noun, thus functioning as a nominal phrase.

(6.1.) 'You can fuckin' bury me before I'll use a stick' (NCCSM)

(6.2.) "There was something I didn't tell you – the day you called here with your old pair and focking what's his face" (PHKU).

⁶¹ My emphasis.

Manual qualitative classifications were also subsequently carried out across all occurrences of intensifying *fucking* to mark them for cases of *standard* and *non-standard tmesis* (see (6.6.1.) for a detailed analysis of *tmesis* cases in CoFIrE). To investigate the pragmatic functionality of intensifying *fucking* (which shall be discussed in (6.7.)), I follow Murphy (2009) by classifying its evaluative meaning in terms of *positive/negative* evaluations. However, this thesis includes the possibility of a *neutral* emotional meaning which represents occurrences where intensifying *fucking* is limited to conveying simple, unemotionally-charged emphasis, like “You know, triplets would make a real *fucking*⁶² mess down there” (PHKU). Furthermore, and as illustrated in Fig. 6.1. above, this chapter also categorizes each occurrence into emotional (sub)functions (e.g. transmission of *anger*, *sexual admiration*, *disdain*, etc.), thus creating a catalogue of emotions which aims at responding to Bednarek’s (2019, p. 17) call for research to use corpora to “identify if specific patterns of usage correspond to specific functions”. While classifying for emotion conveyance is a subjective process, as the emotions may vary depending on the perception of the readers, I believe my familiarity with the context of all the books (as a reader) as well as with all inter-character relationships lends this classification a measure of objectivity.

It also needs to be noted that all occurrences of intensifying *fucking* produced by the character of *Ross O’Carroll-Kelly* in his role as a homodiegetic/autodiegetic (i.e. narrator who participates as a character in his own story) in the two novels by Paul Howard included in CoFIrE were investigated separately from those he produced as a character. Fragment 6.3. below exemplifies both uses well, with coloring signaling if the intensifying *fucking* is used in his narrative role (purple) or when he is a character within his story (green). The occurrences he produced as a character were counted along with those produced by all other characters in the corpus. The reason for this separation is the fact that in previous investigations I found that his narrative production functioned differently with regard to pragmatic emotion conveyance to his production as a character (Terrazas-Calero, 2016a, 2016b). Despite this separation, his narrative production was classified following the same methodology explained above.

⁶² My colored emphasis.

(6.3.) The one thing that keeps going through *my* mind, though, is how Honor will react, if and when it actually happens. Another child in the house? She'll shit her **fucking** molars. I suddenly look at the clock and realize that it's already after two. 'Look at the **fucking** time,' I go. 'I better go and collect Honor from school.'

(Howard, 2013, p. 53)

Finally, the findings are contrasted against two samples of 100 random occurrences of *fucking* from *LCIE* and *BNC2014* which were, subsequently, analyzed for cases of intensifying *fucking* to check the representational validity of the fiction data when compared with real-life use, and for cross-variational (dis)similarities. The same methodology regarding its classification as explained above was applied to the two samples.

6.4. Intensifying *Fucking* in CoFIrE

The quantitative analysis of CoFIrE illustrated in Table 6.1. summarizes the overall count of 1,457 occurrences of intensifying *fucking* (hereafter InF) in CoFIrE. These comprise 506 occurrences which were produced by *Ross O'Carroll-Kelly* (RO'CK) in his role as a homodiegetic/autodiegetic narrator, which, as mentioned in the previous section, were analyzed separately, and 951 occurrences which were produced by all other characters in the corpus. The latter are the occurrences which are examined and contrasted in the following sections.

Corpora	Occurrences	%	Total number of occurrences
All characters ⁶³	951	65.2%	1,457
CoFIrE (Narrative) RO'CK	506	34.7%	

⁶³ These comprise occurrences produced by all character (including Ross O'Carroll-Kelly in his role as a character within the story), but exclude all cases produced by Ross O'Carroll-Kelly in his role as narrator.

LCIE	100	100%	100
BNC2014	97 ⁶⁴	97%	97

Table 6.1. Distribution and percentage of InF occurrences in CoFIrE and *LCIE/BNC2014* randomized samples.

As shown above, the contrastive analysis of the *LCIE* and *BNC2014* samples showed that 100 and 97 of the 100 randomized concordance lines of *fucking* were occurrences of InF in each corpus respectively. This is indicative of the substantial frequency that this particular intensifier reportedly enjoys in natural conversation already outlined by other scholars with respect to taboo language use (see (6.1.)). This prominence seems to extrapolate to its rendition in fictional dialogue, as attested by the large amount of occurrences in CoFIrE.

6.5. The sociolinguistics of Intensifying *Fucking* in CoFIrE

As was already mentioned in section (6.2.2.), at the time of writing there are few studies that evaluate the connection that may exist between InF and external linguistic factors (especially in the context of Ireland), as most researchers focus on the lemma FUCK. Given also how Murphy's (2009, 2010) studies establish a link between the use of InF and young masculinity in her IrE corpora, and to address RQ 2 (see (1.2.)) regarding the what the use of this intensifier may index regarding speaker identity, I investigated the indexical value InF may hold in Ireland, at least according to what its representation in the CoFIrE texts appears to suggest. To do so, each occurrence was manually classified for speaker age, gender, location, and social class, the latter of which was conducted as explained in (3.8.2.). The fiction findings were also contrasted against the *LCIE/BNC2014* samples so as to address RQ 1a concerning how 'authentic' the fictional representation may be when compared with its use in real-life, spoken Irish and British English.

⁶⁴ 3 of the 100 occurrences of *fucking* in the *BNC2014* sample were found to serve the literal meaning of having intercourse and were, therefore, discarded.

6.5.1. Gender and Intensifying *Fucking* across CoFIrE in contrastive samples

Gender was the first factor to be investigated due to the strong association of swearing and InF in particular with masculinity in Ireland (Murphy, 2009, 2010). The quantitative analysis of the InF occurrences in CoFIrE (Table 6.2.) reveals that the majority of fictional InF-users in the corpus are male.

Gender	Occurrences	Male distribution	
Male	688	Male characters ⁶⁵	366
Female	243	RO'CK (character)	322
Unknown	19		
	<i>Total</i>		950

Table 6.2. Gender distribution of InF in CoFIrE.

Given *Ross O'Carroll-Kelly's* outlier nature in the books and in CoFIrE (already mentioned in section (4.2.)), where he produces an elevated number of all linguistic features (i.e. he contributes 322 occurrences of InF in his role as character), one may think that that accounts for the link between InF and masculinity that appears to be evidenced in CoFIrE. However, a more detailed analysis of the occurrences of InF produced by all other male characters in the corpus (i.e. 366 in total) indicates that they continue to be the gender that is more prone to using InF in the fiction dataset, although female characters also produce a rather elevated number of occurrences. Thus, the findings from CoFIrE would seemingly corroborate the link existing between InF and masculinity in the context of Ireland pointed out by Murphy (2009, 2010). The contrastive analysis against the *BNC2014* and *LCIE* randomized samples of *fucking* (illustrated in Fig. 6.2. below) seems to lend further support to this theory, particularly in the case of *LCIE*, which highlights a higher recurrence among male speakers, which further corroborates the findings from the fiction corpus.

⁶⁵ This encompasses all male characters, excluding RO'CK in his roles as both homodiegetic/autodiegetic narrator and as character within his own story.

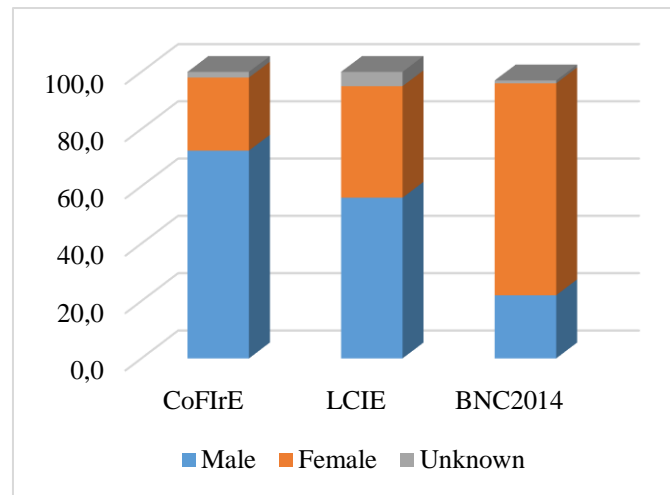


Figure 6.2. Distribution of InF by gender in *LCIE/BNC2014* 100-random sample compared with CoFIrE (normalized to 100 words).

The elevated number of InFs produced by females in both CoFIrE and in the *LCIE* sample attests to the fact that, despite InF still being slightly more recurrent among males, its use continues to be present in the voice of both genders, thus challenging the popular belief that females do not swear as much as male speakers, at least in the context of Ireland. Contrastively, the quantitative analysis of the *BNC2014* sample of *fucking* is interesting, as females are responsible for 71.8% of the overall production in the sample. However, it is worth pointing out that 30 of their 74 occurrences were produced by the same individual in a given conversation (e.g. “I was *fucking* raging it was four in the morning and it had like just *fucking* bust and I was like seconds” (*BNC2014*)). Despite this, the female production is still larger than the male production in the sample. While it cannot be conclusively stated that InF signals femininity in British English, or that it is more frequent among females, it could be hypothesized that this may be indicative of female speakers (at least in the UK) becoming more comfortable with the use of this intensifier. I believe this is also evident in Ireland, as reflected by the elevated female production of InFs recorded by CoFIrE/*LCIE*. Although generalizations should be avoided due to the size of the contrastive samples, and to the fact that CoFIrE offers a fictional rendition of real-life use as perceived by authors, I believe these findings support Coates’ (2004, p. 98) suggestion that “stereotypes of the tough talking male and the pure, never swearing female are false” and warrant, as posited by de Clerk (1992, p. 288), further research in the future.

6.5.2. Age and Intensifying *Fucking* across CoFIrE and contrastive samples

A quantitative analysis of InF production by age cohort reveals that its use appears across all age groups in CoFIrE and in the contrastive samples, as illustrated in Fig. 6.3. below. Nevertheless, it is more prominent among younger speakers across the corpora. This corroborates findings from research into in other varieties (see (6.2.) through (6.2.2.)), where the link between taboo language and youth is also established.

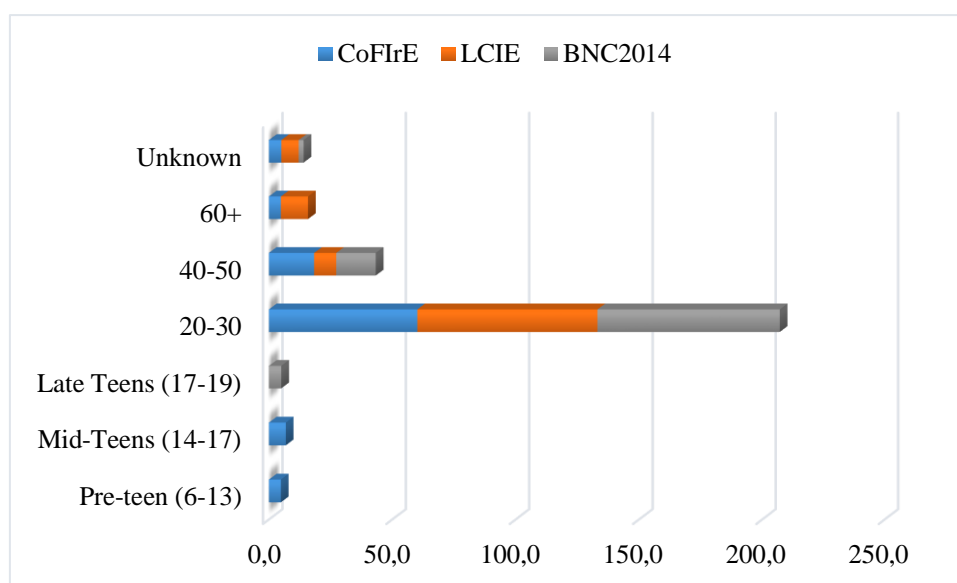


Figure 6.3. Distribution of InF across age cohorts in *LCIE/BNC2014* (100 random samples) with CoFIrE's data normalized to 100 words.

The prominence of InF in the 20-30 age cohort in the case of CoFIrE and *LCIE* also mirrors Murphy's (2009, 2010) findings in the context of Ireland. While Fig. 6.3. evidences a noticeable decrease in the 60+ cohort across the corpora, it cannot be concluded if its use decreases over time. The reduction in use may be due to the generally younger fictional 'population' featured in CoFIrE, while the data from the contrastive corpora could be due to the small samples. Nevertheless, it would be interesting to conduct a future study where the age-grading of this particular intensifier may be tracked cross-generationally.

6.5.3. Social class and Intensifying *Fucking* across CoFIrE and contrastive samples

Class indexation was also quantitatively explored in CoFIrE (as illustrated in Fig. 6.4.), initially revealing a prominence of use among upper-class characters. Upon closer inspection, however, 68.6% (i.e. 322) of the 470 occurrences produced by the *upper-class* in the corpus were found to have been produced by *Ross O'Carroll-Kelly*. Nevertheless, if his input is excluded, then this rank is left with 148 occurrences. This leaves the middle class as the dominating InF-producing class in CoFIrE with an overall occurrence count of 381 InFs.

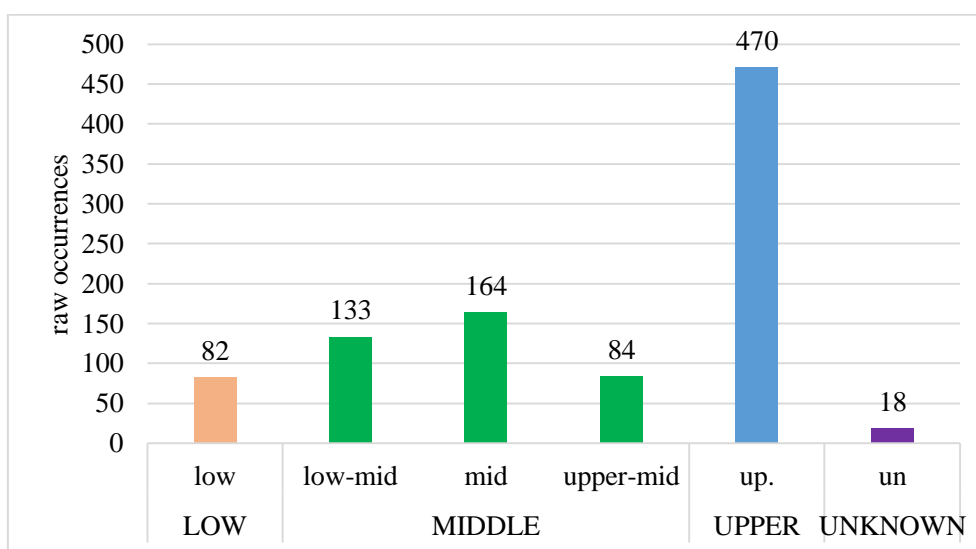


Figure 6.4. Distribution of InF occurrences by social class in CoFIrE.

The high recurrence of InF among the *middle-class* in CoFIrE, however, could also be due to the fact that, as mentioned in chapter (4.4.), this rank (with all its sub-divisions), which produces the largest amount of linguistic items in CoFIrE, is seemingly over-represented in the corpus. Despite this, it is still interesting to note the prominence of InF among this particular social rank in the fiction dataset for it seemingly contradicts, at least in the context of Ireland, Jay's (2000, p. 158) argument that the *middle-class* in other English-speaking countries is reluctant to swearing.

The correlation between InF-use and the *middle-class* in the fiction corpus seems to be substantiated by the use of this intensifier in the *LCIE/BNC2014* samples. As shown in Fig. 6.5. below, where the CoFIrE data has been normalized by 100 for comparative

purposes, the use of InF generally appears across all social classes, but it is overwhelmingly produced by category E speakers (71% in *LCIE* and 52.5% in *BNC2014*), the majority of whom are students. However, and while considerably fewer in number, a significant amount of the *LCIE* occurrences is produced by category A and B individuals (9% and 7% respectively). The British sample shows a higher production by category B speakers (23.7%), followed closely by category C1 (14.4%), both of which are representative of the middle class. In the case of the British sample, this would challenge Hughes' (2006, p. 80) observation regarding the avoidance of British society to swearing, at least with regards to this intensifier.

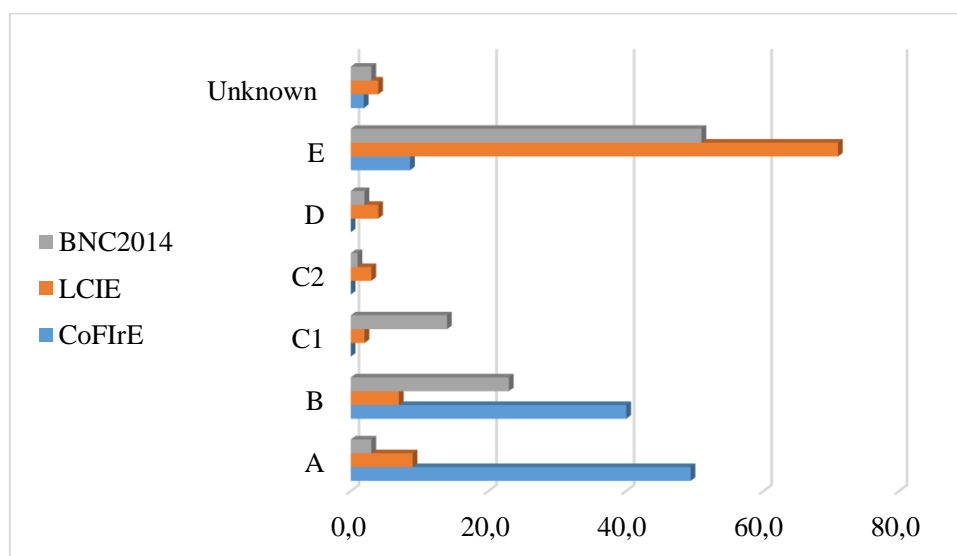


Figure 6.5. Distribution of InF by social class across CoFIrE (data normalized to 100 words) and *LCIE/BNC2014* samples following NRSSG occupation-based classification outlined in section (3.8.2.).

While it cannot be confirmed whether these findings are illustrative of actual use in Irish and British English due to the size of the samples, there still appears to be a clear connection between this particular intensifier and the middle class in both varieties. This would suggest that the portrayal of InF-use offered in CoFIrE is highly realistic, thus answering RQs 1a and 2 (see section (1.2.)), at least regarding the level of realism inherent to the fictional portrayal, as well as the effect of social class over InF use.

6.5.4. Geographic location indexation through Intensifying *Fucking* in CoFIrE

Given the difficult nature of conveying aspects of real-life communication (i.e. speaker emphasis, identity, accent, etc.) via the written medium, I believe orthographic variation is a good tool to construct and display local/non-local speaker identity, which also manipulates the reader's perception of the characters. Orthographic rendition represents a conscious effort on the part of the authors who manipulate the reader's perception of the characters by writing down phonetically *how* they pronounce InF. This allows them to lead readers—at least those familiar with the accents and implications of their use—to gain immediate knowledge as to the characters' regional provenance, while other aspects about the characters' identity may also be indexed, as explained below.

As discussed in section (6.3.), to examine the geographic-location indexation value of these phonetic representations and address RQ2 (see (1.2.)), all InF occurrences were classified by county while the ones produced by Dubliners were divided into North/South. Linguistically, the Northside of Dublin City represents what Hickey (2005a) refers to as 'local' Dublin English. This is spoken by speakers who favor linguistic items which are traditionally and regionally bound to IrE, and who "show strongest identification with the traditional Dublin life of which the popular accent is very much a part of" (Hickey, 2016, p. 22). One of the features of *local* Dublin English includes the use of /ʊ/ in the STRUT lexical set (Hickey, 2005a). This is orthographically rendered in CoFIrE with the variants *fooken* and *fooking*, which are only used in Howard's books, and produced exclusively by Northside characters. Contrastively, the Southside Dubliners are speakers of 'non-local' Dublin English in Hickey's terminology, and are often (stereotypically) connected to what is popularly known as 'D4 accent' or 'DART⁶⁶ speak'. The (normally younger) speakers of *non-local* Dublin English often incorporate 'newer', more globalized linguistic features (i.e. such as the use of New Intensifying *So* or Non-lexical *totally*, both explored in detail in chapter (8)) which help dissociate themselves from "popular Dublin culture" (Hickey, 2016, p. 22). *Dartspeak* or 'dortspeak', is well known for its retracted and rounded vowel, pronouncing words like "car" (*cor*), "park" (*pork*), etc. *Dortspeak* is orthographically represented in the two *Ross O'Carroll-Kelly* novels in CoFIrE through the rendition of InF as *focking*, which is only

⁶⁶ Acronym for *Dublin Area Rapid Transit*; a suburban rail system that commutes the north, south, and coastal areas of the city.

ever produced by Southside Dubliners in that series (for more on language/accent rendition in the RO’CK series, see section (8.2.1.)).

The quantitative and qualitative analyses of all InF occurrences by orthographic rendition summarized in Table 6.3. below display an array of forms, such as standard (*fucking*⁶⁷), and non-standard (*fuckin*⁶⁸, *fucken*) pronunciations, while also highlighting the occurrence of other orthographic renditions of more specific dialectal uses. The latter underline the clear socioeconomic divide that exists in Dublin City between the Northside (with traditionally lower social prestige) and the Southside, popularly known for its postal area code, i.e. D4 (Hickey, 2016, p. 22), and which is often associated with the accommodated class, thus enjoying higher prestige (for a more detailed explanation of the linguistic and non-linguistic stylization of North/Southside stereotypes in the Ross O’Carroll-Kelly books specifically, see section (8.2.)).

InF Variants in CoFIrE			
<i>Fuckin’</i>	306	<i>Fooking</i>	1
<i>Fucking</i>	154	<i>Fooken</i>	57
<i>Focking</i>	426	<i>Fucken</i>	7

Table 6.3. Distribution of phonetically represented variants of InF in CoFIrE.

The quantitative analysis of these renderings, which is illustrated in Fig. 6.6. below, indicates an obvious over-representation of Dublin English InF occurrences in CoFIrE (which is not restricted only to Howard’s *Ross O’Carroll-Kelly* two novels). The data also shows that *fucking* (standard form) is produced by Southside Dubliners double the amount of times than the Northsiders do.

⁶⁷ Including one case of *effing*.

⁶⁸ Including one case of *motherfuckin’*.

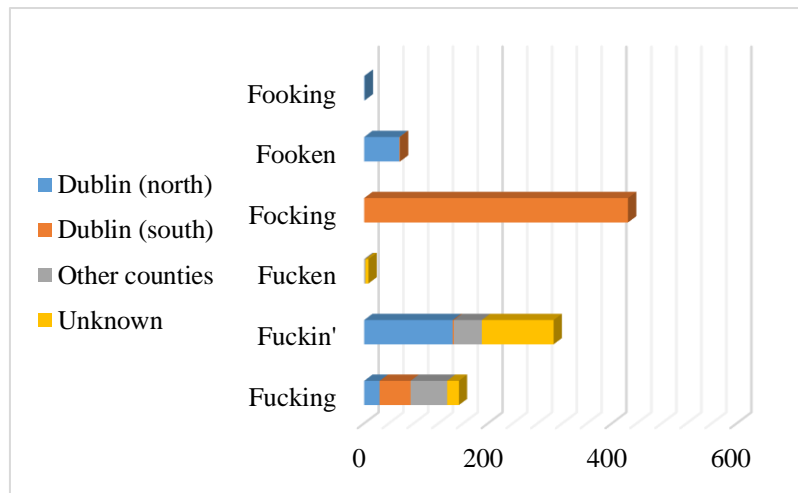


Figure 6.6. Distribution of orthographic renditions of InF by county in CoFIrE.

It should be noted that while the majority of occurrences of InF produced by Southsiders in the two *Ross O'Carroll-Kelly* books are represented with the form *focking*, there are 20 occurrences where it is reproduced as standard *fucking*, in cases like examples 6.4. and 6.5. below.

(6.4.) 'You burned my *fucking* boat' (PHDA)

(6.5.) 'He either doesn't see it or he doesn't *fucking* want to see it' (PHKU)

The Southside-produced, standard *fucking* occurrences are uttered by both male and female characters who are in the 50-60 age cohort. I do not believe this is due to an overlook on the part of the author, as Howard is far too fastidious when it comes to North/South Dublin accent (and non-linguistic) rendition (see (8.2.) for more on the stylization of North/South Dublin stereotypes in his series). Instead, I think that this is a way to represent the fact that (according to his perception), older, upper-class, Southsiders prefer the standard pronunciation, leaving the 'trendier', D4 pronunciation (*focking*) to the younger generation. On the other hand, the use of non-standard *fuckin'* is overwhelmingly produced in the North of the capital (across all books in CoFIrE), with almost no occurrences to the South. This creates a stark North/South Dublin contrast which evidences on paper the socioeconomic and linguistic divide that exists in the city, at least according to the perception of the CoFIrE authors.

As mentioned above, other identifying factors may also be automatically triggered in the readers' mind through orthographic representations. That is the case, for instance, of social class as represented in examples 6.6. through 6.9. below which imbue the less prestigious North Dubliners with the use of *fuckin'/fooken*, whereas the more affluent South Dubliners favor *fucking/focking*.

- (6.6.) “Look wha’ he’s after doin’ te me! Me *fuckin’* leg is in tatters ‘cos o’ him! I jus want te speak te Karina fer a second, it’s not *fuckin’* fair!”(DRTRM)
- (6.7.) “We’re going to Bray please, if my *fucking* wife would ever grace us with her presence.” (DRTRM)
- (6.8.) Look, Ine just aston you to see tings from m...m...moy point of view here. Ine looken at f . . . f . . . *fooken* jail toyum – am I right, feddas?⁶⁹ (PHDA)
- (6.9.) ‘Can you tell that *fucking* woman who you somehow managed to get pregnant that I’m not talking to her at the moment out of pure disgust?’ (PHKU)

Fragments from Ruane’s *Tales in a Rearview Mirror* and Howard’s novels which contrast Northside/Southside characters linguistically were selected here. In 6.6. and 6.8. we see both writers imbuing their Northside speakers with the use of *fuckin’* and *fooken*, both non-standard orthographic variants which appear in conjunction with other characteristic features of *local* Dublin English. These include an ‘after’ perfect (wha’ *he’s after doing* te me!), possessive *me*, and accentual signs in 6.6., while 6.8. includes TH lenition (*tings* for ‘things’) and the centralization of [əɪ] for [aɪ], as in *moy* (my) or *toyum* (time). This centralization is one of the distinctive traits of *local* Dublin English (Hickey, 2007, p. 351), maintained in Dublin by lower and middle-class speakers, as opposed to speakers of *non-local* Dublin English who use [aɪ] or [æɪ] (Hickey, 1999). Contrastively, fragment 6.7. presents us with an exasperated Southsider using *fucking*, while Howard’s 7-year-old, Southside female in 6.9. uses *focking* to express her disdain and sarcasm.

The use of juxtaposing variants in the CoFIrE books, especially in Howard’s two novels, therefore, illustrates very well the value of orthographic rendition in literature. This would answer RQ3 (see (1.2.)), as orthographic rendition of InF would, therefore, be the way through which the authors in the fiction database may be indexing modern Irishness (especially modern Dublin identity) in fictional dialogue. In all, the use of these

⁶⁹ Non-phonetic transcript: ‘Look, I’m just asking you to see things from m...m...my point of view here. I’m looking at f...f...fooken jail time—am I right, fellas?’

orthographic variants not only creates speaker identity within and outside the fictional world, but also manipulates the audience into producing their own opinion about these speakers regarding not just the place of origin (North/South Dublin), but also age (younger speakers' D4-accented *focking* in 6.7. and 6.9.) versus *fucking* which is produced by older speakers in the same fictional world (see examples 6.4. and 6.5. above), or even group identity (i.e. as shall be discussed in (8.2.), Northsiders, for example, are negatively presented by *Ross O'Carroll-Kelly* as scamming 'scumbags').

6.6. Intensifying *Fucking* Colligational Pattern Preference across CoFIrE and contrastive samples

As mentioned in section (6.3.), Mackenzie's (2019) syntactic distribution is followed here to investigate colligational pattern preference. To do so, each pattern was manually classified for colligates (see Appendix 4⁷⁰ for a full list of Mackenzie's colligation patterns contrasted against CoFIrE with examples from the fiction dataset).

The first thing worth noting from the qualitative analysis of pattern preference (Appendix 4) is that most of the categories that Mackenzie identifies have been found both in CoFIrE and in the contrastive *LCIE/BNC2014* samples. This suggests that the fictional representation of InF colligational pattern preference is highly realistic, thus answering RQ 1a pertaining to the level of 'authenticity' present in the fictional rendition of this intensifier in the corpus. The most prominent colligation contexts in CoFIrE and the samples were examined next and are summarized in Figure 6.7.

⁷⁰ The table in Appendix 4 has been color coded. The patterns which agree with Mackenzie's are uncolored, whereas those that disagree or challenge his predictions are in green. Blank spaces have been left in cases where the patterns were not found in CoFIrE (including *Ross O'Carroll-Kelly*'s narrative input) nor in the contrastive *LCIE/BNC2014* samples. However, this does not mean that that specific pattern is inexistent in Irish English or British English, but rather that I have not found it documented in the fiction corpus or the randomized 100-item samples.

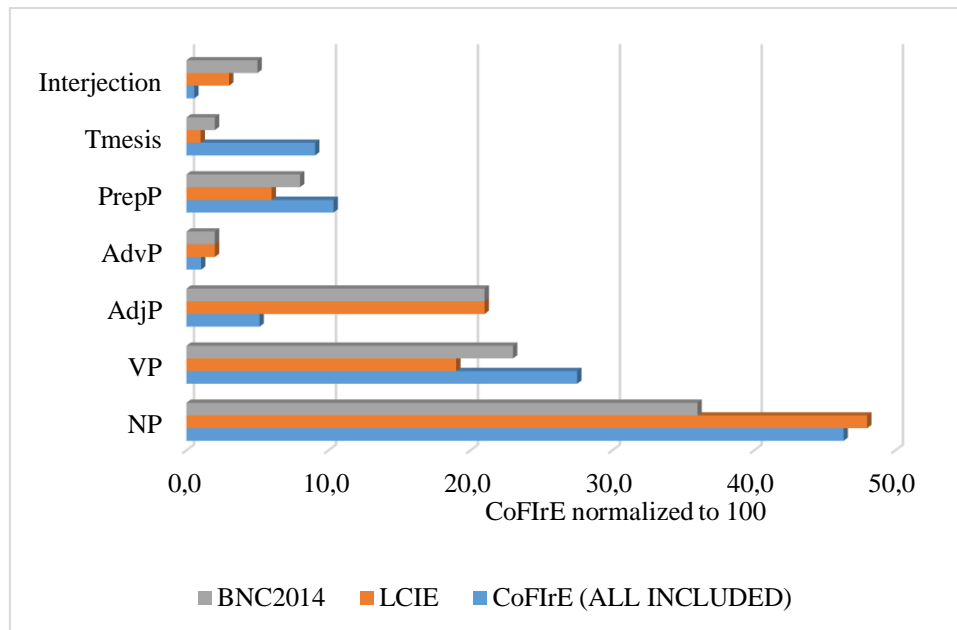


Figure 6.7. Distribution of InF by colligational pattern preference in *LCIE/BNC2014* 100 random item samples, with CoFIrE data normalized to 100 words.

The most noticeable finding evidenced in Figure 6.7. above is the dominance of the *InF+Nominal Phrase* (NP) pattern in CoFIrE, exemplified in 6.10. below, and which includes 14 occurrences of *InF+Pronouns* (see 6.11.), followed by *InF+Verbal Phrase* (VP), which comprises 23 occurrences where *fucking* precedes auxiliaries, as in 6.12. below.

(6.10.) I open my door! I offer ye the full *fuckin'* *courtesy* of my home! (KBDL)

(6.11.) 'Did you talk to um?' [...] I'm like, 'Not yet, no' She goes, '*Fooken* *when* then?' (PHDA)

(6.12.) I don't *fuckin'* *believe* dis! (DRTRM)

The contrast of the CoFIrE data against the comparative *LCIE/BNC2014* samples would further buttress the highly realistic fictional depiction (i.e. RQ 1a), for the *InF+NP* pattern also emerges as the most prominent colligate in *LCIE*, with *InF+VP* also ranking as the second, most recurrent pattern in the Irish sample. These findings are replicated in the *BNC2014* sample which, despite its small size, might be indicative of the prominence of InF use in British English.

A more detailed, qualitative investigation of all of these patterns also retrieved important findings regarding what Mackenzie considers to be patterns which are ‘predicted not to precede’ or are classified as ‘cannot occur’. An example would be InF *preceding possessive pronouns*, which Mackenzie illustrates with, “*That suitcase is *fucking* his”. However, he mentions that contrastive stress might make its use possible in cases like: “That’s *fucking* MINE, not *fucking* YOURS” (Mackenzie, 2019, p. 75). Since I understand all occurrences of InF to carry stress by definition, be it contrastive or non-contrastive, I did not classify for ‘contrastive stress’. The quantitative and qualitative analyses of InF in CoFIrE and the contrastive samples reveal some colligational patterns (summarized in Table 6.4., and colored green in Appendix 4) which contradict Mackenzie’s predictions.

Mackenzie’s ‘Negative’ Predictions				
CASES		CORPORA		
‘Predicted NOT to precede’		CoFIrE	LCIE	BNC2014
	<u>Predicted NOT to precede</u>			
In/within noun phrase	added numerals or determiners	✓	✓	
In/within Verbal Phrase	Only AFTER all auxiliaries (including <i>to</i>)	✓		✓
	It CANNOT precede			
In/within prepositional phrase	Prepositional heads	✓		
	Interrogative pronouns	✓		
With pronouns	Demonstrative pronouns	✓		
	Personal pronouns	✓		
	Possessive pronouns	✓		

Table 6.4. Cases of InF in disagreement with Mackenzie’s ‘predictions’ with checkmarks indicating their appearance in CoFIrE and/or LCIE/BNC2014 randomized samples.

The contradicting patterns include *preceding added numerals/determiners*, in cases like 6.13. below, where speaker 1 asks whether they can have a chocolate mousse and speaker 2 generously offers as many as they want, emphasizing the number with the intensifier. InF is also found to *precede auxiliaries*, like 6.14 and 6.15 which are part of elided, anaphoric verbal phrases. The elided part in 6.14 could be: ‘[...] no sir, it jus’ fuckin’ ain’t [like Wolfie to be in a love muddle]’, while one might expect the speaker in 6.15 to continue: ‘Of course I focking will [look after the shop for you]’.

- (6.13.) <S1> Can I have one? <S2> Yeah have fuckin' | fucking ten if ya | you want.
(LCIE)
- (6.14.) An’ she got my boy Wolfie in a love muddle ’n’ all, and that ain’t like Wolfie, no sir it jus’ fuckin’ ain’t, like. (KBCB)
- (6.15.) ‘Ross, will you look after the shop for me?’ And I’m like, ‘Of course I focking will! You just go and do what you have to do!’ (PHKU)
- (6.16.) ‘Shut fuckin’ up, an’ ged in will ye’ (DRTRM).

There is also a case in CoFIrE (i.e. 6.16. above) where InF can precede a *prepositional head*, although this is a complex case. Its complexity lies in the potential interpretations of *up*. On the one hand, 1) *up* is a particle that is part of a phrasal verb, and, as such, might incidentally be considered an example of *non-standard tmesis*, dividing an otherwise inseparable phrasal verb (see section (6.6.1.) for a detailed analysis of *standard* and *non-standard tmesis* in CoFIrE). On the other hand, and according to Pullum & Huddleston’s (2017, pp. 597-662) innovative take on prepositions, 2) *up* could also be understood as a prepositional phrase. In their *The Cambridge Grammar of the English Language* (2017), they propose a new understanding of prepositions which deviates from traditional grammar in that they are conceived “to be heads of [...] preposition phrases which are comparable in their structure to phrases headed by verbs, nouns, adjectives, and adverbs”, with accompanying nominal phrases (e.g. I saw him before lunch), verbal phrases (e.g. I saw him before he left) or Adverbial Phrases (e.g. I didn’t meet him until recently), while others can stand alone (e.g. I saw him afterwards) (*ibid.*, p. 58; all their examples). According to Pullum & Huddleston’s (2017) new take, the *up* in constructions like “shut the fuck *up*” functions as its own preposition phrase. Although consultation with my supervisors indicated strong disagreement with Pullum & Huddleston’s view, I felt it

necessary to point out this new take on prepositions, as its occurrence in cases like 6.16. above in CoFIrE would also contradict Mackenzie's prediction.

Mackenzie also divides the pattern *fucking+pronouns* into pronouns it CAN and CANNOT precede. CoFIrE documents cases which are identified by Mackenzie as 'cannot precede', as exemplified in 6.17. through 6.21. below.

- (6.17.) 'Did you talk to um?' meaning Kennet. I'm like, 'Not yet, no' She goes, 'Fooken *when* then?,' the absolute chormer that she is (PHDA)
- (6.18.) 'There was something I didn't tell you – the day you called here with your old pair and *fucking what's* his face' (PHKU)
- (6.19.) *Fucking this* again, I'm thinking (PHKU)
- (6.20.) We sit down, then in walks the judge and we have to all, like, stand up again, just for *fucking him* (PHDA)
- (6.21.) I'm thinking to myself, Yeah, remember whose gaff this is now. Here's a hint. It's not *fucking yours*. (PHKU)

As illustrated above, there are *InF* which *precede interrogative pronouns* as in 6.17., where the interrogative pronoun (*fucking when*) is anaphoric in nature (i.e. *fooken when* [will you talk to him]). 6.18. contains another interrogative pronoun ('*what's his face*'), which is part of a (dismissive) idiomatic phrase substituting a proper noun (i.e. '*fucking what's his face*' could be '*fucking Christian Smith*'), and 6.19. illustrates a case of *InF preceding a demonstrative*, which shows the narrator's exasperation. Mackenzie also specifies some cases where *fucking* can only collocate with certain pronouns if they carry 'contrastive stress'. While, as mentioned above, classification for contrastive stress was forgone in this project, instances of *InF+Personal Pronoun* (6.20.) as well as occurrences of *InF+Possessive pronouns* (see 6.21.), which Mackenzie classes as CANNOT occur, were found in the fiction corpus.

It must be noted that the appearance of colligational patterns which were predicted to be 'non-possible' in Mackenzie's classification in CoFIrE does not necessarily mean that these patterns are new. Instead, it is possible that the (contemporary) oral component of CoFIrE and contrastive samples accounts for the larger variety of patterns.

6.6.1. *Tmesis* Pattern Preference across CoFIrE and contrastive samples

The examination of all InF occurrences for colligation patterns (discussed above in section (6.6.)) also retrieved cases of *tmesis* which, as mentioned in section (6.2.), were divided into *standard* (i.e. InF inserted into one word, e.g. ‘Wouldn’t mind but **Immacu-fuckin’ -lata** is the spawn o’ fuckin’ dock trash’ (KBCB)), and *non-standard tmesis* (occurring within compounds, e.g. he was going to be the **Attorney Fucking General!** (DRYTAD)), hereafter ST and NST respectively for short. This section aims at providing a more detailed classification of the types of NST that are possible, at least in the corpus and contrastive samples. As illustrated in Table 6.5., the quantitative study of the 132 raw occurrences of *tmesis* in CoFIrE indicates that while ST and NST are both present in the fiction corpus, NST is overwhelmingly more popular.

TMESIS		
	<i>Standard Tmesis</i>	<i>Non-standard Tmesis</i>
CoFIrE	9.09	90.9
LCIE	-	1
BNC2014	-	2.06

Table 6.5. Distribution of *Standard* and *Non-Standard* InF *Tmesis* across CoFIrE and 100-item *LCIE/BNC2014* samples, all normalized to 100 words.

Further research was, then, conducted into the colligation pattern preference of each type of *tmesis*. To do that, all occurrences were manually classified for colligates. Their qualitative analysis (summarized in Table 6.6.) indicates that *tmesis* can occur, at least mostly in the context of CoFIrE, with(in) compounds, fixed units, with proper names, and with numbers.

TMESIS across the corpora

<p>Standard Tmesis (inserted into one word)</p>		con- <i>fuckin</i> –sortium (DRYTAD); a- <i>fuckin</i> ’ – broad (KBCB); sur- <i>fucking</i> - prise (PHDA); un- <i>fucking</i> -believable; lemon- <i>fucking</i> - ade (PHKU)
	+Compounds	The elevator takes for- <i>fucking</i> -ever. (PHDA)
	+Fixed units	Fine, there's a tenner. Merry <i>fucking</i> Christmas (DRTRM)
<p>Non-standard Tmesis (inserted between compounds)</p>	+Names	Michael <i>fuckin</i> ’/ <i>fucking</i> Owen (LCIE)
	+Numbers	I’m forty- <i>fuckin</i> ’-three and I’m sat around talkin’ fuckin’ gang fights? (KBCB)

Table 6.6. Classification of InF *Tmesis* across CoFIrE and the *LCIE/BNC2014* samples.

Despite the frequent occurrence of *tmesis* in CoFIrE, the contrastive analysis against the samples (Table 6.5.) shows that the use of ST and NST is almost non-existent in either. While generalizations as to whether *tmesis* is more or less prevalent in Irish/British English should be avoided due to the size of the samples, I believe the fact that there is only one occurrence in LCIE could also be due to the fact that these features were emerging at the time of compilation, but gained trendiness in later years as is recorded in contemporary fiction. Similarly, the two occurrences in the British sample are not conclusive to whether or not this feature is widespread in that variety. However, their presence in such small samples warrants future research into my theory that this may be a more recent phenomenon, at least in the context of Ireland.

6.7. The pragmatics of Intensifying *Fucking* across CoFIrE and contrastive samples

Jakobson (1980, p. 82) defined *emotive language* as a “direct expression of the speaker’s attitude toward what he [sic] is speaking about [...which] tends to produce an impression of a certain emotion whether true or feigned”. Since swearing is emotive language at its core, this chapter must look at the pragmatic and emotive value of InF since, to my

knowledge, no other study has investigated this subject in as much depth in the context of IrE (fictional) dialogue.

In her analysis of spoken IrE, Murphy (2009, pp. 96-102) found that while InF transmits emotive meaning, it can also convey what she terms ‘emotive attitude’ in the form of *positive* or *negative* evaluations, further discovering that both genders almost exclusively use it to express *negative* evaluations. Stapleton (2010, p. 294) states that the type of emotive ‘meaning’ conveyed by the intensifier is dependent upon the context and manner in which it was used. Ljung (2011, p. 23) also remarks upon the fact that it is up to the hearer (or reader, in the case of CoFIrE) to interpret those meanings according to both linguistic and non-linguistic information, although a level of uncertainty regarding the feelings conveyed by the intensifier is still guaranteed sometimes. He also provides a list of emotive meanings (2011, p. 22) which several interjections, emphasizees, and expletive slot fillers (i.e. *bloody, fucking, goddamn*, etc.) may have, expressing “surprise, pain, fear, anger, disappointment and even joy”, while others could communicate “emphatic disagreement and/or incredulity”. For emphasizees and expletive fillers he cites the transmission of ‘pure emphasis’ or their functioning as ‘adjectives of dislike’. Despite the variety of meanings his list includes, Ljung remarks upon the incompleteness of the list. Thus, this thesis aims at building on his list of emotive meanings, while also expanding Murphy’s idea of *positive/negative* evaluations by adding a third possibility, i.e. *neutral*, which accounts for Ljung’s ‘pure emphasis’ category.

Thus, to explore the ‘catalogue’ of emotive meanings InF conveys in CoFIrE and the contrastive LCIE/*BNC2014* samples, and as mentioned in section (6.3.), all occurrences were manually classified for *positive*, *negative*, or *neutral* emotions according to the context of speech (including the wider context of the exchange). The *neutral* category was added to account for Ljung’s (2011) ‘pure emphasis’, as I found cases in CoFIrE where InF only transmits a simple, emotionally neutral stress, such as in: “You’re like peaches and fake cream. Polar **fucking** opposites.” (NCCSM). Finally, each occurrence was further sub-classified for emotive meaning (e.g. transmission of *admiration, sarcasm, disdain*, etc.). This process was also carried out on the LCIE and *BNC2014* samples. It must be noted that classifying occurrences for emotive meaning is a subjective process in nature. However, and as mentioned in (6.3.), I believe my familiarity (as a reader) with all the books, contexts of utterance, and inter-character relationships lends objectivity to the emotional catalogue.

The quantitative analysis of CoFIrE regarding *positive/neutral/negative* meanings (Fig. 6.8.), with the corpus and *LCIE/BNC2014* sample data normalized to x100, highlights the dominance of *negative* evaluations/emotions, which corroborates Murphy's (2009) findings. *Neutral* meanings take the second position, while *positive* emotions are substantially fewer in number. The fact that these findings are mirrored in the contrastive *LCIE/BNC2014* sample data, I believe, further reinforce the high level of realism evidenced in the fictional portrayal of InF use offered in CoFIrE, which would answer RQ 1a, with regard to real-life use in the context of Ireland and in its neighboring variety

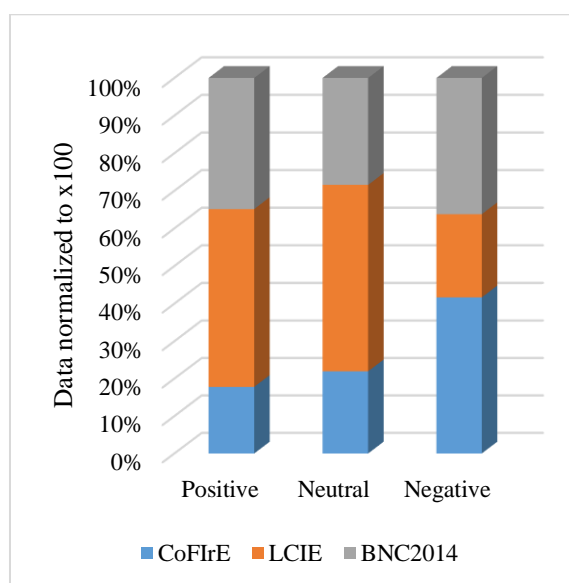


Figure 6.8. Functional distribution of InF across CoFIrE and *LCIE/BNC2014* randomize samples.

From a sociolinguistic perspective (Fig. 6.9.), we see that *negative* emotions are dominant across both genders in CoFIrE and the *LCIE* sample, which also suggest a highly realistic fictional portrayal of use in the context of Ireland. However, the *BNC2014* sample shows a deviation, with females producing more negative emotions. This, however, should be further looked into given the reduced size of the British sample.

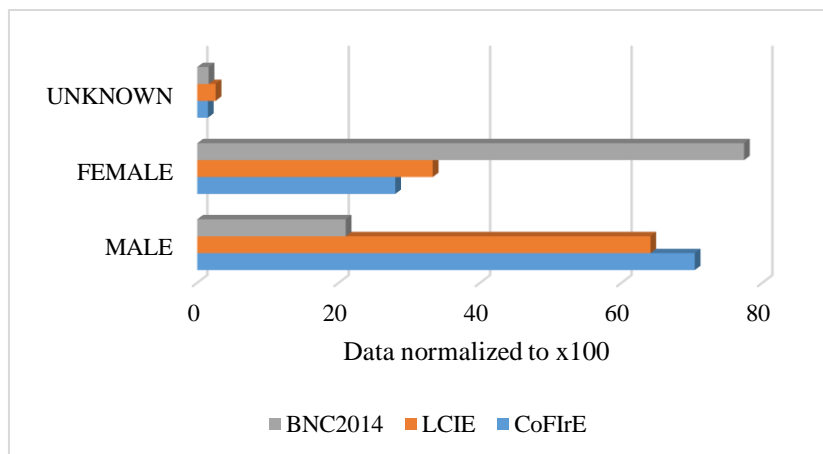


Figure 6.9. Gender distribution of InF across CoFIRE and *LCIE/BNC2014* randomized samples.

While the quantitative analysis of all CoFIRE *negative* occurrences per gender (see Table 6.7.) indicates that *negative* emotions are dominant across both genders, as mentioned above, male characters appear to convey more *positive* feelings than their counterparts, although this could be due to their more elevated production of InFs in the corpus.

	Positive	Neutral	Negative
Male	2.6	26	71.36 ⁷¹
Female	1.7	17.7	80.6
Unknown	15.8	26.3	57.9

Table 6.7. Percentage of distribution of InF by gender in CoFIRE.

Having determined the prominence of *negative* emotion transmission, and in order to expand on to Ljung’s (2011) list of emotions, the next step was to qualitatively examine each occurrence and classify them for emotive meaning. To do this, I designed an emotional classification which is visually illustrated in Fig. 6.10. below. Given the overwhelming frequency of *negative* meanings, I decided to explore this category in

⁷¹ 47.9% (i.e. 236 occurrences) were produced by *Ross O’Carroll-Kelly*, which still leaves a considerable 256 occurrences of *negative* emotions being conveyed by other male characters in CoFIRE.

depth, both quantitative and qualitatively, so as to lay out the *negative* emotive catalogue the (fictional) speakers convey (Appendix 5 lists each negative emotion with examples).

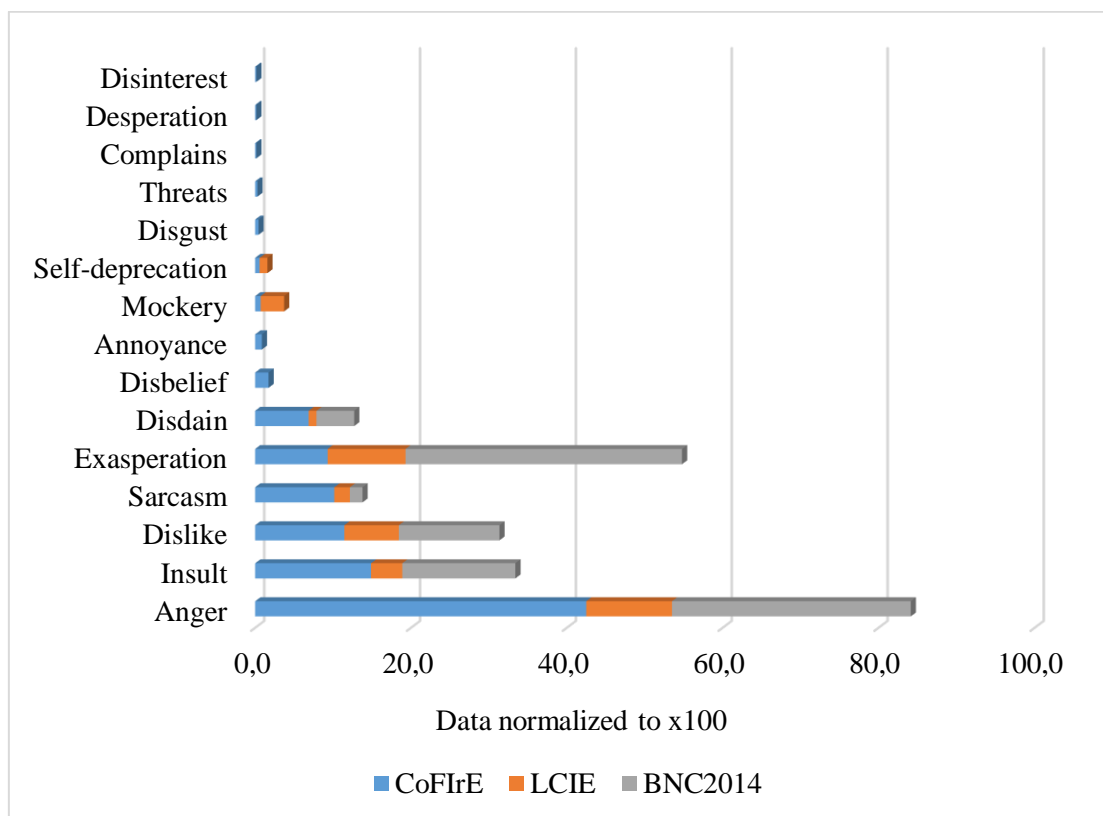


Figure 6.10. InF (Negative) emotion breakdown across CoFIrE and the *LCIE/BNC2014* randomized samples.

The qualitative analysis of the *negative* evaluations transmitted by both genders revealed a wide array of meanings (Fig. 6.10. above), the top three most frequent of which are: *anger* (6.22.), *insults* (6.23.), and *dislike* (6.24).

(6.22.) I want my **fuckin** pension you little prick (DRYS)

(6.23.) ‘I s’pose you know that possessed **fuckin**’ she-devil above in the house will put me in the ground?’ (KBDLI)

(6.24.) ‘Happy? Who’s happy in **fuckin**’ Bohane? Ya’d be a long time scoutin’ for happy in this place.’ (KBCB)

While comparatively smaller, the (negative) emotive list in *LCIE/BNC2014* resembles CoFIrE’s very closely, for the top three emotive meanings in *LCIE* (i.e. *anger*,

exasperation, and *dislike*) and in *BNC2014* (i.e. *exasperation*, *anger*, and *insults*) are almost identical to those found in the fiction corpus. This suggests the highly accurate portrayal of the pragmatic value of InF that the contemporary Irish authors in CoFIrE have managed to record on paper (RQ 1a), which offers a very faithful representation of real (pragmatic) use in the context of Ireland and in the neighboring islands.

6.8. Intensifying *Fucking* and the homodiegetic narrator in CoFIrE: a *Ross O’Carroll-Kelly* case study

Ross O’Carroll-Kelly’s (RO’CK) homodiegetic/autodiegetic narrative production of InF was explored separately from the rest of the fiction cast in CoFIrE. At 506 occurrences, this narrator is responsible for producing 34.7% of all InF occurrences in CoFIrE. Nevertheless, his narrative input regarding pragmatic functions appears to be different to that of all other characters, including his own voice. The first step was to analyze his pragmatic production both qualitatively and quantitatively. As summarized in Table 6.8., his narrative voice uses InF more recurrently to transmit *positive/neutral* emotions rather than *negative* evaluations, which differs from his own production as a character. However, it is *neutral* emotions which are the most frequently conveyed by the narrator in cases such as: “It’s possibly the nicest meal I’ve ever *fucking* eaten” (PHDA).

Positive	Neutral	Negative
12	269	225

Table 6.8. Pragmatic functions of InF as used by narrator *Ross O’Carroll-Kelly* in CoFIrE.

Given the fact that RO’CK ‘reports’ his story in a confessional manner to the author, and ultimately to the readers, (for more on style in the RO’CK series, see section (8.2.1.)), Amador-Moreno (2015, p. 377) makes the point that his voice and linguistic use (at least with regard to discourse pragmatic marker *like* in her study) functions as a mirage of intimacy between the narrator and the reader, as he “appear[s] to be conveying [his] message directly” to the reader (Montoro, 2012, p. 131). I believe the elevated number of

InF occurrences, and the high conveyance of *neutral* emotive meanings, might also be due to RO’CK’s attempt to construct and reinforce an informal, intimate, and close bond (Stenström, 1991; Murphy, 2010, p. 81; Clancy, 2016) with his ‘listener’, which would also be a way through which the author (in the voice of the character) manipulates the reader’s emotional response to the narrator. Another explanation from an authorial and business angle could be the fact that the more familiar the readership becomes with the protagonist/narrator, the stronger the fan base may become, which might increase, or at least consolidate, a stable group of buyers who are interested to hear/read the latest news from their “close friend” in the next installment.

The breakdown of *negative* emotions expressed by the narrator (Table 6.9.) is noticeably similar to the one transmitted by all other characters in CoFIrE (see previous section).

<i>Exasperation</i>	57	<i>Self-</i>	6
<i>Sarcasm</i>	43	<i>deprecation</i>	
<i>Insult</i>	36	<i>Disbelief</i>	6
<i>Disdain</i>	30	<i>Annoyance</i>	5
<i>Dislike</i>	22	<i>Mockery</i>	3
<i>Anger</i>	16	<i>Disinterest</i>	1

Table 6.9. *Negative* emotion breakdown and distribution of InF when produced by narrator *Ross O’Carroll-Kelly* in CoFIrE.

Notice how the three most recurrent emotions conveyed by narrative RO’CK are *exasperation*, *sarcasm*, and *insults* in cases like 6.25. through 6.27.

- (6.25.) We seem to hit every *focking* pothole between Mbale and wherever the fock we’re going (PHKU)
- (6.26.) I’m about as welcome here as a *focking* skin disease (PHDA)
- (6.27.) Oissin slithers to me like the *focking* snake that he is (PHKU)

(6.28.) ‘Jesus focking Christ. Is that it, Ro?’ because this is a great kid we’re talking about, with an IQ of something *focking* astronomical (PHKU)

(6.29.) *Focking* yes! Another bottle – this time of Grey Goose – with a security tag on it. (PHDA)

Finally, there are also two *positive* emotions being transmitted in the form of *admiration* of his son’s intelligence (see 6.28.), and *excitement* (6.29.). In this particular case, the speaker is excited about having timed the order of the items he would purchase at the grocery store so perfectly that it would delay the cashier’s breakneck speed at scanning them, thus granting him extra time to bag all the articles.

6.9. Concluding remarks

This chapter endeavored to investigate the use of InF in CoFIrE aiming to answer the RQs of the thesis (see (1.2.)). In answer to RQ1, pertaining to the features which are more recurrently reproduced in the corpus to render IrE orality, the quantitative analysis of the pragmatic items in CoFIrE indicated the prominence of InF as the second, most produced feature in the corpus, which mirrors the substantially high frequency this particular intensifier enjoys in natural conversation, as also seen in the study of its use in the *LCIE* and *BNC2014* samples.

From a sociolinguistics perspective, and to address RQ2 regarding the indexation of speaker identity, the analysis of InF distribution by age has illustrated the use of this intensifier across all age cohorts in CoFIrE but more prominently among younger speakers, which corroborates findings from other varieties. In particular, InF is more frequently used in the fiction corpus by speakers in the 20-30 cohort, which mirrors Murphy’s (2009, 2010) research based on her corpora of spoken IrE. This would validate the fictional representation of use offered in CoFIrE, underscoring a high level of realism which answers RQ1a.

Further validation of the realism in the fictional rendition was obtained through the analysis of InF production by gender, which also supported Murphy’s theory of a connection between the intensifier and masculinity in the context of Ireland, as it is more prominently used by male characters in CoFIrE. The authenticity of fictional rendition (RQ 1a) is also supported by the fact that InF usage in the contrastive *LCIE* and *BNC2014*

samples was also found to be more prevalent among male speakers in the case of *LCIE*. The *BNC2014* sample, however, indicated that females are responsible for the majority of occurrences. This differing finding, however, may not necessarily indicate that InF signals femininity in British English (at least, this cannot be conclusively asserted given the size of the sample), but rather, I believe, this may be indicative of females (at least in the UK) becoming more comfortable with its use. This could also be extrapolated to Ireland, especially if one considers the elevated number of InF occurrences produced by females in the fiction corpus and *LCIE* sample, which challenges the popular belief that females do not swear as much as male speakers, at least in the context of Ireland.

In terms of social class, and while initial findings indicated a prevalence among the upper-class due to the over-representation of InF by *Ross O'Carroll-Kelly*, closer inspection illustrated a larger production among the middle-class, which is also substantiated in the contrastive samples. This would, at least in the context of Ireland, challenge Jay's (2000, p. 158) argument that middle-class people in other English-speaking countries are reluctant to swear.

Orthographic rendition of InF pronunciation was then explored since the deliberate use of variant forms was found to be a great medium which enables the CoFIrE authors to develop speaker identity within and outside the fiction world. The qualitative analysis of CoFIrE uncovered a wide array of standard and non-standard forms (e.g. *fucking*, *focking*, *fooken*, etc.), also highlighting the occurrence of renditions of more specific dialectal uses. Further analysis also validates my theory as to the usefulness of this tool, which allows the audience to understand the "bigger-picture" regarding speaker identity, for the renderings can not only index place of origin (North/South Dublin), but also age (younger, D4-accented *focking* vs. older generational *fucking* in Southside Dublin), and even group identity. Thus, to answer RQ3, variant orthographic renditions of InF would be one of the ways through which the contemporary authors in CoFIrE may linguistically index modern Irishness.

The study of intralinguistic variables, such as colligational pattern preference, and *tnesis* in particular, further underscore the realism of the fiction corpus (RQ 1a), as it documents its use in the context of Ireland in the majority of contexts identified by Mackenzie (2019), which are also present in the *LCIE/BNC2014* samples. The pattern *InF+NP* was found to be the dominant one in CoFIrE, followed by *+VP*, which was also the case in the *LCIE* and *BNC2014* samples. This which would further validate the level

of authentic representation present in the fiction database (RQ 1a). Subsequent qualitative investigations of the patterns in CoFIrE and the comparative samples also revealed cases which Mackenzie (2019) classifies as ‘predicted not to occur’ and ‘cannot occur’ (e.g. *+possessive pronouns*). While their presence in CoFIrE and the *LCIE/BNC2014* samples does not necessarily imply that their use is new, it is possible that the oral component of CoFIrE and contrastive samples accounts for the larger variety of patterns.

CoFIrE was also found to provide a realistic representation (RQ1a) of the pragmatic versatility of InF, which is demonstrated through the wide range of visceral emotions that this intensifier conveys (across all colligational patterns) in the books, and which this chapter set out to catalogue, adding a *neutral* type of emotion to add to Murphy’s (2009, 2010) *positive/negative* typology. The fact that *negative* emotions are prominent in the fiction corpus, which corroborates previous studies in the context of Ireland, but also in the samples, also illustrates the faithful rendition of InF use in the fiction corpus, which appears to be an effective, written record of real-life use. This is also a testament to the elevated awareness that the contemporary Irish authors in CoFIrE have of the pragmatic value of InF.

Finally, the case study of pragmatic conveyance by the homodiegetic/autodiegetic RO’CK narrator indicates the transmission of *positive/neutral* emotions (as opposed to his own conveyance of negative emotions as a character). This narrator, who ‘speaks’ directly to the author/reader, utilizes a series of direct-speech, linguistic tools, which enable him to develop a closer bond with the audience. I believe his use of InF and conveyance of mostly *positive/neutral emotions* is a part of that linguistic repertoire (RQ3) which allows him to construct and reinforce the narrator-reader bond of intimacy, which was also found to be reinforced through the use of ventriloquizing, non-standard quotatives like GO (see chapter (5)). Finally, I suggest that this may also be understood from a business angle if we consider that the more familiar and ‘closer’ the readership feels to the protagonist/narrator, the stronger the fan base may become, which might increase, or at least consolidate, a stable group of books buyers who are interested to hear/read the latest news from their “close friend”.

The following chapter will explore in detail the third most frequently reproduced pragmatic item in CoFIrE: discourse pragmatic marker *like*.

7. DISCOURSE PRAGMATIC MARKER *LIKE*

Perhaps one of the clearest signs of naturalistic orality portrayal in fiction is the presence of discourse pragmatic markers, illustrated in CoFIrE by the prominence of *like* as the third, most frequent pragmatic item (i.e. RQ 1 (see 1.2.)).

In the last few decades, the field of discourse pragmatic markers has seen a large increase in academic interest. However, the fact that this functional class encompasses so many items (e.g. *like, right, you know, in my opinion, kind of, well, yeah no, I think, so, that is to say, in consequence*, etc.), all varying in length, complexity, and functionality, may have led to disagreement among scholars on a variety of subjects. For example, the definition of the items, their functions, meaning, clausal positioning, inter alia, all seem to differ according to the scholar and/or their research interests. There is even a plethora of labels that have been applied to these items, including ‘discourse connectives’ (Blakemore, 1992), ‘discourse particles’ (Schourup, 1985), ‘discourse markers’, ‘pragmatic markers’ (Schiffrin, 1987; Fraser, 1988), or *discourse-pragmatic features* (Pichler, 2013) to name a few. This thesis will, however, follow Tagliamonte (2012) by using the more inclusive *discourse-pragmatic marker* (hereafter DPM) label.

Carter & McCarthy (2006, p. 208) define DPMs as “a class of items which operate outside the structural limits of the clause and which encode speakers’ intentions and interpersonal meanings”. In her description of their features, Brinton (1996, 2017) remarks upon their prominence in spoken discourse, especially in the informal and colloquial registers (see Fig. 7.1. for Brinton’s (2017, p. 9) revised and updated list of features distributed into the five categories identified by Aijmer & Simon-Vandenberg (2011, p. 226)). And while DPMs may be used in written discourse, their appearance in that medium is often viewed negatively.

Phonological and lexical characteristics

- (a) Pragmatic markers are often “small” items, although they may also be phrasal or clausal; they are sometimes phonologically reduced.
- (b) Pragmatic markers may form a separate tone group, but they may also form a prosodic unit with preceding or following material.
- (c) Pragmatic markers do not constitute a traditional word class, but are most closely aligned to adverbs, conjunctions, or interjections.

Syntactic characteristics

- (d) Pragmatic markers occur either outside the syntactic structure or loosely attached to it.
- (e) Pragmatic markers occur preferentially at clause boundaries (initial/final) but are generally movable and may occur in sentence-medial position as well.
- (f) Pragmatic markers are grammatically optional but at the same time serve important pragmatic functions (and are, in a sense, pragmatically non-optional).

Semantic characteristics

- (g) Pragmatic markers have little or no propositional/conceptual meaning, but are procedural and non-compositional.

Functional characteristics

- (h) Pragmatic markers are often multifunctional, having a range of pragmatic functions.

Sociolinguistic and stylistic characteristics

- (i) Pragmatic markers are predominantly a feature of oral rather than written discourse; spoken and written pragmatic markers may differ in form and function.
 - (j) Pragmatic markers are frequent and salient in oral discourse.
 - (k) Pragmatic markers are stylistically stigmatized and negatively evaluated, especially in written or formal discourse.
 - (l) Pragmatic markers may be used in different ways and in different frequencies by men and women.
-
-

Figure 7.1. Brinton’s (2017, 9) updated list of DPM features based on Aijmer & Simon-Vandenberg’s (2011) categories.

With regard to functionality, research finds that while DPMs may be grammatically optional since their extraction “leaves the sentence structure intact” (Schiffrin, 1987, p. 55), this also hinders the knowledge of the “commitment the speaker makes regarding the relationship between the current utterance and the prior discourse” (*ibid.*, p. 32). In addition, DPMs are largely multi-functional. So much so, that Aijmer & Simon-Vandenberg (2011, p. 229) point out the difficulty of extricating their pragmatic functions since “[they] can have an almost infinite number of functions depending on context” and/or which overlap with other DPMs.

The pragmatic multifunctionality of these items may be due to the process of *grammaticalization* involved in their development as DPMs. Hopper & Traugott (2003, p. xv) describe this as the process “whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions”. In fact, Traugott (1995, p. 15) concludes that the development of DPMs aligns with “prototypical *grammaticalization* in its early stages”. To understand this process, which was already introduced in section (5.5.), and

its connection with the development of DPMs, it is first necessary to outline the three functional components of language (Traugott, 1982):

- 1) **Propositional**: linguistic resources to talk about something.
- 2) **Textual**: linguistic resources that create cohesion.
- 3) **(Inter)Personal**: linguistic resources that transmit speaker attitudes/emotions about the event/topic of conversation.

Initially, a DPM starts its process of development with only propositional functions, but will move down the continuum acquiring textual, and later (inter)personal functions. The DPM may, then, undergo *subjectification* (i.e. indexing speaker attitude), or *intersubjectification* (i.e. encoding meanings centered on the interlocutor/addressee). At this point the DPM will have become *grammaticalized*. This is better explained using my own examples of *like* below:

- a) *I like strawberries*
 - Propositional: Verb. Statement of a fact.
- b) *That looks like a strawberry*
 - Textual: Conjunction.
- c) *You're always, like, eating strawberries*
 - Interpersonal: DPM, focuser function. The focus is placed on the fact that the interlocutor always eats strawberries. This could be pragmatically interpreted as a neutral statement of a fact, or as negative criticism over the repetitive action.

Brinton (2017, pp. 27-30) also outlines other features of *grammaticalization* exhibited by DPMs, which include: a) having originated as nouns, verbs, adjectives, etc. (e.g. *like*, *right*, *see...*), b) changing from open to closed word classes, and c) becoming fixed. They may also d) fuse to some extent (e.g. *sort of* = *sorta*) and/or e) be semantically bleached. The f) acquisition of non-propositional functions is also consistent with *grammaticalization*. Nevertheless, DPMs may also *not* exhibit all features of *grammaticalization*. Examples of the excluded characteristics include their grammatical optionality, their syntactical independence, or the fact that they often do not become phonetically reduced, among others (see Brinton, *ibid.* for a broader discussion).

The occasional exhibition of certain features of *grammaticalization* has led some scholars to consider the need for a new process of development that is distinctive to DPMs: *pragmaticalization*. In their study of *you know* and *ba'* (Swedish DPM), Erman & Kotsinas (1993, p. 79) propose *pragmaticalization* as the process whereby a lexical item in a given context develops “directly into a discourse marker without an intermediate stage of grammaticalization”. Claridge & Arnovick (2010, pp. 179-182) outline the features of *pragmaticalization* as including: a) persistence of original meaning, b) decategorization, c) divergence, d) semantic bleaching followed by pragmatic strengthening, e) scope expansion, f) acquisition of textual and interpersonal functions, and g) subjectification. All of these are, as mentioned above, features of *grammaticalization* as well. The main differences between *grammaticalization* and *pragmaticalization* being 1) the optional nature of DPMs (Aijmer, 1997), and 2) the prominence of pragmatic strengthening and (inter)subjectivity (Claridge & Arnovick, 2010, p. 186; Brinton, 2017, p. 32).

Now that a brief definition of DPMs and their multifunctionality has been provided, this chapter will focus on the study of DPM *like* as the third, most frequently produced pragmatic feature in CoFlrE.

7.1. *Like* in the literature

The use of DPM *like* has arguably become a trending topic among scholars in the last couple of decades (see, e.g. Tagliamonte & Hudson, 1999; Dailey-O’Cain, 2000; Levey, 2006; Tagliamonte & D’Arcy, 2007; Buchstaller, 2008, 2013; Schweinberger, 2014, 2015; Amador-Moreno, 2015; D’Arcy, 2017 *inter alia*). While academic interest in this DPM is relatively recent, perhaps due to its increase in frequency over the second half of the twentieth century (D’Arcy, 2007, 2008), D’Arcy (2017, p. 14) points out its presence in the past, attested by its use in different historical and contemporary media.

The increase in academic interest could also, perhaps, be due to the multifunctionality of *like* (see Jucker & Smith, 1998; Amador-Moreno, 2010b, pp. 121-122; and D’Arcy, 2017, pp. 14-15 for discussion of its various functions). While it can serve grammatical purposes, acting as verb, preposition, conjunction, etc., such as in example 7.1. below

where it functions as a conjunction, it may also have pragmatic functions. For example, research shows that *like* is often perceived as a marker of informality and friendliness (Dailey-O’Cain, 2000, p. 73), which also has inter-personal functions, as it may mitigate the force of a speech act.

(7.1.) ‘Whath a thithy?’ ‘It’s when you’re a boy but you act *like* a girl. You don’t want to get the shit kicked out of you in school, do you?’ (PHDA)

Appearing across all clausal positions, D’Arcy (2017, p. 15) also comments upon its use as a marker of subjective information, signaling the speaker’s stance. Furthermore, research (see Siegel, 2002; Levey, 2003; Kallen, 2005; Amador-Moreno, 2010b, pp. 121-122; Truesdale & Meyerhoff, 2015) indicates that *like* can also function as an exemplifier, approximator with numerals and quantitative expressions (see 7.2. below), or as a hedge (as in 7.3.), marking hesitation, pauses, or indecisiveness/unfamiliarity.

(7.2.) [...] all I was meant to have to do was drive up, he’d hand over the kid and I’d keep him for *like*, a night or some shit. (DRYS)

(7.3.) “Yeah, he is, but you haven’t met him, Amy. He’s... *like*, well when he was...” (DRTRM)

Like may also act as a focusing device (see example 7.4.), emphasizing the speaker’s feelings, introducing explanations or exemplifying (Amador-Moreno, 2010, pp. 121-22). Finally, a more recently ‘acquired’ function (or at least recently noticed in academia) is its use as the non-standard quotative BE LIKE in cases like 7.5., which was already explored at length in sections (5.9.) through (5.9.5.).

(7.4.) ‘This is not actual ceweal,’ she said. ‘This is, *like*, twibute ceweal?’ (KBDLI)

(7.5.) *She’s like*, ‘You know what’s going to happen, of course?’ (PHKU)

The use of *like* also appears to have very strong sociolinguistic associations regarding speaker age, gender, and regional location. For example, although it is present across all varieties of English, research finds a strong general perception that it originated in the

USA (Andersen, 2001, p. 215; D’Arcy, 2017, p. 217), where it is often associated with the negative stereotype of the California “Valley Girl” (Dailey-O’Cain, 2000, p. 70). As already discussed in section (5.9.), this is a persona which typecasts young, middle-class, females from the San Fernando Valley in California (Tagliamonte, 2016, or D’Arcy, 2017 for further discussion, and (7.5.1.) for more on the sociocultural perceptions of this stereotype). The stereotype, which seems to have spread through and been perpetuated in pop culture, also has an associated linguistic repertoire. For example, features other than *like* which are popularly (mis)associated with the Valley Girl, and which are prominently used in CoFIrE and CoROCK, are the use of *high rising terminal*⁷² (i.e. intonational deviation whereby statements are posed with as *interrogatives?*), non-standard quotative BE LIKE (see sections (5.9.) through (5.9.5)), or New Intensifying *So* and Non-lexical intensifier *Total(ly)*, which are examined in detail in chapter (8). Section (7.5.1.) provides a detailed overview of the Valley Girl linguistic repertoire.

Besides the (mis)association of *like* with the Valley Girl stereotype, this DPM is often perceived as a marker of female speech, with studies often highlighting a greater frequency of use among females than their counterparts (see Romaine & Lange, 1991; Andersen, 2001 for London teenage speech; D’Arcy, 2005 for Canadian English). However, D’Arcy (2017) finds that while some of its uses are clearly more prominent among females, others (what she labels ‘particle LIKE’) are more frequent in male speech. With regard to age, research generally suggest a link between *like* and young speakers, especially teenagers (see Tagliamonte & Hudson, 1999, for Britain and Canada; Andersen, 2001, for London; Tagliamonte, 2005 for young Canadians aged 10-19; or Tagliamonte, 2016 for teenage speech in general). Nevertheless, D’Arcy (2017, pp. 14-15) notes that *like* seems to have quickly spread across generations, and that despite popular association with teen voice, historical documentation (*ibid.*, p. 139) evidences its use by older speakers across variations of English (e.g. England, Ireland, New Zealand, and other former colonies). Regardless of the popular (mis) associations, the fact remains that *like* has become a common feature across all varieties of English, as is the case with IrE.

⁷² *High Rising Terminal* or *Uptalk*: written representation of intonational deviation in declarative sentences which are purposely issued as *interrogatives?* Of unknown origin, it is sometimes associated with Australian English, but popularly perceived as having originated in the USA (1970s-80s). It is also typically associated with trendy young/teen California ‘Valley’ girls (Warren 2016, p. 82).

7.1.1. *Like* in IrE

Despite its (mis)association with American English and its worldwide cross-variational presence (Kallen, 2013), *like* has been identified as a frequent item in the IrE pragmatic repertoire (Clancy, 2016; Vaughan & Clancy, 2016; Clancy, 2018; O’Keeffe et al., 2020 inter alia). In fact, Siemund et al. (2009, pp. 21-30) notice a traditionally IrE-distinctive clause-marginal preference as in 7.6. and 7.7. below, which they find tends to outnumber clause-medial occurrences like 7.8.

- (7.6.) ‘Do you think it's possible to ever get over losing someone? *Like*, really get over it?’ (NCCSM)
- (7.7.) ‘What a fucking hypocrite! She nearly raped my father at my granny's funeral. His mother, *like*. (DRYS)
- (7.8.) ‘I think they have, *like*, a whole litter.’ (PHKU)

In his study of IrE and South-Eastern British English, Schweinberger (2015) notices a preference for clause-final *like* with mitigating functions among Irish speakers, which mirrors Kallen’s (2006), and which contrasts with the British preference for clause-medial position. More recent studies on the use of *like* in spoken and written representations of IrE (see, e.g. Amador-Moreno, 2015, p. 376; Schweinberger, 2015, p. 127; or Terrazas-Calero, 2020) have also observed a shift toward clause medial preference within this variety. This shift, however, could be much older, seeing as how Amador-Moreno’s (2019b, p. 112) exploration of pragmatic markers in the historical IrE of the *Corpus of Irish English Correspondence* finds clause-medial *like* to be the most prominent position in letters from 1890s-1900s, followed by clause-final occurrences, with initial position being the least used.

In exploring the prominence of *like*, particularly in clause medial position, in Paul Howard’s *The Curious Incident of the Dog in the Night Dress*, Amador-Moreno (2015, p. 376) argues that the conscious portrayal in fiction of this particular clausal-preference may be representative of *supraregionalization*. In other words, a feature which is “less regionally bound” (Hickey, 2005a, p. 203). If so, then the prominence of clause-medial *like*, Amador-Moreno (2015, p. 376) posits, would indicate a shift from “vernacular to

more globalized uses of *like*". In addition, Amador-Moreno & Terrazas-Calero (2017, p. 257) argue that this shift may be motivated by language external factors such as mass media and speaker attitudes to certain linguistic items (see (7.5.1.) for more on the sociopragmatic associations of *like* in American and Irish English).

With regard to sociolinguistic perceptions, the use of *like* in Ireland appears to diverge slightly from the associations it carries in other varieties. While the connection with younger speakers remains (Schweinberger, 2014, p. 266), studies find that *like* is more significantly produced by female speakers who are in their 20s (Murphy, 2015, p. 131; Schweinberger, 2015). This deviates from the association with teen speech in varieties like American English (Romaine & Lange, 1991) or British English (see Andersen, 2001 for London teenage speech; D'Arcy, 2005), although it supports D'Arcy's (2017, pp. 14-15) belief that its use has spread cross-generationally.

In terms of pragmatic multifunctionality, studies examining *like* in Ireland (Schweinberger, 2012; Amador-Moreno, 2015, 2016; Murphy, 2015) have commented upon its value as a marker of informality in both spoken and written representations of IrE. Others (and Kallen, 2006 for IrE intimate discourse; Amador-Moreno, 2010a; Schweinberger, 2012 for IrE in general; Clancy, 2011) point out its mitigating or downtoning functions, or its use as a hedge in clause final position (Kallen, 2005, p. 137). Amador-Moreno (2015, pp. 376-77) also identifies a wide variety of uses in IrE fiction which correspond with those seen in real, spoken discourse, ranging from *like* functioning as approximator, to serving as a tool to foster inter-personal closeness (see also Murphy, 2010, p. 81) and intimacy. Furthermore, Amador-Moreno (2015) also lists a number of functions it serves in the context of fiction which include the infusion of realism and natural fluidity into the narration and storyline, as well as the linguistic indexing of the complex identities the speakers (i.e. fictional characters) have. This chapter, therefore, builds on previous research on the use of *like* in IrE as represented in CoFIrE, examining in detail clausal positioning preference, pragmatic multifunctionality, and sociolinguistic associations to address the thesis' RQs (see (1.2.)).

7.2. Methodology of analysis

The analyses carried out in this chapter were conducted using a similar methodology as the one outlined in (6.3.). Since every occurrence of *like* was annotated in CoFIRe according to its phrasal positioning, the first step was to retrieve all occurrences of *like* via a search for the tags attributed to its use in clause initial (i.e. <IL>), medial (<ML>), and final (<FL>) positions. All occurrences were, then, extracted onto an Excel File (Fig. 7.2. below), where all cases of clause medial *like*, which was found to be the most recurrent position, were manually classified for: book code, use, function, and (sub)function (to address RQs 1 and 1a), as well as for speaker name, gender, age, region, and class so as to answer RQs 2 and 3 regarding the potential speaker identity indexation value of this DPM. Social class was divided into the ranks all CoFIRe characters were attributed (i.e. upper, middle, low), which are explained in section (3.8), as well as their equivalent in the alphabetical class distribution system used in the *BNC2014* (explained in section (3.8.2.)). It should be noted that *Ross O'Carroll-Kelly's* narrative production (i.e. 816 occurrences) of clause-medial *like* in Paul Howard's books will be excluded from the analyses as I have already explored it elsewhere (Terrazas-Calero, 2017).

BOOK CODE	CON-LINE	USE	FUNCTION(SUB)FUNCTION	HRT	SPEAKER	M/F	AGE	REGION	CLASSASS (BNC2014)	
PHDA	Yes, you smiled at her.' 'It mightn't have bee	foc	neg	exasperation	sorcha	F	30	Southside Dublin (upper	A1	
PHDA	efinitely a twenty-four-hour one.' 'Great.' 'Anc	approx numbe	neutral	emphasis	sorcha	F	30	Southside Dublin (upper	A1	
PHDA	Oh my God,' Claire goes, quickly changing the	foc	neutral	emphasis	claire	F	30	Bray (SD)	upper A1	
PHDA	Ross, you're not getting behind that wheel. O	foc	neg	outrage	sorcha	F	30	Southside Dublin (upper	A1	
PHDA	PHDAR}, 'I met this girl from Kilcoole when I	foc	neutral	emphasis (place)	garret	M	30	Bray (SD)	upper A1	
PHDA	So,' Chloe goes, 'can I just ask – did you say	approp word	neutral	emphasis	chloe	F	30	Southside Dublin (upper	A1	
PHDA	would you think of Saturday, the sixteenth of	foc	neutral	emphasis	sorcha	F	30	Southside Dublin (upper	A1	
PHDA	R} goes {GR}{QGO}{/PHDAR}, 'I can't believ	foc	neg	disbelief	disbelief	sorcha	F	30	Southside Dublin (upper	A1
PHDA	{GR}{QGO}{/PHDAR}, 'I think this thing we'r	foc	neutral	emphasis	sorcha	F	30	Southside Dublin (upper	A1	

Figure 7.2. Sample image of the Excel file illustrating the classification of clause-medial *likes* in CoFIRe.

Once all the clause-medial *like* occurrences were classified for use (e.g. focuser, approximate number, etc.), they were also categorized by pragmatic function (i.e. *positive*, *neutral*, and *negative*). To further examine the pragmatic value of clause-medial *like* in CoFIRe and the contrastive *LCIE/BNC2014* samples, I also classified all occurrences for pragmatic (sub)functions in terms of emotion conveyance following the system designed and applied to the classification of intensifying *fucking* outlined in

section (6.3.). Finally, the occurrences were also distributed into co-occurring with *high rising terminal*⁷³ or not, and these were also further classified for emotive conveyance.

With regard to the contrastive *LCIE/BNC2014* samples, and to better answer the thesis' RQs, 100 randomized occurrences of *like* (including all uses of the word, ranging from verb to conjunction, DPM, etc.) were retrieved via a search for the word *like* in each corpus. These were, subsequently, extracted onto another Excel file where they were manually sorted and classified using the same process and classification explained above.

7.3. Frequency of occurrence of *like* in CoFIrE and contrastive samples

The quantitative analyses of all annotated features in CoFIrE identified *like* as the third, most frequently used pragmatic features at 1,249 occurrences (across all clausal positions). Such a large amount of occurrences may be indicative of the high level of informality that is comprised in the CoFIrE texts, given the connection between this DPM and the indexing of informal spokenness already pointed out in research regarding its use in written and spoken portrayals of IrE (Schweinberger, 2012; Amador-Moreno, 2015; 2016; Murphy, 2015).

The high representativeness of real, spoken informality in the fiction corpus (i.e. RQ 1a) is further supported by the data from the analysis of the randomized contrastive corpora samples of 100 occurrences from the *LCIE/BNC2014* samples. As illustrated in Fig. 7.3., the quantitative analysis of the manual categorization of all *like* occurrences in the samples (including other functions of the word) also evidences the dominance of the DPM function in both corpora, which signals the high frequency of this feature in natural discourse. This would, therefore, underscore the realistic value (RQ 1b) of the fictional portrayal of *like*-use offered in the fiction dataset.

⁷³ Intonational deviation whereby statements are posed with as *interrogatives*?

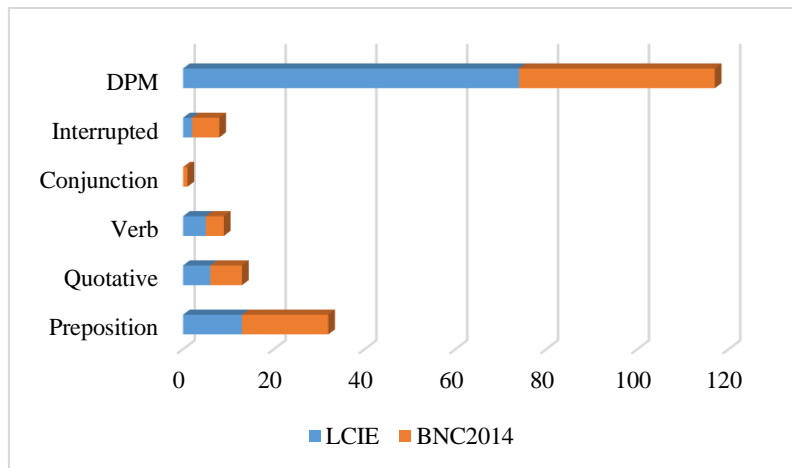


Figure 7.3. Raw functional distribution of *like* across the *LCIE/BNC2014* randomized samples.

The prominence of *like* in CoFIrE along with GO (chapter (5)) and intensifying *fucking* (chapter (6)) may not only be illustrative of the level of highly realistic (RQ 1b) informality portrayed in the fiction texts, but it may also be a testament to the conscious authorial effort made to create a sophisticated representation of contemporary IrE (informal) orality (RQ 3), whose realism and authenticity seems to be validated by the spoken data in *LCIE* sample and in the cross-variational sample.

7.4. Clausal positional preference of *like* in CoFIrE and contrastive samples

Given the different positions *like* may take in a sentence, already outlined in examples 7.6. through 7.8. in section (7.1.1.) above, this factor was also examined. The qualitative analysis of position preference in CoFIrE showed that *like* could appear in clause marginal contexts, including initial and final, such as 7.9.-7.10. below, or in medial positions (see 7.11.).

(7.9.) *Like*, why would he not text or email or Facebook? (DRYS)

(7.10.) ‘Here ye are Louise, yew sort id out wid ‘im, yeah? Yew know ‘im *like*’ (DRTRM)

(7.11.) This is, *like*, typical of my dad (PHKU)

Quantitative examinations (see Table 7.1.) were conducted by means of tag searches across CoFIrE and the 100-item randomized *LCIE* and *BNC2014* samples. The tags used to expedite the retrieval of occurrences from CoFIrE were the *specific linguistic tags* corresponding with its positionings, already explained in section (7.2.), namely <IL> (initial); <FL> (final), and <ML> (medial)).

	MARGINAL		MEDIAL
	Initial	Final	Middle
LCIE	2.70	63.5	33.8
BNC2014	9.3	7	83.7
CoFIrE	2.32	11.45	86.3
			<i>With Narrative Ross</i>
			<i>O'Carroll-Kelly's input</i>
			86.3 (n=1,077)
			<i>Without Narrative Ross</i>
			<i>O'Carroll-Kelly's input</i>
			20.9 (n=261)

Table 7.1. Contrastive clausal-position preference of *like* across CoFIrE and *LCIE/BNC2014* samples with all data normalized to by 100 words.

The findings indicate an apparent preference for clause-medial *like* in CoFIrE. Nevertheless, a closer inspection revealed that 86.22% of the occurrences of clause medial *like* are produced by the character of *Ross O'Carroll-Kelly* in his role as narrator in Paul Howard's books. However, as mentioned in (7.2.), this chapter excludes his narrative production. Despite the removal of his narrative input, clause-medial *like* remains the preferred position at 261 occurrences. These include occurrences produced by *Ross O'Carroll-Kelly* in his role as homodiegetic character as well as by other characters from across all the CoFIrE books. Thus, the quantitative analysis of clause-placement preference in the fiction corpus seems to corroborate the shift toward medial position evidenced in previous IrE studies (see Amador-Moreno, 2015, p. 376; Schweinberger, 2015, p. 127; or Terrazas-Calero, 2020), deviating from the traditional predilection for clause marginal places (Kallen, 2006). This would also attest to the high level of realism (RQ 1a) offered in the fictional dataset, which mirrors shifts in development already documented in the literature.

The traditional inclination for clause-marginal placement, particularly clause final position, is, nevertheless, significantly present in the *LCIE* sample, in cases like, “Shure you know who won everything like so” (*LCIE*). Notice, however, the fact that clause-medial *like* ranks as second most used placement, which contrasts with the very marginal numbers of initial position found in the *LCIE* sample (Table 7.1. above). In contrast, the British sample also evidences a preference for clause-medial *like* in cases like, “yeah like he’s Korean like you ha- like they have to go to the army for like two years” (*BNC2014*). While it is not possible to draw conclusions regarding the seeming preference of *LCIE* speakers for final position based on a the small sample, it could be speculated that this might be due to the time of collection of the corpus data (i.e. 1990s-early 2000s) when the shift toward medial position pointed out in the literature and evidenced in the CoFIrE data might have already been initiated (as evidenced in Table 7.1. above). This move is much more evident in the CoFIrE texts, most of which were published in the first and second decades of the 2000s, at which point clause-medial *like* would have already become established as a feature of supraregional IrE due to, as suggested by Amador-Moreno & Terrazas-Calero (2017, p. 257), “the influence of mass media, or speaker’s attitudes to Irish English in general”.

To answer RQs 2 and 3 (see (1.2.)) regarding the speaker identity indexing value of clause-medial *like* as represented in CoFIrE, the following sections will provide a detailed examination of the sociolinguistic distribution and pragmatic functions of clause-medial *like* in CoFIrE and in the samples (sections (7.5.), (7.6.), and (7.6.1.)), while also exploring the identarian value this DPM has in American English (see (7.5.1.) which appears to extrapolate to the Irish corpus.

7.5. Sociolinguistic distribution of clause-medial *like* in CoFIrE and contrastive samples

The close connection between clause-medial *like* and young (female) speakers that has already been identified in previous studies of IrE (see section (7.1.) for an overview of these) was quantitatively examined in CoFIrE so as to better answer RQ 2 regarding the identity indexical value of this DPM in Ireland according to the corpus’ texts. The data (Fig. 7.4.) indicates the dominance of clause-medial *like* in the 20-30 age cohort in

CoFIRE, which corroborates Murphy’s (2015) and Schweinberger’s (2015, p. 131) findings regarding the use of this DPM among mostly female IrE speakers in their 20s. The fact the fiction findings also mirror the data in the contrastive samples, as illustrated in Fig. 7.4. below, leads me to consider the highly authentic representation (RQ 1a) presented in the CoFIRE texts as regards the indexing of speaker age and gender.

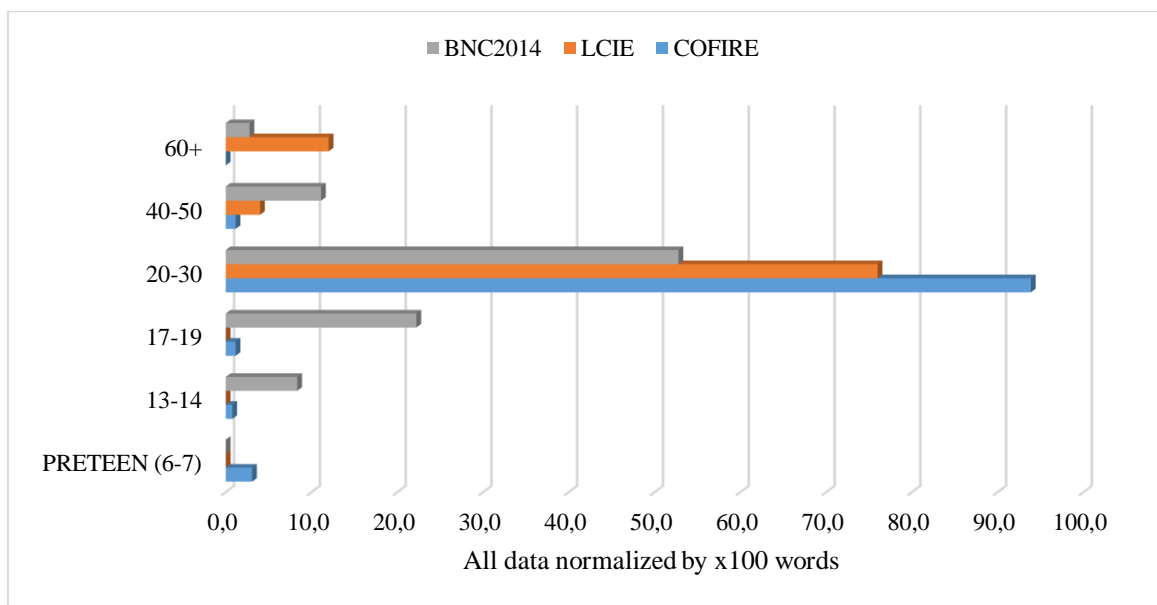


Figure 7.4. Age distribution of medial *like* across CoFIRE and LCIE/BNC2014 samples with all data normalized to by 100 words.

Despite showing a clear prominence among young speakers (i.e. (pre)teenagers and individuals in their 20-30s), both CoFIRE and the contrastive samples also record the use of clause-medial *like* across a wide range of age cohorts, which further corroborates D’Arcy’s (2017, pp. 14-15) assertion that, regardless of clausal positioning, *like* has spread cross-generationally. This evidences the realism inherent to the fictional portrayal (i.e. RQ 1a) of *like* use in CoFIRE. In addition, the CoFIRE data may also be indicative of the fact that this is not only a DPM that has spread cross-generally, but also evidences the predilection for medial position which is now, to a greater or lesser extent, a cross-generational feature in both IrE and British English (as demonstrated in the BNC2014 sample findings in Fig. 7.4. above).

Before analyzing the data with regard to gender in more detail, it is necessary to remind the reader that I will only be examining the 261 occurrences of clause-medial *like* produced by the characters in CoFIRE, which excludes Ross O’Carroll-Kelly’s narrative production. The quantitative analysis of the 261 occurrences summarized in Table 7.2.

below highlights a clear dominance of clause-medial *like* use by female speakers in the corpus. However, it should be noted that a more detailed analysis revealed that the characters of *Ross O'Carroll-Kelly* (RO'CK) and *Sorcha O'Carroll-Kelly* (SO'CK) in Paul Howard's books were responsible for producing 25% and 71% of occurrences in each gender respectively. Nevertheless, if their inputs are discarded (as indicated in the table by 'without RO'CK' and 'without SO'CK'), the remaining occurrences are still much more prominent among female characters than their counterparts. This supports the close association that exists between clause-medial *like* and Irish females already discussed by Murphy (2015) and Schweinberger (2015, p. 131). Furthermore, the analyses of the contrastive samples (Table 7.2.) also highlight this prominence among female speakers, which further underlines the level of realism (RQ 1a) present in the corpus regarding the indexation of gender through this particular DPM. The over-representation of clause-medial *like* in RO'CK's and SO'CK's dialogue will be discussed in more detail below.

CoFIrE		Contrastive Samples			
		LCIE		BNC2014	
Male	Female	Male	Female	Male	Female
92	169	8	17	12	24
<i>Only RO'CK</i>	69	<i>Only SO'CK</i>	120		
<i>Without RO'CK</i>	23	<i>Without SO'CK</i>	49		

Table 7.2. Raw gender distribution of medial *like* across CoFIrE and LCIE/BNC2014 samples.

With regard to social class, a quantitative analysis indicates (Fig. 7.5. below) the presence of clause-medial *like* across all strata in CoFIrE and the LCIE/BNC2014 samples. Despite this, there is a contrast between the fictional portrayal and the natural use. Indeed, clause-medial *like* is noticeably more prominent among class-E speakers in the samples (for a full description of the class distribution system used in BNC2014 and applied to both LCIE and CoFIrE in this study, see section (3.8.2.)). While this class, in the majority of the cases, comprises students (LCIE: 19; BNC2014:14), some clause-medial *like* occurrences are also produced by female homemakers (LCIE: 3), and retired or unemployed speakers (BNC2014: 2 and 1 respectively). Contrastively, the fiction corpus shows Class A speakers being the most common users of clause-medial *like*, producing 247 of the 261 total number of occurrences. It is again necessary to point out that *Ross*

O'Carroll-Kelly's and *Sorcha O'Carroll-Kelly's* inputs make up for 72.4% (i.e.189) of these. Nevertheless, if their inputs are removed, clause-medial *like* is still more prominently produced among Class A speakers (73 occurrences), most of whom are Southside Dubliners.

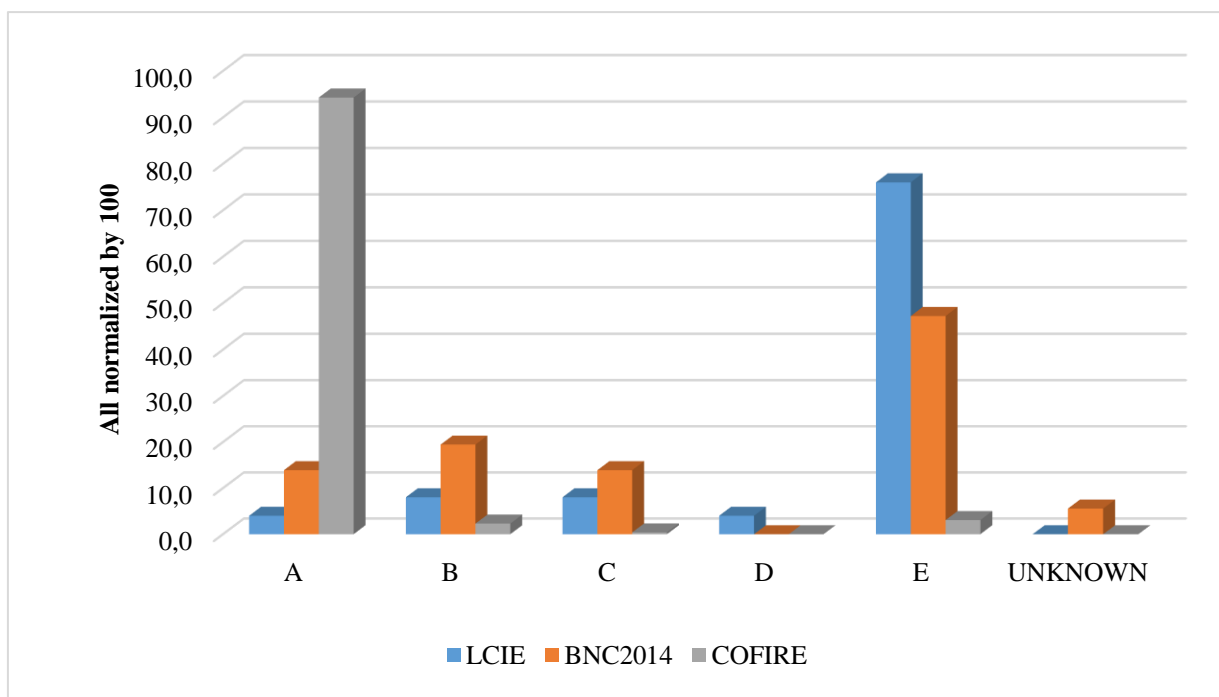


Figure 7.5. Class distribution of medial *like* normalized by x100 words across CoFiRE and LCIE/BNC2014 samples.

The correlation that seems to exist between speaker class and geographical location in CoFiRE (Fig. 7.6.) is worth exploring. To identify this, all occurrences were qualitatively distributed county. In the case of Dublin city, the occurrences were distributed into North/Southside, given the traditional associations between the North (traditionally working class and linguistically regionally-bound) and the South (traditionally affluent, linguistically regionally-unbound) already mentioned in (6.5.4.) and explained in detail in sections (8.2.) and (8.2.1.). Figure 7.6. below summarizes the findings.

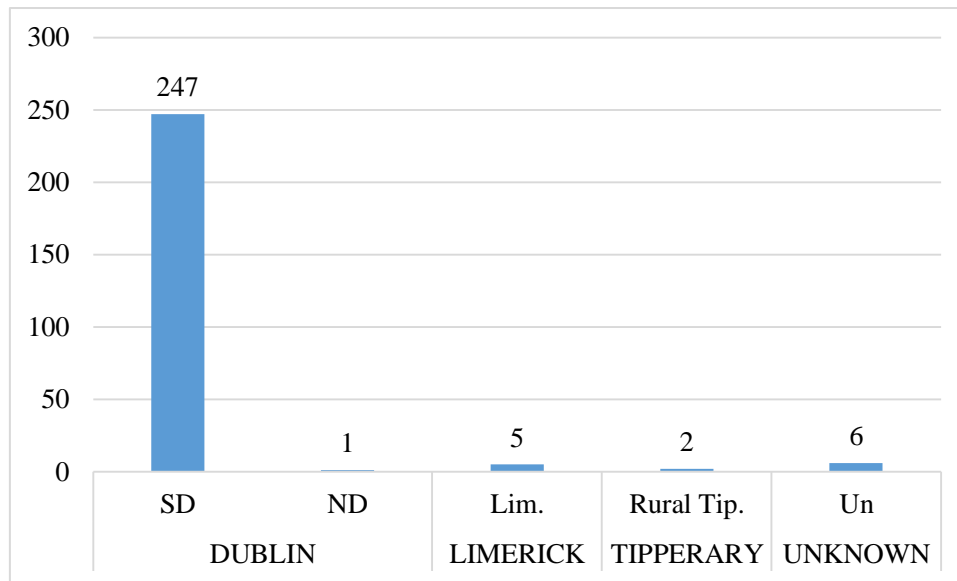


Figure 7.6. Raw regional distribution of medial *like* in CoFIrE.

The data in Fig. 7.6. shows that clause-medial *like* is present in Limerick, rural Tipperary, unspecified counties, and in Dublin city. In the case of the latter, there is a clear distinction between the single occurrence used in an unspecified part of Dublin, which was taken here as produced by a Northsider (see ND 1 in Fig. 7.6. above), and the majority of the occurrences which are produced by Southside Dubliners (247), all of whom are upper-class, Class A speakers. It must be mentioned that 91.57% (i.e. 239 occurrences) of the Southside production of clause-medial *like* appears in Howard’s books which, as stated through this thesis, present a satirical portrayal of the affluent South Dublin society that heavily contrasts with the North Dublin one; a contrast that is also linguistically rendered in the books. However, the use of clause-medial *like* is also present in other texts. Table 7.3. below better illustrates the presence of this DPM also in Barry’s and Ruane’s books, all of which are produced by young, (mostly female) upper-class, Southside, D4-accented⁷⁴ speakers.

⁷⁴ As mentioned in previous chapters, the “D4 accent” corresponds to what Hickey (2005) labels ‘non-local’ Dublin English. See section (8.2.1.) for a detailed explanation of this accent and its sociocultural associations in Dublin.

Author	Occurrences
Howard	239
Ruane	4
Barry	4

Table 7.3. Raw distribution of Southside-produced clause-medial *like* by author in CoFIrE.

The findings drawn from CoFIrE, therefore, would support Amador-Moreno’s (2015) suggestion that Howard’s (and in this case, Barry’s and Ruane’s) conscious authorial reproduction of clause-medial *like* in South Dublin could indicate the *supraregionalization* of this less-vernacular, more globalized feature. Furthermore, and as mentioned in previous sections, Amador-Moreno & Terrazas-Calero (2017, p. 257) argue that the shift in preference toward this ‘more global’ use, which seems to be particularly distinctive to South Dublin English, may also be motivated by language external factors such as mass media and speaker attitudes to certain linguistic items.

In this case, perhaps the speaker attitude may concern the strong sociolinguistic association *like* and clause-medial *like* in particular have regarding speaker identity in other varieties like American English, which are explored in detail in the next section.

7.5.1. Clause-medial *like* identarian associations in American English extrapolated to Irish English?

As discussed in (7.1.), *like*, especially clause-medial *like*, is often perceived as being a distinctive items in the California Valley Girl linguistic repertoire (Dailey-O’Cain, 2000, p. 70, Tagliamonte, 2016, and D’Arcy, 2017) and in the repertoire of its male counterpart, the ‘Jock’ or ‘Surfer Boy’ (Beltrama, 2016). The Valley Girl stereotype spread through pop culture music and movies, and typecasts young females from the San Fernando Valley, Southern California, triggering negative attributes such as ‘air-headed’, ‘vain’, ‘dizzy’. A search for ‘Valley Girl’ on *UrbanDictionary* for popular perceptions often retrieves user definitions such as “white and rich, fasion [sic] icons, ditzy, stupid, blonde, or beach going freaks” (grace1129, 2006) or “girly girls, poppers, and rich spoiled little brats” (Dinah, 2004). Its male counterparts, the ‘Jock’ or ‘Surfer Boy’, do not fare any better. The former retrieves popular perceptions that connect it with young, popular athletes, who normally play American football and/or other types of competitive sports,

yet it also triggers very negative attributes (see definition below) which prompt notions such as ‘unintelligent’, ‘womanizing’, and ‘arrogant bullies’:

A big athlete with a small ass brain. Usually an arrogant asshole, making everyone feel like nothing. [...] Seemingly popular in high school, but goes nowhere after graduation. Ends up either working fast food and/or drunk for a living. The epitome of loser. (Sweetemotion09, 2007)

In contrast, a search for ‘Surfer Boy’ on *UrbanDictionary* retrieves more positive adjectives, such as “extremely good looking”, “awesome hair”, can “easily seduce both genders”, and “can surf” very well (Aybaybay, 2011).

Linguistically, Kim (2019) describes the Surfer accent as the Valley Girl counterpart, and comments on the fact that it is also “tightly associated with Southern California[n] culture”. Features shared by both linguistic stereotypes include the use of *like*, Non-lexical *total(ly)* and New Intensifying *So* (see chapter 8 for analyses of both), and *high-rising terminal*⁷⁵, as well as ‘vocal fry’⁷⁶, and plenty of slang terms in the case of the Surfer (e.g. *dude*). Users also address the Valley Girl repertoire, noticing the “excess[ive] use of words such as “like, duh, tubular, rad, awesome, totally, and oh my god” (grace1129 2006), with others like Hunni (2003) commenting on their predilection for “insert[ing] the word “like” wherever she possibly can”. Hunni illustrates with the following example: “*like*⁷⁷ OMG can you believe that he, *like*, did that!? OMG its [sic], *like*, so, *like*, stupid, *like*, really, *like*, stupid! *like*, yeah”. Notice the excessive amount of occurrences of medial *like* this user displays.

The disproportionate use of clause-medial *like* attributed to the Valley Girl seems to correspond with its ample use by upper-class, Southside Dubliners in CoFIRE, particularly in Howard’s books. It is possible, therefore, that Howard, Barry and Ruane play on the already perceived linguistic stereotype associated with clause-medial *like* in American English (i.e. wealthy, air-headed, fashionable). By infusing their Southside

⁷⁵ Intonational deviation whereby statements are posed with as *interrogatives*?

⁷⁶ *Vocal Fry*, also known as ‘creaky voice’, refers to a quality of the voice by which the speaker uses “lower frequency vocal-tract resonances” (Lieberman, 2013). While frequently associated with (young) women, this is a rather polarizing phenomenon which is often very negatively perceived by other speakers. For example, actress Lake Bell very critically labels it as the “sexy baby vocal virus” and a “plague, where very smart women have taken on this affectation that evokes submission and sexual titillation to the male species” (Bell, 2013). See the following video (<https://www.youtube.com/watch?v=mIBg-w6TNLE> , minutes 2:36-2:59) featuring UK actress Emilia Clarke impersonating “Cali from the Valley”; a young Southern California fictional female through which she represent the Valley Girl repertoire. The video includes vocal fry in cases like “this...She’s, like, this whole, like, *situation*. It was *amazing!*”, “...a 700\$ air *conditioning unit*”, He was *freaking out*. It was *awesome!*”.

⁷⁷ My emphasis.

characters with an over-representation of clause-medial *like*, as well as with other linguistic items distinctive to the Valley Girl repertoire, especially in Howard's case (see chapter (8) for detailed discussions of other globalized, 'Valley-Girl'-perceived features that have been incorporated into the D4 repertoire), the authors may be 1) stylizing their Southside Dubliner cast as linguistically cosmopolitan and 'trendy', which differentiates them from the working/middle-class Northsiders who prefer more regionally-bound features (i.e. RQ 3 regarding the type of modern Irishness that appears to be indexed through the use of pragmatic features in the corpus). In addition, I believe this also enables them to 2) extrapolate the *Valley Girl/ Jock-Surfer Dude* stereotypes to the Irish context through the linguistic stylization of Irish equivalents in the form of what I have come to label the *Southside Socialite*⁷⁸ and its male counterpart, the *Rugby Jock* or *D4 Head* personae. A search for *D4 Head* on *Urban Dictionary* retrieves the following definition:

"Posh dublin southsiders [sic] who have developed the "d4 acsent"[sic], a d4 head doesnt [sic] have to be from d4, just has to sound posh, wear designer clothes, go to a private school, and live on the southside-all things along those lines, the guys tend to play rugby. the girls sometimes play hockey." (Aegirly, 2006)

As illustrated above, the linguistic construction of these personae in the books (particularly in Howard's novels) would trigger a series of *perceived* sociocultural 'must-haves' in the mind of the listener/reader which also correspond with the *Valley Girl-Jock/Surfer Dude* ones in the USA. These correspondences, which are summarized in Table 7.4., include: 1) *socioeconomic affluence*, evidenced by the use of designer clothing and attendance to private schools, 2) *association with a specific geographic location* (i.e. Southside Dublin). 3) *Sports as a marker of in-group* (Southside) *identity* with males favoring rugby and females preferring hockey, and 4) the use of *linguistically distinctive repertoires* (i.e. D4 accent).

⁷⁸ *Southside Socialite* is a term I have created as I have not encountered any existent term for a stereotypical Southside Dublin female neither in academic literature, nor in non-academic sources.

	Moniker	Geographic location	Socioeconomic background	Sports
Female	Valley Girl	SoCal	“rich spoiled little brats” (Dinah 2004); “fashion icon” (grace1129 2006)	‘Cheerleaderish’ (Lii, 2005)
	Southside Socialite	SD	Affluent, “posh”; designer clothes; private schools	Hockey
Male	Rugby Jock/D4 Head	SD	Same as above + Unintelligent athlete; arrogant; spoiled, unemployed ⁷⁹	Rugby
	Jock ⁸⁰ / Surfer Dude	(Surfer) SoCal	Jock: unintelligent athlete; arrogant; unemployed “loser” (Sweetemotion09)	American football/ Surfing

Table 7.4. Corresponding attributes between the American/Irish English stereotypes.

The elevated number of clause-medial *like* occurrences in CoFIrE produced by young Southside Dublin (particularly female) characters who are mostly in their 20-30s could answer RQ 3 (see (1.2.)), potentially suggesting the adoption of a more globalized item into ‘non-local’ Dublin English which, along with the incorporation of other global features (see chapter (8)) for detailed discussions of other incorporations into the D4 repertoire) would ensure the differentiation of their linguistic identity from the more regionally-bound, less prestigious Northsiders (see also (6.5.4.) for distinctive North/Southside Dublin English distinctive orthographic renderings of intensifying *fucking* and their socioeconomic associations). Finally, the fact that clause-medial *like* appears across a range of counties and social class ranks in the corpus, as mentioned before, may also indicate its status as a feature of supraregional IrE.

⁷⁹ These are the characteristics of a *D4-Head* as portrayed and embodied in the parodical character of *Ross O’Carroll-Kelly* in Howard’s books, which has become the stereotypical representation, both linguistically and culturally, of the *D4-Head/Southside Rugby Jock*.

⁸⁰ The *jock* is not a regionally-distinctive stereotype but rather a stereotype regarding high school social groupings.

7.6. Pragmatic functions of clause-medial *like* in CoFIrE and contrastive samples

To examine the pragmatic uses of clause-medial *like*, and thus better answer the RQs of the thesis, a search for its specific tag (i.e. <ML>) was conducted to retrieve occurrences which were manually classified for pragmatic uses. These were then qualitatively examined so as to arrive at functional categorizations which are summarized in Table 7.5. below.

USE		FRAGMENTS
<i>Unclarity</i>	<i>Focuser</i>	(7.12.) ‘We snack hard and we just, <i>like</i> , <i>sway with the kitchen vibe?</i> ’ (KBDLI)
	<i>explanation</i>	(7.13.) ‘And you must be S...S...S...Surrogate.’ Sorcha goes, ‘Yes, but it’s actually pronounced <i>Sorcha?</i> [...] As in, <i>like</i> , <i>Sorcha?</i> ’ (PHDA)
	<i>exemplifying</i>	(7.14.) ‘[...] try and act more like a boy – that’s all I’m saying. Boys aren’t interested in, <i>like</i> , <i>dresses and what their mothers are wearing.</i> They’re interested in shit like rugby.’(PHDA)
	<i>Unfamiliarity with term</i>	(7.15.) ‘She became, <i>like</i> , <i>a fat whisperer or something.</i> ’ (PHKU)
	<i>+time</i>	(7.16.) ‘The gestation period is only, <i>like</i> , <i>twenty-one</i> days.’ (PHKU)
	<i>Approximator+quantifiable units</i>	(7.17.) ‘He’d hand over the kid and I’d keep him for <i>like</i> , <i>a night or some shit.</i> ’ (DRYS)
	<i>+numbers</i>	(7.18.) ‘Yeah, he is, but you haven’t met him, Amy. He’s... <i>like</i> , <i>well when he was...</i> ’(DRTRM)
	<i>Hedge</i>	

Table 7.5. Qualitative classification of all occurrences of clause-medial *like* in CoFIrE.

As illustrated in Table 7.5., clause-medial *like* in CoFIrE can function as a 1) *hedge* in cases like 7.18. above where the speaker hedges and hesitates in her description of the man to the point where she does not even finish the sentence. It may also be used as a term that denotes 2) *lack of clarity with the part of the sentence coming after the discourse pragmatic marker*. In these cases, clause-medial *like* can function as a (2a) *marker of unfamiliarity with the term* being used. For example, in 7.15., the speaker is not familiar with the concept of what a *fat whisperer* may be, and she further reinforces that by following the term with “or something”. Unclear clause-medial *like* may also be used as (2b) *approximator with quantifiable units* (e.g. *time and/or numbers* as in 7.16. and 7.17.), which suggest a (un)conscious effort on the part of the speaker not to commit to the quantity being expressed. Finally, clause-medial *like* may also be used as a 3) *focuser device* (i.e. 7.12. through 7.14.) which, not only emphasizes what comes after the discourse pragmatic marker (DPM) and the speaker’s feelings, but also dramatizes the content of the sentence. For example, the male speaker in 7.12. places the focus of the listener’s attention on the action that follows the DPM (*sway with the kitchen vibe*), and dramatizes the action even further by adding *high rising terminal* intonation at the very end of the sentence.

Furthermore, and as mentioned by Amador-Moreno (2010, pp. 121-22), focuser clause-medial *like* may also be used in CoFIrE to *introduce explanations and/or to exemplify*. For instance, in 7.13., Sorcha O’Carroll-Kelly explains to Kennet, a Northside Dubliner in Paul Howard’s books, how to pronounce her name, as he often mispronounces it. Notice how the focuser medial *like* not only is used to *explain* the pronunciation, but this is further reinforced through the italicization of the name itself and the use of *high rising terminal* at the end. A case of an exemplifying focuser would include 7.14. where Ross O’Carroll-Kelly explains to a child who is interested in women’s clothing and whom he considers to be “a sissy” what he *should* like as a boy. The *exemplifier focuser* medial *like* in this case precedes the items that a boy “should not” like according to Ross. I believe that the fact that these are all uses that have already been identified in the literature (e.g. Kallen 2006, Amador-Moreno, 2010, inter alia in section (7.1.)) lends validity to the fictional representation of natural use comprised in CoFIrE, thus answering RQ 1a.

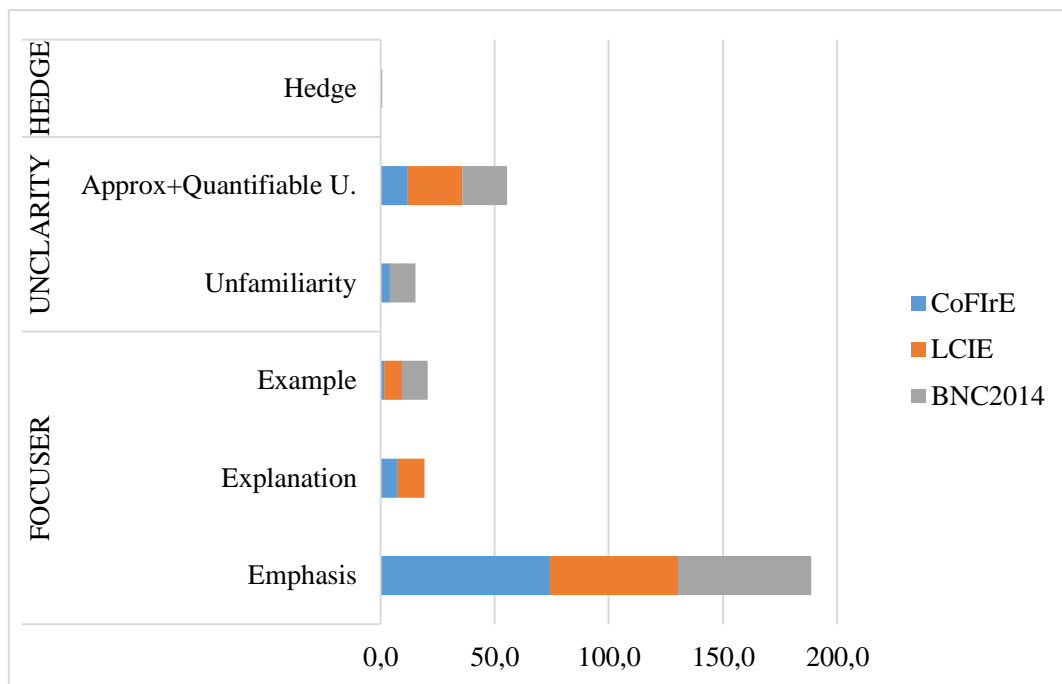


Figure 7.7. Contrastive pragmatic function distribution of medial *like* across CoFIrE and LCIE/BNC2014 samples normalized by 100 words.

After all occurrences of clause-medial *like* were classified for pragmatic functions CoFIrE and in the 100-item LCIE/BNC2014 contrastive samples, they were quantitatively compared. The findings from this analysis, which are summarized in Fig. 7.7., show that the *focusing* use is the most dominant function in CoFIrE, followed by *marking lack of clarity*, with *hedging* being a significantly small occurrence. The fact that these findings are replicated in the LCIE/BNC2014 samples also gives credence (RQ 1a) to the representation of clause-medial *like* functional use offered in CoFIrE and its level of authenticity regarding real-life use.

The pragmatic functions and (sub)functions (i.e. emotion transmission) of clause-medial *like* were examined next. Thus, all occurrences were classified depending on whether the use of clause-medial *like* conveyed *positive*, *negative*, or *neutral* emotions (i.e. transmission of a simple emphasis). The data, summarized in Fig. 7.8., suggests that this particular DPM is more prominently used in CoFIrE to transmit *neutral* evaluations (i.e. conveyance of simple emphasis). This corresponds with the data from the LCIE/BNC2014 samples, particularly with LCIE, which further evidences the high level of realism (RQ 1a) present in the representation of use in the fictional corpus.

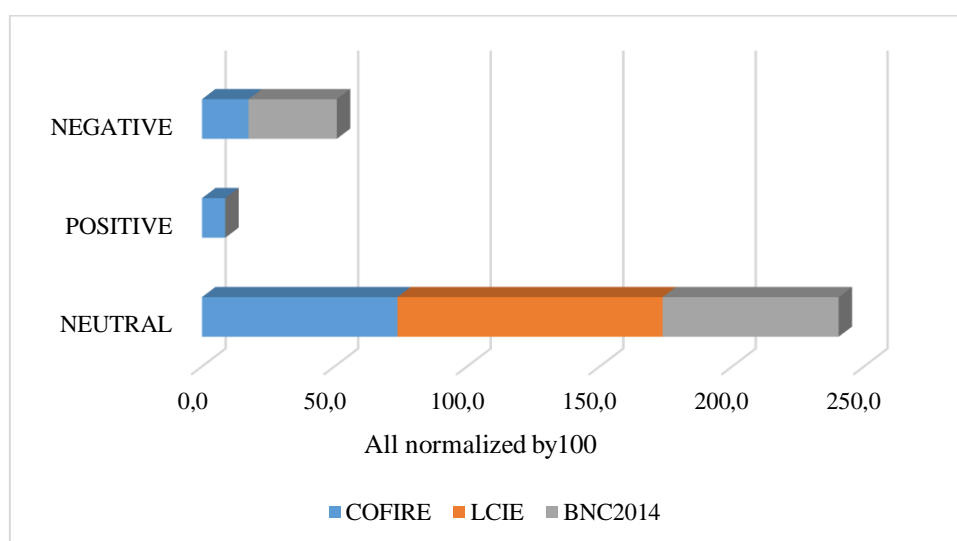


Figure 7.8. Pragmatic distribution of medial *like* across CoFIRE and LCIE/BNC2014 samples normalized to by 100 words.

Furthermore, the cases of *neutral* medial *like* can have two subfunctions, namely a) a simple conveyance of *emphasis to the content of the utterance* such as in 7.19., where the speaker emphasizes with medial *like* how much the *he* in the utterance recounted (i.e. up to the day he died).

(7.19.) He told everything about him *like* you know *up to the day he died* (LCIE)

(7.20.) ‘We need to take her to counselling. Family counselling.’ ‘As in, *like*, *all three of us?*’ (PHDA)

It is interesting to note that while the occurrences of *neutral* medial *like* contrastive LCIE/BNC2014 samples only carry emphasis, the ones in CoFIRE may also be used to b) *seek clarification*. For example, the speaker in 7.20. above is asking his wife whether “we” means that they should *both* go to family counseling with their daughter, which he does by emphasizing with the medial *like* the “all three of us” part.

The pragmatic CoFIRE medial *like* repertoire appears to be broader than LCIE’s in that it features *negative* evaluations as the second, most prominent pragmatic function, which mirrors the British sample’s data. These may include cases like the *exasperation*

conveyed in 7.21., or the clear *sarcasm* shown in 7.22. which is reinforced by the italicization of intensifier *really*.

- (7.21.) I don't like it when people go around in their skinny jeans just hanging off their arse and it it [sic] just looks weird and then it's just like legs like matchsticks (BNC2014)
- (7.22.) George sacked the cleaner.[...] So I'm stuck doing it most of the time anyway. For forty euros a week less than I used to get. But aren't I lucky to have a job? Ya, like, I'm really lucky. (DRY)

CoFIrE also documents the transmission of *positive* evaluations/emotions which are absent from the *LCIE/BNC2014* samples, and which include *excitement*, as in 7.23. above, *admiration*, or *compliments*, such as the praising description of a nun present in 7.24.

- (7.23.) 'Oh my God, I would love it if Honor discovered what an amazing, amazing thing music can, like, be?' (PHDA)
- (7.24.) She was an – oh my God – amazing person with, like, really strong beliefs. She was a big believer in, like, social justice. (PHDA)

Overall, the fact that the CoFIrE pragmatic repertoire is seemingly larger than the *LCIE/BNC2014* samples could be 1) due to the size of the samples. It could also be due to the fact that 2) CoFIrE may document 'newer' pragmatic uses of medial *like* which had not been adopted yet at the time of compilation of *LCIE*. This would, therefore, buttress the fiction corpus as a valid source of linguistic evidence (RQ 1a) in the context of Ireland.

7.6.1. The sociopragmatics of focuser medial *like* in CoFIrE

Given the elevated number of occurrences of focuser medial *like* in CoFIrE, and to better answer RQs 2 and 3 (see (1.2.)), a detailed qualitative analysis of this DPM was, then, undertaken to examine its sociopragmatic functions. The first thing to be noted is the occurrence of focuser medial *like* in examples like 7.25. below. Here, the DPM collocates

with New Intensifying *So*, that is, with intensifying SO in traditionally non-standard contexts, such as, in this case, noun phrases (see Amador-Moreno & Terrazas-Calero, 2017, and chapter (8) for more on this intensifier).

(7.25.) ‘It’d be great to have Ro here. Jesus, imagine the crack?’ ‘It’s, *like, so an amazing idea.*’ (PHDA)

As illustrated above, the speaker in 7.25. uses focuser medial *like* to emphasize what a great idea she believes her interlocutor had, which is further reinforced by the extra layer of intensification that is applied to New Intensifying *So* through its italicization. Thus, focuser medial *like* strengthens the intensifying value of SO. The fact that, as will be detailed in chapter (8), this particular feature is often perceived to be an American English intensifier which is especially prominent among young females (Zwicky, 2011; Irwin, 2014) leads me to consider this particular collocation as one which the Southside Dubliners in CoFIrE may use as a linguistic tool to infuse their identity with more cosmopolitanism (RQ 3), which directly opposes them to the Northsiders.

Similarly, the qualitative analysis also retrieved a variety of instances where focuser ML co-occurs with *high rising terminal*⁸¹. As illustrated in 7.26., these are cases where the speakers pose a statement with rising intonation as though asking a question. For example, the speaker in 7.26. clarifies to her interlocutor what that syndrome means. This is preceded by an explanatory focuser medial *like* whose emphasis is doubled by the rising intonation which also serves a means by which the speaker checks if the interlocutor understands the explanation. Quantitative analyses revealed that 25% (i.e. 55 of 220) of all focuser medial *like* co-occur with *high rising terminal*, whereas the analysis of the contrastive samples indicated the absence of this pattern from either.

(7.26.) ‘Jenny has a form of Münchhausen’s Syndrom, Ross.’ ‘Okay, what’s that?’ ‘It’s a psychiatric illness. It’s, *like, a compulsive need for attention?*’ (PHKU)

⁸¹ As already discussed in previous chapters, *high rising terminal* is an intonational deviation whereby statements are posed as *questions*?

The focuser medial *like+high rising terminal* pattern was, subsequently, classified for pragmatic functions and emotive conveyance following the classification used across this thesis since, to my knowledge, there is no study that has conducted a detailed examination of the pragmatic functions of *high rising terminal*. The findings illustrated in Fig. 7.9. provide a variety of functions which can be divided into *checking/seeking feedback from the interlocutor*, *transmitting emphasis*, and *softening the force of an utterance*.

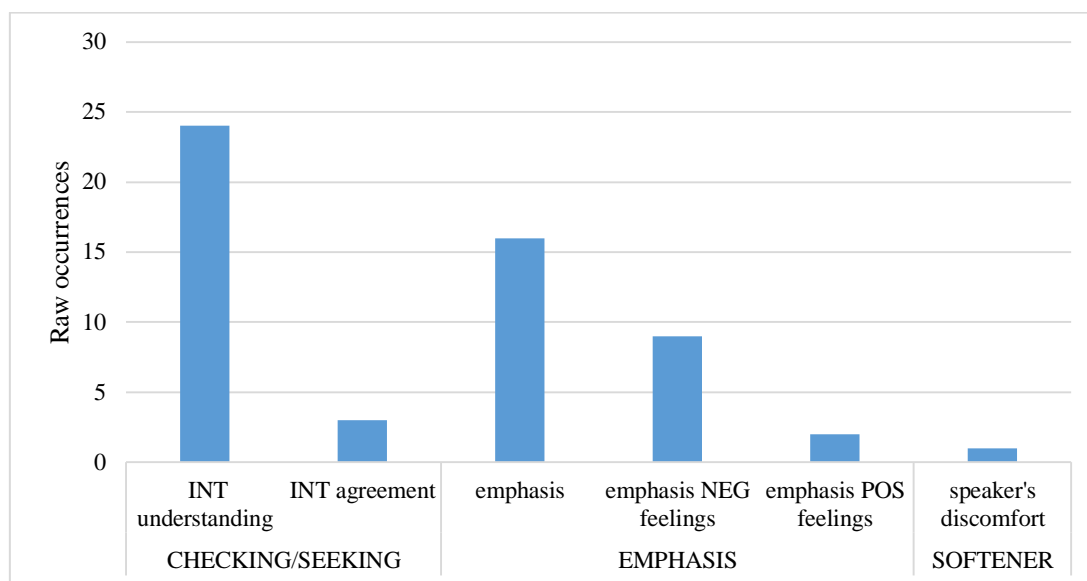


Figure 7.9. Pragmatic functionality of focuser medial *like+high rising terminal* pattern in CoFIrE.

Quantitatively, the most prominent function is the use of focuser medial *like+high rising terminal* to a) *check whether the interlocutor understands the utterance*, as exemplified in 7.26 above. However, the speaker may also use the pattern to b) *check/seek the interlocutor's agreement*. For example, the speaker in 7.27. reports to her husband what the composer of the theme song for their vow renewal ceremony said (the theme should 'have more personality'), which she qualifies by further explaining the meaning (i.e. it should be individualistic). Notice the extra emphasis placed on the qualified term which is italicized. The use of focuser medial *like+high rising terminal* pattern here, therefore, checks whether her husband understands the content of the conversation (i.e. their theme song should be something personal, about them as a couple, and not a romantic song by a popular artist). In cases where this pattern functions this way, the *high rising terminal* could be substituted by DPM *you know* or *do you know what I mean?*.

(7.27.) ‘I was talking to Dechtire [...] and she thinks that because it’s a renewal rather than an actual wedding, we should go for something with more personality – as in, *like, more individualistic?*’ (PHDA)

The focuser medial *like+high rising terminal* pattern is also used in CoFIrE to c) *emphasize the content of the utterance*, as in 7.28. where Ross O’Carroll-Kelly apologizes to his father and step-mother for arriving unannounced. He politely refuses her offer of breakfast by emphasizing through the focuser medial *like+high rising terminal* pattern how short his visit is. Notice the double emphasis through the italicization of *flying*.

(7.28.) ‘[...] I’ll fix you some breakfast--if you can promise to keep it down, that is!’ [...] ‘Yeah, no, I won’t thanks. I’m only here on, *like, a flying visit?*’ (PHDA)

Emphasis can also be applied to d) *positive emotions* like the *excitement* in 7.29., but mostly to transmit e) *negative emotions* such as the speaker’s *outrage* in 7.30. Finally, the focuser medial *like+high rising terminal* pattern was also found to function as a *downtoner*, in cases like 7.31. where it softens the speaker’s discomfort at being asked to talk about a cartoon character he created. Notice how his discomfort is further evidenced by his use of hedges (*er, yeah, no*) and how he downplays the importance of his action by reinforcing that it is only a *character* via italicization of the word.

(7.29.) ‘Oh my God, I would love it if Honor discovered what an amazing, amazing thing music can, *like, be?*’ (PHDA)

(7.30.) ‘There’s, *like*, no way she’s living here without contributing, Ross.’ (PHKU)

(7.31.) ‘Rob the Builder![...]Tell my friend about Rob the Builder’ Oissin’s *like*, ‘Er, yeah, no, he’s, *like, a character I invented?*’ (PHKU)

From a sociolinguistics perspective, and as illustrated in Table 7.6., the focuser medial *like+high rising terminal* pattern is only ever produced by upper-class, Southside Dubliners, the majority of whom are young, female speakers.

Age			Gender	
<i>(Pre)teen</i>	20-30	40+	<i>M</i>	<i>F</i>
3	51	1	10	45

Table 7.6. Raw distribution of the focuser medial *like+high rising terminal* pattern by age/gender in CoFIrE

Given that *high rising terminal* is popularly perceived as having originated in the US (despite its origin being unknown), and being typically associated with trendy young/teen California ‘Valley’ girls (Warren, 2016, p. 82), I argue that the correspondences between these perceptions in the context of American English and the sociolinguistic indexicality this pattern appears to suggest in CoFIrE, as well as its absence from the contrastive *LCIE/BNC2014* samples, may be indicative of the fact that it functions as a marker of urban sophistication (i.e. RQ 3) which may linguistically stylize the voice of the *Southside Socialite-Rugby Jock/D4 Head* personae.

7.7. Concluding remarks

To conclude, the analysis of discourse pragmatic marker *like* in CoFIrE evidences a prominence of the more globalized, clause-medial positioning, which deviates from the traditional preference for clause marginal positions in IrE. Contrastively, the *LCIE* sample data continues to favor marginal (i.e. final) placement, although clause-medial is the second favored position. The contrast between the fiction and spoken data could mirror the shift toward clause-medial *like* already signaled in previous research, which might have already begun at the time of the compilation of *LCIE*, but had become established as a feature of *supraregional* IrE in later years, when the CoFIrE texts were published. This shift, which is evidenced in academic research, would emphasize the level of realism present in CoFIrE (i.e. RQ 1a), underscoring the fiction corpus as a good source that documents linguistic developments (at least as perceived by the authors) in the context of Ireland.

The quantitative analysis of clause-medial *like* pragmatic uses revealed functions already identified in previous research, and highlighted focuser clause-medial *like* as the

most prominent one. Furthermore, an examination of the conveyance of evaluations indicated that CoFIrE mainly used medial *like* to transmit *neutral* meanings, which mirrors the findings from the contrastive samples. I believe this also lends validity to the fictional representation of this DPM offered by the corpus (RQ 1a). Nevertheless, the range of pragmatic conveyance in CoFIrE was larger than the one in the samples, containing *positive* and *negative* evaluations as well. While their absence from the samples might be due to their smaller size, their appearance in CoFIrE could also buttress the value of the corpus as evidence for linguistic developments or, in this case, for the acquisition of ‘new’ pragmatic uses, at least as represented in the texts.

From a sociolinguistic perspective (i.e. RQs 2 and 3), the CoFIrE data also corroborates previous findings in real spoken discourse, indicating a frequency of use among the 20-30 cohort, particularly among female speakers, which is also mirrored in the *LCIE/BNC2014* samples. This also evidences the highly realistic portrayal of use offered in the fiction texts (i.e. RQ 1a). Furthermore, CoFIrE also shows the presence of clause-medial *like* across all cohorts, which also mirrors D’Arcy’s (2017, pp. 14-15) assertion that this marker is cross-generational. I believe the similarities between the fiction and spoken samples, as well as with D’Arcy’s assertion, lends credence to the validity of the fictional representation of clause-medial *like* use in CoFIrE. In addition, the examination of social class-indexation retrieved contrastive results. On the one hand, the samples indicate a prominence of clause-medial *like* use among class E speakers (mostly students). However, CoFIrE highlights its recurrence among Class A speakers, most of whom are Southside Dubliners, which could be due to the over-representation of these characters in the corpus, and in Paul Howard’s books in particular. Class and region were also found to be connected, for although clause-medial *like* appears in different counties in the fiction corpus, it is most salient in Southside Dublin. This is in line with Amador-Moreno’s (2015, p. 376) assertion that the use of medial *like* functions as a marker of the *supraregionalization* of this feature, which is favored by the more cosmopolitan, affluent, and linguistically global Southsiders (i.e. RQ 3).

Finally, I believe that, to answer RQ 3, the shift toward clause-medial *like*, particularly favored by Southside Dubliners in CoFIrE, may be due to the association of this marker with the stereotypes of the *Valley Girl - Jock-Surfer Boy* in the USA, both of whom have very distinctive linguistic repertoires which include clause-medial *like* as a distinctive item, along with New Intensifying *SO* and *high rising terminal*. I posit that

these stereotypes can be extrapolated to the context of Ireland. Indeed, I argue that by drawing on the specific sociocultural values attached to the stereotypes (i.e. affluence, sports, etc.) and to their linguistically distinctive (at least as popularly perceived) items, particularly to the use of clause-medial *like*, medial *like*+New Intensifying SO, and focuser medial *like*+*high rising terminal*, the CoFIrE authors, and Howard in particular, extrapolate those stereotypes to the context of Ireland, and Dublin especially. In doing so, they use clause-medial *like* to create a stylization of contemporary Southside Dublin identity (i.e. RQs 2 and 3) for their intra-cultural audience which could be summarized in the dramatis personae of what I have labeled the *Southside Socialite-Rugby Jock/D4 Head*. Finally, in recording these two stereotypes linguistically on paper, the authors guide the readers into drawing their conclusions about the characters' identities while simultaneously ensuring that they 'perceive' the sociocultural differences (i.e. wealthy, spoilt, cosmopolitan, etc.), which separate the Southside from the Northside cast of characters. Chapter (8) will explore in detail the use of New Intensifying *So* and Non-lexical intensifier *Totally*, which are also part of the Southside repertoire, as represented in CoROCK.

8. ‘New’ *over-the-top* intensifiers in the *Ross O’Carroll-Kelly* corpus

As detailed in Chapters 3 and 4, in the process of curating texts for CoFIrE, it became clear that one of the authors, Paul Howard, was an outlier within the group of authors as his books contributed the largest amounts of items to the corpus. As discussed, it was decided a case study using a separate corpus was merited. For details on the *Corpus of Ross O’Carroll Kelly* (hereafter CoROCK), see section (3.9.). As shall be explained in section (8.1.), preliminary analyses regarding frequency of production retrieved New Intensifying *So* and Non-lexical intensifier *Totally* as the most frequently reproduced *over-the-top* intensifiers. Thus, this chapter explores in detail these two pragmatic items. Section (8.1.) outlines why these two intensifiers were selected for study while sections (8.2.) and (8.2.1.) explore the stylization of North/Southside Dublin stereotypes as well as the connection that exists between language and style in the *Ross O’Carroll-Kelly* series respectively. An overview of intensifiers and *over-the-top* intensification, the category New Intensifying *So* and Non-lexical intensifier *Totally* fall into, is offered in section (8.3.), while sections (8.3.1.) and (8.3.2.) will focus on surveying the literature on the use of these two intensifiers. The methodology of analysis is outlined in section (8.4.), with (8.5.) through (8.6.3.) exploring the use, form, pragmatic functionality, and rendition of *over-the-top* emphasis and prosody evidenced by the use of New Intensifying *So* and Non-lexical intensifier *Totally* in the different CoROCK subcorpora. Finally, their potential identity indexicality value is addressed in section (8.7), with concluding remarks being offered in section (8.8.).

8.1. The selection of New Intensifying *So* and Non-lexical intensifier *Totally*

As already discussed throughout this thesis, the findings drawn during the annotation and analytical stages of the *Corpus of Contemporary Fictionalized Irish English*, CoFIrE, already highlighted Paul Howard’s *Ross O’Carroll-Kelly* books as the texts which contributed the largest amounts of pragmatic linguistic items to CoFIrE (see sections (4.1.)-(4.2.) for a detailed account of the production of feature occurrences per book). Given the outlier nature of Howard’s texts in the fiction database, and to better address the thesis’ research questions (outlined in section (1.2.)), it was decided that this thesis would also compile the first *Ross O’Carroll-Kelly Corpus* (see section (3.9.) for a detailed

description of this corpus, its compilation, and the subcorpora it comprises, among other factors). The compilation of the *Ross O'Carroll-Kelly Corpus* (hereafter CoROCK) would allow for a detailed analysis of two pragmatic items which have not traditionally been associated with IrE, but with American English, and often with the Valley Girl repertoire, as discussed in chapter (7) in relation to DPM *like*. These items are New Intensifying *So* and Non-lexical intensifier *Totally*, both of which are exemplified in 8.1. and 8.2.

(8.1.) Oh my God, Erika **SO** freaked me out (Howard, 2007)

(8.2.) I know. It's, like, **totally** ... duuuhhh!' (Howard, 2003)

The reason New Intensifying *So* and Non-lexical intensifier *Totally* (see sections (8.3.1..) through (8.3.2.)) for a review of the literature on their use) were selected for in-depth study is not only because they are more frequently associated with the American English repertoire but because they arose as salient in case studies independent to this thesis, which drew on preliminary findings from CoFIrE. On the one hand, a *word frequency list* (illustrated in Fig. 8.1. below) was run on the CoROCK-SO subcorpus (see section (3.9.) for a detailed explanation of the CoROCK subcorpora) using *Wordsmith Tools* (Scott, 2012) which highlighted SO (across all its functions, e.g. conjunction, intensifier, etc.) as a prominent item in the subcorpus, ranking in position 45.

N	Word	Freq.	%	Texts	%
26	THEY	5.598	0,49	16	100,00
27	HIM	5.484	0,48	16	100,00
28	BUT	5.417	0,48	16	100,00
29	ALL	5.344	0,47	16	100,00
30	UP	5.314	0,47	16	100,00
31	BE	5.050	0,44	16	100,00
32	OUT	4.937	0,44	16	100,00
33	HAD	4.849	0,43	16	100,00
34	WE	4.839	0,43	16	100,00
35	NOT	4.523	0,40	16	100,00
36	WHAT	4.482	0,39	16	100,00
37	HAVE	4.442	0,39	16	100,00
38	SAID	4.266	0,38	16	100,00
39	FROM	3.892	0,34	16	100,00
40	ONE	3.623	0,32	16	100,00
41	THIS	3.581	0,32	16	100,00
42	WHEN	3.432	0,30	16	100,00
43	THEM	3.428	0,30	16	100,00
44	YOUR	3.404	0,30	16	100,00
45	SO	3.295	0,29	16	100,00
46	NO	3.262	0,29	16	100,00
47	WERE	3.259	0,29	16	100,00
48	ABOUT	3.217	0,28	16	100,00
49	OR	3.216	0,28	16	100,00
50	ARE	3.199	0,28	16	100,00

Figure 8.1. Screen capture of frequency list run on CoROCK-SO with *Wordsmith Tools* (Scott, 2012).

A detailed qualitative analysis would subsequently reveal that its intensifying function was the most prominent in the corpus, while even further qualitative examination determined the appearance of a ‘newer’ intensifying function which Amador-Moreno & Terrazas-Calero (2017, 2022) label ‘New Intensifying *So*’ (this feature will be explained in detail in section (8.3.1.)). Having already identified New Intensifying *So* as one of the highest ranking traditionally non-IrE intensifiers in CoROCK-SO, another *frequency list* was run on this subcorpus which identified *Totally* as the next, most recurrently reproduced intensifier, placing comparatively lower on the list (i.e. position 286) than New Intensifying *So* (ranking as the 45th most frequent item). To undertake a more longitudinal look at the use of *Totally* given its low frequency in CoROCK-SO (see (8.3.1.) for an overview of the literature on this item), it was decided that this subcorpus would be expanded with more texts. This developed into a second subcorpus, i.e. CoROCK-T, with which the analysis would be conducted (see section (3.10) for a detailed explanation of the subcorpus, and Appendix 3 for a list of all the books comprised in each subcorpus).

The following sections will describe the *Ross O’Carroll-Kelly* series with regard to the stylization of North/South Dublin stereotypes (8.2.) and will provide a detailed

explanation of the linguistic style (8.2.1.) evidenced in the series so as to better understand how New Intensifying *So* and Non-lexical intensifier *Totally* (explained in detail in sections (8.3.1.) and (8.3.2.) may function in the series, and what their use could potentially index with regard to modern Irishness, especially to modern Dublin identity.

8.2. Stylizing North/South Dublin stereotypes in CoROCK

As already discussed in section (3.4.8.), the *Ross O'Carroll-Kelly* series is narrated in the first person by *Ross O'Carroll-Kelly* (hereafter RO'CK), who functions as a *homodiegetic/autodiegetic* narrator (Genette, 1980, pp. 244-245), who is both narrator and active participant character in his own story. In addition, it could be argued that he is also a *card* character. In Harvey's (1965, p. 58) terminology, *cards* are people in the fiction who are "larger than life" or "distinguished by some fiction-like idiosyncrasy". Culpeper (2001, p. 88) explains that *cards* are so perceived because "they are an exaggerated prototype of some social category", in the sense that they are "prototypical in some exaggerated way". I would argue that RO'CK is a prime example of a *card* character who is hyperbolically archetypal of the affluent Southside Dublin society that arose during the Celtic Tiger (see section (1.4.) for a description of this period), embodying, as O'Brien (2015, p. 64) posits, the "attitudes of a financial elite to those outside that elite world". RO'CK is an entitled young man, who lives off his parents' wealth and never works. He is self-centered and an unrepentant womanizer, who is continuously unfaithful to his wife. Furthermore, he is ignorant and prejudiced against anybody who is not from Southside Dublin, and particularly against Northside Dubliners, whom he describes in the books as "Dublin's poorer classes" (Howard, 2007b, p. 20) among many other pejorative terms. All of these traits combined illustrate well his sense of grandeur and classist attitude.

The scathing satire that is encapsulated in the character of RO'CK himself would be a product of what Howard (2019, my transcript from communication during interview) describes as his early antagonistic feelings as a working-class (Northside) Dubliner against the Southside upper class. This animosity arose as a consequence of witnessing

and having encounters like the one Howard recounts when trying to get onto the last free carriage of the DART⁸²:

I go to get on, and this guy stops me. Hand on my chest, he says, ‘Sorry, dude. This is a Rock⁸³ carriage’. And [...] so I kind of took a step backwards and apologized to him, you know. The doors close, and, as the train is pulling away, he just did this to me through the doors *pointing finger gun gesture*. So this kid... This rugby...this kind of sense of entitlement. This obnoxious...I hated it. I really, really despised the classism. (Howard, 2019)

This particular encounter would also appear to be responsible for even the ‘Southside’ body language RO’CK is attributed, given the fact that the finger gun gesture became his signature move, appearing in book covers, illustrations, and even in a statue which was temporarily placed in different Dublin City sites⁸⁴. RO’CK’s (and the Southsiders’ by extrapolation) classism against Northsiders is also evidenced in the next fragment. In this scene, RO’CK and his wife, Sorcha, are attending their daughter’s—Honor—recital, which is taking place at a *public* school she attends after being expelled from her private, Southside school for her bullying behavior.

Me and Sorcha stand out among all the other parents, as you can imagine. Firstly, we’re the only ones not wearing either tracksuits or leather Members Only jackets. Secondly, they’re the only ones who keep turning to each other and smiling every time a kid on the stage says something in what I call a *real* Dublin accent. A knacker⁸⁵ accent, in other words. (Howard, 2014, p. 221)

The previous fragment is also useful in illustrating two indexes of contemporary Dubliner identity (which would address RQ 3 regarding the potential indexation of modern Irishness) which this series relies upon: *fashion* and *language* as distinctive of regional and socioeconomic status. For example, with regard to *fashion*, the Southside cast is always described as wearing expensive, designer clothes. Southside men in the books

⁸² Acronym for *Dublin Area Rapid Transit*; a suburban rail system that commutes the north, south, and coastal areas of Dublin City.

⁸³ ‘Rock’ is an abbreviation for *Blackrock*; a suburb in the affluent Southside Dublin area.

⁸⁴ Access the following article on *The Irish Times* (<https://www.irishtimes.com/life-and-style/should-ross-o-carroll-kelly-replace-tort-with-the-cort-1.1966128>) to see the *Ross O’Carroll-Kelly* statue.

⁸⁵ As already explained in section (3.7.), *Knacker* or *Skanger* is a derogatory IrE terms for lower-class “young, uncouth youth” (Dolan, 2006, p. 199) from a high criminal rate area often recognized for their outfit (e.g. casual sports) and poor education level.

wear a very specific type of ‘outfit’ which is associated with this affluent region. Upon recounting an anecdote about a group of Southside male youths who used to gather together to read the first RO’CK book at a bookshop, Howard (2019, my transcript from communication during interview) describes prototypical, Southside male fashion as including: “the chinos, the sailing jackets, the, you know, rugby jerseys with the collar up, the sailing shoes that had never seeing salt water ever”. Notice the use of the definite article (*the*) which makes it all the more evident that these items are *distinctive* to Southside, male fashion. The female cast, however, always wears designer clothing, with fashion brand names peppering all the books: “[Sorcha’s] wearing her Alexander Wang cigarette-leg trouser suit with a white shirt by Saint Laurent and Pigalle 100 pumps by Christian Louboutin” (Howard, 2013, p. 342). The excessive detail in the naming of the brands and the type of clothing the Southsiders wear is a clear index of their affluent status which contrasts heavily with the many scornful descriptions of Northsiders, whom RO’CK also derogatorily refers to as *skangers*⁸⁶. Take, for instance, the following description of his Northside, co-parent-in-law, Dordeen:

[...] in walks Shadden, followed by the famous Dordeen, wearing—there’s no real surprise here—ski pants, humongous Nikes, a tracksuit top, then underneath a T-shirt with the words Psycho Slut on it. ‘Weer is she?’ is her opening line. ‘Weer’s tat gowergeous little granthaughter of moyen?’ She’s like something out of *Mrs Brown’s Boys*⁸⁷. You’d nearly be listening out for the audience laughter.” (Howard, 2014, pp. 217-218)

Notice the contrast that is being drawn in both descriptions between Sorcha’s designer attire, which indexes not only affluence but class as well, and Dordeen’s ‘classless’ use of sports clothing and the ‘tasteless’ T-shirt message. RO’CK’s scorn is also evidenced in his side comment to the reader (‘there’s no real surprise here’), as though that type of outfit (and, therefore, classlessness) were to be expected of a Northsider. His disdain is further accentuated in his rendition of Dordeen’s Northside accent, which includes prototypical IrE phonetic features, such as the merging of the PRICE-CHOICE lexical sets, whereby words containing the diphthong [aɪ] can be realized as [ɔɪ], evidenced in

⁸⁶ *Skanger* is derogatory IrE terms for lower-class “young, uncouth youth” (Dolan, 2006, p. 199) from a high criminal rate area often recognized for their outfit (e.g. casual sports) and poor education level.

⁸⁷ *Mrs Brown’s Boys*: Irish-British sitcom revolving around Brendan O’Carroll’s fictional character of Agnes Brown, a working-class matriarch, and her family. The series is set in Finglas, a suburb in North Dublin.

moyen (mine). Other examples in the RO'CK books of this merger, which Hickey (2007, p. 317) identifies as a feature distinctive to *local* Dublin English (i.e. Northside Dublin), include, for instance, *loike* (like), or *roysh* (right). The following section will focus on exploring the second most recurrent index of regional and class identity in the RO'CK series: *language* representation.

8.2.1. Language and style in the *Ross O'Carroll-Kelly* series

While relatively under-researched in terms of their linguistic richness (with the exception of Amador-Moreno, 2012a, 2015, 2016; Terrazas-Calero, 2016a, 2016a for taboo language; Amador-Moreno & Terrazas-Calero, 2017, 2022 for intensifiers and discourse pragmatic markers; Lillo, 2007, and Gorman, 2011, 2013 for other elements of every-day Dublin English use), academics believe that one of the main trademark stylistic features of Howard's narrative is his acute reproduction of Dublin orality and spontaneity (Amador-Moreno, 2012, p. 23, 2015, 2016). This is masterfully achieved through not just phonological depictions like the orthographic renditions of intensifying *fucking* explored in section (6.5.4.) in this thesis, but also by integrating pragmatic features from every-day, natural interaction, such as discourse pragmatic markers, taboo words, or quotative devices like the ones investigated in chapter 5 (Amador-Moreno, 2016, p. 301; Amador-Moreno & Terrazas-Calero, 2017, p. 257). Take, for instance, the fragment below from *This Champagne Mojito is the Last Thing I Own* (Howard, 2008, p. 1) which illustrates the use of discourse pragmatic markers (*like*), (rhyming) slang (*chicken oriental* 'mental'), swearwords (*focker* 'fucker'), intensifiers (*totally*), or representations of prosody (notice the *high rising terminal*⁸⁸ in ...*where he is and shit?*).

The *focker's*⁸⁹ *chicken oriental*. It's like he's *totally* oblivious to where he is and shit? It's probably, *like*, a defence mechanism. Wouldn't envy him the next few years in here.

⁸⁸ As discussed in (7.1.), *high rising terminal* is an intonational deviation whereby statements are posed as *questions?* While its origin is unknown, it is sometimes associated with Australian English. However, it is more popularly perceived as having originated in the USA (1970s-80s), and it is also typically associated with trendy young/teen California 'Valley' girls (Warren 2016, p. 82). Section (7.5.1.) provides a detailed description of the Valley Girl stereotype and its linguistic construction.

⁸⁹ My emphasis.

Howard's keen linguistic awareness has also garnered him a reputation as a witness to the linguistic changes occurring IrE over the years (Linehan, 2016), which makes his books a great source with which to investigate this thesis' research questions, and has led to continuous critical acclaim for his accurate representation of Dublin orality. He often attributes his success to his predilection for eavesdropping⁹⁰ in public transportation and venues (Howard, 2021). More recently, he has acknowledged the usefulness of social media which grant him access to contemporary linguistic uses across different age groups, enabling him to take note of emerging features and to phase out others which are seem to be going out of fashion (Howard, 2019). Indeed, his acute awareness of language is illustrated in the next fragment, where he was asked about how he manages to keep his books linguistically current:

Language is just evolving all the time, you know. There's been more changes in [...] Hiberno English—the way *we* speak English— in the last 30 years than in the previous two hundred, and a lot of it is being accelerated by technology (Howard, 2019).

This 'silent observer' technique is, therefore, one of his stylistic trademarks, which is evidenced in the use of prosodic items like *high rising terminals*, orthographic renderings of prosodic and sentence stress, but mostly through phonetic representations of North/Southside Dublin accents (like those discussed in section (6.5.4.)) which are illustrative of the geographic and socioeconomic Dublinese divide and which Amador-Moreno & Terrazas-Calero (2017, p. 257) propose are also part of Howard's "hallmark style". Indeed, the RO'CK novels have been acclaimed for their depiction of what Hickey (2005a, pp. 7-8) labels *local* and *non-local* Dublin English, corresponding with the Northside and Southside areas respectively. According to Hickey (*ibid.*), *local* Dublin English is spoken in the Northside by people who "show strongest identification with traditional conservative Dublin life of which the popular accent is very much a part of". Thus, *local* Dublin English includes linguistic forms which are regionally bound to IrE, such as the use of /ʊ/ in the STRUT set (Hickey 2005a). This is orthographically rendered in the RO'CK series with variants like *anudder* (another), *brudder* (brother), etc., which

⁹⁰ Howard has often attributed his representation of the Dublin accent and several storylines to his many rides in the 46A bus and the DART. See: http://www.rte.ie/tv/tubridytonight/av_20090404.html?2520705,null,228

are produced exclusively by Northside characters (for more on the indexation of geographic-location through orthographic renderings of intensifying *fucking* in CoFIrE, see (6.5.4.)).

Contrastively, *non-local* Dublin English is spoken by Southsiders, who linguistically dissociate themselves from “popular Dublin culture” (Hickey, 2005a, p. 7). These (normally young) speakers are often (stereotypically) connected to what is popularly known as ‘D4 accent’ or ‘DART speak’. *Dartspeak* or ‘*dortspeak*’, is well known for its retracted and rounded vowel, pronouncing words like ‘Dartspeak’ [dɔː.tspɪ:k]. This pronunciation, which was popular in the late 1980s and 1990s, eventually fell ‘out of fashion’, although it has been consistently satirized in the RO’CK series with words like *car/park/guys/fuck* being rendered as *cor/pork/goys/fock*. In addition, *non-local* Dublin English also seems to have incorporated newer, more globalized linguistic items, like New Intensifying *So* and Non-lexical intensifier *Totally* which will be explored in this chapter, which aids the speakers to further increase the identarian divide that separates them from *local* Dublin English speakers. Hickey (2005a, p. 203) argues that this *non-local* Dublin English could reveal the *supraregionalization* of the dialect, which would entail the loss or exchange of standardized, locally-specific features for less regionally-bound forms which are acquired through contact with other varieties (Hickey, 2007, p. 309). Hickey (2005a, p. 203) also argues that this *supraregionalization* may have been triggered by external factors like speaker attitude towards Irish (and Dublin) English, or by the influence of the media (see also Amador-Moreno & Terrazas-Calero, 2017, p. 257).

The linguistic North/South divide is masterfully recorded in the RO’CK series, with Northsiders utilizing items which are avoided by Southsiders, like the *after* perfect, sentence tag *so* (Amador-Moreno & Terrazas-Calero, 2017, p. 266) as in ‘Yeah, I haff to show it to you, so I do’⁹¹ (Howard, 2014), or clause-marginal uses of discourse pragmatic marker *like*, which are the traditional clausal positioning preference in this variety (see chapter (7) for a detailed investigation of the use of *like* in CoFIrE). The Southsiders, however, use elements like non-standard quotatives (especially those outlined in chapter (5)), and more globalized items and uses which are never utilized by Northsiders in the books. These include features such as New Intensifying *So* (e.g. Oh my God, that is SO not the case (Howard, 2005)), or the positioning of *like* mid-clause which is more

⁹¹ My emphasis.

commonly associated with American English (chapter (7) for an investigation of this clausal preference in CoFlrE). Although, as mentioned in section (1.1.), one must be aware of the level of artificiality inherent to fictional (literary) dialect-dialogue, Amador-Moreno & Terrazas-Calero (2017, p. 266) contend that the linguistic contrast between North/Southside Dubliners in the RO'CK books is also a stylistic tool in Howard's repertoire with which he indexes "social contexts rais[ing] awareness among readers" as to the socioeconomic background of their fictional identities. All of these reasons made the RO'CK series a great tool with which to investigate the research questions of this thesis, and also led to the compilation of CoROCK.

The use of language for the manipulation of the reader's response to the fiction and its characters is further enhanced via the way the entire series is framed. The books are all narrated in the first person by RO'CK, and presented to the readers as one-sided conversations between RO'CK and Howard, who prefaces the novels with a foreword stating that the content is reproduced "as told to Paul Howard". The confessional style of the novels, which are portrayed as naturally-produced, intimate conversations, creates the illusion of intimacy and turns the readership into a silent observer to what is being said between two friends (Amador-Moreno & Terrazas-Calero, 2017, p. 257), making them feel like they are also part of the conversation. One might also believe their conversations to be real, especially because, when describing his creative process, Howard (2017) admits:

I spend my working life sitting behind a desk thinking in the voice of an idiot. A south Dublin idiot, I spend ten hours a day with this really loud, privileged, bombastic voice in my head and switching that off is often the most difficult thing.

The manipulation of the readers' response is further enhanced by the fact that they are told the story through the lens of a very unreliable narrator. RO'CK, and Howard in reality, consciously utilizes misleading fictional (literary) dialect-dialogue representations and clothing (as discussed in section (8.2.)) to misdirect the readers into perceiving other characters as having a specific (often negative) socioeconomic identity: i.e. *skanger* Northsider. In a volume of faux interviews titled *We Need to Talk About Ross* (Howard, 2009), Paul Howard 'converses' with RO'CK and the fictional cast that

recurrently appear in the stories, ranging from RO'CK's wife, parents, close friends, to his illegitimate Northside son's mother and relatives, all of whom he describes throughout the series as unemployed, scamming *skangers*. In addition, their *local* Dublin English accent is also heavily rendered phonetically, which occasionally seems to hinder some readers' comprehension of their dialogue, as evidenced in the following reader's review:

[It] was tough reading some of the phonetically written parts supposed to mimic accents and I am from Northern Ireland! (Ashfield, 2017)

An example of such rendering and their misdirecting nature is offered in the next fragment⁹². Here, teenage RO'CK describes his trip to North Dublin where he is to stay through a school 'exchange' program. He recounts the moment Tina, the daughter of his Northside exchange 'family', returns from vacation.

She's like, 'Howiya?' Now we're talking big-time CHV⁹³ here, but I could tell straight away, *roysh*, that she had a bit of a thing for yours truly. [...] The taxi-driver brought her luggage into the kitchen. Four *focking* cases. I was like, 'How long were you gone – a year?' and she *went*, 'Only one of *dem's* clothes. *Tree* of *dem's* *cigarettes*,' and I sort of, like, nodded and *went*, 'So isn't that, *like*, smuggling?' [...] Tina went, 'Me mate's *brudder*-in-law sells *dem*. On O'Connell Bridge. I *do bring dem* back wi' me. *Dat's* how I'm able to go away *tree* times a year' (Howard, 2005, p. 7)

Notice how RO'CK uses items (colored blue) which are not distinctive exclusively to IrE, such as non-standard quotatives (*be like*, *be there*, *went*) and discourse pragmatic markers *like*, *roysh* ('right') which renders his Southside accent along with *focking*. In contrast, Tina, who is already pejoratively introduced as CHV, that is, as a working class citizens living in housing schemes, is attributed a dialogue which displays an array of non-standard and traditional IrE items. These include word combinations *howiya* ('how are you?'), demonstrative *them* (*dem clothes*; *dem's cigarettes*), habitual *do* (I *do* bring dem), *me* (my), and accentual features associated with *local* Dublin English like interdental fricatives [θ, ð] realized as alveolar plosives [t, d] in *dem* (them), *dat* (that), *tree* (three), *anudder* (another), *brudder* (brother). The use of all of these items would trigger not only

⁹² Colors are used here to show contrast.

⁹³ CHV (i.e. *Council House Vermin*). Pejorative slang term to refer to lower/working class citizens who live in local authority housing schemes.

her location (i.e. Northside Dublin), but also her social status (lower/working class) in the mind of the readers, further buttressing the stereotype created by the narrator. Since the reader has no access to these characters and the stories are told from RO'CK's point of view, the audience has to believe that his portrayal of Tina, her accent, and, therefore, the stereotype he draws on, as being true. However, Tina's account of the same event in the volume of faux interviews (Howard, 2009, pp. 201-202) automatically evidences the biased and unreliable nature of RO'CK:

'Ross said you were on holidays in, where, Santa Ponsa?'

Tina: That's what he said. He's always made me sound like a knacker in *them books*.

'You weren't in Santa Ponsa?'

Tina: I've never been to Santa Ponsa in *me* life. I was in Santorini. Anyway, I came home, walked in and here's this fella—never saw him before—with his feet up on the kitchen table, watching the little portable. I'd no change left for the taxi—all I had was a twenty—so I asked me ma and she'd nothing. [...] So this fella—Ross—I ask *him*. He gives me a tenner and says, 'Get me a receipt—I'll claim it back from the school.'

While Tina does evidence use of *local* Dublin English features (colored purple) like *me* (my) and a demonstrative *them*, none of the other non-standard grammatical or pronunciation features illustrated in RO'CK's negative description are present. For example, there are no fricatives that are being realized like plosives despite the use of *this*, *the*, *with*, or *nothing*. Notice also how Tina is aware of the type of linguistic and non-linguistic negative stereotype (i.e. a *knacker/skanger*) which RO'CK consciously constructs of her in the books through misleading fictional (literary) dialect-dialogue usage.

As evidenced throughout this section, *language* is used as a 'multilevel' tool in the RO'CK books. On the one hand, the narrator purposely exaggerates the use of language and accent in his report of other characters' voices to stylize specific (often negative) stereotypes within the fiction world. In this case, Northsiders who use *local* Dublin English heavily are negatively stylized as low-working class, classless, *knackers/skangers*, whereas the more affluent, D4-accented Southsiders are stylized as linguistically trendy, upper-class, and cosmopolitan. On the other hand, it is Howard, as the real creator of the books, who purposely utilizes language to trigger these potential stereotypes of modern Irishness in the mind of the reader. Given the linguistic richness of

the RO'CK series and the multifunctionality of language representation in it to encode different contemporary (i.e. Dublin) identities, it was decided that a detailed analysis of CoROCK, and of New Intensifying *So* and Non-lexical intensifier *Totally*, would contribute to answering the research questions of this thesis, in particular RQ 3 pertaining to the potential modern (Dubliner) Irishness indexation value these features may have, at least as perceived and portrayed by Howard. The following sections will overview intensifiers, focusing on the use of *over-the-top* intensifiers (8.3.), also surveying the literature on the use of New Intensifying *So* (see (8.3.1.)) and Non-lexical intensifier *Totally* (8.3.2.) as *over-the-top* intensifiers.

8.3. (*Over-the-top*) intensifiers

From a functional perspective, New Intensifying *So* and Non-lexical *totally* are what Waksler (2009) labels *over-the-top* intensifiers (see below for a full definition of these items). However, before investigating *over-the-top* intensification, one must be aware of what standard *intensifiers* are and what their use may index with regard to speaker identity.

As already discussed in section (6.2.), intensifiers as “adverbs that maximise or boost meaning” (Bolinger, 1972, p. 17). Typical examples include *very*, *really*, *so*, *absolutely*, *fucking*, etc. Research finds a discernible link between these items and certain contexts of use. For instance, Lorenz (2002, p. 143) points out that intensifiers are more recurrent in spoken, informal contexts (especially among younger speakers). In terms of genre, Biber et al. (1999, 2007, p. 564) find intensifiers to be more prevalent in conversations rather than in academic discourse, with a seemingly common repertoire for American and British English. Nevertheless, they find that while *so*, *totally*, *really*, and *real* are more frequent in American English, British speakers favor *bloody* and *absolutely*. In televisual dialogue, Tagliamonte & Roberts (2005) investigate the US sitcom, *Friends*, finding *so* to have taken over *really* especially among female speakers between 1994-2002.

The use of intensifiers is closely linked with sociodemographic factors like *age*, *gender*, and *class*, which makes them a great index of speaker identity (i.e. and great tools with which to examine RQs 2 and 3 (see (1.2.)) in this thesis). For example, with regard to *gender*, these items have often been perceived as/associated with female discourse (see

Jespersen, 1922; Lakoff, 1975; Macaulay, 2006 for Scottish English; Ito & Tagliamonte, 2003; and Hancock et al., 2015 for British English; Tagliamonte & Roberts, 2005; and Schweinberger, 2020 for American English; Tagliamonte, 2008 for Canadian English; and D’Arcy, 2015, and Saarenpää, 2016 for New Zealand English). This association is probably due to the fact that intensifiers are part of ‘emotive language’, as they tend to reflect the speaker’s emotional state, transmitting emphasis to their *anger*, *sarcasm*, *compliments*, etc. The close association between intensifiers and *gender* is examined, for example, by Xiao & Tao (2007) who looked at intensifiers in the *British National Corpus*, finding they were more frequently produced by females. Fuchs (2017, pp. 362-363) examines a list of previously identified intensifiers in the spoken component of the *British National Corpus 1994* and *BNC2014*, finding male speakers use fewer intensifiers than females from the same *age* cohort and *social class*. Fuchs (*ibid.*) also proposes that such larger intensifier use can “contribute to a stereotypically feminine communicative style”. Nevertheless, his analysis indicates that both male and female middle class speakers show a similar frequency of use over time.

In IrE, Schweinberger (2020) points out a scarcity of research concerning intensification in this variety when compared to others which this chapter aims to address. In his study, he looks at the IrE intensifying system in the spoken component of *ICE: Ireland* and compares it to Tagliamonte’s (2008) Canadian findings. Through part-of-speech tagging searches, Schweinberger retrieves a list of items, which is topped in frequency by *very*, *really*, and *so*. He notices that they are all more prominent among female speakers, with *very* being more noticeable among older speakers (50+), although he remarks upon a noticeable use among the youngest speakers. *Really* shows no decrease in terms of *age* although he notices it becomes more delexicalized among young speakers, and *so* appears not to be an on-the-rise intensifier in this variety. Murphy (2010, p. 132) also investigated a list of 13 intensifiers and their connection to the variables of *age* and *gender* in IrE, finding they are more prevalent in female speech across all age groups than in male discourse. Males showed a continuous decrease over time, and only older females used fewer intensifiers.

Research seems to have also identified a connection between intensifier-use and younger speakers. For example, Ito and Tagliamonte (2003, p. 266) outline 11 intensifiers with adjectival heads in their corpus of York English, finding *very*, *really*, and *so* to be the most recurrent items. In addition, they notice *really* surpasses *very*, especially among

young and middle-aged speakers. Similarly, Lorenz (2002, p. 154) looks at intensifiers in the *British National Corpus* and the *Bergen Corpus of London Teenage Language*, finding *really* (adjective modifier) to be more frequent in the latter corpus, especially among younger speakers. Lorenz (*ibid.*) suggests that this corroborates the *grammaticalization* of *really*. This rise in popularity of *really*, which seems to be mirrored in Canadian English (Tagliamonte, 2008), is also encountered by Aijmer (2018) in her study of *fucking*, *super*, *dead*, *real*, *well* (*good*), and *so* (+Noun Phrase) in the spoken part of the *BNC2014*. She restricts her examination to their use in collocation with adjectival heads, noticing a clear increase in the use of *really* which is not reflected in the use of *real*. Furthermore, *so* continues to rise in frequency, as is also the case for *fucking* (*ibid.* 64-65), which is slightly more recurrent than in the spoken component of the *British National Corpus* (for a detailed investigation of the use of intensifying *fucking* in CoFIrE, see chapter (6)). Contrastively, Macaulay (2006, pp. 269-270) examines a corpus of working-class, Glasgow adolescent recorded speech from 1997 to 2004, and finds that *really* is not used recurrently with adjectival heads. Instead, he notices the absolute prominence of *pure* as the most frequent intensifier in the corpus (especially among females) followed with a noticeably decreasing tendency over time by *dead*, and an increasing rise of *so* (see Barbieri, 2008; and Calude, 2017 *inter alia* for more examples). It is important to point out that some studies have also noticed that intensifiers are less recurrent among teenagers than in the speech of speakers in their 20s (see Tagliamonte, 2008; Núñez Pertejo & Palacios Martínez, 2014).

Finally, the connection between intensifiers and *social class* has also been examined in research. Tagliamonte & Ito (2002) looked at their use in York English, finding them to be more abundant in the speech of educated speakers as opposed to uneducated individuals. Similarly, Macaulay (2002, 1995) looks at data collected from Ayr and Glasgow (Scotland) and notices that *-ly* intensifying adverbs are more recurrent among middle-class speakers, and less frequent among the working-class, also finding *very* (2002, p. 404) to be almost exclusively a middle-class intensifier in his study.

Aside from their strong indexical value, intensifiers are described as the linguistic group of items within which the most grammatical change and renewal takes place (Brinton & Arnovick, 2006, p. 441). Indeed, one of the reasons for such renewal may be the fact that they are part of the colloquial register, which makes them subject to linguistic trends, so they are in a constant, innovatively fluctuant state. Due to their optional nature,

their use is, therefore, driven by interpersonal and discourse pragmatic factors (D’Arcy 2015, p. 458), as their continuous innovation and renewal is also partly triggered by the speaker’s wish for (linguistic) creativity (Stoffel, 1901, p. 2; Quirk et al., 1985, p. 590). Thus, the ‘newer’/‘trendier’ the intensifier is, the more linguistically fashionable and original their interlocutor may perceive the speaker to be (Peters, 1994, p. 270). *Delexicalization* is also believed to be another active trigger for renewal (Sinclair, 1992; Partington, 1993, p. 183; Lorenz, 2002). *Delexicalization* is the process whereby the original word gradually loses its meaning (or is ‘bleached’) as it turns into an intensifier. Once the intensifier becomes popular and grows in frequency of use, then it gains more lexical unboundness, collocating with a larger variety of words (Partington 1993, p. 183; Lorenz 2002, p. 144), and becoming ‘delexicalized’.

With regard to their context of collocation, scholars note that intensifiers often collocate with (gradable) adjectives (Bäcklund, 1973, p. 279; Rickford et al., 2007). Nevertheless, a number of studies have recently found intensification may also occur in non-gradable contexts (see Waksler, 2009; Bylinina, 2011; McCready & Kaufmann, 2013; Irwin, 2014; Beltrama, 2015b, 2016; Beltrama & Staum Casasanto, 2017; Amador-Moreno & Terrazas-Calero, 2017; 2022 inter alia), as is the case with New Intensifying *So* and Non-lexical intensifier *Totally* in CoROCK. Quirk et al. (1985, p. 429), for example, remark upon the use of intensifying adjectives that enhance the noun they premodify. These may “scale upwards [...] and are central adjectives if they are inherent and denote a high or extreme degree [...] However] when they are noninherent, [they] are attributive only”. They exemplify this through the use of *complete* in cases like (a)-(b).

- (a) a *complete* victory— The victory was *complete*
- (b) a *complete* fool— *The fool is *complete*.

Waksler (2009, p. 17) also identifies a number of intensifiers which fall under this non-traditional category, which she labels as *over-the-top* intensification. In her study, she investigates cases where *super*, *uber*, *so*, and *totally* occur in non-traditional contexts in a corpus of two sets of naturally-occurring language. She finds that these *over-the-top* intensifiers are placed in a syntactic context “usually unavailable to that intensifier”, often occurring with non-gradable targets, which marks their subjective nature and highlights the “speaker’s particular point of view regarding the intensification target” (*ibid.*, p. 19-

23). The examples she provides are: “This place is *so* San Francisco to me” or “Sean Connery was totally the best James Bond” (*ibid.*, p. 3) However, Waksler (2009, p. 19) proposes that subjectivity is not embedded in the intensifiers in general, but is actually an outcome of *over-the-top* intensification. Finally, she also finds that, pragmatically, *over-the-top* intensifiers may transmit a range of affects, including *positive*, *negative*, and *neutral* connotations (*ibid.*, p. 22), which is why, as will be mentioned in section (8.4.), the occurrences of New Intensifying *So* and Non-lexical *totally* in CoROCK have been manually annotated for *positive*, *negative*, or *neutral* connotation conveyance.

Given the innovative and varied pragmatic functionality of intensifiers in general, and of *over-the-top* intensifiers in particular, CoROCK was used as database with which to explore the use of New Intensifying *So* and Non-lexical *totally*, the two intensifiers which, as mentioned in section (8.1.)) were found to be most recurrent *over-the-top* intensifiers in this fictional dataset. The following sections will provide an overview of the literature on New Intensifying *So* (8.3.1.) and Non-lexical *totally* (8.3.2.) while section (8.4.) will explain the methodology used for their analysis in this chapter.

8.3.1. New Intensifying *So* in the literature

As observed in Amador-Moreno & Terrazas-Calero (2017, 2022), a difference must be made between general intensifying *so* (which is illustrated in fragment 8.3. below) and what they label New Intensifying *So* (represented in 8.4.). The latter seems to be a variant which has risen in popularity in the last few decades across different varieties of English, and is used in traditionally ungrammatical contexts (e.g. colligating with verbs like in example 8.4. below).

(8.3.) It was, like, SO⁹⁴ romantic. (SHGSP, general intensifier)

(8.4.) I am SO not bluffing, Ross. (SHGSP)

Although different names have been attributed to this ‘new’ intensifier, all refer to different elements of the structure. Take, for instance, Pott’s (2004) ‘Speech Act *So*’, Zwicky’s (2006, 2011) ‘GenX *So*’, which alludes to its popularity among Generation Xers

⁹⁴ Capitalized emphasis in the originals.

(particularly in the US), Irwin's (2009, 2014) 'Drama *So*', which highlights its prosodic emphasis and elevated emotional transmission, or Aijmer's (2018, p. 87-90) self-explanatory 'Pragmatic *So*' (in reference to occurrences of *so*+noun phrases). However, this chapter uses Amador-Moreno & Terrazas-Calero's (2017, 2022) label 'New Intensifying *So*' (hereafter NISo), which provides a general umbrella for this particular construction, which we take as an example of intensifier *delexicalization* in IrE (Amador-Moreno & Terrazas-Calero, 2022, p. 522). Its *delexicalization* would be buttressed by the fact that NISo is now able to modify not only gradable predicates, but to colligate with contexts that would have been traditionally ungrammatical. These include "verbs, (proper) nouns and pronouns, negatives, time expressions, dates, and prepositions" (Amador-Moreno & Terrazas-Calero, 2017, p. 262).

As it is often the case with some linguistic developments, NISo has been frequently perceived as having originated in the USA, where research documents its recurrent status among young females (Tagliamonte & Roberts, 2005; Zwicky, 2011). While its real origin is unknown, its use has already been documented in other varieties like Canadian English. Tagliamonte (2005), for example, examines the use of *so* (including cases of NISo) in a corpus of conversations among young, Canadian speakers (aged 10-19). Tagliamonte (2005, 2008) finds this intensifier to be gender biased, as it is overwhelmingly favored by females, and also notices age-grading, for it is more recurrently produced at secondary school level and decreases over time. Amador-Moreno & Terrazas-Calero (2017, 2022) are the first to attest the use and popularity of NISo in IrE as represented in three *Ross O'Carroll-Kelly* novels. And while *so* is considered not to be an on-the-rise intensifier in Schweinberger's (2020) examination of the IrE intensifying repertoire in *ICE: Ireland*, this may be due to the fact that no distinction seems to be made between the use of *so* and NISo in his study. The present chapter will expand on Amador-Moreno & Terrazas-Calero (2017, 2022) by exploring the use, colligational preference, pragmatic multifunctionality (8.5.) through (8.5.3.) respectively), and potential identity indexical value (section (8.7.)) of NISo in CoROCK-SO.

8.3.2. Non-lexical intensifying *Totally* in the literature

Similarly to New Intensifying *So*, the use of Non-lexical intensifying *Totally* (hereafter NLT) occurs in non-standard contexts where it modifies non-gradable

predicates, as opposed to the traditional colligation of *totally* with bounded/gradable adjectives-adverbs. For example, the intensifier in example 8.5. below would be a case of standard colligation where *totally* modifies a gradable adjective (i.e. the evening can be more or less *different*). This chapter, however, will explore non-standard cases where *totally* precedes non-gradable items like 8.5., where *totally* precedes a verb (non-gradable), modifying it by lends emphasis to the speaker's proposition (i.e. he *strongly* recommends it).

- (8.5.) I had a totally different kind of evening planned this particular night. (OMCFY 2003)
- (8.6.) 'I have to say, roysh, having made the trip down the aisle myself, I can **totally** recommend it' (CIDN 2005)

The terminology used to label this particular non-standard intensifier *totally*, which, like NISO, falls within Waksler's (2009) *over-the-top* intensifier category (for a definition, see (8.3.)), varies according to the author and which factors they study. Irwin (2014, pp. 37-39), for example, labels it 'speaker-oriented *totally*'. According to Irwin, 'speaker-oriented *totally*' is an adverb which is similar to "definitely, absolutely, [and] frankly". Like NISO, this *totally* carries sentence pitch/stress in almost all cases (*ibid.*, p. 37), and shows commitment to the proposition. Semantically, Irwin (*ibid.*, pp. 60-65) outlines a series of features distinctive to this intensifier, some of which were already pointed out by McCready & Kauffman (2013). For instance, it is 1) a positive polarity item which cannot occur with denials or be "commanded by a negative subject or by sentence negation" (*ibid.*, p. 62). It 2) cannot be paraphrased with *completely* (see also Beltrama, 2015b), and is 3) generally unacceptable with *wh*-questions, although occurrence with *yes/no* questions is more acceptable. When used in reported speech, 4), it can transmit the speaker's attitude or the attitude of the person being reported.

Building on Irwin's research, and to further differentiate between standard and non-standard uses of *totally*, Beltrama (2015a, 2015b, 2016, and Beltrama & Staum Casasanto 2017) creates a distinction between what she labels 'lexical *totally*' and 'non-lexical *totally*'. Beltrama (2015b, p.23) describes lexical *totally* as an intensifier which colligates with adjectives or "verbs which encode a bounded scale as part of their meaning", and which can occur with negation. An example she provides is (*ibid.*, p. 19): *the bus is **totally***

full and *I totally support this movement*⁹⁵. Contrastively, Non-lexical *totally* colligates with unbound adjectives/verbs, such as *You should totally click on that link*, or *Dude, this is a totally deep hole* (*ibid.*; emphasis in original). The main function of Non-lexical *totally* is to modulate the speaker's attitude and "degree of commitment [...] towards the assertion" (2015a, p. 137) and the proposition (Potts, 2004; Irwin, 2014, p. 39; Beltrama & Staum Casasanto, 2017, p. 162). Beltrama (2015b, p. 17), therefore, proposes that its intensification "operates over a scale that is introduced via pragmatic reasoning [...] or by shifting the focus on the speaker's commitment towards the sentence". Furthermore, she indicates that Non-lexical *totally* "features a complex interaction with discourse structure and subjectivity" (2016, p. 22) which is not present in lexical *totally*.

In her diachronic study of the use of *totally* as a degree modifier, slack regulator and of Non-lexical *totally* in the *Corpus of Historical American English*, Beltrama (2015a, p.128) argues that Non-lexical *totally* is a good representation of *subjectification*. Following Traugott's (1989, 1995b) view, *subjectification* refers to the process whereby the speaker's attitudes/beliefs towards the proposition progressively become the basis of meaning. Beltrama's (2015a, p. 134) study reveals an expansion of the distribution of *totally* over the years covered by the corpus (i.e. 1810-2010), as well as a progression from degree modifier/slack regulator (which is solely grounded on the proposition), to Non-lexical *totally*, which modulates "speaker-oriented scales" (*ibid.*, p. 138). It is attested to colligate with non-gradable predicates, negated constituents, relative adjectives, noun phrases, and is also predicted to occur in assertions but banned with direct imperatives (Beltrama, 2015a, p. 135) or *wh*-exclamatives (Beltrama, 2016, p. 24). She finds that Non-lexical *totally* may also occur in standalone position to transmit agreement with a previous statement. This standalone use was already documented by Quaglio (2009) in his study of the language used in a corpus of the sitcom *Friends*, which he contrasted against a corpus of spoken American English. Quaglio (*ibid.*) finds that the use of standalone *totally*, which he claims to be a "linguistic innovation of American English conversation" (*ibid.*, p. 98-143), is used to express 1) emphatic agreement, and 2) mark informality. He notes that it only occurs in turn-initial position, always functioning as an *affirmative/positive*, non-minimal response, or as backchannel. Semantically, Quaglio posits that it shows speaker stance and "agreement without any restrictions" (*ibid.*, p. 99). In his contrastive analysis, Quaglio finds that standalone *totally*

⁹⁵ Emphasis in the original.

is three times more frequent in the fictional corpus than in the spoken dataset, suggesting that its high frequency in the fiction set “illustrates an apparent process of language change in progress (or at least the addition of a new function to the word)” (*ibid.*, pp. 113-6) which marks emotionality and informality in (fictional) speech.

Taking all of the existent labels into consideration, this chapter follows Beltrama’s more recent term, as it builds and expands on previous research on the uses, functions, and collocation patterns of Non-lexical *totally*. Although the *Oxford English Dictionary* includes a 2005 addition to the main entry of *totally* which specifies that such colloquial intensifying use in non-standard contexts is ascribed “originally and chiefly [to the] U.S.” (“Totally”, OED, 2022), Non-lexical *totally* is also present in other varieties of English (Irwin, 2014, p. 37). For example, Anderson (2006) documents its use in Scottish English, as attested in the spontaneous spoken language component of the *Scottish Corpus of Texts and Speech* corpus. In her study, Anderson looks at intensifying adverbs, noticing that *totally* may colligate in contexts other than with adjectives/adverbs, such as modifying an entire proposition, a verb, or a noun. Interestingly, she also points out the occurrence of standalone *totally*, which had the function of transmitting emphatic agreement with previous content offered by the interlocutor (Anderson, 2006, pp. 13-14). Although, as mentioned above, Quaglio (2009, p. 98) defined this particular use as an innovation of American origin, Anderson (2006) documents this particular use in Scottish English prior to Quaglio’s study, which could not only be indicative of its presence in other varieties of English, but could also undermine Quaglio’s belief that this could be an item distinctive to the American repertoire.

Various studies have commented on the association of Non-lexical *totally* with informality and close social distance between interlocutors (Quaglio, 2009, pp. 113-16; McCready & Kaufmann, 2013; Irwin, 2014; Beltrama & Staum Casasanto, 2017, p. 162). Sociolinguistically speaking, and despite its attestation in other varieties, this feature is often believed to have originated in the USA, where it is prominently associated with younger speakers (Zwicky, 2011; Beltrama, 2015b; Irwin, 2014; McCready & Kaufmann, 2013). Demographically, Non-lexical *totally* seems to trigger specific speaker attributes, “evok[ing...] ideologically-linked identity-based features” (Beltrama & Staum Casasanto, 2017, p. 159). For instance, Thorne’s (2014, p. 445) *Dictionary of Contemporary Slang* defines standalone Non-lexical *totally* as an exclamation often used by “female adolescents in the US [which] frequently, but not necessarily, indicates

approbation”. Beltrama (2016, p. 122) notes that it is often associated with/perceived to be part of the stereotypical linguistic repertoire of the ‘Valley Girl’ (for more on this particular stereotype and its linguistic (mis)associations, see section (7.5.1.)). Indeed, a quick online search for popular perceptions of Non-lexical *totally* use on *Urban Dictionary* retrieves descriptions like it being used by “ditzzy” (grace1129, 2006), “girly girls, poppers, and rich spoiled little brats” (Dinah, 2004) who sound “kinda ‘cheerleaderish’” (Lii, 2005). Beltrama (2016) also points out that despite the recurrent association with female speech, she finds no consistent gender distinction, although she warns that Non-lexical *totally* may also trigger the male stereotype of the “Surfer Dude, Frat Boy” or Jock which was explained at length in (7.5.1.).

The association of Non-lexical *totally* with the Valley Girl/Jock stereotypes is also illustrated in Reichelt & Durham’s (2017, p. 75) study of intensifiers in their *Buffy, the Vampire Series* corpus, where Non-lexical *totally* ranks fifth in terms of frequency of use. They suggest that intensifier distribution in televisual dialogue is more connected to characterization than to natural-use portrayal, arguing that while televisual intensifier use may not reflect real use, “within the fictional world language [, it] is employed to replicate some commonly held stereotypes in order to create identifiable character roles” (*ibid.*, pp. 76-84). For instance, they find that *totally* is mainly used by young female speakers, which they suggest might be a conscious effort by the scriptwriters to stereotype these characters by using “a Valley Girl speech pattern” (*ibid.*, p. 72), which includes the use of Non-lexical *totally* as well as of NISo.

Since, to my knowledge, there currently are no studies that investigate the use of Non-lexical *totally* in the context of IrE, and given its high frequency as an over-the-top intensifier in CoROCK-SO already discussed in section (8.1.), this chapter will examine its use in CoROCK-T as pertains to the thesis’ RQs in order to determine 1) its colligational pattern preference, 2) pragmatic multifunctionality, and 3) to discern what type of sociodemographic modern (Dubliner) identity Non-lexical *totally* could index according to Howard’s fictional portrayal.

8.4. New Intensifying *So* and Non-lexical *totally* in CoROCK: Methodology of analysis

To analyze the use of New Intensifying *So* in IrE as represented through contemporary IrE fiction, I explore its use in the CoROCK-SO subcorpus, which includes 8 (un-annotated) novels, spanning 13 years (i.e. 2005-2018), and comprising 935,458 words. CoROCK-T was subsequently used to Non-lexical *totally* in IrE according to Howard’s portrayal. This sub-corpus comprises 1,114,948 words, and includes 11 full-length un-annotated novels, spanning 17 years (2001-2018). Appendix 3 offers a list of the texts which are comprised in each subcorpus.

Quantitative and qualitative corpus methods were used to study the frequency of use, prosodic representation, pragmatic value and identity indexicality of both intensifiers in CoROCK. *Wordsmith Tools* (Scott, 2012) concordancer was utilized to run individual searches of SO (including all functions) across the books, retrieving concordance lines like those illustrated in Fig. 8.2. below. These were subsequently manually sorted for occurrences of intensifying uses, and those were also classified for regular intensifying *so* (e.g. He is *so* polite) and NISo cases. Following Waksler (2009), occurrences of regular *so* were excluded from this chapter. The same process of analysis was used in the sorting of cases of Non-lexical *totally* in CoROCK-T, whereby all other functions of *totally* as well as cases of lexical intensifier *totally* (e.g. *totally* out of the blue, *totally* unbiased, *totally* bored, etc.) were automatically excluded from the analysis. It must be noted that the occurrences of Non-lexical *totally* produced by *Ross O’Carroll-Kelly* in his narrative role were *not* taken into consideration in this particular case study.

N	Concordance	Set	Tag	Word #	Sent	Sent	Para	Para	Hea	Hea	Secl	Secl	File	Date	%
1	go ite for thot meal we talked abite so. ’ And of course I can’t say anything,			107.5{9.410}	0	73			0	73			2011-NAMA	2017/ago/14	73%
2	even allowed to have a bank account. ’ So I’ll pay you in cash. I’m coining it in			39.54{3.429}	0	27			0	27			2011-NAMA	2017/ago/14	27%
3	, and I’d no jobs this afternoon, so I said I’d bring him in the van.			99.03{8.778}	0	67			0	67			2011-NAMA	2017/ago/14	67%
4	that the lights have turned green again, so I put my foot down and, with a			106.4{9.439}	0	72			0	72			2011-NAMA	2017/ago/14	72%
5	. Looks like it’s Kate Thornton again so. ‘Oh my God,’ I suddenly hear a			139.5{12.10}	0	95			0	95			2011-NAMA	2017/ago/14	95%
6	economic miracle of which we were all so proud – it didn’t come about without			6.174	44	37			0	4%			2011-NAMA	2017/ago/14	4%
7	to, like, hurry the conversation along. ‘ So the upshot of all this is ...’ ‘Well,			9.738	75	7%			0	7%			2011-NAMA	2017/ago/14	7%
8	O’Loughlin – slick. Continue.’ Alreet, so ten minutes into it, I steerts getting			65.93{5.823}	0	45			0	45			2011-NAMA	2017/ago/14	45%
9	with her, Ross. She’s always so calm.’ ‘We may never know. I think			133.7{11.80}	0	91			0	91			2011-NAMA	2017/ago/14	91%
10	across as, like, patronizing, but I am so fucking proud of you at this moment			65.65{5.865}	0	45			0	45			2011-NAMA	2017/ago/14	45%
11	each other. And she says I am so an amazing person. I’m in the sack			118.9{10.70}	0	81			0	81			2011-NAMA	2017/ago/14	81%

Figure 8.2. Screen capture of a concordance line search for SO in CoROCK-SO.

Once all the NISo and Non-lexical *totally* occurrences had been sorted from the concordance lines, they were extracted onto two separate Excel files (Fig. 8.3.) where they were manually classified for 1) book code and year of publication, 2) occurrence and preceding content, 3) orthographic prosodic representation (i.e. capitalized, italicized, or unmarked) due to these *over-the-top* intensifiers' distinctive intonational and phrasal stress (Waksler, 2009, p.19; Zwicky, 2006; see also section (8.3.)), and 4) colligate pattern. 5) Pragmatic functions were also coded, including a) general functions (i.e. *positive*, *neutral*, and *negative* emotions as identified by Waksler (2009, p. 22)), and 2) sub-function (e.g. emotion transmission, e.g. *anger*, *compliments*, *sarcasm*, etc.). This analysis is novel in classifying for emotion conveyance, particularly the conveyance transmitted through Non-lexical *totally*, as to my knowledge, no other studies identify what type of pragmatic emotive content Non-lexical *totally*, especially in the context of IrE. It is worth restating, as already discussed in previous chapters (e.g. in section (6.3)), that while the manual annotation of emotions is a subjective process, a measure of objectivity was lent to the classification of pragmatic sub-functions given my familiarity with all contexts of utterance and inter-character relationships.

1	BOOK CODE	PRE	TOTALLY	COLLIGATE	FUNCTION	SUBFUNCTION	PROS.	GENDER	SPEAKER	AGE
2	RHINO (2009)	nshine in Malibu.	Brittany Snow also v		positive	emphasis + admirati	x	f	Sorcha	27
3	CIDN (2005/6)	horts to break...'	Christian's in my ee v		neutral	emphasis (emphatic	x	m	Christian rep Ros	23
4	CHAMPAGNE (2008)	ing to eat them...	, then roast loin of vNP		negative	dislike	x	f	Erika R	26
5	CIDN (2005/6)	n a hug and I go,	'I have to say, roys v		neutral	emphasis	x	m	Ross	23
6	CIDN (2005/6)	enism,' and I go,	'Totally,' cracking (anaphoric)	STANDALONE	neutral	emphasized agreem	x	m	Ross	23
7	EDUCATION (2001)	am this year,' and	I nod and go, 'Tota (anaphoric)	STANDALONE	neutral	emphasized agreem	x	m	Ross	18
8	CIDN (2005/6)	I'm there going,	'Please, God, let m v (imperative)		neutral	emphasis	x	m	Ross	23

Figure 8.3. Screen capture of MS Excel documenting the classification of occurrences of Non-lexical *totally* in CoROCK-T.

Furthermore, I also coded for 6) intensifying 'booster' items which may occur with or in close proximity to NISo and Non-lexical *totally* given the potential appearance of other elements in the neighboring utterances that highlight the *over-the-top* intensifying nature of both NISo and Non-lexical *totally* (Waksler, 2009, p.19). In addition, the manual analysis of the occurrences of NISo and Non-lexical *totally* indicated that these other 'booster' items were intensifiers which co-occurred with or appeared in close proximity

to them. The cases where NISo or Non-lexical *totally* co-occurred with booster items were classified according to the various contexts of proximity to the *over-the-top* intensifiers. These include a) immediate to the left; b) farther to the left; c) immediate to the right, and d) farther to the right. Examples of these are analyzed in sections (8.5.2.) and (8.6.2.). Finally, and in order to address RQs 2 and 3 regarding the potential identity indexation value of these items in CoROCK, 7) speaker *name*, *gender*, *age*, and *geographic location* were also coded.

The following sections will focus on analyzing the use and pragmatic multifunctionality, of NISo (see (8.4.) through (8.4.3.) in CoROCK-SO and of Non-lexical *totally* in CoROCK-T (see (8.5.) through (8.5.3.)) while section (8.6.) explores their potential identity indexical value as represented in CoROCK.

8.5. NISo: Distribution of occurrences in CoROCK-SO

A *Wordsmith Tools* (Scott, 2012) *frequency list* was run to gauge the recurrence of the word SO (in all its functions) in CoROCK-SO. This includes cases of SO functioning as a subordinate conjunction, as a substitute, as pragmatic marker, as tag *so*, as part of set phrases, or as an intensifier in traditionally grammatical contexts. To determine the number of occurrences of NISo present within SO, individual concordance line searches were run using *Wordsmith Tools* (Scott, 2012). As discussed in the previous section, these were extracted onto an Excel file, where they were manually classified for cases of general intensifier *so* (e.g. He is still *so* young (PHDA)) and NISo (e.g. You are *so* a good person (NAMA)). The data from the quantitative analysis of NISo in CoROCK-SO is summarized in Table 8.1. below.

Book Codes	Freq. rank order SO	NISo occurrences
<i>The Curious Incident of the Dog in the Nightdress</i> (2005)	47	36
<i>Should Have Got off at Sydney Parade</i> (2007)	41	38

<i>Rhino What You Did Last Summer</i> (2009)	50	11	
<i>NAMA Mia!</i> (2011)	56	13	
<i>Downturn Abbey</i> (2013)	60	31 ⁹⁶	
<i>Keeping up with the Kalashnikovs</i> (2014)	66	9	
<i>Game of Throw-ins</i> (2016)	54	2	
<i>Dancing with the Tsars</i> (2018)	60	1	
		<i>Total</i>	141

Table 8.1. Rank order of SO and NISo occurrences across CoROCK-SO.

As illustrated in Table 8.1., the corpus data documents not only a drop in the position of SO in the frequency ranking, featuring lower over time, but also a progressive decrease in term of NISo-use in the span of 13 years covered by the CoROCK-SO books. A longitudinal analysis of NISo production was, then, undertaken to investigate its use across time, which is visually represented in Fig. 8.4.

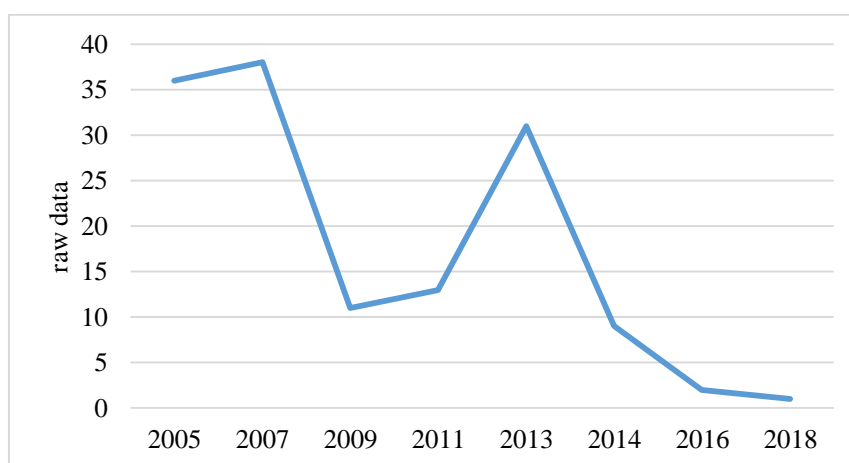


Figure 8.4. Distribution of NISo occurrences in CoROCK-SO by year of publication.

⁹⁶ These include 3 occurrences of duplicate NISo, such as “there’ll be some—oh my God—so, so romantic story involved.”

The longitudinal analysis of the 141 occurrences of NISo in CoROCK-SO and their distribution across the books by year of publication illustrated in Fig. 8.4. is very telling, suggesting that NISo reached its peak of popularity in Dublin (where all the books are set) in 2005/2007, at least according to the fictional portrayal. The data also shows a striking drop in use in 2009/11 which, despite the unusual increase in use experienced in 2013, continues to fall dramatically from 2014 onwards. Possible explanations for this drop may include a shift in Howard's perception of how NISo is used in Dublin, which may also hint at the status of NISo among Dublin English speakers. As posited by Amador-Moreno & Terrazas-Calero (2017, p. 261), if Howard's rendition of Dublin English is taken as a reflection of real-life use, then CoROCK-SO highlights an accelerated decrease in use of NISo which could also echo a potential loss of popularity among Dubliners. This would mirror the natural progression of intensifiers, and discourse pragmatic markers, which tend to rapidly grow in popularity and fall out of fashion just as quickly (Amador-Moreno & Terrazas-Calero, 2022, p. 522). A contrastive analysis with real, spoken Dublin English data, however, would be required so as to validate these theories. The next section examines the colligational pattern preference NISo appears to have as represented in CoROCK-SO.

8.5.1. NISo: Colligational pattern preference in CoROCK-SO

The manual quantitative and qualitative investigation of the most frequent contexts of occurrence of NISo, which are summarized in Table 8.2.⁹⁷, identifies a variety of traditionally ungrammatical colligates (see also Amador-Moreno & Terrazas-Calero 2017, 2022) in which NISo can occur in IrE, and in Dublin English in particular, according to Howard's books.

⁹⁷ All emphases in the original.

Colligational Pattern	Occurrences	Example
<i>Preceding verbal phrase</i>	81	(8.7.) ‘Oh my God, Erika SO freaked me out with all that talk about my perineum.’ (SHGSP)
<i>Preceding noun phrase</i>	39	(8.8.) ‘Oh my God,’ she goes, ‘that’s lollers. That is, like, so lollers.’ (NAMA)
<i>Preceding prepositional phrase</i>	11	(8.9.) She’s so like Rachel Bilson, you’d actually do a double-take. (RYDLS)
<i>Preceding adverbial phrase</i>	10	(8.10.) ‘The contractions are supposed to be - oh my God! - SO less painful’ (SHGSP)

Table 8.2. NISo contexts of occurrence in CoROCK-SO. All emphases in the original.

Quantitatively, the most frequent context of occurrence is *NISo+Verbal Phrase*, with cases like 8.1. in Table 8.2. above. Here, the speaker uses NISo (doubly emphasized through capitalization) to amplify the action of the verb, thus magnifying how ‘freaked out’ she was by what her friend said. The second most recurrent colligational pattern is *NISo+Noun Phrase* in examples like 8.2. above, where the speaker expresses how funny she finds something by emphasizing ‘lollers’; a nominalized version of the texting abbreviation LOL (*laughing out loud*). While less prominent, the last two colligational contexts merit attention. Thus, NISo is also used in CoROCK-SO with *+Prepositional Phrases* like 8.3., which places the emphasis on how much the woman resembles another person, and can co-occur with *+Adverbial Phrases* in cases like 8.4., where NISo modifies a phrase which would have standardly been intensified by *much* (i.e. *so much less painful*).

A longitudinal analysis of colligate pattern frequency by year of publication was then undertaken. The findings (illustrated in Fig. 8.5. below) reveal that *NISo+Verbal Phrase* is not only the most prominent colligational pattern in CoROCK-SO, but also the only one used recurrently throughout all the texts.

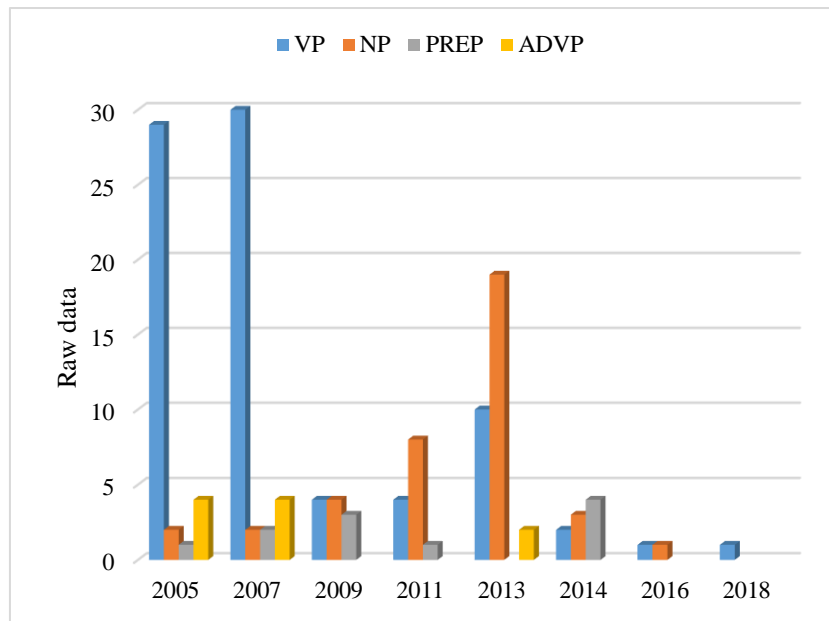


Figure 8.5. Colligational pattern distribution of NISo across CoROCK-SO by year of publication.

While *NISo+Noun Phrase* is, as illustrated in Fig. 8.5. above, also present across almost all novels, its use is substantially less frequent than *NISo+Verbal Phrase*. Contrastively, the remaining contexts of colligation (i.e. *+Preposition*; *+Adverb*) are sporadically utilized in certain books. In Amador-Moreno & Terrazas-Calero (2017, p. 263; 2022) we hypothesized that the three *Ross O’Carroll-Kelly* books examined in that particular study (i.e. CIDN, PHDA, and PHKU) could reveal a shift in colligational pattern preference from *NISo+Verbal Phrase* to *NISo+Noun Phrase*. However, the present longitudinal analysis, which incorporates other novels published before and after 2014, seemingly disproves our initial hypothesis, as not only does the *NISo+Verbal Phrase* pattern reemerge, but is also the only remaining colligational pattern by 2018, with examples like 8.11. and 8.12.

- (8.11.) ‘I am SO not bluffing, Ross. If that cot isn’t together by the time I get back, our marriage is over.’ (SHGSP)
- (8.12.) ‘She is SO not going to the gym anymore, that’s for sure. I heard she piled it on when she went on the pill.’ (CIDN)

A qualitative collocation analysis of the *NISo+Verbal Phrase* pattern uncovered the existence of a subpattern which accounts for 34.5% of cases: *NISo+Negative Verbal Phrase*, where the force of the intensifier, which often conveys (inter)personal, negative emotions, appears to be strengthened by the negative particle as in the two examples above. For instance, the angry speaker in 8.11. above reinforces how serious she is, with *NISo* being doubly emphasized by its capitalization. This also amplifies the content of her utterance and its pragmatic force, conveying her *anger* and issuing a *threat*. Similarly, the judgmental and disdainful speaker in 8.12. above doubly emphasizes how evident it is that the person of whom she speaks is clearly not exercising. The orthographic emphasis is also used here to convey the speaker’s *sarcasm* and *disdain* toward the person she speaks of.

Colligate variability was also found within the *NISo+Noun Phrase* pattern where qualitative analyses revealed a number of nominal contexts *NISo* may modify. As illustrated in Table 8.3. below, the most recurrent context was *NISo+(Positive) Adjective + Noun* in cases like 8.13. or 8.14., where the intensifier, which could be substituting for intensifiers *such* and *very* (e.g. ‘you are (*such*) a (*very*) good father’, ‘some *very* romantic story’), strengthens the force of the speaker’s utterance (i.e. *compliments*).

Pattern Variation	Occurrences	Examples
(pos)Adjective+Noun	22	(8.13.) ‘You are so a good father . [...] Anyone who says you’re not is, like – oh my God – so wrong’ (RHINO)
		(8.14.) ‘Well, if I know Charles, there’ll be some—oh my God— so, so romantic story involved.’ (PHDA)
Determiner +Noun	12	(8.15.) ‘It was so, so a pleasure to meet you’ (PHDA)
+Adjectivized Noun	4	(8.16.) ‘She’s this, like, Chinese girl in my class? She’s, like, so book .’ (NAMA)
(neg)Adjective+Noun	1	(8.17.) ‘The only reason I would smoke [...] is to, like, keep my weight down when I get older? But even though I hate to say it, I actually agree with Mom – it’s, like, so a disgusting habit .’ (NAMA)

Table 8.3. *NISo+Noun Phrase* pattern variation in CoROCK-SO.

NISo may also co-occur with *Determiner+Noun* such as 8.15. in Table 8.3. In this case, NISo emphasizes the content of the utterance (i.e. it has been a true pleasure), which receives an extra layer of emphasis due to the repetition of the intensifier (‘so, so a pleasure’). To a lesser extent, NISo can also modify an *Adjectivized Noun* like 8.16., where the speaker shows admiration for a schoolmate who she thinks is “so book”. I believe the noun in this context does not refer to a book, but rather may be slang for *cool*, hence its adjectivized nature. Finally, occurrence with (*Negative*) *Adjective+Noun* is also possible, as reflected in 8.17., although this is almost non-existent. As mentioned above, it must be noted that NISo seems to also replace other traditionally grammatical intensifiers like *such* (8.15.), (8.17), and *very* (8.16.).

A longitudinal look at these innovative nominal contexts of colligation (illustrated in Fig. 8.6.) indicates that while *NISo+(Pos)Adjective+Noun* is the most recurrent pattern, this was restricted to a span of 6 years (2009-14), whereas *NISo+Noun Phrase* spreads from the beginning through 2014. It could, therefore, be hypothesized that perhaps, and according to Howard’s perception, 2009/14 may be the span of time when NISo gained lexical unrestrictedness in Ireland, that is, colligating in more non-standard contexts, thus becoming *delexicalized* and more established in the context of Dublin. If this hypothesis is found to be valid when contrasted against a corpus of real spoken Dublin English (which does not exist to date), then these findings would evidence the high level of realism (RQ 1a in section (1.2.)) present in CoROCK.

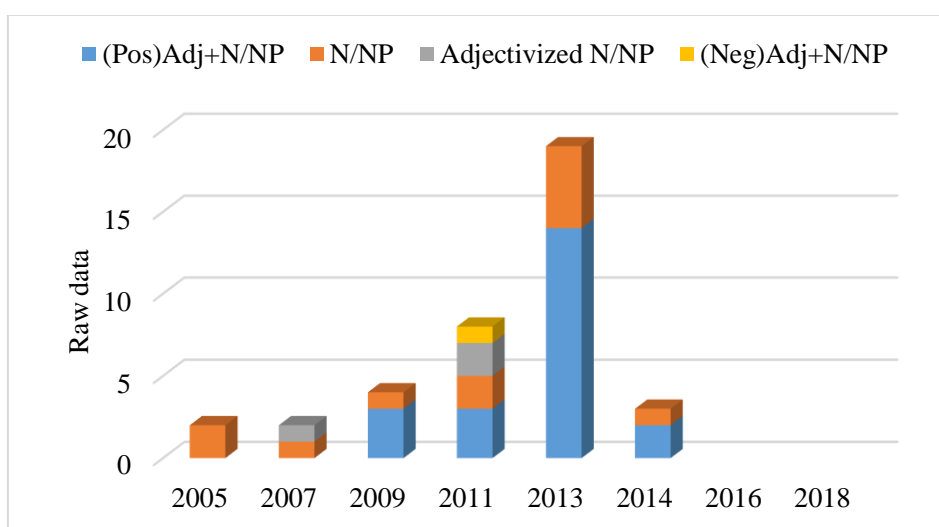


Figure 8.6. Distribution of NISo+Noun colligational pattern variation in CoROCK-SO by year of publication.

Along with the aforementioned colligational unrestrictedness of NISo, the CoROCK-SO texts also document in writing the specific intonational and phrasal stress that is characteristic to NISo due to its *over-the-top* intensifying nature. In CoROCK-SO, as well as in CoROCK-T (as shall be discussed in section (8.6.2.)), emphasis and unrestrictedness are represented in two forms: through 1) orthographic prosodic rendition, and via 2) proximity to other intensifying ‘booster’ items. The following section will explore in detail how the *over-the-top* intensification value of NISo is rendered in CoROCK-SO.

8.5.2. NISo: Marking prosody and *over-the-top* emphasis in text

Intonation is, according to Potts (2004), the main characteristic that distinguishes NISo from *so*, as the former “must receive the highest pitch accent in the sentence” (Irwin 2014, p. 41). However, the natural absence of paralinguistic elements of communication from the written medium hinders authors from conveying elements of natural interaction, such as pitch, intonation or prosody, to name a few. As observed in Amador-Moreno & Terrazas-Calero (2017, p. 265; 2022), Howard’s novels overcome this hurdle by using typographic *capitalization* (see 8.18.) and *italicization* (8.19.) to alert the reader to the prosodic, sentence stress carried by NISo.

(8.18.) Oh my God, I **SO**⁹⁸ have to have a cedilla (SHGSP)

(8.19.) Ross, you are *so*⁹⁹ a snob (PHDA)

A quantitative analysis of the typographic portrayal of NISo reveals that 80.1% (i.e. 113) occurrences are prosodically stressed in CoROCK-SO, which suggests that Howard consciously seeks to call the reader’s attention to the distinctive phrasal stress which marks its *over-the-top* intensifying nature. A quantitative, frequency analysis of both types of prosodic renditions initially suggests *capitalization* as the most recurrent phenomenon (68 occurrences), followed by *italicization* (45 occurrences), with 28 cases of *unmarked* NISo. Nevertheless, the longitudinal examination of prosodic rendition of

⁹⁸ Capitalized emphasis in the original.

⁹⁹ Italicized emphasis in the original.

NISo illustrated in Fig. 8.7. reveals that *italicization* is actually the most prevalent manner of marking *over-the-top* intensification and prosody in CoROCK-SO.

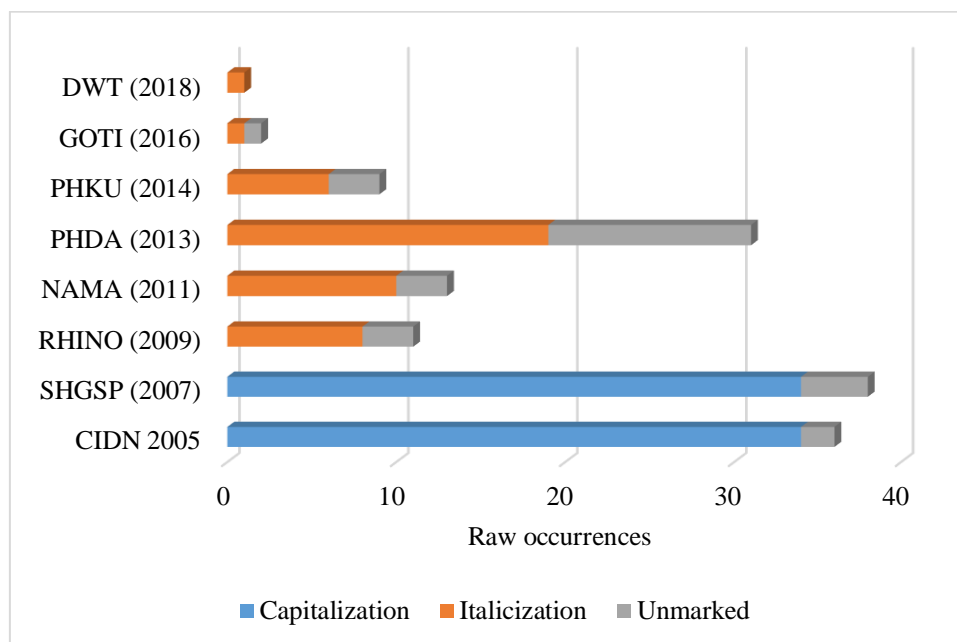


Figure 8.7. Longitudinal analysis of NISo orthographic prosodic representation across CoROCK-SO by year of publication.

As illustrated above, the corpus data Fig. 8.7. also documents a shift in the prosodic rendition of NISo in CoROCK-SO. For example, while *capitalization* was the dominant representation of *over-the-top* intensification in the first two novels, this type of orthographic rendering completely disappears after 2007. It appears it was instead replaced by *italicization*, which despite its absence in 2005/07, increases exponentially in 2009, seeming to claim dominion over the prosodic representation of the over-the-top intensifying value of NISo from then on (see also Amador-Moreno & Terrazas-Calero, 2017, 2022). This overtake could be due to *italicization* being a less ‘obvious’ manner for the author to convey extra emphasis than using capital letters. Furthermore, there is a clear increase of *unmarked* occurrences beginning to gain prominence from 2009 onwards, which is also mirrored in the use of Non-lexical *totally*, as shall be discussed in section (8.6.2.). This evolution towards *unmarkedness* could respond perhaps to the Howard’s perception of NISo as an established, non-standard, *over-the-top* intensifier in

Dublin, and to his assumption that, by then, readers were familiar enough with its use that marking its *over-the-top* intensifying value was/is no longer needed.

The *over-the-top* intensification of NISo is further strengthened by its occurrence with/in proximity to other intensifying ‘booster’ items in CoROCK-SO. A quantitative analysis of *NISo+Intensifier Boosters* reveals a total amount of 44 cases of this pattern, comprising 4 main intensifiers, which are outlined in Table 8.4., and which, as mentioned in section (8.4.), were classified depending on the context of proximity to NISo. Thus, these intensifiers were found to appear a) immediately to the left of NISo (see 8.20 and 8.22. in Table 8.4 below); b) farther to the left of NISo (examples 8.22. and 8.23.); c) immediately to the right of NISo (see 8.21.), and d) farther to the right of NISo (8.23).

Other Intensifiers		
Item	Occurrences	Examples
<i>Oh my God</i>	40	(8.20.) ‘I know you’ve never given her a chance, Ross – but Giovanna is, oh my God , <i>so</i> an amazing person’ (NAMA)
<i>Fucking</i>	2	(8.21.) ‘That’s it, I am SO <i>fucking</i> out of here.’ (SHGSP)
<i>Hello?</i>	1	(8.22.) Melanie who – Oh! My God! –sits in the sauna for, like, forty-five minutes before she goes to Weight Watchers, which is, like – HELLO? – SO cheating, even if she is only fooling herself (SHGSP)
<i>High Rising Terminal</i>	1	(8.23.) ‘ Oh my God , there is SO <i>not</i> ages to go? Of course, you wouldn’t know that. You don’t

have to carry this baby around with
you everywhere you go. (SHGSP)

Table 8.4. Distribution of NISo+intensifying booster items across the CoROCK-SO texts.

As is clear from examples 8.20. through 8.23 above, the use of *NISo+Intensifier Boosters* lends additional force to NISo. For example, the speaker in 8.20. uses *oh my God* to show her admiration but also to reinforce the strength of the compliment issued through the *NISo+(Pos)Adjective+Noun* pattern that follows. Further reinforcement is lent by *fucking* in 8.21. which provides bonus intensification to the fact that the speaker is intent on leaving. Incidentally, it also helps to convey more vividly his own exasperation at the situation he finds himself in. The intensification in fragment 8.22. is multilayered. Here, the speaker uses *hello* as added emphasis for NISo. However, *hello* is set apart from the rest of her utterance by hyphens, thus marking a pause in the speech of this individual. In addition, *hello* is doubly emphasized, not just by the insertion of *high rising terminal*¹⁰⁰, but also through its *capitalization*, through which the author relays to the reader the fact that *hello* carries extra intensification. It would seem that, in this case, the use of *HELLO?* not only provides additional emphasis to NISo and the criticism it amplifies, but also seeks the interlocutor's participation (perhaps looking for their agreement), which is indicated through the *high rising terminal*. Finally, the speaker in 8.23. above is a heavily pregnant lady who confronts her husband who believes they still have time to prepare before the baby arrives. In this case, three other intensifiers are found in the proximity of NISo. The initial 1) *oh my God* appears to render her disbelief and exasperation while intensifying the force of the 2) NISo which is followed by the negative particle, thus doubly emphasizing how little time there is in fact. Finally, there is 3) *high rising terminal* at the very end of the utterance, which may convey her *exasperation* and *sarcasm*.

A quantitative longitudinal and sociolinguistic examination of the distribution of occurrence of *NISo+Intensifier Boosters* illustrated in Fig. 8.8. below indicates that the presence of these intensifying booster items was more marked in the first two novels, with *Oh my God* being the only booster item that is used across all CoROCK-SO texts, although its use decreases over time.

¹⁰⁰ Posing a statement *as a question?*

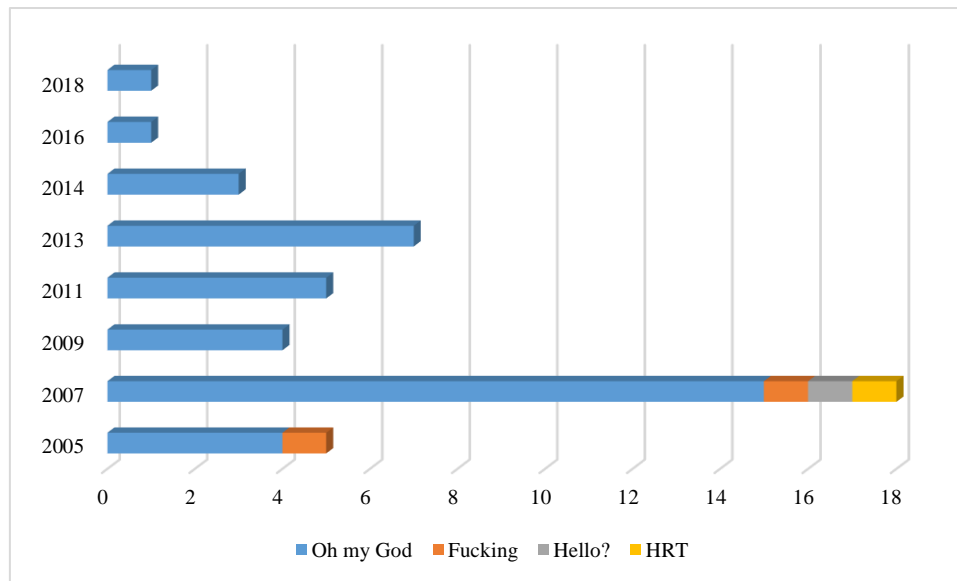


Figure 8.8. Distribution of *NISO+Intensifier Boosters* distribution across CoROCK-SO by year of publication.

I believe the progressive disappearance of these intensifying booster items, which may have functioned to signal to the reader the *over-the-top* intensification of the ‘newly’ acquired NISO in the earlier novels, mirrors the evolution towards prosodic *unmarkedness* and could, therefore, also be due to a shift in Howard’s perception whereby he might consider that his readers no longer needed to be reminded of its *over-the-top* intensification force.

Since *Oh my God* appears to be the intensifying booster that co-occurs most frequently with NISO in CoROCK-SO, its use was quantitatively analyzed with regard to gender production over time. The findings, which are illustrated in Figure 8.9. below, indicate that the *NISO+Oh my God* pattern is heavily prominent among female characters in the texts as opposed to their counterparts (41 to 3 occurrences respectively).

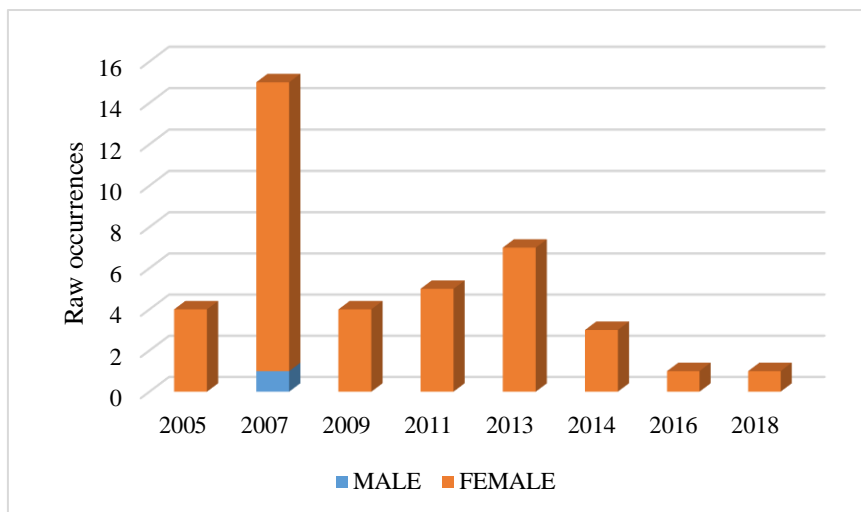


Figure 8.9. Gender distribution of occurrences of *NISo+Oh my God* pattern across CoROCK texts by year of publication.

While clear conclusions cannot be drawn regarding the use of NISo with other intensifying booster items like *fucking*, *hello?*, or with *high rising terminal* due to the small amount of occurrences of these found in the corpus (see Table 8.4. above), we could hypothesize that the use of the *NISo+Oh my God* pattern may be perceived as a female-distinctive feature in Dublin from 2007 onwards, at least according to Howard’s rendition. Nevertheless, 69% (i.e. 27) of *NISo+OMGs* are produced by the same character, *Sorcha O’Carroll-Kelly*, which turns it into an almost idiolectal feature. Stylistically, it could be a way for the author to make that character easily recognizable to the audience through the creation of her own idiolect. Thus, the readers automatically identify it as her voice, even when her speech is being reported by others, like in example 8.24. below, where her husband reports what he was expecting her to say once he returned to their house.

(8.24.) When I walk through the door, roysh, it’s like I’ve never been away. There’s no, *Welcome home*. No, *Oh my God*, *I SO missed you*. (SHGSP)

The seeming gender-distinctiveness of NISo to female discourse evidenced in the corpus, which answers RQs 2 and 3 in this thesis (see section (1.2.)) would further be supported by the fact that there is only one occurrence of *Oh my God* which is produced by a male character in 2007 as well. The complete inexistence of any male-produced *oh my God* prior to 2007 and its disappearance from then on from male speech, coupled with the lack

of occurrences produced by males of *NISo+Intensifier Boosters* other than *+fucking*, leads me to consider the possibility that this one male-produced occurrence of *NISo+Oh my God* could be due to the Howard's experimentation with the sociolinguistic indexicality value of *NISo*, and with the *NISo+Intensifier Boosters* pattern more specifically.

The number of intensifying boosters (especially *oh my God*) that co-occur with *NISo* in CoROCK-SO may also attest to the potential *delexicalization* that *NISo* may have undergone in the context of Dublin, as recorded by Howard. Another possible explanation for the *NISo+Intensifier Boosters* pattern may be its need for intensifying 'aid'. This is supported by Tagliamonte (2008, p. 391), who points out that "overuse, diffused use, long-time use, will lead to a diminishment in the intensifier's ability to boost and intensify". If this is the case with *NISo*, and with Non-lexical *totally*, as shall be discussed in section (8.6.2.), then its pronounced occurrence with other intensifier boosters may also 1) indicate that speakers have become very used to *NISo* and it has lost its linguistic 'novelty' and intensifying power, which could also 2) forecast its impending fall in use.

The following section will explore the pragmatic functionality of *NISo*, as represented in CoROCK-SO, examining if it conveys *positive*, *neutral*, or *negative* emotions, and providing a detailed classification of the emotions it may convey in Dublin English, as perceived and portrayed by Howard.

8.5.3. The pragmatic functionality of *NISo* in CoROCK-SO

The pragmatic functionality of *NISo* as represented in the CoROCK-SO texts was examined next. As discussed in section (8.3.), the pragmatic connotations conveyed by *NISo* were divided into: 1) general pragmatic connotation (i.e. *positive*, *neutral*, and *negative*), and 2) sub-functions or transmission of emotions (e.g. *anger*, *excitement*, etc.). The quantitative examination of general functions which is illustrated in Table 8.5. reveals that *neutral* connotations are the most frequently transmitted type of connotations, with only two sub-functions: conveying *simple emphasis* (56 occurrences) or *amplified determination* (13 occurrences).

Positive	Neutral	Negative
36	69	36

Table 8.5. Raw distribution of NISo by pragmatic connotation in CoROCK-SO.

The conveyance of *simple emphasis* is illustrated in 8.25. Contextual information is required to understand the example. The Northside Dublin, teenage, (step) daughter-in-law of the speaker in 8.25 below is to be enrolled at Holy Child Killiney, an all-girl, private secondary school in the affluent suburb of Killiney in Dún Laoghaire–Rathdown (i.e. Southside Dublin). The speaker expresses her hope that she be introduced to the school’s motto, which she finds inspirational. In this case, NISo, which carries double emphasis through *italicization*, simply strengthens her opinion (i.e. *simple emphasis*) as to the inspirational nature of the quote.

- (8.25.) ‘I hope whoever shows her around mentions the thing in their mission statement about exploring one’s talent to realise one’s potential. I just think that’s, like, *so an inspiring quote.*’ (PHKU)

Amplified determination is illustrated in 8.26. below, where the speaker reports a female who mentions her eagerness (i.e. *amplified determination*), or speaker intent (Waksler 2009, p. 23), to watch a specific movie. In this case, the determination is stressed through the *capitalization* of NISo.

- (8.26.) So this Emily bird mentions that she is *SO going to go to see the new Mel Gibson movie when it comes out* (CIDN)

As regards the conveyance of *positive* and *neutral* connotations through the use of NISo in the texts, Table 8.5. above illustrated a balanced distribution of occurrences of both in CoROCK-SO. A longitudinal analysis of the conveyance of *positive* or *negative* connotations was, then, undertaken to check for potential pragmatic developments within these two categories. As illustrated in Fig. 8.10., while *positive* functions appear to be used rather consistently throughout the years, *negative* ones were more frequent in the

first two novels, beginning their marked decrease in 2009. The significance of 2009 as the year in which these functions begin to fade away may support the hypothesis mentioned in section (8.5.) whereby 2009-onwards marks the period in which NISo could have started to/become an established feature in Dublin English, at least as evidenced in CoROCK-SO.

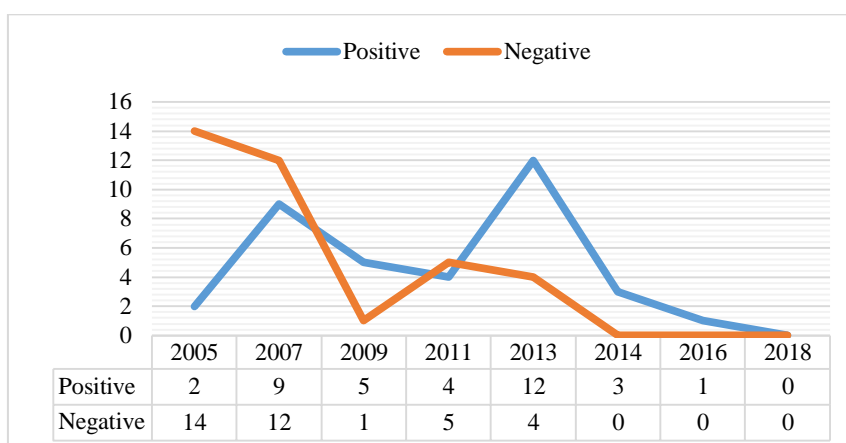


Figure 8.10. Distribution of the conveyance of *positive/negative* connotations through NISo across CoROCK-SO by year of publication.

All of the *positive* and *negative* pragmatic occurrences were manually classified for emotion connotation, which are visually represented in Fig. 8.11. below.

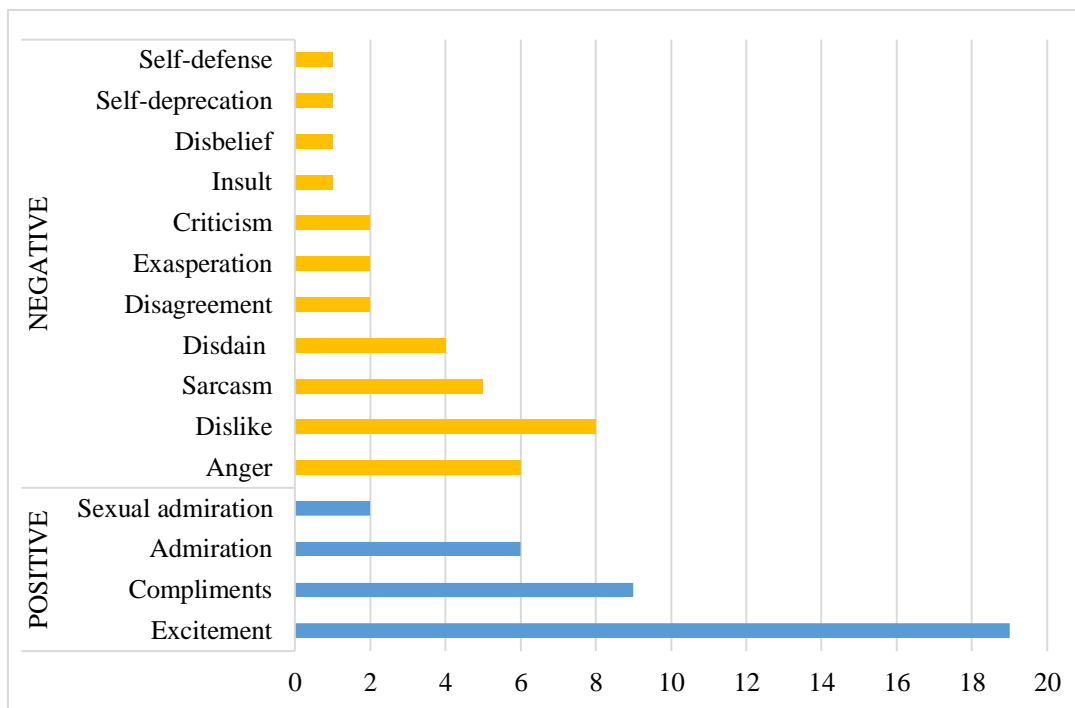


Figure 8.11. Raw distribution of *positive/negative* emotion catalogue conveyed through NISo in CoROCK-SO.

The quantitative analysis of the conveyance of *positive* and *negative* connotations by emotion transmission represented in Fig. 8.11. above indicates that despite there being the same amount of *positive/negative* connotations, the *negative* emotive spectrum is much larger than the *positive* one, where the two most frequently conveyed emotions are: *excitement* and *compliments*. *Excitement* is illustrated in 8.27. below, where the speaker's husband (who reports the exchange) has asked her for an aftershave recommendation. The speaker's reply via text emphasizes her *excitement* with the use of the *NISo+Oh my God* pattern and by its *capitalization*, which stresses the fact that the aftershave she proposes is undoubtedly the best option. This fragment also shows that NISo is not only a feature of spoken language, but may also be utilized in texting and other online sources, lending a sense of orality and spontaneity to those media.

(8.27.) 'OMG, it SO has to be Jean Paul Gaultier 2. It's got a kind of soft, woody smell.'
(SHGSP)

An example of a *compliment* is offered in 8.28. where the speaker praises her mother-in-law, an erotica author, on how aware she is of the latest trends. The *compliment* is extended through her mention that the genre she writes “is so huge right now”. Notice how *NISo+Oh my God* is also used in this case.

(8.28.) ‘Oh my God, [...] you always have your finger so on the pulse, Fionnuala. Mummy porn is so huge right now.’ (PHDA)

As for *negative* emotive meanings, *anger* and *dislike* are the most frequently transmitted emotions, as exemplified in 8.29. and 8.30. below.

(8.29.) ‘He’s fucking lying. You are bet-down¹⁰¹. And you are SO not making a holy show of me’ (CIDN)

(8.30.) ‘Vans are, like, so not book. Malorie says they’re for skanks.’ (NAMA)

Prior to the exchange in 8.29. above, the speaker has insulted his mother, who has been asked to participate in a naked, charity calendar to raise breast cancer awareness. After having insulted her looks to discourage her from participating, his father reassures his mother of her beauty. Angry, the speaker tries to harshly ‘disabuse’ her of her husband’s reassurance (‘He’s fucking lying.’), insults her again (‘You are bet-down’), and uses *NISo* to reinforce his *anger* and *amplified determination* not to let her ridicule him. Finally, the speaker in 8.30. above uses *NISo* to overtly show her *dislike* for Vans, which is further amplified by the negative particle that follows it (‘so not book’). The following sections will analyze the occurrence distribution (8.6.), colligational pattern preference (8.6.1.), prosodic emphasis representation (8.6.2.), and pragmatic multifunctionality (8.6.3.) of Non-lexical *totally* in CoROCK-T.

¹⁰¹ IrE slang for ‘very ugly’.

8.6. Non-lexical *totally*: distribution of occurrences in CoROCK-T

The quantitative analysis of CoROCK-T summarized in Table 8.6. below reveals a total of 342 occurrences of Non-lexical *totally* (hereafter NLT). However, a more detailed examination of production by character revealed that the majority of occurrences were produced by *Ross O'Carroll-Kelly* in his narrative role. Nevertheless, and as stated in section (8.4.), his narrative input was not taken into consideration in this particular study. Thus, the total amount of NLT occurrences examined in this study is 52.

Distribution of Occurrence	
<i>RO'CK</i> (narrator)	290
<i>Other characters</i> ¹⁰²	52
<i>Total</i>	342

Table 8.6. Raw distribution of occurrences of NLT in CoROCK-T.

A longitudinal analysis was subsequently undertaken to track the adoption /development of this particular intensifier in the IrE repertoire by examining the production of occurrences per book. As illustrated in Fig. 8.12., the corpus data shows a relatively constant use of NLT between 2001-2005. However, there is a noticeable drop in 2008/09, with NLT completely vanishing in 2011. While NLT does appear in the 2013-2018 books, its use is substantially lower in comparison to the number of occurrences produced in the first decade.

¹⁰² This includes occurrences produced by RO'CK in his role as character within his story.

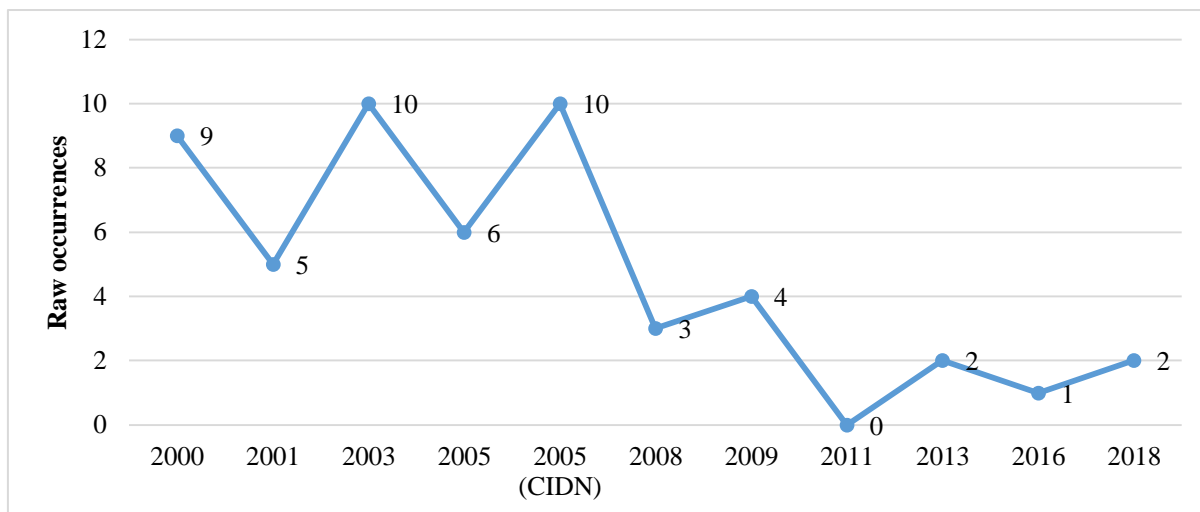


Figure 8.12. Longitudinal distribution of NLT occurrences across CoROCK-T texts by year of publication.

Socioeconomic factors may help to explain NLT’s trajectory of distribution. For instance, the relatively sustained use of NLT takes place between 2001 and 2005, a span which coincides with the economically wealthy times of the Celtic Tiger (see section (1.4.) for an overview of this period). Wealth, worldliness, and cosmopolitanism may have been linguistically flaunted by speakers (as represented in the corpus) through the acquisition of linguistic items (e.g. particularly NISo and NLT) popularly connected with other varieties where they are also part of commonly held stereotypes. For example, both NISo and NLT are, as mentioned in previous sections, often connected with American English and with the stereotypes of the *Valley Girl/Surfer Boy* (see section (7.5.1.) for a detailed description of this stereotype and its linguistic stylization in the context of the USA). With regard to the considerable decrease in use of NLT by 2008/09 in CoROCK-T, it is worth noting that these years approximately mark the end of the Celtic Tiger and the onset of the economic recession that followed it. Perhaps the noticeable decrease in NLT use in the books after the Tiger years could be a consequence of the recession and of the author’s desire to discontinue the characters’ linguistic ‘association’ with those stereotypes (i.e. affluent, cosmopolitan, ‘Valley Girl’-sounding speaker). On the other hand, the fact that NLT seems to be an age-graded feature in other varieties (see Reichelt & Durham 2017, p. 80) may account for its reduction in use as the main fictional Dubliner characters in the book series grow older, while the author’s own perception of NLT (dis)use may also play

a role in its fictional portrayal. The following section explores the colligational pattern preference NLT appears to have in Dublin English as represented in CoROCK-T.

8.6.1. NLT: Colligational Pattern Preference in CoROCK-T

All occurrences of NLT in the CoROCK-T were qualitatively classified for non-gradable colligates taking into consideration its colligational pattern preferences in American English (see Beltrama, 2015a, p. 134) and Scottish English (see Anderson, 2006, pp. 13-14). A quantitative examination summarized in Table 8.7. below reveals a rather small amount of patterns. Like NISo (see section (8.5.1.)), the most recurrent colligational pattern is *NLT+Verb Phrase* in cases like 8.31. in Table 8.7. Here, Sorcha’s mother apologizes in advance for having to miss their vow renewal ceremony. Sorcha accepts the apology and attempts to mitigate the threat to her mother’s negative Face (in green) by telling her that she understands the reasons why she will miss that day (i.e. her father hates her husband for ‘stealing’ his daughter and home from him, and has even decided to move away to the UK). This mitigation is further stressed through *NLT+Verb Phrase* (i.e. “I totally do [understand]”).

Pattern	Occurrences	Example	Book Code /Year
+Verb Phrase	28	(8.31.) ‘I’m so sorry [...] to be missing your big day.’ Sorcha’s like, ‘ I understand. I totally do. ’	PHDA 2013
Anaphoric Standalone	21	(8.32.) I hang up and high-five all of the goys in turn and tell them that these are going to be the best five weeks of our lives. And everyone just goes, ‘ Totally. ’	PSSB 2005
+Noun Phrase	3	(8.33.) ‘[...] If she, like, takes a few fries off your plate or has, like, half your dessert, it’s like she thinks it doesn’t count.’ Aoife’s	OMFY 2003

there, ‘That is SO, like ... aaagggghh!’ and
Sophie goes, ‘I know. It’s, like, totally ...
duuuhhh!’

Table 8.7. Non-gradable NLT colligational pattern preference across CoROCK-T.

Another example of the *NLT+VP* pattern is evidenced in 8.34. below, which also evidences the sensitivity of NLT to speech acts. As mentioned in section (8.3.2.), McCready & Kauffman (2013) suggested that NLT is predicted to occur with assertions but banned from occurring with direct imperatives (see also Beltrama, 2015a, p. 135). Nevertheless, such a case was found in CoROCK-T in the example of *NLT+Verb Phrase* offered in 8.34.

(8.34.) ‘Please, God, let me live. Oh my God, totally let me live,’ (CIDN 2005)

In the fragment above, *Ross O’Carroll-Kelly* is being threatened by his father-in-law. He has a gun to Ross’ head and reminds him of the warning he gave him regarding cheating on his daughter again. A terrified Ross closes his eyes and begins to pray. In this case, the direct imperative is emphatically expressed through *NLT+Verb Phrase* pattern (*let me live*). However, it could also be argued that while NLT emphasizes a direct imperative, this is also a plea to God. Regardless of the blurry duality of this particular speech act, it is still a case of *NLT+Verb Phrase* worth pointing out, as its use in CoROCK-T suggests the usefulness of the fiction corpus to study linguistic developments in progress.

The second most recurrent colligational pattern preference evidenced in Table 8.7. above is NLT in ‘anaphoric’ *standalone position*. While Quaglio (2009) described this as innovatory of American conversation, its appearance in CoROCK-T provides further evidence to its geographic unrestrictedness (see also Anderson (2006) for its presence in ScoE). Studies have described *standalone* NLT as a marker of *emphatic agreement* with previous content, as an *affirmative/positive*, non-minimal response or as backchannel. (Quaglio, 2009, pp. 98-143; Beltrama, 2015a; Anderson, 2006, pp. 13-14), with Thorne (2014, p. 445) mentioning that it frequently, although not always, may show approbation. The qualitative and quantitative analyses of anaphoric *standalone NLT* in CoROCK-T reveal the use of previously outlined functions while also identifying a ‘new’ one, all of which are illustrated in Table 8.8.

Function	Occurrences	Example	Year of Publication
<i>Emphatic agreement</i>	17	(8.35.) I hang up and high-five all of the goys in turn and tell them that these are going to be the best five weeks of our lives. And everyone just goes, ‘ Totally. ’	2005
<i>Affirmative response</i>	2	(8.36.) ‘So you’re really taking the plunge? ’ Oisinn goes to me [...] and I’m there, ‘ Totally, ’ trying to sound a bit more enthusiastic than I basically am.	2005
<i>Speaker emphasis</i>	2	(8.37.) ‘What happened?’ ‘ Nothing, Ronan. And we’re talking totally ’	2005/6

Table 8.8. Anaphoric *standalone* NLT functions across CoROCK-T.

The CoROCK-T data summarized in Table 8.8. above shows that the most frequent pragmatic function of *standalone* NLT in the corpus is *showing emphatic agreement with previous content offered by the interlocutor*, as exemplified in 8.35. in Table 8.8. Here, Ross O’Carroll-Kelly and his friends are about to go on holidays, so he expresses his prediction (in green) that they will have the best time of their lives. His friends’ group response (*standalone totally*), therefore, conveys their *emphatic agreement* with his prediction. To a lesser extent, NLT was also found to function as an *affirmative response token* (Quaglio, 2009), in cases like 8.36. above where Ossin asks Ross whether he is really prepared to marry Sorcha. Ross’ NLT provides a seemingly *positive response* whose emphatic, affirmative nature is belied by the ensuing narratorial description (in purple) which evidences he is *not* as ready as he is making himself out to be. Finally, the corpus records a previously undocumented function: *lending emphasis to the speaker’s previous statement*. For example, in 8.37. above, Ross and his illegitimate son, Ronan, are at the hospital where Ronan’s mother was rushed after experiencing a severe allergic reaction. Ronan confronts his father as to the state in which he found them both: his

mother in a nightgown and passed out on top of an underwear-clad Ross. Ronan’s question (‘what happened?’) does not refer to what triggered her allergic reaction but he is rather asking if they had intercourse. Ross’ reply (in green) states that *nothing* happened (this is true only because she lost consciousness). In this particular case, the use of anaphoric, *standalone* NLT lends *emphasis to the speaker’s prior statement* (i.e. there was no intercourse; nothing happened), but also seems to amplify the veracity of it. This is further strengthened by the preceding intensifying booster *we’re talking* (for more on co-occurrence with intensifier boosters, see section (8.6.2.)). The fact that this function has not been documented in other varieties yet lends validity to CoROCK-T as a great source with which to examine pragmalinguistic developments.

Finally, the least recurrent patterns in CoROCK-T is *NLT+Noun Phrase*. An example is fragment 8.33. in Table 8.7. above, where Aoife and Sophie are disparaging their friend, Keera, who is following a specific diet. After Keera leaves for the bathroom, they begin to criticize her weight, her food/drink choices, with Sophie describing Keera’s apparent inability to follow the diet. In this fragment, Aoife expresses her disdain, irritation, and agreement with Sophie’s description of Keera’s dietary habits, which is emphasized with the onomatopoeic, capitalized *NISo+Noun Phrase* (“SO, like...aaaggghhh!”). Sophie not only agrees with Aoife’s *exasperation* (“I know”) but also contributes her own *emphatic disdain* and *criticism* with the *NLT+Noun Phrase* pattern (‘totally...duuuhhh!’).

A longitudinal, quantitative analysis was also undertaken to evaluate colligational pattern distribution over time. The findings summarized in Table 8.9. below highlight the prominence of the *NLT+Verb Phrase* pattern which is substantially present across all books.

Colligates	Year of Publication									
	2000	2001	2003	2005	2005	2008	2009	2013	2016	2018
	(CIDN)									
<i>Verb Phrase</i>	7	4	0	1	6	2	4	2	1	2
<i>Anaphoric standalone</i>	4	1	8	4	3					
<i>Noun Phrase</i>			2			1				

Table 8.9. Raw NLT colligational pattern distribution across CoROCK-T texts by year of publication.

Contrastively, *NLT+NP* is almost only used in two novels, while anaphoric, *standalone* NLT is extensively used across a specific span of time (i.e. 2001-2005). It is worth noting that *standalone* NLT disappears completely in 2008, which, as mentioned in section (8.6.1.), coincides with the onset of the recession. While its disappearance might be explained by a change in Howard’s perception of the use of NLT in Dublin, it could also respond to a conscious disuse due to loss of popularity among speakers in the real world.

The following section will focus on examining the way in which the *over-the-top* intensifying value and distinctive phrasal stress of NLT is marked in the CoROCK-T texts.

8.6.2. NLT: Marking prosody and *over-the-top* emphasis in CoROCK-T

Just like NISo, all occurrences of NLT in CoROCK-T were analyzed for their *prosodic rendition* in text (i.e. *capitalization*, *italicization*, or *unmarkedness*), while its potential co-occurrence with other *intensifying booster items* was examined so as to analyze whether its *over-the-top* intensification was strengthened or not in the texts. The same manual classification for context of proximity used with NISo (discussed in section (8.5.2.)) was implemented to the analysis of the *NLT+Intensifier Boosters* pattern.

The analysis of the distribution of occurrences of *NLT+Intensifier Boosters* summarized in Table 8.10. below shows a total amount of 26 boosters comprised by the 4 patterns outlined below.

Pattern	Occurrences	Fragment	Book Code/Year
Oh my God	12	(8.38.) ‘[...] She was telling me she tied her hair up, so as to emphasize her—’ and I’m there, ‘I	CIDN 2005
DPM like	7	don’t want to fucking hear it. I swear to God, I seriously think I’m going to spew...’ ‘You are, like,	

totally overreacting, Ross. It's
like, OH! MY! GOD!'

		(8.39.)	'And you are SO Joey from 'Dawson's Creek' as well,' and Sophie's like, 'Oh my God, TOTALLY.'	OMFY 2003
NISo	4			
		(8.40.)	She says Sorcha's a total Bridezilla and I go, 'Yeah, and we're talking totally here,'	PSSB 2005
'We're talking'	3			

Table 8.10. NLT+*Intensifier Boosters* distribution across CoROCK-T texts.

While a quantitative exam shows that +*Oh my God* is the most frequent cluster, just like this was also the preferred booster colligate pattern with NISo (see section (8.5.2.)), discourse pragmatic marker *like* is actually the most recurrent item occurring always one place to the left of NLT. Colligation with both boosters is illustrated in example 8.38. in Table 8.10. above. The context of this particular exchange includes an argument between *Ross O'Carroll-Kelly* and his wife, Sorcha, over a nude photograph his mother has been invited to contribute to a charity calendar. Ross absolutely opposes the idea, believing that she will make a laughing stock of him among his friends. However, Sorcha defends her mother-in-law's choice, alluding to her beauty and the good intentions of the calendar. When Ross threatens to become violently ill over the idea, Sorcha accuses him of overreacting, which is emphatically delivered by means of the *NLT+Verb Phrase* pattern ('*totally overreacting*'). Her accusation is reinforced by the preceding focuser *like*, which centers the emphasis on the content that follows it. Finally, further strengthening is provided by the use of *Oh my God*, which itself contains double emphasis in the form of capitalization and the separation of the elements of the compound into individual, exclamatory units, signaling the speaker's pauses in between elements (i.e. *Oh! My! God!* vs *Oh my God!*).

NISo is also found in proximity to NLT in cases like 8.39. above. In this exchange, Sophie, Aoife, and Keera are vilifying a girl for flirting with one guy but sleeping with

his best friend on the same weekend. After sharing their common disdain, Aoife self-deprecatingly claims to be similar, calling herself a “total Samantha”, while her friends disagree, proposing that she is more like Ally McBeal given her undecidedness regarding whom she wants to be with. At this point in 8.39., Keera suggests that Aoife is also very much like Joey from *Dawson’s Creek*. Contextual information is required here since this exchange has many references to American, pop culture TV characters from the late 1990s to early 2000s. These include Samantha Jones from HBO’s *Sex and the City* (1998-2004), Ally McBeal from the eponymous show (1997-2002), and finally, Joey Potter, from WB’s TV drama *Dawson’s Creek* (1998-2003). In this particular fragment, Keera uses *NISo+Noun Phrase* to emphatically express her own opinion as to what type of character Aoife is more like. Here, *NISo* would trigger the features of a particular character trait (i.e. Joey Potter’s) in the mind of the interlocutor. According to the *Urban Dictionary*’s entry for “Dawson and Joey”, this *Noun Phrase* refers to “[h]aving a very long drawn-out romance in which you break up and get back together on a regular basis because the couple in question thinks they’re soulmates” (Mur11, 2006). It is logical, therefore, to infer that Joey Potter would be a character that triggers items such as “faithful”, “romantic”, or “committed”. Thus, in using “SO a Joey”, Keera’s utterance indicates that she strongly believes that Aoife is also faithful and committed when in a relationship. Sophie’s anaphoric *standalone* “TOTALLY” expresses her emphatic agreement with Keera’s comparison, which is doubly strengthened through *capitalization*.

Finally, example 8.40. in Table 8.10. above illustrates how *we’re talking* may function not as exemplifier, nor as a speech verb, but as a focusing discourse pragmatic marker akin to *like* or *I mean* (e.g. Yes, *like/I mean*, totally a Bridezilla), which acts as intensifying booster. Here, Erika complains to Ross about his soon-to-be bride and the fact that she is acting irrationally when it comes to wedding preparations (i.e. *a total Bridezilla*). Ross openly agrees and uses *standalone* *NLT* to stress his agreement with the description. His reiteration of Erika’s description is further reinforced by *we’re talking*, which, in effect, focuses and/or dramatically sets the ‘scene’ for the elided content: ([Sorcha is] *totally* [a Bridezilla]).

To examine the potential ‘loss’ of the over-the-top intensification of *NLT* that is distinctive to these type of intensifiers, a quantitative, longitudinal analysis of all *NLT+Intensifier Boosters* in CoROCK-T was conducted, which was compared to the

number of NLT occurrences produced in each book. As evidenced in Table 8.11. below, the majority of *NLT+Intensifier Boosters* occurrences are concentrated in the span between 2001-2005, which represent the height and last years of the Celtic Tiger, as well as the peak of NLT-use in the corpus.

Pattern	Year of Publication									
	2000	2001	2003	2005	2005 (CIDN)	2008	2009	2013	2016	2018
<i>NLT</i> occurrences	10	5	10	6	10	2	5	2	1	2
<i>+Intensifying</i> <i>Boosters</i>	5	2	6	12	0	0	0	0	0	1

Table 8.11. Distribution of NLT and NLT+intensifying booster items occurrences across the CoROCK-T texts by year of publication.

It is worth noting that, with the exception of 2 occurrences in 2018, there are no *NLT+Intensifier Boosters* colligations from 2008 onwards, which is also the time at which the number of NLTs drops dramatically (as already discussed in section (8.6.1.)). As mentioned previously, 2008 also corresponds with the fall of the Tiger and start of the economic crisis, and could be taken as indicative of a drop in popularity among speakers. Finally, if, as posited in the case study of NISo in section (8.5.2.), we understand that the recurrence of these intensifying booster patterns may signal the *delexicalization* and loss of *over-the-top* intensification, then the prominence of these patterns between 2001-05 in CoROCK-T could 1) indicate its *delexicalization*, its b) loss of *over-the-top* intensification, and its c) fall from popularity in the context of Dublin portrayed in the books. If this is correct, then it is logical to hypothesize that NLT was already in use and enjoying perhaps much more popularity in Ireland in the years prior to 2000. However, contrastive analyses should be undertaken against corpora of spoken Dublin English documenting its use in this variety prior to this date to validate this hypothesis.

Just like NISo (see section (8.5.2.)), the *over-the-top* intensification value of NLT is further reinforced in the books through orthographic, prosodic rendition. A quantitative analysis of all occurrences of NLT in CoROCK-T indicates that 21 of the 52 occurrences received overt prosodic stress, with a very balanced distribution between *capitalization*

and *italicization* (10 and 11 occurrences respectively), as opposed to 31 occurrences which were devoid of typographic stress. As argued in (8.5.2.), the fact that 40.4% of the occurrences of NLT in the corpus do receive typographic, prosodic stress highlights the fact that Howard purposely calls the reader's attention to NLT's distinct phrasal stress which is a consequence of its *over-the-top* intensification value. Examples 8.41. through 8.43. below exemplify each type of prosodic rendition in the corpus.

- (8.41.) He goes, 'You must be totally bummed out.' I'm like, '**TOTALLY.**' (OMFY 2003)
(8.42.) 'Oh my God, I heard JP's *totally* flipped out' (CIDN 2005)
(8.43.) 'It's, like, oh my God, you have no idea how much of a shock that was to my system? But I *totally* over-reacted and I'm sorry.' (RHINO 2009)

A longitudinal analysis was, then, conducted to examine the distribution of the prosodically stressed occurrences across the texts. The findings, which are illustrated in Fig. 8.13. below, indicate an uneven distribution of typographic stress rendition across the corpus, with most being concentrated in the same timespan: i.e. 2001-08. Contrastively, *unmarked* NLTs are largely present across all books. In fact, from 2009 onward all NLT occurrences are unmarked. It must be pointed out that the trajectory of over-the-top intensification rendering in text follows that evidenced in the analysis of NISo (see section (8.5.2.)), in that *capitalization* seems to be the most common type of typographic stress-rendering tool in the initial books, disappearing in 2005 and being surpassed by *italicization* in subsequent years. Finally, *italicization* disappears by 2009, with *unmarkedness* becoming the sole manner of NLT-stress representation in CoROCK-T.

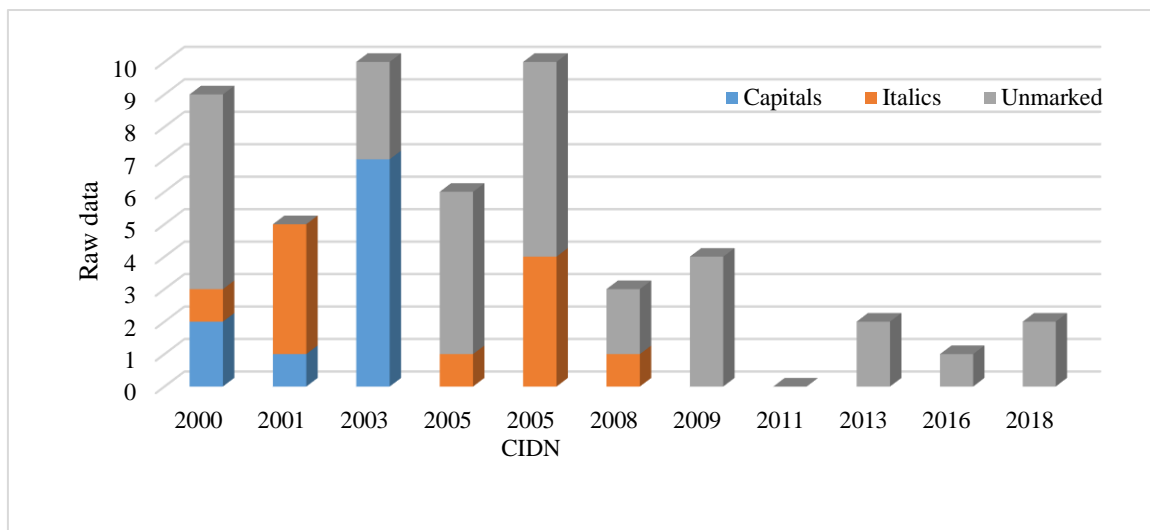


Figure 8.13. Distribution of NLT prosodic representation in CoROCK-T by year of publication.

The fact that explicit and variant typographic stress representation is more recurrent from 2001-03 and begins to steadily disappear from then on could be due to different reasons. On the one hand, this could imply that readers in 2001-03 still needed reminding of the ‘novelty’ and required stress of this *over-the-top* intensifier as opposed to its use as lexicalized *totally* or as standard *totally*. On the other hand, if the hypothesis that the recurrence of *NLT+Intensifier Boosters* outlined above in the first decade evidences its loss of intensifying value and fall from popularity is correct, then the elevated use of typographic emphasis in 2001-03 and steady disappearance thereafter may represent a change in authorial perception. This would entail that, as already argued in (8.5.2.), Howard may have perceived NLT to be a feature with a use so entrenched among Dubliners by 2009 that overt, stress representation was no longer required. The following section explores the pragmatic multifunctionality of NLT in CoROCK-T, examining its emotional conveyance in detail.

8.6.3. The pragmatic functionality of NLT in CoROCK-T

Given the speaker-orientedness of NLT intrinsic to *over-the-top* intensifiers (already explained in section (8.3.)), an examination of its pragmatic connotation conveyance was undertaken. The quantitative analysis of the connotation conveyance through NLT in the

corpus outlined in Table 8.12. below indicates that *neutral* meanings are the most recurrent ones, which mirrors findings from the analysis of the pragmatic functionality of NISo (see section (8.5.3.) for a discussion).

Positive	Neutral	Negative
3	34	15

Table 8.12. Distribution of NLT connotation conveyance in CoROCK-T.

The prominence of *neutral* connotations is also evidenced in the findings gathered from the quantitative, longitudinal analysis of pragmatic connotation conveyance distribution per book summarized in Table 8.13. which illustrate their consistent occurrence throughout the texts. Contrastively, and unlike NISo, which showed a balanced distribution of *positive/negative* emotions (see section (8.5.3.)), *positive/negative* connotations only make brief appearances in certain books. For example, *negative* emotions, which are the second, most transmitted ones, were most recurrent in the first part of the decade (i.e. 2001-08). On the other hand, *positive* emotions appear sporadically in the latter books (i.e. 2009 and 2018).

Connotation	Year of Publication									
	2000	2001	2003	2005	2005 (CIDN)	2008	2009	2013	2016	2018
<i>Neutral</i>	4	4	6	6	7	1	2	2	1	1
<i>Negative</i>	5	1	4	0	3	2	0	0	0	0
<i>Positive</i>	0	0	0	0	0	0	2	0	0	1

Table 8.13. NLT Connotation conveyance distribution across CoROCK-T books.

With regard to *neutral* meanings, previous studies (outlined in section (8.3.2.)) describe NLT as an intensifier which may show 1) *speaker attitude and degree of commitment towards an assertion or proposition* (Irwin, 2014, p. 39; Potts, 2004; Beltrama, 2015a, p. 137; Beltrama & Staum Casasanto, 2017, p. 162), while it may also function, when in *standalone* form (Quaglio, 2009, pp. 98-143), as a tool to show 2) *emphatic agreement*

with previous content, or to 3) give an affirmative response. The qualitative and quantitative analyses of all *neutral* occurrences of NLT illustrated in Table 8.14. below reveal the existence of all of the already documented sub-functions in CoROCK-T, which underscores the high level of realism (i.e. RQ 1b) with regard to the fictional representation of NLT offered in the texts.

Function	Occurrences
<i>Speaker emphasis</i>	21
<i>Emphasized agreement</i>	9
<i>Affirmative response</i>	3

Table 8.14. NLT pragmatic functionality and distribution of occurrences in CoROCK-T.

As illustrated above, *speaker emphasis* (i.e. showing the speaker’s level of commitment to their own utterance) is the most frequent *neutral* function in CoROCK-T in cases such as 8.44. below. Here, Honor asks her father, *Ross O’Carroll-Kelly*, if he knows the whereabouts of an electric heater which could be used in trial to charge her grandmother with the murder of her late, second husband. Although Ross has it stored in a safe place, he pretends he does not know. Honor, however, knows he is lying. Her use of NLT in 8.44. emphasizes her ‘suggestion’ that he use the heater to extort his mother.

(8.44.) ‘Do you know where it is? [...] Because if you did have it [...] you could, like, totally blackmail her’ (DwT 2018)

The second, most transmitted meaning is *emphasized agreement*. It is important to note that this was only ever conveyed through the use of anaphoric *standalone* NLTs in cases like 8.45. below. After having given his phone number to a girl he had a one-night stand with, JP jokes that he is a danger. All of his friends strongly agree with that description, as is exemplified by the use of anaphoric, *standalone* NLT, which further shows extra emphasis through *capitalization*.

(8.45.) 'I'm a danger to myself and to wider society,' and we're all like, 'Totally.' (PSSB 2005)

Finally, *affirmative responses* may also be transmitted through *standalone* NLTs, as in example 8.46. where Christian assumes that Ross is upset over his ex-girlfriend having traveled to Australia where she now lives with her new boyfriend. Ross simply confirms this with a NLT, which functions as a strong affirmative response that is further reinforced through *capitalization*.

(8.46.) 'You must be totally bummed out.' I'm like, 'TOTALLY. By the way, my congratulations.' (OMFY 2003)

The remaining *positive/neutral* meanings of NLT were further classified for emotion transmission following the emotive catalogue used in the examination of NISo (outlined in detail in section (8.5.3.)). In terms of *positive* meanings (only 3 occurrences), CoROCK-T indicates that characters use *positive* NLT to transmit *emphatic admiration*, in cases like 8.47. below, where Sorcha profusely praises her mother-in-law's fashion style.

(8.47.) 'She was wearing this amazing Galliano babydoll dress. It was, like, lace? And, oh my God, she totally pulled it off...' (RHINO 2009)

Finally, the range of *negative* meanings was found to be wider, which also mirrors the findings from the study of NISo (see section (8.5.3.)). Since, to my knowledge, there are no studies that delve into the types of *negative* emotional meanings that may be transmitted through NLT, I endeavored to provide a list of its emotive range in IrE as represented in CoROCK-T. It must be mentioned that these meanings were distributed into two categories: 1) *Negative Speaker Emphasis* (i.e. emphasized *negative* meanings expressed by the speaker), and 2) *Negative Emphatic Agreement* (i.e. *negative* meanings expressed by interlocutor which the speaker emphatically agrees with). As outlined in Table 8.15., *Negative Speaker Emphatic* emotions are slightly more frequently conveyed

in the corpus than *Negative Emphatic Agreement* emotions (with 9 and 6 occurrences respectively).

<i>Negative Speaker Emphasis</i>		<i>Negative Emphatic Agreement</i>	
Emotion	Occurrences	Emotion	Occurrences
<i>Dislike</i>	3	<i>Outrage</i>	4
<i>Anger</i>	2	<i>Criticism</i>	1
<i>Criticism</i>	3	<i>Dislike (of someone)</i>	1
<i>Disbelief</i>	1		
<i>Outrage</i>	1		

Table 8.15. Distribution of occurrences by NLT *Negative Speaker Emphasis / Negative Emphatic Agreement* emotion conveyance in CoROCK-T.

The quantitative and qualitative study of all *Negative Speaker Emphasis / Negative Emphatic Agreement* emotions outlined in Table 8.15. above indicates that the most frequent *Negative Speaker Emphatic* emotion conveyed is *dislike*, with *outrage* being the most recurrent *Negative Emphatic Agreement* emotions, in cases like 8.48. and 8.49. respectively.

(8.48.) *Oh my God, Bewley’s just, like, totally attracts the wrong type of people to Grafton Street.* (CIDN 2005)

(8.49.) ‘*Oh my God. That is like, OH! MY! GOD!*’ and Sophie goes ‘*TOTALLY. It is, like, SO not a cool thing to do.* [...]’ (OMFY 2003)

Fragment 8.48. exemplifies *dislike* well. Bewley’s is a Dublin café, which has sometimes been thought of as a city ‘landmark’ set in one of the capital’s main shopping streets. Earlier in the exchange, Ross O’Carroll-Kelly cynically expresses his profound dislike of the café’s customers, derogatorily describing them as “boggers¹⁰³, septic tanks and coffin-

¹⁰³ Derogatory terms. *Bogger* is “an unsophisticated person, often from the countryside or simply from outside of Dublin” (Howard, 2007a, p. 223), while *septic tanks* is rhyming slang in the *Ross O’Carroll-Kelly* universe for ‘Yank’ or “a citizen of the United States of America” (*ibid.*, p. 246). *Coffin Dodgers*,

dodgers” (Howard, 2005, p. 285), with Sorcha publicly calling for the opening of a Starbucks in its place. Immediately after, Aoife, the speaker in 8.48., supports Sorcha’s idea, mentioning and emphasizing through NLT her derisive opinion of the customers (i.e. they are the “wrong type of people”). It is clear by this exchange that these speakers dislike Bewley’s because they associate it with a specific type of identity (i.e. unsophisticated, provincial, and lower/working class) which heavily contrasts with the identity these Southside Dublin speakers associate with (i.e. cosmopolitan, globalized, and upper-class). For more on the stylization of North/Southside Dubliners in the *Ross O’Carroll-Kelly* series, see sections (8.2.) and (7.5.1.).

An example of *outrage* is found in 8.49. above. Here, Aoife has just learnt that Ross’ mother was attacked on the street with paint by an animal-rights activist when she was walking out of a fur shop. She expresses her *outrage* in the form of two *Oh my God*’s, the second of which is *capitalized* and separated for further emphasis. Sophie shows her *emphatic agreement* with Aoife’s *outrage* by using anaphoric *standalone* NLT, which is doubly emphasized through *capitalization* as well. Her own *outrage* is also evident in the coda of the sentence which includes a NISo (“SO not a cool thing to do”). The following section will investigate the potential speaker identity indexation value the use of both NISo and NLT may have in CoROCK.

8.7. Identity indexicality of NISo and NLT in CoROCK

In order to answer RQs 2 and 3 in this thesis (see section (1.2.)) pertaining to the potential speaker identity both NISo and NLT may index in the *Ross O’Carroll-Kelly* series as represented in the CoROCK texts, both intensifiers were examined for sociodemographic indexical value in detail.

With regard to number of occurrences of NISo and NLT produced by social class and geographic region, a quantitative analysis reveals that both intensifiers are only produced by upper-class, Southside Dublin characters in CoROCK. According to Howard’s rendition, the reader will, therefore, perceive the use of NISo and NLT in the corpus as an identifying feature within the *non-local* Dublin English linguistic repertoire, which

however, is a derogatory slang term for elderly people, perceived to be a nuisance, and to be ‘avoiding’ death.

coexist with other Southside Dublin-bound items already explored in this thesis, like the use of non-standard quotatives (see chapter (5) for a detailed analysis of quotative GO and BE LIKE), discourse pragmatic marker, clause medial *like*, or *like+high rising terminal*, among many others. Given the common stereotypical (mis)associations NISo and NLT have in the USA, being popularly connected, as already discussed at length in section (7.5.1.), with the stereotypes of the *Valley Girl/Jock-Surfer Boy*, it is possible that Howard uses NISo and NLT as a tool with which he stylizes the Southside Dubliners' identity as one of *perceived* linguistic and socioeconomic prestige, as well as linguistic 'coolness' and 'cosmopolitanism'. The use of these intensifiers, which are so heavily connected with the distinctive American stereotypes, coupled with the non-linguistic stylization of the North/Southside stereotypes offered in the books which was discussed at length in section (8.2.), may also allow Howard to stylize the stereotype of the cosmopolitan, linguistically trendy, *Southside Socialite/Rugby Jock-D4 Head* personae (see Table 7.4. in section (7.5.1.) for a more detailed description and distinguishing traits of each stereotypes). The fact that Northsiders in the corpus do not use NISo nor NLT but stick to using more locally found features of IrE which are distinctive to *local* Dublin English, such as the use of tag *so* in cases like, 'It's lucky Ross was there, *so it was.*' (Howard, 2013) instead of NISo is significant. This further reinforces the fact that both NISo and NLT are used in the books to 1) create a linguistic divide between the wealthier, linguistically 'trendy' Southsiders who use global features, and the *local* Dublin English-speaking Northsiders, who retain regionally bound items, making their identity much more locally-restricted. Thus, in stylizing the linguistically stereotypes *Southside Socialite/Rugby Jock-D4 Head* on paper through the representation of NISo and NLT, Howard ensures that his readers automatically identify who the characters are, while at the same time ensuring that they are aware of the perceived sociocultural differences that separate the Southside/Northside cast of characters.

In terms of age, the link that the literature identifies between the use of NISo and NLT and young speakers in other varieties (see section (8.3.) through (8.3.2.) for an overview) seems to be mirrored in the CoROCK data which is summarized in Table 8.16. below, with both over-the-top intensifiers appearing to index more frequently the voice of young speakers in the 20-30 age cohort. The fact that these findings not only corroborate the literature, but also correspond with previous findings in the context of fictionalized IrE (see Amador-Moreno & Terrazas-Calero, 2017, 2022) highlights the level of authenticity

(i.e. RQ 1a) present in the fictional rendition of NISo and NLT use offered in the CoROCK texts, and identifies these items as potential indexes of young voice in the context of Dublin.

	AGE			GENDER	
	Preteen ¹⁰⁴	20-30	60+	Male	Female
<i>NISo</i>	10	130	1	42	99
<i>NLT</i>	16	36		25	27

Table 8.16. Distribution of NISo & NLT occurrences by age and gender across CoROCK.

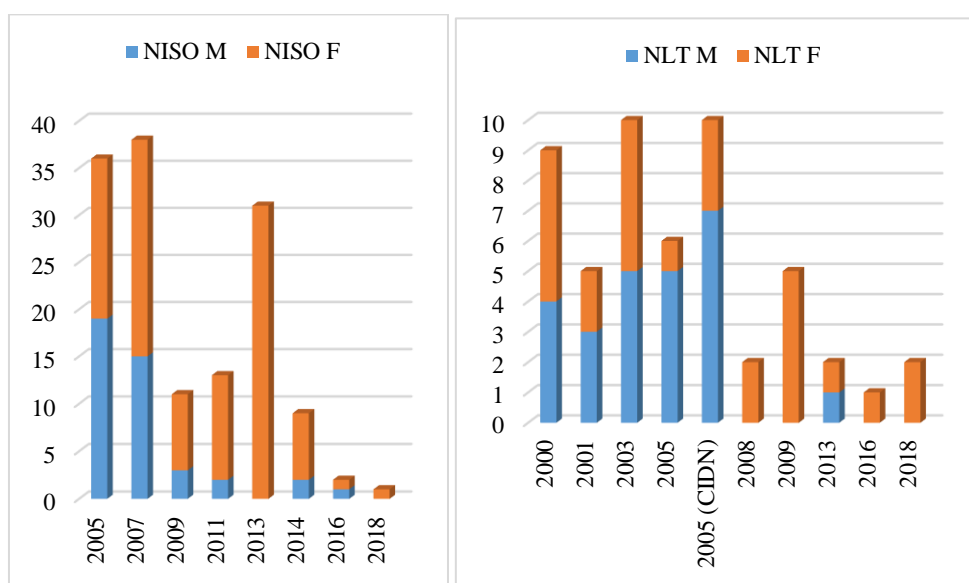
In terms of gender production, the quantitative analysis of the occurrences of NISo and NLT in CoROCK summarized in Table 8.16. above highlights the connection that appears to exist between these intensifiers and female characters in Howard’s fictional world. This matches findings drawn from fictionalized IrE (Amador-Moreno & Terrazas-Calero, 2017, p. 265) but also from other varieties (see Tagliamonte & Roberts, 2005, and Zwicky, 2011, for American English, and Tagliamonte, 2005, 2008, for Canadian English. For more on the gender associations of NISo and NLT in other varieties, see sections (8.3.1.) and (8.3.2.)). Nevertheless, it must be noted that the characters of *Ross O’Carroll-Kelly* and his wife, *Sorcha O’Carroll-Kelly*, are responsible for producing a large percentage of both NISo (88% and 73.7% respectively) and NLT (64% and 33.33% respectively) in the subcorpora as illustrated in Table 8.17 below.

¹⁰⁴ All cases of preteen use are produced by one character, *Honor O’Carroll-Kelly*.

NISO				NLT			
Male		Female		Male		Female	
42		99		25		27	
<i>Ross O'Carroll-Kelly</i>		<i>Sorcha O'Carroll-Kelly</i>		<i>Ross O'Carroll-Kelly</i>		<i>Sorcha O'Carroll-Kelly</i>	
37	73	16	9				
Other Males		Other Females		Other Males		Other Females	
5	26	9	18				

8.17. Raw distribution of occurrences of NISO & NLT by gender across CoROCK-SO and CoROCK-T.

If Ross' and Sorcha O'Carroll-Kelly's production of NISO and NLT is excluded from the analysis, female characters continue to produce substantially more occurrences of both intensifiers than their counterparts. In answer to RQs 2 and 3, then, this may suggest that both NISO and NLT could function as indexes of female discourse in the context of Dublin that is present in the books. To explore this seeming gender preference, longitudinal analyses illustrated in Figs. 8.14. and 8.15. below were conducted on the distribution of NISO and NLT occurrences by gender across the subcorpora.



Figures 8.14. and 8.15. Raw distribution of NISO and NLT by gender across CoROCK-SO and CoROCK-T respectively.

With regard to NISo use, the findings outlined in Fig. 8.14. above indicate that while there seemed to be a balanced distribution of occurrences among both genders in 2005, with male characters using it slightly more often than females, by 2007 that balance tips to females. There is also a noticeably decreasing number of male-produced occurrences from 2009 onwards, with NISo disappearing completely in 2018 from male speech in CoROCK-SO. Meanwhile, female characters continue to use it recurrently, peaking in 2013 (which heavily contrasts with the complete absence of male-produced NISo occurrences that year). This appears to be indicative of a clear development, underscoring the progressive association of NISo with female voice in *non-local* Dublin English as portrayed in the corpus, despite such connection being inexistent in earlier *Ross O'Carroll-Kelly* books (for more on this, see Amador-Moreno & Terrazas-Calero, 2017, p. 265). A similar progression appears to take place in the use of NLT in CoROCK-T. Despite the few occurrences of NLT in this subcorpus, Fig. 8.15. also evidences a rather balanced distribution between both genders during the first half of the decade, at least as represented in the CoROCK-T texts. While we are analyzing very low frequencies, the balance also appears to tip in favor of female characters by 2008/09, like NISo, being almost exclusively used by female characters in the corpus thereafter. In response to RQs 2 and 3, the noticeable shift in terms of gender-preference evidenced by both NISo and NLT in the subcorpora, particularly after 2009, may perhaps suggest that by the time the Celtic Tiger ended, both *over-the-top* intensifiers had been adopted as indexical features of the female repertoire in Ireland (and Southside Dublin in particular), or at least that is what Howard's perception seems to indicate. The pronounced reduction of female NISo and NLT use illustrated in the later novels could perhaps respond to a) the natural progression of intensifiers, which tend to rise in popularity very rapidly and fall out of it just as quickly. However, it could also be due to the fact that b) Howard may be under the assumption that both intensifiers are now well known by his readers and so is phasing out its use.

Although a more in-depth contrastive analysis with real, spoken IrE should be undertaken in the future to compare findings, it is clear that the preference for NISo and NLT by female characters in the corpus coincides with the perception of use these intensifiers have in varieties like American English (Throne 2014, p. 445; Beltrama 2016; Reichelt & Durham 2017, p. 72), which further strengthens the level of realism (RQ 1a) offered in the CoROCK texts.

8.8. Concluding remarks

Although this thesis did not set out to look at one author, as explained in chapters (3) and (4), during the stages of annotation and analysis of CoFIrE, it became apparent that Howard's representation of orality was on a level that merited a case study. This chapter, therefore, set out to investigate the use of NISo and NLT in IrE fiction as represented in CoROCK because these two, non-traditional IrE, *over-the-top* intensifiers were found to be the most frequently reproduced ones in the corpus. Thus, and to supplement the answers to this thesis' RQs (section (1.2.)) drawn from the analyses of the three most reproduced pragmatic items in the *Corpus of Contemporary Fictionalized Irish English* (see chapters (5) through (7)), both NISo and NLT were explored in detail to examine their use, form, pragmatic multifunctionality, and potential identity indexation value.

The quantitative and longitudinal study of NISo frequency of production by year of publication (section (8.5.)) in CoROCK-SO indicates a progressive decrease, which seems to commence in 2009/11. This could be indicative the *supraregionalization* of this feature in the context of Ireland, and/or of the loss of its linguistic 'trendiness' status among Dubliners which would correspond with the natural tendency of intensifiers to quickly rise to popularity and fall out of 'fashion' just as fast.

In all, and according to Howard's perception and portrayal, 2009 seems to be a deciding year in the evolution of NISo in Dublin English with regard to its *delexicalization*, *lexical unrestrainability* and its establishment as an identifiable marker of *non-local* Dublin English identity (i.e. RQ 3). For example, the analysis of colligational pattern preference offered in section (8.5.1.) revealed a series of patterns, the most prominent of which is *NISo+Verb Phrase* (including *NISo+Negative Verb Phrase*), which is consistently used across all CoROCK-SO books. This disproves Amador-Moreno & Terrazas-Calero's (2017) hypothesis that this pattern could have been replaced by *NISo+Noun Phrase* in this series. However, and while featuring in most of the CoROCK-SO texts, the *NISo+Noun Phrase* pattern *not* consistently used throughout all. A more detailed qualitative study showed pattern variability, with NISo colligating *Nouns*, *adjectivized Nouns*, *positively adjectivized Nouns*, or *negatively adjectivized nouns*, which are cases where NISo appears to replace traditionally grammatical intensifiers like *such* or *very*. The longitudinal examination of these 'newer' patterns dates their first appearance in 2009, which could be indicative of the fact that that was the period

of time when, according to Howard's perception, NISo may have become lexically unrestrained and delexicalized in Ireland.

The analysis of frequency of occurrence (section (8.6.)) and colligational pattern preference (8.6.1.) of NLT in CoROCK-T drew similar findings. Thus, the longitudinal analysis of NLT occurrence production by year of publication suggests that the peak of NLT use in Ireland (and Dublin in particular) was 2001-05. This coincides with the height of the Celtic Tiger when linguistic items traditionally not associated with IrE, such as NLT and NISo, may have been utilized by speakers to project the notions of wealth and cosmopolitanism so closely connected to the Tiger (i.e. RQs 2 and 3). Just like the analysis of NISo, the CoROCK-T data also illustrates a marked decrease in use of NLT toward 2008-09, which runs parallel to the end of the Tiger and to the onset of its subsequent recession. This decrease could be explained due to Howard's wish to discontinue some characters' linguistic 'association' with the stereotypes connected with the use of both NISo and NLT (i.e. affluent, cosmopolitan, 'Valley Girl'-sounding).

With regard to colligation pattern-preference (section (8.6.1.)), the findings from the analysis of NLT coincide with the ones observed in real, spoken interaction in other varieties of English, highlighting the prominence of the *NLT+Verb Phrase* pattern, as well as recording the use of *standalone* NLT, and the *NLT+Noun Phrase* pattern. This would buttress the highly realistic portrayal of real language use that is offered in these texts. In addition, the corpus data also documents some developments in progress, which evidences the usefulness of the fiction dataset as a source with which to examine real linguistic developments (i.e. RQ 1a). Thus, and while NLT was found to colligate primarily with *NLT+Verb Phrase* patterns, the data also reveals cases NLT was predicted 'not to occur' with, such as with directives. *Standalone* NLT was also present in CoROCK-T, evidencing the geographic unrestrictedness of this intensifier, and refuting Quaglio's (2009) claim that it is innovatory of American English conversation. A more detailed analysis of *standalone* NLT revealed functions already documented in previous studies (i.e. *emphatic agreement* and giving an *affirmative response*), which attests to the realism of the fictional rendition (i.e. RQ 1a), but also identified a previously undocumented use: *lending emphasis/veracity to the speaker's previous statement*.

The rendition in text of the extra intensification *over-the-top* intensifiers like NISo and NLT carry was also examined in this chapter. Both were found to show *over-the-top* intensification in two manners: 1) through orthographic prosodic rendition, and 2) by

appearing in proximity to other intensifier ‘booster’ items. While 80.1% of NISo occurrences were found to be orthographically stressed, their longitudinal evaluation (section (8.5.2.)) suggests that *capitalization*, which was the preferred method of prosodic rendition in the first two CoROCK-SO novels, was replaced by *italicization* from 2007 onward, becoming the favored type of rendering by 2009. Similarly, 40.4% of the occurrences of NLT in CoROCK-T (see (8.6.2.)) also received overt prosodic stress, which highlights a purposeful call for the attention of the readers to the distinctive phrasal stress of this intensifier. Overt prosodic rendition of NLT, however, decreases from 2008 onward, when *unmarkedness* becomes the only form of representation of NLT in the texts. This coincides with the increase of *unmarkedness* NISo experiences from 2009 onward. While the overt rendition of the distinctive prosodic stress of both NISo and NLT evidenced at the beginning of both subcorpora could be a purposeful authorial effort at reminding the reader of the specific *over-the-top* intensification value of both intensifiers, the steady disappearance of the typographic forms of stress and the rise of unmarkedness may be due to 1) a shift in Howard’s perception, or to the fact that 2) by 2009, readers did not need reminding of their distinctive stress due to their familiarity with this intensifier.

With regard to the occurrence of NISo and NLT with/in proximity to other intensifying ‘booster’ items (sections (8.5.2.) and (8.6.2.)), the analyses uncovered a variety of booster colligates with NISo (*Oh my God, Fucking, Hello, high rising terminal*) and NLT (i.e. *Oh my God, like, NISo, I mean*) which appear to be female-distinctive in the *Ross O’Carroll-Kelly* series, according to the Howard’s portrayal. My contention is that while their colligation with these other intensifying boosters could be due to the characters’ familiarity with NISo and NLT, which may have lost their over-the-top intensification over time, this pattern could also attest to the *delexicalization* of NISo and NLT in the context of Dublin.

Pragmatically, the chapter finds *neutral* connotations to be the most frequently transmitted by NISo and NLT in CoROCK-SO and CoROCK-T respectively, with most subfunctions mirroring those identified in the literature for NISo (i.e. conveyance of *simple emphasis* or *amplified determination*), and NLT (i.e. *speaker emphasis, emphatic agreement, and giving affirmative responses*). This would validate the fictional representation (RQ 1a) of the pragmatic functionality of these intensifiers use offered in the books. The subcorpora also document pragmatic developments in process. For example, while NISo was found to convey an equal number of *positive* and *negative*

connotations, the former were transmitted consistently throughout the books, yet *negative* ones began to decrease dramatically after 2009. This could also suggest that 2009 would be the point in time when, according to Howard, this intensifier may have become an established feature in the Southside Dublin English repertoire (i.e. RQ 3). With regard to NLT developments, CoROCK-T records the conveyance of *negative* and *positive* emotions (i.e. *emphatic admiration*), the latter of which appears to be a rather recent acquisition given that this is only found in the final books.

Qualitatively, this chapter is novel in classifying all occurrences of NISo and NLT for emotion transmission, identifying a variety of *positive* and *negative* emotions which may be transmitted by these items. For example, the analysis of the type of emotions conveyed by NISo indicates that while *positive* emotions (i.e. *excitement, compliments, admiration, and sexual admiration*) are consistently used throughout all books, the *negative* emotion repertoire is far larger, with *anger* and *dislike* being the most conveyed emotive meanings.

Finally, RQs 2 and 3 are addressed through the examination of the potential identity indexation value of NISo and NLT according to their representation in CoROCK. The findings highlight both intensifiers as distinctive items to upper-class, young (often female) Southside Dubliners and to the *non-local* Dublin English repertoire in the series, and perhaps in real life. My contention is that the Southsiders' exclusive use of NISo and NLT may be two-fold. On the one hand, it could serve as a tool through which the Southsiders construct an identity of *perceived* cosmopolitanism and socioeconomic prestige which separates them from the more regional, less prestigious linguistic identity of the Northsiders. On the other hand, by infusing his Southside cast with the use of NLT, Howard may be attempting to create very identifiable characters which replicate the stereotypes of what I have come to label the *Southside Socialite-Rugby Jock/D4 Head* personae.

While all these findings provide illuminating answers to the RQs of this thesis, further studies should be conducted on the use of NISo and NLT in real, spoken Dublin English in order to further check the validity of their fictional rendition and to address the hypotheses outlined in this chapter.

9. CONCLUSIONS

This chapter begins with a review of the aims of the thesis, restating the thesis' research questions in section (9.1.). Section (9.2.) overviews the innovations this thesis has contributed to the field of IrE fictionalized representations, and the interdisciplinary nature of the thesis itself, which uses Corpus Stylistics as a main investigative framework supplemented by Sociolinguistics and Pragmatics techniques. The corpora that have been created to answer the RQs will also be discussed with regard to their innovative nature, as well as the novel and original annotation system designed and applied to the corpora. Section (9.3.) will summarize the findings from CoFIrE with regard to RQ 1 pertaining to the way IrE orality was represented in the corpus, while sections (9.4.) and (9.4.1.) will discuss the findings pertaining to the use, form, developments, and level of realism inherent to the fictional portrayal of the top three most recurrent pragmatic items in CoFIrE as well as of the two case studies conducted in CoROCK when contrasted against the *LCIE/BNC2014* samples. The pragmatic multifunctionality of the most frequently reproduced pragmatic items in both corpora will be explored in sections (9.5.) and (9.5.1.) respectively, while section (9.6.) will provide answers to RQs 2 and 3 pertaining to the potential identity indexical value of these items with regard to modern Irishness as represented in the corpora. The limitations of this study will be overviewed in (9.7.). Finally, section (9.8.) will offer future directions.

9.1. Review of thesis' aims and objectives

Taking fictionalized (literary) dialect-dialogue as valid source of linguistic use, this thesis set out to investigate the use of fictionalized IrE in an original corpus of contemporary Irish fiction with the goal of investigating not only the level of realism inherent to the fictional portrayal, but also to examine the potential type of modern Irishness the most frequently reproduced (pragmatic) features may encode when represented in fiction. Thus, the thesis revolves around three main research questions:

1. Which are the most frequently (re)produced features of spoken IrE in contemporary IrE fiction?

- a. How ‘authentic’ or valid is the fictional representation when contrasted against real, spoken IrE and British English?
2. What does the (re)production of these features in fiction index with regard to speaker identity (i.e. age, gender, geographic location, and social class)?
3. How is modern ‘Irishness’ indexed in contemporary IrE literature through the use of pragmatic items in fictional dialogue?

As outlined above (and in section (1.2.)), RQ 1 aims at determining the most prominently reproduced linguistic items of IrE orality in a corpus of contemporary Irish fiction. To answer this question, I compiled the first corpus of *contemporary* IrE fiction: the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE). Given the artificial nature of fictionalized (literary) dialect-dialogue, RQ1 led to the development of a sub-question (i.e. RQ 1a) which focused on testing the level of authenticity or realism of the fictional representation by contrasting it against natural uses. As explained in section (3.8.), the corpora selected for comparison were the *Limerick Corpus of Irish English* and the *BNC2014*. In order to compare the fictional data against the corpora of real, spoken IrE and British English use, randomized samples of 100-item from each corpus were utilized. The process of selection, methodology of comparative analysis and method of sociolinguistic cross examination are explained at length in sections (3.8.) through (3.8.2.). The second research question of this thesis (RQ 2) is concerned with exploring the type of speaker identity these items indexed in the texts. Similarly, RQ3 aimed at analyzing what type of modern Irishness may be encoded through the use of pragmatic items in fiction given the fact that the literature (see Barron, 2017; Hickey & Amador-Moreno, 2020, p. 15) identifies pragmatic features of IrE as distinctive and crucial to the construction of “Irishness”. The findings regarding each RQ are discussed at length in sections (9.3.) through (9.6.) below.

9.2. Novel contributions of the thesis to the study of fictionalized IrE representations

The survey of the literature which is offered in chapter (2) highlighted the fact that while much work has been conducted on individual (often grammatical/lexical) features of IrE, and/or on individual author's style, there is a considerable dearth of Corpus Stylistics research (see (2.4.5.)) regarding the study of fictionalized IrE in a large corpus of *contemporary* fiction. Furthermore, the number of studies looking into pragmatic features IrE and their potential identarian indexical value in a large corpus is even fewer. This thesis is, therefore, innovative with regard to the methodological frameworks it uses, taking mostly a Corpus Stylistic approach to the study of the RQs. However, to contribute to the growing field of Variational IrE Pragmatics and to the study of present-day portrayal of Irishness in fiction, this thesis also makes use of Corpus Linguistics, Sociolinguistic, and Pragmatic analytic methods (for a detailed explanation of what type of analytic methods of each discipline are used here and which are not, see section (1.3.)).

Another contribution of the thesis is the compilation of two original corpora of *contemporary* fictionalized IrE. Given the lack of existing corpora (outlined in section (3.1.)) which comprise *contemporary* IrE fiction or which fulfilled the requirements of the thesis (see (3.2.) through (3.2.2.) for a detailed account of the design criteria used in CoFIrE to ensure its maximum utility), a new corpus had to be compiled to answer the RQs: the *Corpus of Contemporary Fictionalized Irish English* (CoFIrE). In total, CoFIrE contains 1,123,601 words and comprises 16 works of fiction published in the Republic of Ireland by 8 authors (an overview of each author and the books included is offered in sections (3.4.) through (3.4.8.)). Another novelty offered in this thesis is the novel and original system of annotation that was designed and implemented in CoFIrE (explained at length in sections (3.7.) and (3.7.1.)). This system has proven to be of invaluable use, as it facilitated a faster retrieval of pertinent data. Thanks to this system, faster, more rigorous, quantitative and qualitative analyses could be conducted on issues such as frequency of occurrence, use, form, pragmatic functionality, and potential identity indexicality value of the features explored in the thesis (these shall all be summarized in (9.3.) through (9.6.) below).

The annotation system was also instrumental in highlighting Paul Howard as an outlier within the CoFIrE author group. At corpus annotation and analytical stage, his two

Ross O'Carroll-Kelly novels were found to be the books that contributed the largest amount of pragmatic items to CoFIrE (see (9.6.) below for the summarized conclusions regarding the indexation of modern Irishness through the CoFIrE and CoROCK pragmatic items). Given the fact that the *Ross O'Carroll-Kelly* series has been praised for its detailed chronicling of modern Ireland and has garnered critical acclaim for its depiction of IrE orality, it was considered that this series merited a detailed study which would supplement the answers to the thesis' RQs gathered from CoFIrE. This led to the compilation of the first *Ross O'Carroll-Kelly Corpus* (CoROCK), which is explained at length in section (3.9.). New Intensifying *So* and Non-lexical intensifier *Totally* were selected for analysis given their high frequency as *over-the-top* intensifiers in CoROCK (as discussed in (8.1.)). Neither of these items has traditionally been associated with IrE, yet they are rather popularly perceived as belonging to the American English repertoire and with the construction of the negative stereotypes of the *Valley Girl* and *Jock/Surfer Boy* in particular (see (7.5.1.) for an overview of these). Thus, their detailed study in CoROCK aided the answering of RQ3 regarding the potential existence of identarian developments and/or the indexation of 'new' Irishness through pragmatic items as portrayed in fiction, and in the *Ross O'Carroll-Kelly* series in particular.

The following sections will discuss the findings regarding each individual RQ gathered from both CoFIrE and CoROCK, as well as will include comments on other innovations this thesis contributes to the field of fictionalized renditions of IrE.

9.3. RQ 1: Orality reproduction in CoFIrE

To answer RQ1, on the most frequently (re)produced features of spoken IrE in contemporary IrE fiction, CoFIrE was first analyzed quantitatively via the searches for annotated items. From this analysis, 202 individual items were identified with a total of 19,569 occurrences. This represents 1.74% of the total corpus word count, which is a considerable amount for a small, specialized corpus, and which, I believe, underscores the value of CoFIrE as a rich source for linguistic investigation.

The quantitative analysis of CoFIrE by frequency of occurrence of the items (see section (4.1.)) also revealed that while all books contribute more or less occurrences of features, a number of books (i.e. Ruane's *Tales in a Rearview Mirror*, Ryan's *The*

Spinning Heart and *The Thing about December*, and Howard's *Ross O'Carroll-Kelly* novels) were found to have contributed the most amount of items and occurrences. These books, which have garnered acclaim for their portrayal of IrE orality, are all narrated in the first person, and begin with either an implicit (in the case of Ruane's and Ryan's books) or explicit premise (i.e. Howard's novels are prefaced "as told to Paul Howard") that the narrated events are rooted in reality. Thus, I propose that the elevated amount of IrE distinctive features as well as the inclusion of pragmatic items which are not necessarily traditionally distinctive to IrE but common across varieties (e.g. *fuck*, *like*, *be like*, etc.) contained in these books is multifunctional. On the one hand, this represents 1) a conscious, authorial, stylistic choice which 2) builds the characterization in the books, and which is 3) grounded in their claim to orality. On the other hand, the faithful portrayal of IrE speech offered in these books, therefore, would also 4) legitimize the veracity of the story, using speech portrayal as a tool akin to the storytelling claim "based on a true story/real events". Finally, this would subvert any disbelief the readers may have regarding the (in reality) fictitious events and characters in the books.

Quantitative analyses of the three linguistic categories (grammar, lexis, pragmatics) offered in sections (4.2.) and (4.3.) revealed an abundance of pragmatic elements in CoFIrE. As mentioned above, more detailed investigations identified Howard's *Ross O'Carroll-Kelly* novels as responsible for contributing the largest amount of pragmatic occurrences to CoFIrE, which is perhaps to be expected given this series' premise (i.e. first person narration by Ross) as well as the longer length of the books when contrasted against other texts in the corpus (i.e. Sinclair (2005) points out the possibility that including longer texts may unbalance a corpus). This prominence of pragmatic items in the *Ross O'Carroll-Kelly* books also led to the compilation of CoROCK. Despite the over-representation of features of IrE orality in the *Ross O'Carroll-Kelly* novels, if their pragmatic contribution were discarded from analysis, pragmatic items remain the largest category in CoFIrE. This lends credence to the distinctiveness of the IrE pragmatic profile proposed by Barron (2017), which Hickey & Amador-Moreno (2020) present as a key element through which Irish identity is indexed. In addition, the prominence of pragmatic elements in CoFIrE may also signal a shift in the representation of the Irish through speech in literature which deviates from traditional literary renditions (embodied in the linguistic stereotype of the Stage Irishman explained at length in section (1.4.)) which centered on reflecting the use of mostly grammatical and lexical items in literature (see

(9.6.) below for a discussion of the findings regarding the potential indexicality of modern Irishness through pragmatic items in CoFIrE and CoROCK). Furthermore, the dominance of pragmatic elements could also underscore a stylistic shift among authors who 1) emphasize the ‘reality’ value of their books by utilizing the representation of these items, which are frequently found in every-day interaction. In prioritizing pragmatic features, the authors also 2) document linguistic developments in the pragmatic repertoire (see sections (9.5.) and (9.5.1.) below), thus 3) potentially indexing contemporary Irishness through fictional speech (as will be discussed in (9.6.)).

The value of pragmatic items, coupled with the traditionally neglected study of IrE pragmatics, led this thesis to contributing to the growing field of IrE Variational Pragmatics by focusing on analyzing the three most frequently reproduced pragmatic items: non-standard quotative GO (chapter (5)), intensifying *fucking* (chapter (6)) and clause-medial, discourse pragmatic marker *like* (chapter (7)). Although, as mentioned before, this thesis focuses on exploring pragmatic items, the fact that grammatical and lexical features were also annotated leaves the door open for future synchronic studies to explore their form and use, as well as their potential identity indexical value. Other studies could also conduct longitudinal, diachronic analyses against older publications to check for potential linguistic, pragmatic, and identarian developments.

9.4. RQs 1 and 1a: Use, form, developments, and realism in CoFIrE

Having already answered RQ 1 by identifying the use of non-standard quotative GO, intensifying *fucking* and clause-medial discourse pragmatic marker *like* as the most frequently (re)produced, pragmatic features of spoken IrE in contemporary IrE fiction, the thesis, then, explored their use, form, pragmatic functions, and indexicality in detail. Comparative analyses against *LCIE* and *BNC2014* random samples were conducted so as to answer RQ 1a pertaining to the level of realism the fictional portrayal of their use may have when contrasted against real use.

Chapter (5), therefore, examined in detail the use of non-standard quotative GO, also analyzing the use of the second most recurrent non-standard quotative: BE LIKE (which is explored in sections (5.9.) through (5.9.5.)). The thesis contends that the preference for these innovative non-standard ‘ventriloquizing’ devices over standard

quotatives enables the CoFIrE authors to include the reader into the narration, strengthening the (fantasy) bond of familiarity that exists between author-narrator-reader, which is further reinforced through the use of the second and third more frequently reproduced pragmatic markers in the corpus: *fucking* and *like*. The study of GO with regard to verbal tense coincides with the literature, highlighting the dominance of the *Historical Present Tense*, which was also mirrored in the *LCIE* sample (with the *Present Tense* also featuring prominently in the British sample). I believe this highlights the realism of the fictional portrayal, validating CoFIrE as a good source for the representation of real, linguistic use.

Chapter (5) also lends credence to this thesis' contention that fiction corpora are good sources for the analysis of linguistic developments. For example, the contrastive study of CoFIrE against the *LCIE* sample appears to indicate a morphosyntactic development regarding *grammatical subject preference* whereby GO shifts from reproducing male (in *LCIE* sample) to female voice (in CoFIrE). Further validation of the usefulness and highly realistic portrayal of GO-use in the fiction corpus is provided by the study of the pragmaticity of GO in CoFIrE, which reveals that it functions as a tool for the construction of dialogue in a dramatized, animated manner. This a) infuses the dialogue with a sense of immediacy (achieved through the use of *Historical Present Tense*), and b) can create intimacy between narrator-reader, c) making readers feel as though they are part of the scene. CoFIrE has also shed light on the acquisition of 'new' pragmatic contexts in GO, such as *reporting imagined speech* and *reconstructing written content*, which are present in the *BNC2014* sample, yet absent from the *LCIE* sample. I believe the fact that these uses which are present in CoFIrE also appear in the British sample could indicate that the fiction corpus documents the *grammaticalization/pragmaticalization* of this quotative in the context of Ireland, which may have begun after the compilation of *LCIE*, thus further evidencing the high level of realism inherent in the fictional rendition of GO.

The exploration of intensifying *fucking* in CoFIrE offered in chapter (6) also indicates the usefulness of the fiction corpus, registering the majority of *colligation patterns* previously documented in the literature in other varieties. The fact that they are also found in the contrastive samples answers RQ 1a by highlighting the realistic nature of the fictional portrayal in CoFIrE. The high level of authenticity is further reinforced by the fact that the fiction corpus shows a preference for the *Intensifying Fucking+Nominal*

Phrase pattern, followed by *Intensifying Fucking+Verb Phrase*, which were also found to be the dominant patterns, in that order, in the *LCIE/BNC2014* samples. A novel finding gathered from CoFIrE (and the samples) was the occurrence of intensifying *fucking* in cases (e.g. *+possessive pronouns*) which Mackenzie (2019) proposed as ‘predicted not to occur’ and ‘cannot occur’ in other varieties of English. While it cannot be assumed that their presence in the fiction corpus and random contrastive samples examined here indicates their occurrence is ‘new’ or restricted to Ireland, it is possible that the highly oral component of CoFIrE and the samples may explain the wider variety of patterns. However, their presence in CoFIrE and in the contrastive samples further validates the value of the fiction corpus as a representative source of linguistic use and developments.

Finally, chapter (6) also contributed to examining the largely under-research subject of *tmesis* in IrE and other varieties (with the exception of Mackenzie’s (2019) study, which also looks at other variants of the lemma FUCK) by providing a detailed classification of the types of *tmesis* present in CoFIrE. Thus, the occurrences of *tmesis* were separated into *standard tmesis* (i.e. splitting 1 word into 2 units) in cases like *un-fucking-believable*, and *non-standard tmesis* (e.g. *Merry fucking Christmas*). All occurrences of *tmesis* were, then, examined for non-standard colligation contexts in CoFIrE. Although *tmesis* was virtually non-existent in the contrastive samples (1 and 3 occurrences in the *LCIE* and *BNC2014* samples respectively), this could, perhaps, be due to the emerging nature of this feature at the time of compilation of *LCIE*, which would have developed later in time, as recorded by CoFIrE. While generalizations cannot be made due to the small (100 item) size of the samples from *LICE* and *BNC2014*, their presence warrants future research into their use and pragmatic functionality of *tmesis*.

The highly authentic/realistic of representation of IrE use offered by the CoFIrE texts with regard to documenting linguistic developments was also evidenced in chapter (7). Here, the quantitative and qualitative analyses of discourse pragmatic marker *like* appeared to signal a shift in terms of *clause-position preference* from traditionally IrE-distinctive clause marginal to more globalized, clause-medial *like* position. The shift suggested by the fiction corpus is validated by the study of *like* in the *LCIE* sample. While clause final was found to be the preferred placement in the *LCIE* sample, the findings reveal medial *like* as the up-coming position, outranking clause-initial as the second, most prominent placement in the sample. I believe this positional shift, which has already been noted in the IrE literature, could have already been in progress at the time of the

compilation of *LCIE* (as evidenced by the status of medial *like* in the sample), but became established as a feature of *supraregional* IrE in subsequent years when the CoFIrE texts were published. Future research should test this theory by examining the use of *like* in a corpus of real-life, spoken, *contemporary* IrE covering the end of the compilation of *LCIE* to the present. This could be carried out using the more updated version of *LCIE*, the *Limerick Corpus of Irish English 2.0.*, whose compilation is being led by Dr O’Sullivan from Mary Immaculate College (University of Limerick, Ireland). Future researchers may also wish to trace the development of this potential contemporary shift by conducting longitudinal and diachronic analyses against historical corpora of IrE speech, such as McCafferty & Amador-Moreno’s (2012a) *Corpus of Irish English Correspondence*, corpora of historical IrE texts like Fitzgerald’s (2021) *Corpus of Irish Historical Narratives*, or fiction corpora that contain historical representations of IrE, such as Hickey’s (2003) *Corpus of Irish English*, Cesiri’s (2012) *Corpus of Irish Fairy and Folk Tales*, or Connell’s (2014) *Corpus of Hiberno- English Literary Dialect*).

Further evidence validating fictional renditions of IrE speech with regard to realistic representations and the documentation of linguistic developments was offered by the study of the portrayal of Dublin English (and IrE in general) in the CoROCK case studies, as will be discussed in the next section.

9.4.1. RQs 1 and 1a: Studying use, form, and developments in CoROCK

The case studies conducted on CoROCK demonstrate Howard’s keen ear for linguistic developments by documenting the adoption of ‘new’, more global discourse pragmatic elements which ranked as frequent *over-the-top* intensifiers. These items are: New Intensifying *So* (NISo) and Non-lexical intensifier *Totally* (NLT). While their studies retrieved findings that align with the previous literature on these features in other varieties, thus belying the high level of authenticity present in the fictional rendition of the CoROCK texts, the corpus also documents noteworthy pragmalinguistic developments that are either first documented here, and/or appear to disprove previous theories. For example, the longitudinal analyses of the use of NISo and NLT (discussed in chapter (8)) in their respective subcorpora (i.e. CoROCK-SO and CoROCK-T) appear to track the rise and decline in popularity of these pragmatic items in Ireland. Thus, the longitudinal analysis of NISo production marks a significant decrease in use from 2009

onward. This seems to be a decisive year in its development in Dublin English (according to Howard's perception), where the analysis of *colligational pattern preference* also appears to mark the *delexicalization* of NISo in two forms. On the one hand, *NISo+Verbal Phrase* is found to be the preferred pattern which is consistently used throughout the corpus. This disproves Amador-Moreno & Terrazas-Calero's (2017) theory that this pattern may have been replaced by *NISo+Noun Phrase* which, despite featuring prominently as second, most produced pattern, does not appear across all CoROCK-SO books. On the other hand, *delexicalization* is also highlighted by the identification of a variety of colligation contexts within the *NISo+Noun Phrase* pattern (i.e. *NISo+Noun Phrase*; *+Adjectivized Nouns*; *+(Positive) adjectivized Nouns*, *+(Negative) Adjective+Noun*), some of which seem to replace standard intensifiers *such* or *very*. The fact that these patterns begin to be documented in 2009 could also indicate that this is the year when NISo may have gained lexical unrestrainability in Dublin, at least as perceived and portrayed by Howard.

The examination of the trajectory of NLT-use in CoROCK-T offered in chapter (8), which is the first to be carried out in the context of IrE, also retrieved noteworthy findings pertaining to its use and development in Dublin. The corpus data highlighted 2001-05 as the peak of popularity of NLT in Dublin English (according to Howard's perception) which decreased considerably after 2008. This thesis contends that the fact that the peak and decrease in use of NLT mirrors the height of the Celtic Tiger and the onset of its subsequent recession (see (1.4.) for an exploration of the Celtic Tiger), just like the analysis of NISo suggested, could be indicative of the fact that Howard represents both NLT and NISo as a linguistic tools with which characters may have linguistically projected the ideals of affluence and cosmopolitanism intrinsic to the Tiger years, and which they (and Howard in reality) stopped using to discontinue those identarian 'associations' once the economic recession began.

The findings from the quantitative and qualitative analyses of NLT *colligation pattern preference* retrieved colligational patterns (i.e. *NLT+Verb Phrase*, *standalone*, and *NLT+Noun Phrase*) which also coincide with the literature on the use of this intensifier in other varieties of English, which underscores the realistic value of the fictional rendition offered in the CoROCK-T books. However, the corpus also highlights developments in progress, such as the existence of the *NLT+Verb Phrase* pattern with *directives* which had been predicted 'not to occur with' NLT in the literature.

Furthermore, the documentation of *standalone* NLT in CoROCK-T also suggests the geographic unrestrictedness of this particular pattern, thus refuting Quaglio's (2009) claim that this is an innovation originated in American English conversation. While the analysis of *standalone* NLT underscored previously documented functions in other varieties (i.e. *emphatic agreement*, *affirmative responses*), it also recorded previously undocumented uses (i.e. *lending emphasis/veracity to speaker's statement*). I believe this further endorses the realism of the fictional portrayal, while also highlighting CoROCK as a great source for the investigation of realistic IrE linguistic uses.

The findings with regards to the pragmatic multifunctionality and identity indexical value of NISo and NLT shall be discussed in sections (9.5.1.) and (9.6.) respectively.

9.5. RQ 1 and 1a: Pragmatic multifunctionality of top 3 pragmatic items in CoFIrE

From a pragmatic standpoint, this thesis is novel in utilizing an original system of pragmatic, manual emotion annotation both in the CoFIrE and CoROCK studies which divides the occurrences of the pragmatic items examined here into 1) type of pragmatic connotation (i.e. *positive*, *neutral*, and *negative*), as well as 2) sub-function/emotion conveyance (e.g. transmission of *anger*, *sarcasm*, *compliment*, *simple emphasis*, *derision*, etc.) While emotion annotation is subjective in nature, objectivity was achieved through my familiarity with all the books, contexts of interactions, and interpersonal relationships. This annotation proved to be crucial to evaluating the pragmatic functionality of these items and to shedding light on new developments which appear to have taken place in the use of intensifying *fucking* (see section (6.7.)) and clause medial discourse pragmatic marker *like* as represented in CoFIrE (see (7.6.)), as well as in the use of NISo and NLT-use in CoROCK (sections (8.5.3.) and (8.6.3.) respectively).

Section (6.7), therefore, contributes to Murphy's (2009, 2010) distribution of the pragmatic functionality of lemma FUCK in IrE into *positive* and *negative* emotions, by adding a *neutral* category which often implies the fact that the occurrence of intensifying *fucking* is used to lend emphasis to the utterance. The analysis of the pragmatic functionality of intensifying *fucking* in CoFIrE indicated the prominence of *negative*

emotion conveyance. This aligns with the findings from the *LCIE/BNC2014* samples and with previous literature on the use of this item in other varieties, thus highlighting the value of the reproduction of natural intensifying *fucking* usage portrayed in the CoFIrE texts. Contrastively, the analysis of the emotion conveyance through intensifying *fucking* at the hands of the homodiegetic/autodiegetic narrator in the *Ross O'Carroll-Kelly* novels in CoFIrE (explored in section (6.8.)) reveals a prominence of *positive/neutral* emotions which deviates from the *negative* emotions he transmits as character within his story. Based on previous studies (i.e. Stenström, 1991; Murphy, 2010, p. 81; Clancy, 2016), I propose that this deviation may suggest a conscious use on the part of the narrator (and the author in reality), who 'speaks' directly to his audience, of a series of direct-speech, linguistic tools (e.g. intensifying *fucking* and non-standard quotatives) with which he constructs and reinforces the narrator-reader bond of intimacy. I also propose that this narrator-reader bond, which is linguistically created and strengthened, may also be understood from a business angle, whereby the closer/more familiar readers *feel* to the narrator/protagonist, the stronger the fan base and pool of stable book buyers will be, as they are likely to be interested in *hearing* more about their 'close friend(s)'. Finally, the analysis of the pragmatic functionality of clause medial discourse pragmatic marker *like* offered in sections (7.6.) and (7.6.1.) also retrieved previously documented functions, highlighting *focuser* medial *like* as the most prominent use. In addition, *neutral* evaluations were found to be more frequent in CoFIrE, including the conveyance of 1) *emphasis to the content of the utterance*, and 2) the *seeking of clarification*. While the former was also the most prominent subfunction in the contrastive *LCIE/BNC2014* samples, which validates the realism of the CoFIrE representation, *seeking clarification* is only present in the fiction corpus, which could underscore the usefulness of this database as a source for the documentation of pragmatic extensions. The following section will discuss the findings pertaining to the pragmatic multifunctionality of New Intensifying *So* and Non-lexical intensifier *Totally* as represented in CoROCK.

9.5.1. RQs 1 and 1a: Pragmatic multifunctionality of New Intensifying *So* and Non-lexical intensifier *Totally* in CoROCK

From a pragmatic perspective, this is the first study that provides an in-depth investigation of the pragmatic functionality and emotion transmission of New Intensifying *So* (NISo) and Non-lexical intensifier *Totally* (NLT) in the context of IrE using a large corpus of fiction. Each individual case study also retrieved important findings which further endorse the level of realism of their fictional portrayal with regard to their real-life usage, while buttressing the usefulness of CoROCK as a valid source for linguistic research.

The findings from the analysis of the pragmatic functionality of NISo and NLT (explored in detail in sections (8.5.4.) and (8.6.3.)) in CoROCK-SO and CoROCK-T correspond with those identified in the literature of their use in other varieties. For example, both were found to convey mostly *neutral* meanings. In the case of NISo, these appeared in the form of *simple emphasis* or *amplified determination*, while the analysis of NLT retrieved previously documented sub-functions such as *speaker emphasis*, *emphatic agreement*, and *giving affirmative responses*. Pragmatic developments regarding the conveyance of *positive* and *negative* meanings are also documented in the respective subcorpora. For instance, NISo showed a balanced distribution of *positive/negative* emotions, yet their longitudinal evaluation highlighted *positive* meanings (i.e. *excitement*, *compliments*, *admiration*, and *sexual admiration*) as the most consistently transmitted ones despite the larger catalogue of *negative* emotions, whose use dramatically decreases after 2009. The study of NLT also indicated the conveyance of *negative* and *positive* emotions (i.e. *emphatic admiration*), although the latter seems to be a recent acquisition given its exclusive use in the final book (i.e. 2018) of CoROCK-T. While these findings are based on a fiction corpus, they do warrant future research into the existence of these and other pragmatic developments which may have taken place in Dublin English, and IrE by extension, perhaps contrasting against a corpus of real, contemporary, spoken Dublin English (which does not exist to date at the time of writing), or comparing against other spoken varieties of English to test the validity of the fictional rendition.

9.6. RQ 2 and 3: Identity and modern Irishness indexation in CoFIrE and CoROCK

The examination of the fictional representation of IrE speech present in CoFIrE with regard to the potential fictionalization of speaker identity and modern Irishness, which addresses RQs 2 and 3, appears to document a shift from traditionally negative stereotypes (i.e. the Paddy or Stage Irishman) which were often portrayed through grammatical and lexical items in literature. This shift would imply the (re)construction of ‘new’ Irish *identities* in fiction which are reflective of major socioeconomic events that have affected 21st-century Ireland. These *identities* would, therefore, be indexed through the use of the different pragmatic items examined in the thesis, and may be distributed into two themes:

a) *Post Celtic-Tiger (recession) Irishness*

b) *Identity in globalized Ireland*

The findings from the quantitative and qualitative analyses of the general production of grammatical, lexical, and pragmatic feature occurrences in CoFIrE by social class offered in section (4.4.) highlight the *middle-class* as the most prominent, occurrence-producing rank in the corpus. While this class comprises a variety of occupations, we find many speakers who are former builders and/or unemployed construction workers. This could be reflective of Irishness in a post-Celtic-Tiger Ireland which has been affected by the burst of the property bubble, the collapse of the banks, and which is struggling to recover from the economic recession that came in the aftermath of the Tiger.

The more detailed analyses of the most prominent items in CoFIrE (i.e. non-standard quotative GO, intensifying *fucking*, and clause medial discourse pragmatic marker *like*), as well as the case studies conducted with CoROCK, pertaining to their potential indexicality value and the type of modern Irishness they may encapsulate, seem to suggest a type of fictionalized Irishness corresponding with theme (b) above, which is perhaps grounded in Ireland’s participation in a globalized world. It must be noted that the great majority of these features appear to be represented as being more commonly produced by Dubliners, and by Southside Dubliners in particular. This is perhaps to be expected, given the status of Dublin as the capital and largest city of the country, which

is also a center for multiculturalism where all types of developments and influences take place. In addition, and as pointed out by Hickey (2020, p. 79), Dublin city has historically been the center for the changes in (linguistic) identity that have taken place in Ireland, which were particularly driven by the more affluent, educated, Southside population, who distanced itself from the more regionally-bound, less socioeconomically prestigious of the Northsiders (for a detailed explanation of the North/Southside Dublin linguistic (and non-linguistic) differences as represented in the Howard's books, see sections (8.2.) and (8.2.1.)).

The study of quotative GO (see (5.8.4.)) identifies this verb as a distinct marker of identity in (fictionalized) Dublin English, which indexes the voice of *upper-class*, male speakers in their early 30s. The distinctiveness to South Dublin (not exclusively in Howard's novels, but also in other CoFIrE texts) suggests this may be a conscious stylistic marker of regional and socioeconomic identity in the context of Dublin, which actively separates the linguistically cosmopolitan, affluent Southside speakers of *non-local* Dublin English (in Hickey's (2005) terminology) from the working-class, Northside speakers of *local* Dublin English and users of more locally-bound IrE items.

The analysis of the second, most represented non-standard quotative in CoFIrE, i.e. BE LIKE, finds that this verb is also distinctive to South Dubliners in the corpus. Nevertheless, this is not a quotative that has traditionally been associated with IrE but is more popularly connected to the American English quotative repertoire. Furthermore, it is also another one of the linguistic items (along with *high rising terminal*, New Intensifying *So*, Non-lexical *totally*, or clause medial *like*, inter alia) which is often (mis)associated with the construction of the negative (cultural and linguistic) stereotypes of the young, affluent, *Valley Girl – Jock/Surfer Boy* in the context of the USA. Thus, I believe the adoption of this item (and other features aforementioned) into IrE as represented in the fiction texts would support Hickey's (2005, p. 351) theory of the *supraregionalization* of this variety, and of Southside Dublin English in particular.

From a sociolinguistic perspective, the analysis of intensifying *fucking* (see (6.5.) through (6.5.4.)) suggests the encoding of the voice of young, (mostly) *middle-class* speakers. Indeed qualitative and quantitative analyses found that despite its presence across all age-cohorts in CoFIrE, intensifying *fucking* is especially prevalent among young characters, particularly by those in their 20-30s, which coincides with the literature in IrE and other varieties. Furthermore, intensifying *fucking* was more prominent among

CoFIrE male characters, which aligns with the findings from the *LCIE* sample, and supports Murphy's (2010) proposition as to the link between intensifying *fucking* use and masculine voice/masculinity in real spoken discourse in Ireland. However, the comparative analysis against the *BNC2014* sample pointed out females as the more frequent users of intensifying *fucking* in the sample. I hypothesize that while this could be due to the small size of the random sample, it could also be indicative that (UK) females' feel comfortable with the use of this intensifier. If extrapolated to the context of Ireland, this could explain the high number of occurrences of intensifying *fucking* produced by females in CoFIrE and *LCIE* sample, which effectively subverts the popular belief that females do not swear or prefer euphemisms. I believe the fact that CoFIrE renders findings that mirror and corroborate previous studies of real spoken discourse on this subject, as well as the data from the contrastive samples, further highlights the level of realism (i.e. RQ 1a) offered by the books.

The validation of CoFIrE as a useful tool for linguistic research is also supported by the analysis of the orthographic rendition of the pronunciation of intensifying *fucking* in the texts. Indeed, this thesis proposes that not only is the fictional representation of spoken IrE a tool through which characterization is built in the books, but the orthographic manner in which the fictional language is portrayed also aides it, functioning as a multifunctional tool. On the one hand, orthographic renditions serve the obvious purpose of 1) marking prosody and sentence stress, which may or may not be required by the feature itself, and which the studies of orthographic in CoFIrE and of New Intensifying *So* (NISo) and Non-lexical intensifier *Totally* (NLT) in CoROCK have shown can be done through *capitalization*, *italicization*, or *unmarkedness*. By longitudinally analyzing the orthographic rendition of these items, the thesis proposes that these renderings can also function as 2) indicators of the potential *grammaticalization/pragmaticalization* of the items in the Irish context. For example, the analysis of the conveyance of the *over-the-top* intensification carried by NISo and NLT showed this was portrayed in two manners: a) through orthographic prosodic rendition, and/or b) by appearing with/in proximity to other intensifying 'booster' items (e.g. *Oh my God*, *high rising terminal*, *like*, etc.). The longitudinal study of orthographic rendition of NLT and NISo in CoROCK showed that while both intensifiers were initially portrayed carrying overt prosodic rendition in the form of *capitalization* and *italicization*, by 2008/2009 respectively these items began to be *unmarked*. I posit that this trajectory could signal the

grammaticalization/pragmaticalization of these intensifiers in Dublin, whereby the characters (and the readers) no longer needed to be reminded of the *over-the-top* intensification of these ‘newer’ items. A more diachronic corpus would be required in order to test the potential *grammaticalization/pragmaticalization* of these items over time.

The thesis also finds that their *over-the-top* intensifying status may be conveyed through their appearance with/in proximity to other intensifier ‘booster’ items, which may account for the potential loss of intensification *over-the-top* intensifiers are liable to once they are *delexicalized*. Thus, the study of NISo revealed its use in close proximity to other boosters (i.e. *Oh my God, Fucking, Hello, and high rising terminal*), which appears to be female-distinctive in the CoROCK-SO books, whereas the NLT intensifier booster repertoire includes *Oh my God, like, NISo, and I mean*. While these are innovative findings which have not been identified before in the literature of these two intensifiers, and seeing as the peak in popularity of both appears to be prior to 2009, by which they are *delexicalized*, these results should be cross-examined against a corpus of real-spoken IrE that documents language use prior to 2009.

The analysis of the orthographic rendering of intensifying *fucking* in CoFIrE, which retrieved an array of forms, suggests that these portrayals could also 3) function as indexes of speaker/character identity. Indeed, in the context of Dublin English offered in the CoFIrE books, non-standard forms such as *fooking, fooken, fuckin’*, or *fucking* (to a lesser extent) were found to only ever be used by Northside Dubliners, whereas *fucking* and *focking* were restricted to affluent Southside speakers of *non-local* Dublin English. In addition, the use of *fucking* and *focking* was also found to be age-distinctive within the Southside group in CoFIrE, with *fucking* being exclusively used by younger Southsiders, whereas *focking* appeared to be favored by older characters.

The portrayal of clause-medial *like* in CoFIrE (analyzed in section (7.6.)) also seems to be highly realistic from a sociolinguistic perspective, indicating its presence across all cohorts, yet marking its prominence among young (20-30), (mostly) female characters in CoFIrE, which corroborates both the literature and the findings from the *LCIE/BNC2014* samples. With regard to class, the comparative analysis of medial *like* in the samples and in CoFIrE retrieves differing findings, with the samples indicating its prominence among class E speakers (most of whom are students). This could be due to the small size of the samples and demographics of *LCIE*, which warrants future

investigations with larger and longer samples, yet it may also correspond with the young age-range this discourse pragmatic marker is favored by in the literature. In contrast, the occurrences of medial *like* in CoFIrE are more recurrent among Class A, mostly Southside Dubliners, which could be due to a potential over-representation of these characters. In addition to class, medial *like* also seems to be indexical of region, for while it is used by characters from various counties in CoFIrE, it is most prominently produced by Southside Dubliners (in Howard's and other authors' books). This could support Amador-Moreno's (2015, p. 376) assertion regarding the use of medial *like* as a feature of *supraregionalized* IrE. Finally, I theorize that the shift in preference from traditional clause-marginal positioning in IrE to clause-medial placement may have been partly driven by the common (mis)association of medial *like* with the negative stereotypes of the *Valley Girl* and the *Jock-Surfer Boy* in the USA. Indeed, these cultural typecasts have very distinctive linguistic repertoires which, as mentioned above, include the use of medial *like*, along with other features such as *high rising terminal*, NISo, NLT, *Be like*, etc., most of which are found in CoFIrE and CoROCK, and which, as explained throughout the thesis, are seemingly distinctive to South Dubliners. These findings, therefore, would correspond, not to Irishness *in general*, but to *one* type of Irishness in particular: i.e. Dublin Irishness. Notice the already discussed contrast that exists between the type of *general* Irishness that was retrieved from the analysis of *all* linguistic items in CoFIrE as pertains to social class (i.e. Post Celtic-Tiger, *middle class*, mostly made up of former builders and unemployed construction workers), and the more individual, Dublin-focused, cosmopolitan identity that arises from the study of the most prominent pragmatic elements in CoFIrE, and the non-traditionally IrE, pragmatic items in CoROCK.

This thesis, therefore, posits that the linguistic and sociocultural notions attached to these stereotypes in the context of the USA can be extrapolated to the cultural and linguistic landscape of Ireland, and of South Dublin in particular, in the form of what I propose to label the *Southside Socialite*, and its male counterpart, the *Rugby Jock/D4 head*, personae. In section (7.5.1.) I identify attributes that are popularly associated to the American stereotypes which are also present in the IrE context regarding *socioeconomic background* (wealthy; snobbish; private schools), *personality traits* (fashionable; airheaded; dizzy girls; unintelligent and arrogant athletes, unemployed in adulthood for Jocks-D4 Heads; attractive and seductive for Surfer Boys), and even *sports-preferences* (cheerleading-hockey for the females; football-rugby for the male counterparts).

The thesis, therefore, proposes that these analogous stereotypes are linguistically (re)constructed in fiction through pragmatic items (summarized in Table 9.1. below) which are popularly (mis)associated with the American stereotypes and have never been attributed to IrE. I believe these features include, in the case of CoFIrE, the use of clause-medial *like*, in particular in *like+NISo* patterns (e.g. It's, like, *so* an amazing idea), as well as *+high rising terminal* (e.g. 'we should go for something with more personality – as in, like, more *individualistic?*'), all of which are prominently produced by young, (mostly) female, affluent, Southside Dubliners in the books. They are also constructed through the use of *NISo* and *NLT* as evidenced in CoROCK. Indeed, these items are exclusively used in the corpora by young, affluent, Southside Dubliners and are completely absent from the Northsider repertoire.

<i>Southside Socialite-Rugby Jock/D4 Head Repertoire</i>	
Non-standard quotatives	<i>Be Like; Be all</i>
Discourse pragmatic markers	Clause-medial <i>Like</i>
<i>Over-the-top</i> intensifiers	<i>NISo; NLT</i>
Pronunciation	<i>High Rising Terminal</i>

Table 9.1. Associated Valley-Girl items present in the *Southside Socialite-Rugby Jock/D4 Head* repertoires in CoFIrE and CoROCK.

The study of *NISo* and *NLT* in CoROCK with regard to sociolinguistic factors further supports my contention regarding their fictional rendition functioning as stylistic tools which (re)construct these two stereotypes in Howard's fiction, seeing as how they are both exclusively used by the wealthy, snobbish, and young, Southsiders, whereas Northsiders prefer the use of traditionally IrE-bound items like tag-*so*.

Finally, this thesis proposes that the use of all of the pragmatic items outlined in Table 9.1. above both in CoFIrE and CoROCK illustrates Howard's conscious authorial use as stylistic resources with which he 1) builds and refines the characterization in his series. He 2) constructs the recognizable, regionally and socioeconomically-distinctive stereotypes of the *Southside Socialite-Rugby Jock/D4-Head* personae, 3) manipulating and triggering the specific attributes associated to these in the minds of the readers through the fictional rendition of these items, which 4) effectively creates a linguistic Northside-Southside divide that is rooted in the 'perceived' sociocultural/sociolinguistic differences between those areas (for more on these differences, see section (8.2.) and (8.2.1.)).

This thesis is novel in its identification of how authors perceive and construct ‘new’ *contemporary* Irish identity. Potentially rooted in Ireland’s participation in a globalized world, and in answer to RQ 3, the study of these pragmatic items has led this thesis to draw a parallel between cross-variational stereotypes, culminating in the (re)construction of the *Southside Socialite-Rugby Jock/D4-Head* personae whose distinctive pragmastylistic repertoire is summarized in Table 9.1. above. Despite the novelty of this potential finding, it is necessary to reiterate that all of these items are restricted to the (fictional) Southside Dublin repertoire offered in CoFIrE and CoROCK.

The potential construction of these (fictional) identities in CoFIrE and CoROCK calls for contrast with actual spoken IrE corpora to test their validity, but also warrant future research into other *types* of modern Irishness which could be gathered from the study of, perhaps, the grammatical and lexical items annotated in CoFIrE which were excluded from this thesis. Future studies could also conduct diachronic investigations, comparing the use of pragmatic elements in literature in previous centuries to these *contemporary* representations to check for developments. It would also be interesting to contrast the pragmatic findings from CoFIrE/ CoROCK against corpora of *contemporary*, 21st-century, spoken IrE (at the time of writing, no such corpora large exist) to check the validity of the hypotheses made in this thesis.

9.7. Limitations of this study

While the study of the fictional representation of IrE offered in CoFIrE and CoROCK has been invaluable in gathering information about the use, pragmatic multifunctionality, and identity indexicality value of the features explored here, as discussed in previous section, there are limitations to this thesis. In particular, the limitations are connected with the use of *contemporary* fiction. It appears that contemporary Irish authors, at least the ones included in CoFIrE, have focused on using Dublin English to reproduce fictionalized, contemporary IrE spokenness in text. Indeed, Dublin city features prominently across the CoFIrE books and is the setting for all CoROCK novels. Thus, utilizing a corpus of novels that are set in other Irish cities or are written by other Irish authors could have, perhaps, included linguistic representations which may have led to different results.

Another limitation of this study is the fact that the findings are based on fictional portrayals. While, as mentioned in (1.1.), researchers must be aware that fictional

representations of real language use are based on the authors' perceptions, the fact that the findings gathered from the studies conducted on CoFIrE and CoROCK not only mirror the findings from the *LCIE/BNC2014* samples, but also corroborate the literature, attests to the high level of realism present in the fictional portrayals, highlighting both corpora as great sources for the linguistic investigation of *contemporary* IrE. A final limitation of the thesis, beyond the control of the researcher, which has already been noted in section (3.2.3.), is the challenge of accessing texts due to *copyright restrictions*. While at present neither CoFIrE nor CoROCK are publicly available due to copyright constraints, it is hoped that at least some parts of them may be made available in the future for research purposes after consultation with authors and publishers. This would allow future investigators interested in examining developments in IrE or cross-variational studies to use the fiction datasets as sources for investigation, while others may wish to conduct larger scale contrastive studies against corpora of *contemporary* or historical IrE. Future directions are offered in the next section.

9.8. Future directions

This thesis would like to shine a light on a pressing need that exists in the field of IrE. Unlike other varieties such as American or British English, both of which have numerous (and very large) corpora of spoken and written texts which allow for the study of real-life and/or fictional language use, at present, IrE only has a comparatively small number of (often) specialized spoken corpora (e.g. *ICE: Ireland*, *SPICE: Ireland*, *LCIE* or Murphy's (2010) *CAG-IE*, and others mentioned throughout the thesis). In addition, and as is the case with CoFIrE and CoROCK, some of these are not publicly available, at the time of writing¹⁰⁵. Furthermore, and with the exception of the modern version of *LCIE* which is currently under construction, most contain data from the late 1990s to the early 2000s. I would, therefore, like to draw attention to the current gap in the number of corpora that comprise *contemporary*, 21st-century, spoken (and written) IrE.

While aware of the large demands inherent to the compilation of spoken corpora (e.g. ethical considerations, time-consuming process, need for substantial funding which first-time researchers often find hard-to-get, to name a few), these could be mitigated in various manners. Unlike in the past, where researchers would have to conduct interviews

¹⁰⁵ *LCIE* is currently in the process of being made public via *Sketch Engine*.

using often expensive (tape) recorders and/or sound-proofed facilities, the rapid evolution of the ever-present smartphone, which now produces high quality recordings, gives everybody the opportunity to personally participate as an informant and/or as an interviewer in the compilation of a spoken corpus of IrE. Provided that all participants fill in a form that is in accordance with ethical standards, both the forms and the recordings could, then, be sent to the lead researcher, and/or uploaded to a cloud service devoted to the corpus. In addition, it would be necessary to take into consideration the fact that communication is not only restricted to and/or accessed through voice recordings. On the contrary, *contemporary* communication includes other types of multimodal *voices* which can be found in (online) written discourses, such as tweets, online posts, and/or memes.

Another possibility would be the inclusion or compilation of 21st-century, telecinematic, IrE material, which would complement Walshe's (2009) corpus of 50 movies produced in the Republic of Ireland between 1939 and 2007, by including different televisual genres like soap operas, talk shows, late night shows, sitcoms, YouTube videos, recorded public readings with authors, etc. Furthermore, and as has been proven in this thesis, fictionalized (literary) dialect-dialogue is based on and creates a realistic portrayal of true spokenness. Thus, I believe a corpus of *contemporary* IrE would also benefit from adding contemporary fiction.

To that purpose, I plan on expanding on CoFIrE and converting it into a larger, diachronic corpus which comprises more contemporary fiction (including poetry and stage plays) by other writers, but which also include sub-corpora of fiction published in the 19th and 20th centuries in the Republic of Ireland¹⁰⁶. This larger corpus, which would be annotated using the annotation system I have designed and implemented on CoFIrE, would allow me to test various hypotheses which have been developed in this thesis with regard to pragmatic items. It would also allow for the investigation of any grammatical and lexical developments that may have taken place over the span of centuries covered by the larger CoFIrE, perhaps tracing the rise and decrease in use of traditionally Stage Irishman features and documenting a potential shift in the stylistic representation of Irish identity through speech reproduction. In order to also investigate the representation of the Irish through speech in literature, this larger CoFIrE could also include fiction published

¹⁰⁶ I would like to acknowledge Dr Carolina P. Amador-Moreno as the compilation of a diachronic CoFIrE comprising books published in the Republic of Ireland from the 19th century through the 21st century was her original idea for this thesis. Time restrictions, however, led to this thesis focusing only on *contemporary* fiction.

by authors outside of Ireland so that we could use the subcorpora to contrast the level of realism inherent to the non-Irish linguistic representation when contrasted against the representation offered by local authors. Non-fiction discourse (e.g. newspaper articles, magazines, true crime books, autobiographies, etc.) could also be a rich source for linguistic representations of past and modern-day IrE use, which is why the larger CoFIrE would include non-fictional subcorpora. A much larger corpus of fictionalized representations of IrE speech could also be created collaboratively with researchers who have already looked into fictional representations of IrE, such as Shane Walshe, María Palma-Fahey, Bróna Murphy, or Sara Díaz-Sierra (2022). Finally, the creation of CoROCK means that past and future books written by Paul Howard (e.g. unpublished *Ross O'Carroll-Kelly* stage plays) can also be added and annotated, leading to future work on stylistic change, as well as on linguistic developments and potential identity indexicality in *local* and *non-local* Dublin English as perceived and represented by Howard.

To conclude, while I do plan on expanding on CoFIrE and CoROCK, as stated above, I would also like to encourage future researchers (including myself) to also consider the possibility of building a corpus which could be the first-ever, unified, multigenre, contemporary *Corpus of Irish English*. Similar to larger corpora like the *Corpus of Contemporary American English* or the *British National Corpus*, this Irish corpus would grant access to both spoken *and* written representations of contemporary IrE speech, enabling researchers to conduct more empirical studies of this variety and/or cross-examine it by type of speech, genre, or cross-variationally, thus validating or disproving findings that can sometimes only be hypothesized (for now).

BIBLIOGRAPHY

- Abercrombie, N., Hill, S., & Turner, B. S. (Eds.). (2006) *The Penguin Dictionary of Sociology*. London: Penguin Books.
- Adolphs, S. (2006). *Introducing Electronic Text Analysis: A practical guide for language and literary studies*. New York: Routledge.
- Adolphs, S., & Knight, D. (2009). Building a spoken corpus: what are the basics? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 38-52). London and New York: Routledge.
- Aegirly. (2006). *Urban Dictionary*. Accessed at:
<https://www.urbandictionary.com/define.php?term=d4+head>.
- Aijmer, K. (1997). I think — an English modal particle. In T. Swan & O. J. Westvik (Eds.), *Modality in Germanic Languages* (pp. 1-47). Berlin/NY: De Gruyter Mouton.
- 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* (pp. 60-95). London/NY: Routledge.
- Aijmer, K., & Simon-Vandenberg, A.-M. (2011). Pragmatic Markers. In J. Zienjowski, J.-O. Östman & J. Verschueren (Eds.), *Discursive Pragmatics* (pp. 223-247). Amsterdam: John Benjamins.
- Alvarez, L. (2005). Suddenly rich, Poor Old Ireland Seems Bewildered, *New York Times*.
- Allan, K., & Burrige, K. (2006). *Forbidden Words: Taboo and the Censoring of Language*. Cambridge: CUP.
- Amador-Moreno, C. P. (2005). Discourse Markers in Irish English: an example from Literature. In A. Barron & K. P. Schneider (Eds.), *The Pragmatics of Irish English* (pp. 73-100). Berlin/New York: De Gruyter Mouton.
- . (2006). *An Analysis of Hiberno-English in the Early Novels of Patrick MacGill. Bilingualism and Language Shift from Irish to English in County Donegal*. Lewiston, NY: The Edwin Mellen Press.

--. (2010a). How can corpora be used to explore literary speech representation? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (Vol. 531-544). London: Routledge.

--. (2010b). *An Introduction to Irish English*. London and Oakville: Equinox.

--. (2010c). Writing from the margins: Donegal English invented/imagined. In R. McColl Millar (Ed.), *Marginal dialects: Scotland, Ireland and Beyond*. Aberdeen: Accessed at: <https://www.abdn.ac.uk/pfrlsu/documents/PFRLSU%201%20title%20pages.pdf>

--. (2012a). A corpus-based approach to contemporary Irish writing: Ross O'Carroll-Kelly's use of 'like' as a discourse marker. *International Journal of English Studies*, 12(2), 19-38.

--. (2012b). The Irish in Argentina: Irish English transported. In B. Migge & M. Ní Chiosáin (Eds.), *New Perspectives in Irish English* (pp. 289-309). Amsterdam: John Benjamins.

--. (2015). 'There's, like, total silence again, roysh, and no one says anything'. Fictional representations of 'new' DMs and quotatives in Irish English. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic Markers in Irish English* (Vol. 370-389). Amsterdam: John Benjamins.

--. (2016). The language of Irish writing in English. In R. Hickey (Ed.), *Sociolinguistics in Ireland*. Basingstoke: Palgrave-Macmillan.

--. (2019a). 'Matt & Mrs Connor is with me now. They are only beginning to learn the work of the camp': Irish emigrants writing from Argentina. In R. Hickey (Ed.), *Keeping in Touch: Familiar Letters across the English-speaking World* (pp. 139-162). Amsterdam/Philadelphia: John Benjamins.

--. (2019b). *Orality in Written Texts: Using Historical Corpora to Investigate Irish English 1700-1900*. Abingdon/NY: Routledge.

--. (fc). Discourse-Pragmatic Markers in Irish English. In R. Hickey (Ed.), *The Oxford Handbook of Irish English*. Oxford: Oxford University Press.

Amador-Moreno, C. P., & Avila-Ledesma, N. E. (2020). Migration Experiences and Identity Construction in Nineteenth-century Irish Emigrant Letters. In C. P. Amador-Moreno & R. Hickey (Eds.), *Irish Identities. Sociolinguistic Perspectives* (pp. 283-302). Berlin/Boston: De Gruyter Mouton.

- Amador-Moreno, C. P., & McCafferty, K. (2011). Fictionalising Orality: An Introduction. *Sociolinguistic Studies*, 5(1), 1-13.
- . (2015). '[B]ut sure its only a penny after all': Irish English discourse marker sure. In M. Dossena (Ed.), *Transatlantic perspectives in Late Modern English. Advances in Historical Sociolinguistics* (pp. 179-198). Amsterdam/Philadelphia: John Benjamins.
- Amador-Moreno, C. P., McCafferty, K., & Vaughan, E. (Eds.). (2015). *Pragmatic Markers in Irish English*. Amsterdam/Philadelphia: John Benjamins.
- Amador-Moreno, C. P., & Nunes, A. (2009). *The Representation of the Spoken Mode in Fiction: How Authors Write how People Talk*. Lewiston: The Edwin Mellen Press.
- Amador-Moreno, C. P., & O'Keeffe, A. (2018). He's after getting up a load of wind: a corpus-based exploration of be +after + V-ing constructions in spoken and written corpora. In D. Villanueva Romero, C. P. Amador-Moreno & M. Sánchez García (Eds.), *Voice and Discourse in the Irish Context* (pp. 47-73). Cham: Palgrave Macmillan.
- Amador-Moreno, C. P., & Terrazas-Calero, A. M. (2017). Encapsulating Irish English in literature. *World Englishes*, 36(2), 254-268.
- . (2022). Using corpus linguistics to explore literary speech representation: non-standard language in fiction. In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 517-531). London/NY: Routledge.
- Andersen, G. (2001). *Pragmatic Markers and Sociolinguistic Variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam/NY: John Benjamins.
- Anderson, W. (2006). "Absolutely, totally, filled to the brim with the Famous Grouse". *English Today*, 22(1), 10-16.
- Andersson, L.-G., & Trudgill, P. (1990). *Bad Language*. London: Penguin Books.
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141-161.
- Anthony, L. (2018). *AntConc 3.5.7*.
- Archer, D., Culpeper, J., & Rayon, P. (2009). Love-'a familiar or a devil'? An exploration of key domains in Shakespeare's Comedies and Tragedies. In D. Archer (Ed.),

What's in a Word-List? Investigating Word Frequency and Keyword Extraction. (pp. 136-157). London: Ashgate.

Ashfield, R. (Ed.). (2017). *Goodreads*. Accessed at: https://www.goodreads.com/review/show/2089112372?book_show_action=true.

Asián, A., & McCullough, J. (1998). Hiberno-English and the Teaching of Modern and Contemporary Irish Literature in an EFL Context. *Links and Letters*, 5, 37-60.

Atwell, E., Hughes, J., & Souter, C. (1994). AMALGAM: Automatic Mapping among Lexico-grammatical Annotation Models. In J. Klavans & P. Resnik (Eds.), *The Balancing Act: Combining Symbolic and Statistical Approaches to Language. Proceedings of the Workshop in Conjunction with the 32nd Annual Meeting of the Association for Computational Linguistics* (pp. 21-28). New Mexico: New Mexico State University, Las Cruces, NM.

Avila-Ledesma, N. E. (2019). *A Historical, Ethnopragmatic Study of the Conceptualisation of Emotions in Irish Emigrants' Personal Correspondence*. (Doctor of Philosophy), Universidad Autónoma de Madrid, Unpublished.

Avila-Ledesma, N. E., & Amador-Moreno, C. P. (2016). 'The More Please [Places] I See the More I Think of Home': On Gendered Discourse of Irishness and Migration Experiences. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2016: Global Implications for Society and Education in the Networked Age* (pp. 85-105). Berlin/Heidelberg: Springer.

Aybaybay, A. (2011). *Urban Dictionary*. Accessed at: <https://www.urbandictionary.com/define.php?term=surfer%20boys>.

Ayto, J. (2010). *Oxford Dictionary of English Idioms*. Oxford: Oxford University Press.

Bäcklund, U. (1973). *The Collocation of Adverbs of Degree in English*. Uppsala: Uppsala University Press.

Baird, S. (2001). How 'to be like' a Kiwi: Verbs of quotation in New Zealand English. *NZEJ*, 15, 6-19.

Baker, P. (2006). *Using Corpora in Discourse Analysis*. London/New York: Continuum.

Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bakhtin, M. M. (1981). *The Dialogic Imagination: Four Essays by M.M. Bakhtin* (C. Emerson & M. Holquist, Trans.). Austin: University of Texas Press.

- Barbieri, F. (2005). Quotative Use in American English: A Corpus-Based, Cross-Register Comparison. *Journal of English Linguistics*, 33(3), 222-256.
- . (2008). Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics*, 12(1), 58-88.
- Barron, A. (2017). The speech act of offers in Irish English. In R. Hickey & E. Vaughan (Eds.), *Irish English* (Vol. Special Issue of World Englishes, pp. 224-238). Malden: Wiley.
- Barron, A., & Schneider, K. P. (Eds.). (2005). *The Pragmatics of Irish English*. Berlin/New York: De Gruyter Mouton.
- Barry, K. (2011). *City of Bohane*. London: Vintage Books.
- . (2012). *Dark Lies the Island*. London: Vintage.
- Bartley, J. O. (1954). *Teague, Shenkin and Sawney. Being an Historical Study of the Earliest Irish, Welsh and Scottish Characters in English Plays*. Cork: Cork University Press.
- Bauer, L. (1983). *English Word-Formation*. Cambridge: CUP.
- Bednarek, M. (2011 a). The language of fictional television: A case study of the 'dramedy' Gilmore Girls. *English Text Construction*, 4(1), 54-84.
- .(2011b). The stability of the televisual character: A corpus stylistic case study. In R. Piazza, M. Bednarek & F. Rossi (Eds.), *Telecinematic discourse: Approaches to the language of films and television series* (pp. 185-204). Amsterdam/Philadelphia: John Benjamins.
- . (2012). Constructing 'nerdiness': Characterisation in The Big Bang Theory. *Multilingua* 31, 199-229.
- . (2015). "Wicked" Women in Contemporary Pop Culture: 'Bad' Language and Gender in 'Weeds', 'Nurse Jackie' and 'Saving Grace'. *Text & Talk*, 35(4), 431-451.
- . (2017). The role of dialogue in fiction. In M. A. Locher & A. H. Jucker (Eds.), *Pragmatics of Fiction* (pp. 129-158). Berlin/Boston: De Gruyter Mouton.
- . (2019). The multifunctionality of swear/taboo words in television series. In J. L. Mackenzie & L. Alba-Juez (Eds.), *Emotion in Discourse*. Amsterdam/Philadelphia: John Benjamins.

- Beltrama, A. (2015a). From Totally Dark to Totally Old. The Formal Semantics of Subjectification. *Proceedings of Sinn und Bedeutung 19*, 19, 125-142.
- . (2015b). Intensification and sociolinguistic variation: a corpus study. In A. E. Jurgensen, H. Sande, S. Lamoureux, K. Baclawski & A. Zerbe (Eds.), *UC Berkeley: Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (pp. 15-30). Berkeley, CA: Berkeley Linguistics Society.
- . (2016). *Bridging the Gap: Intensifiers between Semantic and Social Meaning*. (Doctor of Philosophy), University of Chicago, Accessed at: <https://knowledge.uchicago.edu/record/1099>
- Beltrama, A., & Staum Casasanto, L. (2017). 'Totally' tall sounds 'totally' younger: Intensification at the socio-semantics interface. *Journal of Sociolinguistics*, 21(2), 154-182.
- Bell, L. (2013). Interview 'In a World...' film director Lake Bell. In K. Monk (Ed.), *Canada.com*. <https://o.canada.com/2013/08/14/0815-film-world-bell/>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999, 2007). *Longman Grammar of Spoken and Written English*. Harlow (England): Longman.
- Blackwell, N. L., Perlman, M., & Fox Tree, J. E. (2015). Quotation as a Multimodal Construction. *Journal of Pragmatics*, 81, 1-7.
- Blakemore, D. (1992). *Understanding Utterances. An Introduction to Pragmatics*. Oxford: Blackwell.
- Bliss, A. J. (1979). *Spoken English in Ireland 1600-1740*. Dublin: Dolmen Press.
- . (1984). English in the South of Ireland. In P. Trudgill (Ed.), *Language in the British Isles* (pp. 135-151). Cambridge: CUP.
- Blyth, C. J., Recktenwald, S., & Wang, J. (1990). "Say What?!": A New Quotative in American Oral Narrative. *American Speech*, 65(3), 215-227.
- Boland, J. (2013). Review: Donal Ryan 'The Thing About December', *Independent.ie*. Retrieved from <https://www.independent.ie/entertainment/books/review-donal-ryan-the-thing-about-december-29672764.html>
- Bolinger, D. (1972). The Intensification of Nouns *Degree Words* (pp. 58-87). The Hague, Paris: Mouton.
- Brinton, L. J. (1996). *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/NY: De Gruyter Mouton.
- . (2017). *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: CUP.
- Brinton, L. J., & Arnovick, L. (2006). *The English Language: A Linguistic History*. Oxford: Oxford University Press.

The British National Corpus 2014: User Manual and Reference Guide (Version 1.1) (2018). Retrieved from <http://corpora.lancs.ac.uk/bnc2014/doc/BNC2014manual.pdf>

Buchstaller, I. (2002). *'He goes and I'm like': The new quotatives re-visited*. Paper presented at the University of Edinburgh Postgraduate Conference.

--. (2006a). Diagnostics of age-graded linguistic behaviour: the case of the quotative system. *Journal of Sociolinguistics*, 10, 3-30.

--. (2006b). Social stereotypes, personality traits and regional perception displaced: attitudes towards the 'new' quotatives in the UK. *Journal of Sociolinguistics*, 10, 362-381.

--. (2008). The localization of global linguistic variants. *English World-Wide*, 29, 15-44.

--. (2011). Quotations across the generations: a multivariate analysis of speech and thought introducers across 5 decades of Tyneside speech. *Corpus Linguistics and Linguistic Theory*, 7, 59-92.

--. (2013). *Quotatives: New Trends and Sociolinguistic Implications*. Oxford: Wiley-Blackwell.

Buchstaller, I., & D'Arcy, A. (2009). Localized globalization: a multi-local, multivariate investigation of quotative 'be like'. *Journal of Sociolinguistics*, 13, 291-331.

Butters, R. R. (1980). Narrative Go 'Say'. *American Speech*, 55(4), 304-307.

Bylina, L. (2011). This is so NP!. In B. H. Partee, M. Glanzberg & J. Skilters (Eds.), *Formal Semantics and Pragmatics: Discourse Context and Model*. Riga: New Prairie Press.

Cahill, S. (2011). *Irish Literature in the Celtic Tiger Years 1990-2008: Gender, Bodies, Memory*. NY/London: Bloomsbury.

Calude, A. (2017). The use of 'heaps' as quantifier and intensifier in New Zealand English. *English Language and Linguistics*, 1-26.

Carey, S. (2018). Irishisms in City of Bohane. Retrieved from <https://stancarey.wordpress.com/?s=irishisms+in+city+of+bohane>

Carleton, G. W. (1864). *The Supressed Book about Slavery!* NY: Carleton, Publisher.

Carter, R. (2014). Stylistics as applied linguistics. In P. Stockwell & S. Whiteley (Eds.), *The Cambridge Handbook of Stylistics* (pp. 77-86). Cambridge: CUP.

Carter, R., & McCarthy, M. (2006). *Cambridge Grammar of English: A comprehensive guide: spoken and written English grammar and usage*. Cambridge: CUP.

- Cesiri, D. (2012). *Nineteenth-Century Irish English: A Corpus-Based Linguistic and Discursive Analysis*. Lewiston; Queenston; Lampeter: The Edwin Mellen Press.
- Clancy, B. (2010). Building a corpus to represent a variety of a language. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 80-92). London/NY: Routledge.
- . (2011). "Do you want to do it yourself like?" Hedging in Irish Traveller and Settled Family Discourse. In B. L. Davies, M. Haugh & A. J. Merrison (Eds.), *Situated Politeness* (pp. 129-146). London: Continuum.
- . (2016). *Investigating intimate discourse: Exploring the spoken interaction of families, couples and friends*. London: Routledge.
- . (2018). Conflict in corpora: Investigating family conflict sequences using a corpus pragmatic approach.
- Claridge, C., & Arnovick, L. (2010). Pragmaticalisation and discursisation. In A. H. Jucker & I. Taavitsainen (Eds.), *Historical Pragmatics* (pp. 165-192). Berlin/NY: De Gruyter Mouton.
- Cleary, L. (2018). [Interview with Donal Ryan] *Donal Ryan How I Write*. Transcript of audio interview from *How I Write, Ireland* session, held at the University of Limerick (Ireland), June 7th, 2018. Accessed at: https://ulsites.ul.ie/rwc/sites/default/files/rwc_Donal_Ryan_How_I_Write%2BIn%2Binterview_transcript.pdf
- Coates, J. (2004). *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language* (3rd ed.). London/NY: Routledge.
- Connell, M. (2014). *It's in the details the devil is: Corpus linguistics and Irish English literary dialect*. (Doctor of Philosophy), National University of Ireland, Galway, Unpublished Thesis.
- Corrigan, K. P. (2010). *Irish English vol.1: Northern Ireland*. Edinburgh: Edinburgh University Press.
- Coupland, N. (2007). *Style: Language Variation and Identity*. Cambridge: CUP.
- Cuckor-Avila, P. (2002). 'She say, she go, she be like': Verbs of quotation over time in African American Vernacular ENGLISH. *American Speech*, 77, 3-31.
- Culpeper, J. (2001). *Language and Characterisation: People in Plays and Other Texts*. Harlow: Longman.
- . (2002). Computers, language and characterisation: An analysis of six characters in *Romeo and Juliet*. In U. Melender-Martalla, C. Otman & M. Kyto (Eds.), *Conversation in Life and in Literature: Papers from the ASLA Symposium* (Vol. 15, pp. 11-30). Uppsala: Association Suedoise de Linguistique Appliquee.

- . (2009). Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*. *International Journal of Corpus Linguistics*, 14(1), 29-59.
- Chambers, J. K., & Trudgill, P. (1998). *Dialectology* (2nd ed.). Cambridge: CUP.
- Chen, X., Yuanle, Y., & Hu, J. (2019). A Corpus-Based Study of Hillary Clinton's and Donald Trump's Linguistic Styles. *International Journal of English Linguistics*, 9(3), 13-22.
- Christensen, L. (1996). *A First Glossary of Hiberno-English*. Odense: Odense University Press.
- D'Arcy, A. (2004). Contextualizing St. John's Youth English within the Canadian Quotative System. *Journal of English Linguistics*, 32, 323-345.
- . (2005). *Like: Syntax and Development (unpublished PhD thesis)*. (Doctor of Philosophy), University of Toronto, Accessed at: <http://web.uvic.ca/~adarcy/web%20documents/DArcy%202005.pdf>.
- . (2007). LIKE and Language Ideology: Disentangling the Fact from Fiction. *American Speech*, 82(4), 386-419.
- . (2008). Canadian English as a Window to the Rise of 'Like' in Discourse. *Anglistik (International Journal of English Studies)*, 19(2), 125-140.
- . (2010). Quoting ethnicity: constructing dialogue in Aotearoa/New Zealand. *Journal of Sociolinguistics*, 14, 60-88.
- . (2012). The diachrony of quotation: Evidence from New Zealand English. *Language Variation and Change*, 24, 343-369.
- . (2015). Stability, stasis and change: The *longue durée* of intensification. *Diachronica*, 32(4), 449-493.
- . (2017). *Discourse-Pragmatic Variation in Context. Eight hundred years of LIKE*. Amsterdam: John Benjamins.
- Dailey-O'Cain, J. (2000). The Sociolinguistic Distribution of and Attitudes to Focuser 'Like' and Quotative 'Like'. *Journal of Sociolinguistics*, 4(1), 60-80.
- de Clerk, V. (1991). Expletives: Men only? *Communications Monographs*, 58, 156-169.
- . (1992). How taboo are taboo words for girl? *Language in Society*, 21, 277-289.
- . (2005). Slang and swearing as markers of inclusion and exclusion. In A. Williams & C. Thurlow (Eds.), *Talking Adolescence: Perspectives on Communication in the Teenage Years* (pp. 109-206). Frankfurt: Peter Lang.

- Díaz-Sierra, S. (2022). Produced and perceived authenticity in the Northern Irish TV show *Derry Girls*. *English World-Wide*, 40(1).
<https://doi.org/10.1075/eww.20012.dia>
- Dinah. (2004). *Urban Dictionary*. Accessed at:
<https://www.urbandictionary.com/author.php?author=Dinah>
- Dolan, T. P. (1990). The Language of Dubliners. In A. Martin (Ed.), *James Joyce: The Artist and the Labyrinth* (pp. 25-40). London: Ryan Publishing.
- . (1991). Language in *Ulysses*. In J. Genet & W. Hellegouarch (Eds.), *Studies on Joyce's Ulysses* (pp. 131-142). Caen: Presses universitaires de Caen.
- . (2006). *A Dictionary of Hiberno-English: The Irish Use of English*. Dublin: Gill & MacMillan.
- . (2020). *A Dictionary of Hiberno English: The Irish Use of English* (3rd ed.). Dublin: Gill Books.
- Dorgan, S. (2006). *How Ireland Became the Celtic Tiger: How Ireland Became the Celtic Tiger*. The Heritage Foundation.
- Dörk, M., & Knight, D. (2015). WordWanderer: A navigational approach to text visualisation. *Corpora*, 10(1), 89-94.
- Doyle, R. (1993). *Paddy Clarke ha ha ha* London: Secker & Warburg.
- . (1996). *The Woman Who Walked into Doors*. London: Jonathan Cape.
- Duggan, G. C. (1937). *The stage Irishman: A history of the Irish play and stage characters from the earliest times*. Dublin: Talbot Press.
- Dunne, E. (2014). Review of *Downturn Abbey* by Paul Howard. *Goodreads*. Accessed at:
https://www.goodreads.com/review/show/1067958198?book_show_action=true&from_review_page=1
- Durham, M., Haddican, B., Zweig, E., Johnson, D. E., Baker, Z., Cockerman, D., Danks, E., & Tyler, L. (2012). Constant linguistic effects in the diffusion of 'be like'. *Journal of English Linguistics*, 40(4), 316-337.
- Eckert, P. (2000). *Linguistic Variation as Social Practice*. Oxford: Blackwell.
- Electric Literature*. (2016). Accessed at: <https://electricliterature.com/>
- Elkins, J. (2014). Review of *The Spinning Heart* by Donal Ryan. *Goodreads*. Accessed at:
https://www.goodreads.com/review/show/1335011579?book_show_action=true
- Engel, D. M., & Ritz, M.-E. (2000). The use of the present perfect in Australian English. *Australian Journal of Linguistics*, 20(2), 119-140.

- Ermann, B., & Kotsinas, U.-B. (1993). *Pragmaticalization: The Case of 'Ba' and 'You Know'* (Vol. 10). Stockholm: Almqvist and Wiksell.
- Fairclough, N. (2010). *Critical Discourse Analysis: The critical study of language* (2nd ed.). London/NY: Routledge.
- Farr, F., & Murphy, B. (2009). Religious references in contemporary Irish English: For the varietal comparison of institutionalised spoken interaction. In R. Reppen, S. Fitzmaurice & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation* (pp. 25-48). Amsterdam: John Benjamins.
- Farr, F., Murphy, B., & O'Keefe, A. (2004). The Limerick Corpus of Irish English: Design, description and application. *Teanga (Yearbook of the Irish Association for Applied Linguistics)* 21, 5–29.
- Ferguson, S. (1998). Drawing fictional lines: Dialect and narrative in the Victorian novel. *Style*, 32(1), 1-17.
- Ferrara, K., & Bell, B. (1995). Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be + like. *American Speech*, 70, 265-290.
- Filppula, M. (1999). *The Grammar of Irish English: Language in Hibernian Style*. London/New York: Routledge.
- Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.
- Fischer-Starcke, B. (2010). *Corpus Linguistics in Literary Analysis: Jane Austen and Her Contemporaries*. London: Continuum.
- Fludernik, M. (2012). Narratology and Literary Linguistics. In R. I. Binnick (Ed.), *The Oxford Handbook of Tense and Aspect* (pp. 75-101). Oxford: OUP.
- Fogarty, A. (2003). Preface *Midwife to the Fairies: New and Selected Stories* (pp. 1-17). Cork: Attick Press Ltd.
- Fowler, R. (1989). *Linguistics and the Novel*. London/New York: Routledge.
- Fraser, B. (1988). Types of English Discourse Markers. *Acta Linguistica Hungarica*, 38(1-4), 19-33.
- Fuchs, R. (2017). 'Do women (still) use more intensifiers than men?': Recent change in the sociolinguistics of intensifiers in British English. *International Journal of Corpus Linguistics*, 22(3), 345-374.
- Fussell, P. (1983). *Class: A guide through the American status system*. NY: Summit.
- Garside, R. (1987). The CLAWS Word-tagging System. In R. Garside, G. Leech & G. Sampson (Eds.), *The Computational Analysis of English: A Corpus-Based Approach* (pp. 30-41). London: Longman.

- Garvin, J. (1977). Anglo-Irish idioms in the works of major Irish writers. In D. Ó' Muirithe (Ed.), *The English Language in Ireland* (pp. 100-114). Dublin/Cork: The Mercier Press.
- Genette, G. (1980). *Narrative Discourse: An Essay in Method*. Ithaca (NY): Cornell University Press.
- Gibbons, A., & Whiteley, S. (2018). *Contemporary Stylistics: Language, Cognition, Interpretation*. Edinburgh: Edinburgh University Press.
- Goodreads. (2016). Accessed at: www.goodreads.com
- Gorman, C. (2011). *Towards the Undecidable: A Reading of the Texts of James Joyce, Sean O'Casey and Paul Howard through the Deconstructive lens of Jacques Derrida*. Mary Immaculate College (University of Limerick).
- Gorman, C. (2013). 'No border is guaranteed, inside or out': A Reading of the Speech/Writing opposition within Paul Howard, a.k.a Ross O'Carroll-Kelly's Fictional Series. *Otherness: Essays and Studies*, 3(2), 1-20.
- grace1129. (2006). *Urban Dictionary*. Accessed at: <https://www.urbandictionary.com/author.php?author=grace1129>.
- Greaves, C., & Warren, M. (2010). What can a corpus tell us about multi-word units? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 212-226). London/NY: Routledge.
- Greenbaum, S. (1996). *The Oxford English Grammar*. Oxford: OUP.
- Hancock, A. B., Stutts, H. W., & Bass, A. (2015). Perceptions of Gender and Femininity Based on Language: Implications for Transgender Communication Therapy. *Journal of Language and Speech*, 58(3), 315-333.
- Harris, A. C. (2006). Revisiting Anaphoric Islands. *Language*, 82(1), 114-130.
- 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* (pp. 139-186). London/NY: Longman.
- Harte, L., & Parker, M. (Eds.). (2000). *Contemporary Irish Fiction: Themes, Tropes, Theories*. London/NY: Palgrave Macmillan.
- Harvey, W. J. (1965). *Character and the Novel*. London: Chatto & Windus.
- Hegedűs, I. (2013). Unlikely infix-like elements in English: Critical remarks on the use of the term infix(ation). *Argumentum*, 9, 162-177.
- Hickey, R. (1999). Dublin English: Current changes and their motivation. In P. Foulkes & G. Docherty (Eds.), *Urban Voices* (pp. 265-281). London: Edward Arnold.
- . (2003). *Corpus Presenter: Software for Language Analysis with a Manual and "A Corpus of Irish English" as Sample Data*. Philadelphia: John Benjamins.

- . (2005a). *Dublin English: Evolution and change*. Amsterdam: John Benjamins.
- . (2005b). Irish English in the Context of Previous Research. In A. Barron & K. P. Schneider (Eds.), *The Pragmatics of Irish English* (pp. 17-46). Berlin/New York: De Gruyter Mouton.
- . (2007). *Irish English: History and present-day forms*. Cambridge: Cambridge University Press.
- . (2010). Irish English in early modern drama: The birth of a linguistic stereotype. In R. Hickey (Ed.), *Varieties of English in Writing: The written word as linguistic evidence* (pp. 121-138). NY: John Benjamins.
- . (2016a). English in Ireland: Development and Varieties. In R. Hickey (Ed.), *Sociolinguistics in Ireland* (pp. 3-40). London/NY: Palgrave Macmillan
- . (2018). 'Yes, that's the best': Short Front Vowel Lowering in English Today. *English Today*, 34(2), 9-16.
- . (2020). Adjusting language identity: Twentieth-century shifts in Irish English pronunciation. In R. Hickey & C. P. Amador-Moreno (Eds.), *Irish Identities: Sociolinguistic Perspectives* (pp. 69-83). Boston/Berlin: De Gruyter Mouton.
- Hickey, R. (Ed.). (2016b). *Sociolinguistics in Ireland*. London/New York: Palgrave Macmillan.
- Hickey, R., & Amador-Moreno, C. P. (Eds.). (2020). *Linguistic identities in Ireland- Contexts and issues*. Berlin/Boston De Gruyter Mouton.
- Ho, Y. (2011). *Corpus Stylistics in Principles and Practice: A Stylistic Exploration of John Fowles' The Magus*. London: Continuum.
- Hodson, J. (2014). *Dialect in Film and Literature*. London: Palgrave Macmillan.
- Höhn, N. (2012). "And they were all like 'What's going on?'" : New quotatives in Jamaican and Irish English. In M. Hundt & U. Gut (Eds.), *Mapping Unity and Diversity World-Wide: Corpus-Based Studies of New Englishes*. Amsterdam/Philadelphia: John Benjamins.
- Holmes, J., & Wilson, N. (2017). *An Introduction to Sociolinguistics* (5th ed.). London/NY: Routledge.
- Hoover, D. L., Culpeper, J., & O'Halloran, K. (Eds.). (2014). *Digital Literary Studies: Corpus Approaches to Poetry, Prose, and Drama*.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge: CUP.
- Howard, P. (2003). *The Orange Mocha-Chip Frappuccino Years*. Dublin: Penguin.
- . (2005). *The Curious Incident of the Dog in the Nightdress*. Dublin: Penguin Ireland.

- . (2007a). *Ross O'Carroll-Kelly's Guide to South Dublin: How to Get By On, Like, € 10,000 a Day*. Dublin: Penguin Ireland.
- . (2007b). *Should Have Got Off At Sydney Parade*. Dublin: Penguin.
- . (2008). *This Champagne Mojito is the Last Thing I Own*. Dublin: Penguin.
- . (2009). *We Need to Talk about Ross*. Dublin: Penguin.
- . (2011). *NAMA Mia!* Dublin: Penguin Ireland.
- . (2013). *Downturn Abbey*. Dublin: Penguin Ireland.
- . (2014). *Keeping up with the Kalashnikovs*. Dublin: Penguin Ireland.
- . (2017). 'I spend my working life sitting behind a desk thinking in the voice of a south Dublin idiot'. In A. Barry (Ed.), *thejournal.ie*. <https://www.thejournal.ie/paul-howard-ross-ocarroll-kelly-interview-3616476-Oct2017/>.
- . (2019). *In Conversation with Paul Howard*. Interviewed by A. M. Terrazas-Calero. Lime Tree Theatre, Limerick (Ireland): Unpublished interview and public event.
- . (2021). Paul Howard: Ten Things I've Learned Writing Ross O'Carroll-Kelly, *The Irish Times*. Accessed at: <https://www.irishtimes.com/culture/books/paul-howard-ten-things-i-ve-learned-writing-ross-o-carroll-kelly-1.4518120>
- Huddleston, R., & Pullum, G. K. (2017). *The Cambridge Grammar of the English Language*. Cambridge: CUP.
- Hughes, G. (2006). *An Encyclopedia of Swearing: The social history of oaths, profanity, foul language, and ethnic slurs in the English-speaking world*. Armonk (NY)/London: M.E. Sharpe.
- Hughes, S. E. (1992). Expletives of lower working-class women. *Language in Society*, 21, 291-303.
- Hunni, M. (Ed.) (2003) Urban Dictionary. Accessed at: <https://www.urbandictionary.com/define.php?term=valley%20girl>.
- Hunt Mahoney, C. (1998). *Contemporary Irish Literature: Transforming Tradition*. NY: St Martin's Press.
- Ingman, H. (2009). *A History of the Irish Short Story*. Cambridge: Cambridge University Press.
- The Irish Times Book Club*. (2022). 2016. Accessed at: <https://www.irishtimes.com/culture/books/the-book-club>
- Irishslang.info*. (2016). Accessed at: <https://www.irishslang.info/>
- Irwin, P. (2009). Hijacked predicate raising in an AAE attributive construction. In P. Irwin & V. Vázquez-Rojas Maldonado (Eds.), *NYU Working Papers in Linguistics (NYUWPL: Papers in Syntax)* (Vol. 2). Accessed at:

https://as.nyu.edu/content/dam/nyu-as/linguistics/documents/nyuwpl/irwin_09_hijacked_predicate_raising_aae_attri_butive_nyuwpl2.pdf

- . (2014). SO [TOTALLY] speaker-oriented: An analysis of 'Drama SO'. In R. Zanuttini & L. Horn (Eds.), *Micro-syntactic Variation in North American English* (pp. 29-70). Oxford/NY: OUP.
- Ito, R., & Tagliamonte, S. A. (2003). 'Well' weird, 'right' dodgy, 'very' strange, 'really' cool: Layering and recycling in English intensifiers. *Language in Society*, 32(257-279).
- Ives, S. (1950). A Theory of Literary Dialect. *Tulane Studies in English, New Orleans*, 2, 137-182.
- Jakobson, R. (1980). *The Framework of Language*. Ann Arbor, Michigan: Michigan Studies in the Humanities.
- Jay, T. (1992). *Cursing in America: A psycholinguistic study of dirty language in the courts, in the movies, in the schoolyards and on the streets*. Philadelphia/Amsterdam: John Benjamins.
- . (2000). *Why we curse: A neuro-psycho-social theory of speech*. Philadelphia/Amsterdam: John Benjamins.
- . (2009). The Utility and Ubiquity of Taboo Words. *Perspectives on Psychological Science*, 4(2), 153-161.
- Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research*, 4, 267-288.
- Jeffries, L. (2007). *Textual Construction of the Female Body*. Basingstoke: Palgrave.
- . (2010). *Critical Stylistics*. Basingstoke: Palgrave.
- Jeffries, L., & McIntyre, D. (2010). *Stylistics*. Cambridge: CUP.
- Jespersen, O. (1922). *Language: Its Nature, Development, and Origin*. London: Allen & Unwin.
- Johnstone, B., Andrus, J., & Danielson, A. E. (2006). Mobility, Indexicality, and the Enregisterment of 'Pittsburghese'. *Journal of English Linguistics*, 34(2), 77-104.
- Joseph, J. E. (2010). Identity. In C. Llamas & D. Watt (Eds.), *Language and Identities* (pp. 9-17). Edinburgh: Edinburgh University Press.
- Joyce, P. W. (1910). *English as We Speak It in Ireland*. London: Longmans, Green and Co.
- Jucker, A. H., & Smith, S. W. (1998). "And people just you know like 'wow'": Discourse Markers as Negotiating Strategies. In A. H. Jucker & Y. Ziv (Eds.), *Discourse Markers. Description and Theory* (pp. 171-201). Amsterdam/Philadelphia: John Benjamins.

- Kallen, J. (2005). Politeness in Modern Ireland: 'You know the way in Ireland, it's done without being said'. In L. Hickey & M. Stewart (Eds.), *Politeness in Europe* (Vol. 130-44). Clevedon: Multilingual Matters.
- . (2006). *Arrah, Like, You Know: The Dynamics of Discourse Marking in ICE-Ireland*. Paper presented at the Paper presented at Sociolinguistics Symposium 16, University of Limerick.
- . (2013). *Irish English, volume 2: The Republic of Ireland*. Berlin/Boston: De Gruyter Mouton.
- Kallen, J., & Kirk, J. M. (2001). Aspects of the Verb Phrase in Standard Irish English: A Corpus-based Approach. In J. M. Kirk & D. P. Ó Baoill (Eds.), *Language Links: The Languages of Scotland and Ireland* (pp. 59-79). Belfast: Cló Ollscoil na Banríona.
- Kallen, J. L., & Kirk, J. M. (2008). *ICE-Ireland: A User's Guide*. Belfast: Cló Ollscoil na Banríona.
- Kallen, J. L., & Kirk, J. M. (2012). *SPICE-Ireland: A User's Guide Documentation to accompany the SPICE-Ireland Corpus: Systems of Pragmatic annotation in ICE-Ireland*. Belfast: Cló Ollscoil na Banríona.
- Keegan, C. (2007). *Walk the Blue Fields*. London: Faber and Faber.
- . (2010). *Foster*. London: Faber & Faber.
- Kelly, A. (2017). The Re-education of Ross O'Carroll-Kelly. *Éire-Ireland*, 52(1&2), 49-77.
- Kelly, G., & Kelly, S. (Eds.). (2007). *Overheard in Dublin Again. Dublin Wit from overheardindublin.com*. Dublin: Gill & Macmillan.
- Kiberd, D. (1979/1993). *Synge and the Irish language*. London: Macmillan.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*(1), 7-36.
- Kim, E. (2019). Dude, Where Does the 'Surfer Voice' Actually Come From? *MEL Magazine*. Accessed at: <https://melmagazine.com/en-us/story/surfer-voice-california-hawaii-history>
- King, B. (2010). "All us girls were like euuh": Conversational work of 'be like' in New Zealand adolescent talk. *New Zealand English Journal*, 24, 17-36.
- Kohn, M. E., & Askin Franz, H. (2009). Localized Patterns for Global Variants: The Case of Quotative Systems of African American and Latino Speakers. *American Speech*, 84(3), 259-297.
- Kozloff, S. (2000). *Overhearing Film Dialogue*. Ewing, NJ: University of California Press.

- Krempels, H. (2012). Dark Lies the Island by Kevin Barry-review. *The Guardian*. Accessed at: <https://www.theguardian.com/books/2012/apr/08/dark-lies-island-kevin-barry-review>
- Lakoff, R. (1973). Language and Woman's Place. *Language in Society*, 2(1), 45-80.
- . (1975). *Language and Woman's Place*. NY: Octagon Books.
- Lancashire, I., Bradley, J., McCarty, W., Stairs, M., & Wooldridge, T. R. (1996). *Using TACT with Electronic Texts: A Guide to Text-Analysis Computing Tools, Version 2.1 for MS-DOS and PC DOS*. New York: MLA.
- Lee, J. (2013). Jumping Off a Cliff: An Interview with Kevin Barry. *The Paris Review*. Accessed at: <https://www.theparisreview.org/blog/2013/11/12/jumping-off-a-cliff-an-interview-with-kevin-barry/>
- Leech, G. (1991). The state of the art in corpus linguistics. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics* (pp. 8-29). London: Longman.
- . (1997). Introducing Corpus Annotation. In R. Garside, G. Leech & T. McEnery (Eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora*. London/New York: Routledge.
- . (2007). New Resources, or Just Better Old Ones? The Holy Grail of Representativeness. In M. Hundt, N. Nesselhauf & C. Biewer (Eds.), *Corpus Linguistics and the Web* (pp. 133-150). Amsterdam/NY: Rodopi.
- Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: The tagging of the British National Corpus *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94)* (pp. 622-628). Kyoto: Japan.
- Leech, G., & Short, M. (2007). *Style in Fiction: A Linguistic Introduction to English Fictional Prose* (2nd ed.). London: Pearson.
- Leerssen, J. (1996). *Mere Irish and Fíor-Ghael: Studies in the Idea of Nationality, its Development and Literary Expression prior to the Nineteenth Century*. Cork: Cork University Press.
- Levey, S. (2003). "He's like 'Do it now!' and I'm like 'No!'": Some innovative quotative usage among young people in London. *English Today*, 73(19), 1.
- Levey, S. (2006). The Sociolinguistic Distribution of Discourse Marker 'Like' in Preadolescent Speech. *Multilingua* 25, 4, 413-444.
- Liberman, M. (2013). Sexy baby vocal virus. *Language Log*. Accessed at: <https://languagelog.ldc.upenn.edu/nll/?p=5842>
- Lii. (2005). *Urban Dictionary*. Accessed at: <https://www.urbandictionary.com/author.php?author=Lii>
- Lillo, A. (2007). Turning Puns into Names and Vice Versa. *SKY Journal of Linguistics*, 20, 429-440.

- Linehan, H. (2016, Feb 20, 2016). How the Irish, like, speak now?, *The Irish Times*. Retrieved from <https://www.irishtimes.com/culture/tv-radio-web/how-the-irish-like-speak-now-1.2541111>
- Ljung, M. (1986). *Om Svordomar*. Stockholm: Akademilitteratur.
- Locher, M. A., & Jucker, A. H. (Eds.). (2017). *Pragmatics of Fiction*. Berlin/Boston: De Gruyter Mouton.
- Lonergan, J. (2016). Real and perceived variation in Dublin English. In J. Cramer & C. Montgomery (Eds.), *Cityscapes and Perceptual Dialectology: Global Perspectives on Non-Linguists' Knowledge of the Dialect Landscape* (pp. 233-256). Boston/Berlin: De Gruyter Mouton.
- Lorenz, G. (2002). 'Really worthwhile' or 'not really significant?': A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English. In I. Wischer & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 143-162). Amsterdam/NY: John Benjamins.
- Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, 22(3), 319-344.
- Lucek, S., & Garnett, V. (2020). Perceptions of Linguistic identity among Irish English Speakers. In R. Hickey & C. P. Amador-Moreno (Eds.), *Irish Identities: Sociolinguistic Perspectives* (pp. 104-130). Boston/Berlin: De Gruyter Mouton.
- Lucy, B., Maher, E., & O'Brien, E. (Eds.). (2019). *Recalling the Celtic Tiger*. Oxford: Peter Lang.
- Lupino, C. (2014). The Old and the New in Claire Keegan's Short Fiction. *Journal of the Short Story in English*, 63(Autumn 2014), 1-14.
- Lynch, P. A. (2006). Hiberno-English in the plays of Marina Carr. *Études Irlandaises*, 31(2), 109-123.
- Macaulay, R. (1995). The adverbs of authority. *English World Wide*, 16, 37-60.
- . (2001). 'You're like "why not?": The quotative expression of Glasgow adolescents. *Journal of Sociolinguistics*, 5, 3-21.
- . (2002). Extremely interesting, very interesting, or only quite interesting: Adverbs and social class. *Journal of Sociolinguistics*, 6, 398-417.
- . (2006). Pure grammaticalization: The development of a teenage intensifier. *Language Variation and Change*, 18, 267-283.
- Mackenzie, J. L. (2019). The syntax of an emotional expletive in English. In J. L. Mackenzie & L. Alba-Juez (Eds.), *Emotion in Discourse* (pp. 55-86). Amsterdam/Philadelphia: John Benjamins.

- Mahlberg, M. (2007). Corpus stylistics: bridging the gap between linguistic and literary studies. In M. Hoey, M. Mahlberg, M. Stubbs & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 219-246). London/NY: Continuum.
- . (2013). *Corpus Stylistics and Dicken's Fiction*. Abingdon: Routledge.
- . (2014). Corpus Stylistics. In M. Burke (Ed.), *The Routledge Handbook of Stylistics* (pp. 378-392). Abingdon: Routledge.
- . (2016). Corpus Stylistics In V. Sotirova (Ed.), *The Bloomsbury Companion to Stylistics* (Vol. 139-156). London: Bloomsbury.
- Mahlberg, M., & McIntyre, D. (2011). A case for corpus stylistics: Ian Fleming's Casino Royale. *English Text Construction*, 4(2), 204-227.
- Mahlberg, M., Stockwell, P., de Joode, J., Smith, C., & Brook O'Donnell, M. (2016). CLic Dickens: novel uses of concordances for the integration of corpus stylistics and cognitive poetics. *Corpora* 2016, 11(3), 433-463.
- Mahlberg, M., & Wiegand, V. (2018). Corpus stylistics, norms and comparisons: Studying speech in *Great Expectations*. In R. Page, B. Busse & N. Nørgaard (Eds.), *Rethinking Language, Text and Context: Interdisciplinary Research in Stylistics in Honour of Michael Toolan* (pp. 123-143). London: Routledge.
- Mastropierro, L. (2018). *Corpus stylistics in Heart of Darkness and its Italian translations*. London: Bloomsbury.
- Mathis, T., & Yule, G. (1994). Zero Quotatives. *Discourse Processes*, 18(1), 63-76.
- McCafferty, K. (2005). William Carleton between Irish and English: Using literary dialect to study language contact and change. *Language and Literature*, 14(4), 339-362.
- . (2008). 'On the trail of "intolerable Scoto-Hibernic jargon": Ulster English, Irish English and dialect hygiene in William Carleton's *Traits and stories of the Irish peasantry* (First Series, 1830). In M. Dossena, R. Dury & M. Gotti (Eds.), *English Historical Linguistics 2006: Selected papers from the fourteenth international conference on English historical linguistics (ICEHL 14), Bergamo, 21-25 August, 2006. III. Geo-Historical Variation in English* (pp. 171-184). Amsterdam: John Benjamins.
- . (2009). "Preserv[ing] every thing Irish?": The Hiberno-English dialect of William Carleton's peasants. In C. P. Amador-Moreno & A. Nunes (Eds.), *The Representation of the Spoken Mode in Fiction* (pp. 67-99). Lewinston, NY: The Edwin Mellen Press.
- McCafferty, K., & Amador-Moreno, C. P. (2012a). CORIECOR – A Corpus of Irish English Correspondence, c. 1700 – 1900. Compiling and using a diachronic corpus to study the evolution of Irish English. In B. Migge & M. Ní Chiosain (Eds.), *New Perspectives in Irish English* (pp. 265-288). Amsterdam: John Benjamins.

- . (2012b). A corpus of Irish English correspondence (CORIECOR). In B. Migge & M. Ní Chiosáin (Eds.), *New Perspectives on Irish English (Varieties of English Around the World)* (pp. 265-287). Amsterdam/Philadelphia: John Benjamins.
- . (2012c). 'I will be expecting a letter from you before this reaches you'. A corpus-based study of shall/will variation in Irish English correspondence. . In M. Dossena (Ed.), *Letter writing in Late Modern Europe* (pp. 179-204). Amsterdam: John Benjamins.
- . (2019). 'But a[h] Hellen d[ea]r sure you have it more in your power in every respect than I have'. Discourse marker sure in Irish English. In S. Jansen & L. Siebers (Eds.), *Processes of Change: Studies in Late Modern and Present-Day English* (pp. 73–94). Amsterdam: John Benjamins.
- McCarthy, J. J. (1982). Prosodic structure and expletive infixation. *Language*, 58(3), 574-590.
- McCready, E., & Kaufmann, M. (2013). *Maximum Intensity*. Paper presented at the Semantics Group, Keio University.
- McEnery, A., & Xiao, Z. (2004). Swearing in modern British English: the case of 'fuck' in the BNC. *Language and Literature*, 13(3), 235-268.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics Method, Theory and Practice*. Cambridge: CUP.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction* (2nd ed.). Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book*. London: Routledge.
- McIntyre, D. (2015). Towards an integrated corpus stylistics. *Topics in Linguistics*, 16(1), 59-68.
- McIntyre, D., & Walker, B. (2019). *Corpus Stylistics: A Practical Introduction*. Cambridge: CUP.
- . (2022). Using corpus linguistics to explore the language of poetry: a stylometric approach to Yeats' poems. In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics, 2nd edition* (pp. 499-516). London/NY: Routledge.
- McKeon, C. (2003). Down those mean streets a cabby must go, *Independent.ie*. Retrieved from <https://www.independent.ie/lifestyle/down-those-mean-streets-a-cabby-must-go-26236491.html>
- McMillan, J. B. (1980). Infixing and Interposing in English. *American Speech*, 55(3), 160-183.
- Mianowski, M. (2017). *Post Celtic Tiger Landscapes in Irish Fiction*. London/NY: Routledge.

- The Millions*. (2016). Accessed at: <https://themillions.com/>
- Montoro, R. (2012). *Chick Lit: The Stylistics of Cappuccino Fiction*. London/New York: Continuum.
- . (2015). Style in popular literature. In V. Sotirova (Ed.), *The Bloomsbury Companion to Stylistics* (pp. 673-689). London: Bloomsbury.
- . (2019). 'Investigating Syntactic Simplicity in Popular Fiction A Corpus Stylistics Approach. In R. Page, B. Busse & N. Nørgaard (Eds.), *Rethinking Language, Text and Context: Interdisciplinary Research in Stylistics in Honour of Michael Toolan* (pp. 63-78). London/NY: Routledge.
- Moss, L. (2014). *Corpus Stylistics and Henry James's Syntax*. (Doctor of Philosophy), University College London.
- Mur11. (2006). *Urban Dictionary*. Accessed at: <https://www.urbandictionary.com/define.php?term=dawson%20and%20joey>
- Murdoch, J. (1892). *From Australia and Japan*. London: Walter Scott Ltd.
- Murphy, B. (2009). 'She's a F**king Ticket': The pragmatics of f**k in Irish English-an age and gender perspective. *Corpora* 4(1), 85-106.
- . (2010). *Corpus and Sociolinguistics: Investigating age and gender in female talk*. Amsterdam/Philadelphia: John Benjamins.
- . (2015). A corpus-based investigation of pragmatic markers and sociolinguistic variation in Irish English. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic Markers in Irish English* (pp. 65-88). Amsterdam/NY: John Benjamins.
- Murphy, B., & Palma-Fahey, M. E. (2018). Exploring the construction of the Irish Mammy in 'Mrs Brown's Boys'. *Pragmatics and Society*, 9(2), 297-328.
- Mustanoja, T. F. (1960). *A Middle English Syntax*. Helsinki: Société néophilologique.
- National Readership Survey. (2016). *Social Grade*. Accessed at: <http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/>
- Nestor, N., Ní Chasaide, C., & Reagan, R. (2012). Discourse 'like' and social identity—a case study of Poles in Ireland. In B. Migge & M. Ní Chiosáin (Eds.), *New Perspectives on Irish English* (pp. 327-354). Amsterdam/NY: John Benjamins.
- Nevalainen, T., & Ramoulin-Brunberg, H. (2017). *Historical Sociolinguistics: Language Change in Tudor and Stuart England* (2nd ed.). London/NY: Routledge.
- Ní Chonchúir, N. (2010a). An interview with Nuala Ní Chonchúir about her novel, *You*. In E. Baines (Ed.), *Elizabeth Baines: How to be a writer without ending up sozzled, behind bars or insane blogspot*. <http://elizabethbaines.blogspot.com/2010/07/interview-with-nuala-ni-chonchuir-about.html>
- . (2010b). *You*. Dublin: New Island.

- . (2013). Interview: Nuala Ní Chonchúir. In R. Rennicks (Ed.), *A Trip to Ireland: Your personal guided-tour through the best of Ireland, Irish culture, and history*. Blogspot. <http://atriptoireland.com/2013/11/06/interview-nuala-ni-chonchuir/>
- . (2014). *The Closet of Savage Mementos*. Dublin: New Island Publishing.
- Ní Dhuibhne, É. (2000). *The Dancers Dancing*. Belfast: Blackstaff Press.
- . (2003). *Midwife to the Fairies: New and Selected Stories*. Cork: Attic Press Ltd.
- . (2007). Give a Thing and Take it Back. *Dublin Review of Books*.
- Ní Eigeartaigh, A. (2009). Invented Irishness: The Americanization of Irish Identity in the Works of Joseph O'Connor. *Irish Journal of American Studies*(1).
- Notes and Queries: A Medium of Intercommunication for Literary Men, General Readers, etc.* (1921). (Vol. 12 (9)). London: The Times Publishing Company.
- Núñez Pertejo, P., & Palacios Martínez, I. M. (2014). 'That's absolutely crap, totally rubbish': The use of the intensifiers 'absolutely' and 'totally' in the spoken language of British adults and teenagers. *Functions of Language* 21(2), 210-237.
- O'Brien, E. (2015). 'Tendency-wit': the cultural unconscious of the Celtic Tiger in the writings of Paul Howard. In E. Maher & E. O'Brien (Eds.), *From prosperity to austerity: a socio-cultural critique of the Celtic Tiger and its aftermath*. Manchester: Manchester University Press.
- . (fc.). *Reading Paul Howard: The Art of Ross O'Carroll Kelly*. London/NY: Routledge.
- O'Byrne, N. (2006). The O'Byrne Files-Dublin Slang Dictionary and Phrase-Book. from <http://homepage.tinet.ie/~nobyrne/a-zcompact.htm>
- O'Donovan, F. (2009). Irish Identity is Far From 'Ideal'. *SOCHEOLAS: Limerick Student Journal of Sociology*, 2(1), 95-115.
- O'Keeffe, A. (fc.). Irish English Corpus Linguistics. In R. Hickey (Ed.), *The Oxford Handbook of Irish English*. Oxford: Oxford University Press.
- O'Keeffe, A., Clancy, B., & Adolphs, S. (2020). *Introducing Pragmatics in Use* (2nd ed.). Abingdon/NY: Routledge.
- O'Malley, J. P. (2013). Roddy Doyle: I'm a middle class person commenting on working class life *The Spectator*. <https://www.spectator.co.uk/article/roddy-doyle-i-m-a-middle-class-person-commenting-on-working-class-life>
- O'Sullivan, J. (2013). Advanced Dublin English in radio advertising. *World Englishes*, 32(3), 358-376.
- . (2015). Pragmatic markers in contemporary radio advertising in Ireland. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic Markers in Irish English* (pp. 318-347). Amsterdam: John Benjamins.

- . (2018). Advanced Dublin English as audience and reference design in Irish radio advertising. The “initiative” role of advertising in the construction of identity. . *English World Wide*, 39(1), 60-84.
- . (2020). Constructing identity in radio advertising in Ireland. In R. Hickey & C. P. Amador-Moreno (Eds.), *Irish Identities: Sociolinguistic Perspectives* (pp. 220-251). Berlin/Boston: De Gruyter Mouton.
- O'Toole, F. (2010). *Enough is Enough: How to Build a New Republic*. London: Faber & Faber Ltd.
- Oxford Advanced Learner's Dictionary*. (2010) (8th edition ed.). Oxford: Oxford University Press.
- Page, N. (1988). *Speech in the English Novel* (2nd edition; first edition 1973 ed.). Atlantic Highlands, NJ: Humanities Press International.
- Palma-Fahey, M. E. (2005). *A cross-cultural discourse and pragmatic analysis of two soap operas: Fair City and Amores de Mercado compared*. (Doctor of Philosophy), Mary Immaculate College.
- . (2015). “Yeah well, probably, you know I wasn’t that big into school, you know”: Pragmatic markers and the representation of Irish English in Fictionalised Dialogue. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic Markers in Irish English* (pp. 348–369). Amsterdam/Philadelphia: John Benjamins.
- Park, D. (2012). *The Light of Amsterdam*. London Bloomsbury.
- Partington, A. (1993). Corpus evidence of language change: The case of intensifiers. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and Technology: In honour of John Sinclair* (pp. 177-192). Amsterdam/NY: John Benjamins.
- Peach, L. (2004). *The Contemporary Irish Novel. Critical Readings*. Houndmills: Palgrave Macmillan.
- Peters, H. (1994). Degree Adverbs in Early Modern English. In D. Kastovsky (Ed.), *Studies in Early Modern English* (pp. 269-288). Berlin/NY: Mouton de Gruyter.
- Piazza, R. (2011). Pragmatic deviance in realist horror films: A look at films by Argento and Fincher. In R. Piazza, M. Bednarek & F. Rossi (Eds.), *Telecinematic Discourse: Approaches to the language of films and television series* (pp. 85-104). Amsterdam/Philadelphia: John Benjamins.
- Pichler, H. (2013). *The Structure of Discourse-Pragmatic Variation*. Amsterdam/Philadelphia: John Benjamins.
- Plag, I., Braun, M., Arndtlappe, S., & Schramm, M. (2009). *Introduction to English Linguistics*. Berlin/Boston: De Gruyter Mouton.
- Poplack, S., & Tagliamonte, S. A. (2001). *African American English in the Diaspora*. Oxford: Basil Blackwell.

- Potts, C. (2004). Lexicalized intonational meaning. In S. Kawahara (Ed.), *UMPO 30: Papers on prosody* (pp. 129-146). Amherst, MA: GLSA.
- . (2007). The Expressive Dimension. *Theoretical Linguistics*, 33(2), 165-197.
- Price, J. (2015). "*Oh Jesus Christ!*" *The Use of Bad Language in Contemporary American Television Series (Honours Thesis)*. (Honours Thesis), University of Sydney.
- Pullum, G. K., & Huddleston, R. (2017). Prepositions and preposition phrases. In R. Huddleston & G. K. Pullum (Eds.), *The Cambridge Grammar of the English Language*. Cambridge: CUP.
- Quaglio, P. (2009). *Television Dialogue: The sitcom Friends vs natural conversation*. Amsterdam/Philadelphia: John Benjamins.
- Queen, R. (2015). *Vox Popular: The Surprising Life of Language in the Media*. Malden, MA: Wiley-Blackwell.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. NY: Longman.
- Rayson, P. (2009). Wmatrix: A web based corpus processing environment. from <http://ucrel.lancs.ac.uk/wmatrix/>
- Reichelt, S., & Durham, M. (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.
- Reppen, R. (2010). Building a corpus: What are the key considerations? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 31-37). London/NY: Routledge.
- Rickford, J. R., Wasow, T., Zwicky, A., & Buchstaller, I. (2007). Intensive and quotative ALL: Something old, something new. *American Speech*, 82(1), 3-31.
- Rodríguez Louro, C. (2013). Quotatives down under: 'Be like' in cross-generational Australian English speech. *English World Wide*, 34(1), 48-76.
- Rohan, E. (2011). Review of *City of Bohane* by Kevin Barry. *Goodreads*. Accessed at: <https://www.goodreads.com/review/show/246965065>
- Romaine, S., & Lange, D. (1991). The Use of 'Like' as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech*, 66(3), 227-279.
- Ruane, D. (2003). *Tales in a Rearview Mirror*. Dublin: Gill & Macmillan.
- . (2004). *I'm Irish Get Me Out of Here!* Dublin: Gill & Macmillan.
- Ruano San Segundo, P. (2018). A Corpus-Based Approach to Waiting for Godot's Stage Directions: A Comparison between the French and the English Version. In D. Villanueva Romero, C. P. Amador-Moreno & M. Sánchez García (Eds.), *Voice and Discourse in the Irish Context* (pp. 139-167). Cham: Palgrave Macmillan.

- Rühlemann, C. (2007). *Conversation in Context: A Corpus-Driven Approach* London: Continuum.
- Rühlemann, C. (2008). Conversational grammar--bad grammar?: A situation-based description of quotative 'I goes' in the BNC. *ICAME Journal*, 32(157-177).
- Ryan, D. (2012). *The Spinning Heart*. Dublin: Doubleday Ireland/Lilliput Press.
- . (2013). *The Thing about December*. Dublin: Doubleday Ireland/The Lilliput Press.
- Saarenpää, T. (2016). *A Comparative Corpus Study on Intensifiers in Canadian English and New Zealand English (MA thesis)*. Master's Thesis. English Language and Literature. School of Language, Translation and Literary Studies, University of Tampere. Retrieved from <https://trepo.tuni.fi/bitstream/handle/10024/98915/GRADU-1462354862.pdf?sequence=1&isAllowed=y>
- Scott, M. (2010). What can corpora software do? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 136-151). London/NY: Routledge.
- . (2012). *Wordsmith Tools*, version 6. Stroud: Lexical Analysis Software.
- Schiffrin, D. (1981). Tense variation in narrative. *Language*, 57(1), 45-62.
- . (1987). *Discourse Markers*. Cambridge: CUP.
- Schneider, E. W. (2002). Investigating variation and change in written documents. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (pp. 67-96). Oxford: Blackwell.
- Schourup, L. (1982). Response to Ronald R. Butters, Quoting with go "say". *American Speech*, 57, 148-149.
- . (1985). *Common Discourse Particles in English Conversation*. NY: Garland.
- Schweinberger, M. (2012). LIKE in Irish English. In B. Migge & M. Ní Chiosáin (Eds.), *New Perspectives on Irish English* (pp. 179-202). Amsterdam/NY: John Benjamins.
- . (2014). *The Discourse Marker LIKE: A Corpus-Based Analysis of Selected Varieties of English*. (Doctor of Philosophy), University of Hamburg, Unpublished. Retrieved from <https://d-nb.info/1058681893/34>
- . (2015). A comparative study of the pragmatic marker like in Irish English and in south-eastern varieties of British English. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic Markers in Irish English* (pp. 114-134). Amsterdam/NY: John Benjamins.
- . (2020). Analyzing change in the American English amplifier system in the fiction genre. In P. Rautionaho, A. Nurmi & J. Klemola (Eds.), *Corpora and the*

- Changing Society: Studies in the Evolution of English* (pp. 223-249). Amsterdam/NY: John Benjamins.
- Semino, E. (2011). Stylistics. In J. Simpson (Ed.), *The Routledge Handbook of Applied Linguistics* (pp. 541-544). Abingdon/NY: Routledge.
- Semino, E., & Short, M. (2004). *Corpus Stylistics: Speech, writing and thought presentation in a corpus of English writing*. London/New York: Routledge.
- Sheidlower, J. (2009). *The F-Word*. NY: OUP.
- Siegel, M. E. A. (2002). Like: The Discourse Particle and Semantics. *Journal of Semantics*, 19(1), 35-71.
- Siemund, P., Maier, G., & Schweinberger, M. (2009). Towards a more fine-grained analysis of the areal distributions of non-standard features of English. In H. Palausto & E. Penttilä (Eds.), *Language Contacts Meet English Dialects. Studies in Honor of Markku Filppula* (pp. 19-46). Cambridge: Cambridge Scholar Publishing.
- Simpson, J. A., & Weiner, E. S. C. (Eds.). (2021) Oxford English Dictionary (2nd ed., Vols. 18, Thro-Unelucidated). Oxford: Clarendon Press.
- Simpson, P. (1997). *Language through Literature: An Introduction*. London: Routledge.
- . (2004). *Stylistics: A Resource Book for Students*. London/NY: Routledge.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- . (1992). Trust the Text: The Implications are Daunting. In M. Davies & L. Ravelli (Eds.), *Advances in Systemic Linguistics* (pp. 5-19). London: Pinter.
- . (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- . (2005). Corpus and text--basic principles. In M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice* (pp. 1-16). Oxford: Oxbow.
- Smith, N., Hoffmann, S., & Rayson, P. (2008). Corpus Tools and Methods, Today and Tomorrow: Incorporating Linguists' Manual Annotations. *Literary and Linguistic Computing*, 23(2), 163-180.
- Stapleton. (2010). Swearing. In M. A. Locher & S. L. Graham (Eds.), *Interpersonal Pragmatics* (pp. 289-306). Berlin/NY: De Gruyter Mouton.
- Stenström, A.B. (1991). Expletives in the London-Lund corpus. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 239-253). London: Longman.
- . (2006). Taboo words in teenage talk: London and Madrid girls' conversations compared. *Spanish in Context* 3(1), 115–138.
- Stenström, A.-B., Andersen, G., & Hasund, I. K. (2002). *Trends in Teenage Talk: Corpus compilation, analysis and findings*. Amsterdam/Philadelphia: John Benjamins.

- Stewart, B. (2011). A Brief History of Ricorso. Retrieved from <http://ricorso.net/rx/about/index.htm>
- Stockwell, P., & Mahlberg, M. (2015). Mind-modelling with corpus stylistics in David Copperfield. *Language and Literature*, 24(2), 129-147.
- Stockwell, P., & Wynne, M. (2006). *Corpus Stylistics: a public inquiry*. Paper presented at the 3rd International IVACS Conference, University of Nottingham, England.
- Stoffel, C. (1901). *Intensives and Down-toners*. Heidelberg: Carl Winter.
- Sweetemotion09. (2007). *Urban Dictionary*. Accessed at: <https://www.urbandictionary.com/define.php?term=Jock&defid=2428668>
- Tagliamonte, S. A. (2005). 'So' who?' Like' how? 'Just' what?: Discourse markers in the conversations of Young Canadians *Journal of Pragmatics*, 37, 1896-1915.
- . (2008). "So different and pretty cool!" Recycling intensifiers in Toronto, Canada. *English Language and Linguistics*, 12(2), 361-394.
- . (2012). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.
- . (2016). *Teen Talk: The Language of Adolescents*. Cambridge: CUP.
- Tagliamonte, S. A., & D'Arcy, A. (2004). 'He's like, she's like': The quotative system in Canadian youth. *Journal of Sociolinguistics*, 8, 493-514.
- . (2007). Frequency and variation in the community grammar: tracking a new change through the generations. *Language Variation and Change*, 19, 119-217.
- Tagliamonte, S. A., D'Arcy, A., & Rodríguez Louro, C. (2016). Outliers, impact, and rationalization in linguistic change. *Language*, 92(4), 824-849.
- Tagliamonte, S. A., & Hudson, R. (1999). Be like et al. beyond America: the quotative system in British and Canadian youth. *Journal of Sociolinguistics*, 3(2), 147-172.
- Tagliamonte, S. A., & Roberts, C. (2005). "So weird"; "So cool"; "So innovative": The Use of Intensifiers in the Television Series 'Friends'. *American Speech*, 80(3), 280-300.
- Taniguchi, J. (1972). *A Grammatical Analysis of Artistic Representations of Irish English with a Brief Discussion of Sounds and Spelling*. Tokyo: Shinozaki Shorin.
- Tannen, D. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (Ed.), *Direct and Indirect Speech* (Vol. 311-32). Berlin: Mouton.
- . (2007a). Talking the Dog: Framing Pets as Interactional Resources in Family Discourse. In D. Tannen, S. Kendall, & C. Gordon (Eds.), *Family Talk: Discourse and Identity in Four American Families* (pp. 49-70). Oxford: OUP.




- . (2007b). *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse* (2nd ed.). Cambridge: CUP.
- Terraschke, A. (2010). “And I was like ‘ah yeah, what are they talking about?’” – The use of quotatives in New Zealand English. In L. de Beuzeville & P. Peters (Eds.), *From the Southern Hemisphere: parameters of language variation: Proceedings of the 2008 Conference of the Australian Linguistics Society* (pp. 1-14). Sydney: Department of Linguistics and e-Scholarship Repository, University of Sydney.
- Terrazas-Calero, A. M. (2016a). ‘*The dude’s like me – Charmin focking Ultra*’: *The Development of Taboo Language in Paul Howard’s Fictionalized Irish English*. Paper presented at the 8th International, Biennial Inter-Varietal Applied Corpus Studies Association (IVACS) Conference, Bath Spa University, Bath, England.
- . (2016b). ‘*You know what I’m focking talking about*’: *Taboo Language in Paul Howard’s Fictionalized Irish English*. Paper presented at the Annual Inter-Varietal Applied Corpus Studies Association (IVACS) symposium, Universidad Pompeu Fabra, Barcelona (Spain).
- . (2017). ‘*That’s when I, like, seriously seriously flip?*’: *Fictionalized Irish English and the pragmatics of ‘like’*. Paper presented at the 15th International Pragmatics Conference, Belfast, 16-21 July 2017.
- . (2020). ‘These kids don’t even sound...Irish anymore’: Representing ‘new’ Irishness in contemporary Irish fiction. In R. Hickey & C. P. Amador-Moreno (Eds.), *Irish Identities: Sociolinguistic Perspectives* (pp. 252-282). Berlin/Boston: De Gruyter Mouton.
- Thorne, T. (Ed.) (2014) *Dictionary of Contemporary Slang* (4th ed.). London/NY: Bloomsbury.
- Timmis, I. (2015). Evaluation: Tails. In K. Aijmer & C. Rühlemann (Eds.), *Corpus Pragmatics: A Handbook* (pp. 304-331). Cambridge: CUP.
- Todd, L. (1989). *The Language of Irish Literature*. Basingstoke: Macmillan Education Ltd.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.
- Traugott, E. C. (1982). From Propositional to Textual and Expressive Meanings: Some Semantic-Pragmatic Aspects of Grammaticalization. In W. P. Lehmann & Y. Malkiel (Eds.), *Perspectives on Historical Linguistics* (pp. 245-271). Amsterdam: John Benjamins.
- . (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1), 31-55.
- . (1995). *The Role of the Development of Discourse Markers in a Theory of Grammaticalization (version of 11/97)*. Paper presented at 12th International Conference on Historical Linguistics, Manchester, UK, August 1995. Available at: <https://web.stanford.edu/~traugott/papers/discourse.pdf>

- Traugott, E. C. (1995b). Subjectivisation in grammaticalisation. In D. Stein & S. Wright (Eds.), *Subjectivity and Subjectivisation* (pp. 31-55). Cambridge: CUP.
- Truesdale, S., & Meyerhoff, M. (2015). Acquiring some like-ness to others. *Te Reo*, 58, 3-28.
- “Totally”. (2022.) *Oxford English Dictionary*. Accessed at: <http://www.oed.com>
- Vandelanotte, L., & Davidse, K. (2009). The emergence and structure of 'be like' and related quotatives: A constructional account. *Cognitive Linguistics*, 20(4), 777-807.
- Vaughan, E., & Clancy, B. (2013). Small corpora and pragmatics. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2013* (Vol. 1). Dordrecht: Springer.
- Vaughan, E., & Clancy, B. (2016). Sociolinguistic Information and Irish English Corpora. In R. Hickey (Ed.), *Sociolinguistics in Ireland* (pp. 365-388). London: Palgrave.
- Vaughan, E., & Moriarty, M. (2018). Voicing the 'knacker': Analysing the Comedy of the Rubberbandits. In D. Villanueva Romero, C. P. Amador-Moreno & M. Sánchez García (Eds.), *Voice and Discourse in the Irish Context* (pp. 13-45). Cham: Palgrave Macmillan.
- Wade, E., & Clark, H. H. (1993). Reproduction and Demonstration in Quotations. *Journal of Memory and Language*, 32(6), 805-819.
- Waksler, R. (2009). Super, uber, so, and totally: Over-the-top intensification to mark subjectivity in colloquial discourse. In N. Baumgarten, I. Du Bois & J. House (Eds.), *Subjectivity in language and discourse* (pp. 17-32). Bingley: Emerald.
- Walshe, S. (2009). *Irish English as Represented in Film*. Frankfurt am Main: Peter Lang.
- . (2011). 'Normal people like us don't use that type of language. Remember this is the real world'. The language of Father Ted: representations of Irish English in a fictional World. *Sociolinguistic Studies*, 5(1), 127-148.
- . (2012). 'Ah, laddie, did ye really think I'd let a foine broth of a boy such as yerself get splattered...?': Representations of Irish English speech in the Marvel Universe. In F. Bramlett (Ed.), *Linguistics and the Study of Comics* (Vol. 264-290). New York: Palgrave Macmillan.
- . (2013). "Irish accents drive me nuts": The representation of Irish speech in DC comics. *Cahiers de L'ILSL*, 38, 93-121.
- . (2016). Irish society as portrayed in Irish films. In R. Hickey (Ed.), *Sociolinguistics in Ireland* (pp. 320-343). Basingstoke: Palgrave Macmillan.
- . (2017). The language of Irish films. *World Englishes*, 36(2), 283-299.
- Warner, C. (2014). Literary pragmatics and stylistics. In M. Burke (Ed.), *The Routledge Handbook of Stylistics* (pp. 362-377). London: Routledge.

- Warren, P. (2016). *Uptalk: The Phenomenon of Rising Intonation* by Paul Warren. Cambridge: Cambridge University Press.
- Winter, J. (2002). Discourse Quotatives in Australian English: Adolescents Performing Voices. *Journal of Linguistics*, 22(1), 5-21.
- Wolfson, N. (1979). The conversational Historical Present alternation. *Language*, 55(1), 168-182.
- Wynne, M. (2005). Archiving, Distribution and Preservation. In M. Wynne (Ed.), *Developing Linguistic corpora: A Guide to Good Practice*. Oxford: Oxbow
- Xiao, R., & Tao, H. (2007). A corpus-based sociolinguistic study of amplifiers in British English. *Sociolinguistic Studies*, 1(2), 241-273.
- Zwicky, A. (2006). So in style at the NYT *Arnold Zwicky's Blog: A blog mostly about language*. Accessed at: <http://itre.cis.upenn.edu/~myl/languageblog/archives/002994.html>
- . (2011). GenX So *Arnold Zwicky's Blog: A blog mostly about language*. Accessed at: <https://arnoldzwicky.org/2011/11/14/genx-so/>
- Zwicky, A., & Pullum, G. K. (1987). Plain morphology and expressive morphology. *Berkeley Linguistic Society Papers*, 13, 330-340.

APPENDIX 1

COLOR-CODING SYSTEM

-  Feature distinctive to IrE
-  Non-standard IrE feature
-  (Non-standard) universal feature

TOKEN DISTRIBUTION BY TYPE / CATEGORY

	Distinctive	Non-standard IrE	Universals
GRAMMAR	2,214	545	627
VOCABULARY	3,036	206	84
PRAGMATICS	2,181	1,378	9,298
Pragmatic Sub-categories			
Quotatives	-	1,101	4,298
Taboo language	941	23	3,631
Discourse-pragmatic markers	677	195	1,230
Slang	505	-	92
Boosters	58	59	47
Totals			
Totals	7,431	2,129	10,009

[Please, continue on to next page to see all annotated grammatical, lexical, and pragmatic items with their tags, number of occurrences, and examples from CoFIrE]

FEATURE/TAGS/TOKENS BY CATEGORY

1. GRAMMAR

GROUPS	ITEMS	TAGS	TOKENS	EXAMPLES
IrE DISTINCTIVE				
	2nd p.pron.: <i>Ye</i> (+vars. <i>Yiz,</i> <i>yous(e)</i>)	YE	853	“She’ll ate <i>ye</i> out of house and home” (CKF)
OVERUSE INDEF. ARTICLE <i>THE</i>	+Uncountable Noun	THEUNC	54	Saturday is busy too of course we have to go into Bray and do <i>the bit of shopping</i> like; and do <i>the bit of hoovering and washing</i> (ENMF)
	+Body part	THEBOD	49	I remember Nicola twirling around and pulling <i>the arm</i> off me (RDWD)
	+Relatives	THEREL	41	He was now dreading going home to get an earful from <i>the wife</i> (DRII)
	+Quantifiers	THEQN	26	The milking would be getting light in October. You might be down to <i>the one</i> milking a day (DRYTAD) ‘Did you see the price of the drink in this place?’ ‘Is it dear?’ ‘Fuckin’ desperate. Still though, it’s only <i>the once</i> . Y’ enjoying yourself? (RDWD)
	+Time span/period	THETIME	17	We always came on the first Saturday, didn’t we? John only got off on <i>the Friday</i> , sure (ENMF) It is not that evening or the following one but the evening after, on <i>the Sunday</i> , that I am taken home (CKF)
	+Geographic location	THEGEO	16	Garrett and I have been trying to open an organic bakery and coffee shop on <i>the Quinnsboro Road</i> (PHKU)
	+Generic reference	THEGEIMP	8	‘Women,’ he says, shaking his head. ‘ <i>The women</i> always have an answer.’ (CKWBF)
	+Languages	THELANG	1	I was as smart as any of the posh lads in school. I was well able for <i>the English</i> and geography and history (DRYS)
	+Occupations	THEOCC	1	She could have gone with any of them smart boys that got the real money out of the boom: <i>the</i>

				<i>architects</i> , solicitors, auctioneers (DRYS)
PAST/PERFECT TENSES	<i>After Perfect</i>	AP	107	There's <i>after bein</i> a terrible bomb (ENDD) I didn't say a word about what <i>was after happening</i> to anyone, excepting of course to Joe (ENMF)
	<i>Be Perfect</i>	BEP	79	<i>Is</i> Anthony <i>gone</i> back to Galway? (NCCSM) Mona has one of her kids going to college. I haven't a clue what he's doing; some college up in Dundalk. He's <i>there</i> two years now. (RDWD)
	<i>Medial Object Perfect</i>	MOP	34	The daylight is burning, and I've yet <i>the spuds to spray</i> (CKF)
	<i>Indefinite Anterior</i>	INDANT	7	I <i>never rang</i> him <i>yet</i> or anything (DRYS) I <i>only know of it the last week</i> , Mr. Hartnett...I haven't slept (KBCB)
VB CONTRACTIONS	<i>Contracted Auxiliary ('have')</i>	CONAUX	192	I sold my taxi plate a couple of months ago so <i>I've no</i> recent first-hand accounts of drunken debauchery (DRII)
	<i>Amn't</i>	AMNT	11	Do you think I haven't noticed? <i>Amn't</i> I tripping over you? (CKWBF)
	<i>Contracted Will+Not</i>	CONWINEG	10	There <i>'ll not</i> be a man in the parish will catch you without a long-handled net and a racing bike. (CKF)
	<i>Embedded questions</i>	EQ	205	I didn't <i>know should</i> I go up near him after he got out on bail (DRYS)
FOCUSING DEVICES	<i>Fronting</i>	FR	151	Just cos <i>I've</i> none of my own doesn't mean I'd see the rain falling in on anyone else's. (CKF) I sold my taxi plate a couple of months ago so <i>I've no</i> recent first-hand accounts of drunken debauchery (DRII)
	<i>Clefting</i>	CLEF	34	<i>It is a godsend, the Irish college.</i> (ENDD) <i>It's on the stage you should be!</i> (CKWBF)

PREP. ON	Dative of disadvantage	ONDD	38	No doubt she's got a lot on her mind, what with having just done the dirt <i>on her</i> boyfriend with a complete stranger (DRII)
	Possession	ONP	72	Everyone seems happy, in fact, except for Garret and Claire. They're both just sitting there with faces <i>on them</i> (PHDA)
SUBORDINATION	Subordinating <i>And</i>	SUBAND	72	I swing around the car [...] giving a quick look back to make sure she gets in okay. Lovely arse <i>on her</i> by the way. (DRTRM)
	Subordinating <i>Till</i> (syn. 'so that')	TILL	10	Now c'mere <i>till</i> ya hear the latest. (KBCB) Come out you bollocks <i>till</i> I kill you. (DRTRM)
HABITUAL TENSES	Habitual <i>Do Be</i>	HDB	42	That's what all them wankers <i>do be feeling</i> when they're going around crying over women (DRYS)
	Habitual <i>Do</i>	HD	13	[...] on Fridays I usually <i>do have</i> some little treat I get on the way home from work (ENMF)
	Habitual <i>Be</i>	HB	8	Mammy said don't <i>be trying</i> to bring her down to your level (DRYS)
	Habitual <i>Go</i> + Future reference	HGO	1	Don't <i>go getting</i> used to it (NCCSM)
INFINITIVES	<i>For To</i>	FT	2	An' we're goin' ad it goodo when Anto comes down <i>fer to go</i> fer a piss, an' he bleedin' flips. (DRTRM)
	Unmarked Infinitives	UNINF	48	They wouldn't go in goal if they weren't <i>allowed score</i> goals (RDPC) Mother said to <i>tell</i> that old biddy <i>make</i> her own tarts (DRYTAD)
EXISTENTIALS	<i>There</i>	EXTHERE	17	<i>There were too many years gone</i> by for me to care (DRYS)
	<i>In It</i>	EXINIT	7	The last thing I needed at this hour of the morning was a talker, especially on the day that was <i>in it</i> . (DRTRM) We was all sent off different places when we were small. There was six of us <i>in it</i> . (DRYS)
	Unbound reflexives	UREFL	22	She'd seen him look at other women during Mass but she knew he'd never lay his hand on anyone, only <i>herself</i> . (CKWBF)

	Non-standard progressive (double gerund)	NSPROG	21	He asked the question the very same way you might ask a lad is he <i>going making</i> a mug of tea. (DRYS) Though not wholly surprised at the outcome of Billy's drunken endeavours, I too was <i>roaring laughing</i> . (DRTRM)
NON-STANDARD IrE ITEMS				
	<i>Me</i> (for 'my')	ME	165	I live with <i>me</i> granny actually (ENDD)
	Lack of <i>Do</i> support	LDOSUP	117	<i>Have</i> you any money, Emma? (DRTRM)
VB REDUCTION	Past participle for Preterite	VRPPPR	67	How can we put up with the things <i>we done</i> , choices we made? (KBCB)
	Preterite for Past Participle	VRPRPP	48	He <i>'d have gave</i> the devil himself a good run for his money on his best days (DRYS)
RELATIVES	Relativization (0 subject)	REL	88	The woman at the social said there was a caravan park <i>in Sligo was</i> taking temporaries all the year round if they were stuck (KBDLI) There's <i>a small port of me actually wants</i> this to happen (PHKU)
	<i>What</i> (+inanimate referents)	WH	5	Fancy was the lads <i>what</i> did the follyin' o' the hoss business, check? (KBCB)
NON-STANDARD VB FORMS	Non-standard Past Participles	NSPP	21	Your da's heart <i>is broke</i> with him (NCY)
	Non-standard Preterites	NSPRET	2	Mumbly Dave said I <i>writ</i> a letter to them newspapers (DRYTAD)
	<i>Will</i> for <i>Shall</i>	NSWILL	18	<i>Will</i> I turn on the light for you? (NCY)
	Singular time reference <i>Never</i>	STRNEV	3	Fellas that <i>never done</i> a day's work in their lives, besides spouting shite about how everyone is wrong except them (DRYS)
	Time adverbial <i>But</i> (for <i>for</i>)	NSBUT	3	His wife was dead <i>but</i> years (DRYTAD)
	Comparative <i>Nor</i>	COMPNOR	1	Josephine is minded better <i>nor</i> any woman in Ireland (CKWBF)
NON-STANDARD UNIVERSALS				

NON-STANDARD VB CONCORD	Plural Subject + -s VB	VCPS	139	You'd think <i>there was no doctors</i> or something (ENMF)
	Singular Subject + inflected VB	VCSS	55	Stow the love-eyes, she said. <i>I gots</i> enough on me fuckin' noodle, check? (KBCB) 'I see,' <i>says I</i> . And I did. (ENMF)
	Demonstrative <i>Them</i>	DTHEM	181	<i>Them</i> that says they know her are liars and thieves. (DRYS)
	Unmarked adverbs	UADV	119	God rest him. Didn't he go <i>quick</i> in the end? (CKF) Den he got <i>real</i> possessive (DRTRM)
	Lack of Preposition	LPR	45	These windows [...] have been designed expressly to <i>prevent</i> children <i>looking out</i> of them (ENDD)
	Failure of negative attraction	FNA	36	I <i>ain't</i> got <i>no</i> brud <i>no more</i> (KBDLI) I <i>never</i> went near <i>nobody</i> (DRYS)
	Unmarked Plurals	UPLU	4	<i>Twenty-five year</i> pass and leaves nothing at all hardly in your hand (KBCB)

2. VOCABULARY

GROUPS	ITEMS	TAGS	TOKENS	MEANING
IrE DISTINCTIVE				
TERMS OF ADDRESS/ENDE ARMENTS	<i>Ma</i>	MA	1083	<i>Mother</i>
	<i>Da</i>	DA	552	<i>Father</i>
	<i>Lad(s)</i>	LAD	402	<i>Boy; Guy</i>
	<i>Dote</i>	DO	9	Endearment for a child
	Lexicalized <i>Auld</i>	AU	225	<i>Old</i>
COLLOQUIAL, QUASI- LEXICALIZED REFERENTS YOUR+	<i>Your Man</i>	YM	101	The/a man
	<i>Your Wan</i> (see Prag.Slang <i>Wan</i>)	YW	18	The/a woman
	<i>Your Woman</i>	YWOM	5	
	<i>Your Ma</i>	YMA	1	The/a mother
	<i>Grand</i>	GRA	110	<i>Great</i>
	<i>Bog</i>	BOG	67	

	<i>Quare</i>	QU	67	<i>Strange; very</i>
	<i>Yoke</i>	YO	50	Wet, muddy ground area
	<i>Aye</i>	AY	45	<i>Yes</i>
	<i>Wee</i>	WE	44	<i>Small; little</i>
	<i>Craic</i>	CRA	39	<i>Fun, news, entertainment</i>
	<i>Eejit</i>	EEJ	39	<i>Idiot</i>
	<i>Gom</i>	GOM	27	
	<i>Spud(s)</i>	SPU	27	<i>Potato(es)</i>
	<i>Gob</i>	GOB	19	<i>Mouth</i>
	<i>Yon(der)</i>	YON	10	demonstrative equivalent to obsolete distal deictics (Hickey 2007, 109)
	<i>Boreen</i>	BO	7	Country lane
	<i>Banshee</i>	BAN	3	“Female spirit whose wailing presages death in a family” (Dolan 2006, 15)
	Word combinations (e.g. <i>howya</i>)	COMB	48	I <i>dunno</i> (CKWBF) <i>Howiya</i> . Baggot Stree’ please (DRTRM)
	Nouns ending in - <i>o</i>	NEO	19	Jim, the cute <i>bucko</i> (DRII) Well, <i>Liamo</i> , will we go fishing? (NCY)
	Diminutive suffix - <i>Een</i>	EEN	14	You got the gift for talk, <i>girleen</i> . (KBCB)
	Clause-final <i>At All</i>	ATA	5	What happened <i>at all</i> ? (CBF) What happened to you <i>at all</i> ? (RDPC)
NON-STANDARD IrE ITEMS				
REDUPLICATION-REDUNDANCY	<i>Inside in</i>	REDININ	113	I destroyed the boy by seeing too early the man <i>inside in</i> him. (DRYS)
	<i>Big Huge</i>	REDBH	9	One fella [...] had a <i>big huge</i> house (DRYTAD)
	<i>Like As If</i>	REDLIAI	8	It’s <i>like as if</i> your eyes don’t see things quick enough (NCY)
	<i>Huge Big</i>	REDHB	3	She has to work so hard to pay her <i>huge big</i> mortgage (DRYS)
	<i>Certain Sure</i>	REDCERS	3	He won’t ever go to Mass out there, you can be <i>certain sure</i> (DRYS)
	<i>At all, at all</i>	REDATA	2	He’s some buck alright now, wouldn’t be behind the door <i>at all at all</i> when it comes te getting his end away. (DRTRM)
	<i>Filthy Dirty</i>	REDFD	1	But they’re <i>filthy rich</i> , love (RDPC)

	<i>Only</i> for 'But, except for'	ONLY	30	Margaret lay supine, wearing nothing <i>only</i> a man's trousers (CKWBF)
	Adverbial <i>Once</i> (combination of logical and causal conditionality, syn. 'as long as')	ONCE	1	I asked him where would I stay [...] I don't know in the fuck, he said, and I don't care, <i>once</i> you're here in the morning at seven. (DRYS)
	Non-standard vocabulary (e.g. <i>give</i> for 'spend'; <i>leave</i> for 'let'; <i>same as</i> for 'as if')	NSV	33	The news hadn't bother Johnsey as much as it had Mumbly Dave, who <i>gave</i> the rest of that day giving out stink about the great injustice that had been perpetrated (DRYTAD) How is it at all we <i>left</i> them run the country to rack and ruin? (DRYS) He'd pass her with a bare nod <i>same as</i> she was a shadow of what she had been (CKWBF)
	Non-standard plural <i>Childer</i> (children)	NSPD	3	Sure haven't you the finest <i>childer</i> that ever walked through the chapel gates? (CKWBF)
IRISH LOANS				
TERMS OF ADDRESS/ENDEARMENTS	<i>Craytur</i>	CRAY	14	Ir. <i>Creátur</i> : creature; someone you feel sympathy for
	<i>A leanbh</i>	AL	2	(My) <i>child/baby</i>
	<i>Peata</i>	PE	1	'Pet'; favorite child
	<i>Shebeen</i>	SHE	18	Ir. <i>Sibín</i> : illicit bar where alcohol is sold without a license
	<i>Lough</i>	LO	17	<i>Lake</i>
	<i>Mar dhea</i>	MD	12	Interj. 'as if'; 'yeah, right'
	<i>Plámásing</i>	PLA	3	<i>Plámás</i> (n. & v.) 'empty flattery, sweet-talk'
	(A) <i>Grá</i>	AG	2	<i>Love, darling, sweetheart</i> ; 'to have no love, interest, faith, respect for something'
	<i>Cigire</i>	CIG	2	<i>Inspector</i>
	<i>Fáilte</i>	FA	2	<i>Welcome!</i>
	<i>Piseog</i>	PI	2	a spell, superstitious practice, witchcraft
	<i>Pucán</i>	PU	2	"(male) buck goat" (Dolan 2006); also linked to <i>pooka</i> :

				“a often malignant goblin that generally appears in the form of a horse, but mischievous and sometimes as a bull, a buck-goar” (Joyce 1910)
	<i>Gombeen</i>	GOMB	2	Ir. <i>Gaimbín</i> : a usurer, a loan shark
	<i>Cíteog</i>	CIT	1	<i>Left-handed</i>
	<i>Gaelgór</i>	GAL	1	A fluent speaker of Irish
	<i>Garsún</i>	GAR	1	<i>Boy</i>
	<i>Poteen</i>	PO	1	Ir. <i>Poitín</i> : privately distilled Irish whiskey
	<i>Sliotar</i>	SLI	1	A hurling ball

3. PRAGMATICS

GROUPS	ITEMS	TAGS	TOKENS	Examples
SLANG				
	<i>Gaff</i>	GAF	86	<u>House</u> I can't go out wit me bird fer de nite an den arrive back in de <i>gaf</i> on me tod now, can I? Wud ye cop on? (DRTRM)
	<i>To cop on</i>	CO	51	<u>Come to one's sense; to get a grip on; to notice/realize/understand</u> He must have <i>copped on</i> then that Johnsey was like a dog with him for making a laugh of his text to Siobhan (DRYTAD) Mr Munier laughs, then he <i>cops</i> the expression on my face and realizes that JP isn't joking. (PHKU)
	<i>Skin</i>	SKI	6	<u>Fellow; lad; friend</u> Realtin whon't leave me inside the door, and her father, who's a quare sound auld <i>skin</i> , says I'm as well off leave her be for a while. (DRYS)
	(To be) <i>Up the duff/the Damien/the spout</i>	DU	5	<u>Pregnant</u> I think of the ma and the da [...] and how they were so let down when I got Realtin <i>up the duff</i> and they not even having met her (DRYS)
	<i>Rapid</i>	RA	3	<u>Amazing</u> The biggest mistake I made when I was younger was getting tattoos all over my face. [...] I done it for a woman. She told me I'd look rapid with a <i>spider</i> on my cheek. (DRYS)
	<i>Gatch</i>	GAA	2	<u>Gait</u> Watched him go:

				A big unit, with the splay-footed <i>gaatch</i> of an old slugger, and he turning down a Trace wynd; the carry, the burliness, the country shoulders rolling. (KBCB)
	<i>Mot</i>	MOT	1	<u>Girlfriend</u> Only just married and his <i>mot</i> is already pregnant. (RDWD)
	<i>Segotia</i>	SE	1	<u>Close friend</u> The bus conductor is <i>gas</i> ; he calls you his ould <i>segotia</i> , even though you've never met him before. (NCY)
	<i>Gas</i>	GAS	30	<u>Fun; funny</u> (Same as above)
	<i>Gick</i>	GI	13	<u>Excrement</u> Then they're covered in black <i>gick</i> and the gravel sticks to it and it all becomes a big mess that's impossible to get off (NCY)
<i>Slang: Derogatory Terms of Address</i>				
	<i>Bird</i>	BI	114	<u>Girl, Girlfriend</u> 'When my brother started courting Kate here, we all said he'd never pull as fine a <i>bird</i> .'
HOOR	<i>Hoor</i>	HOP	41	<u>Prostitute</u> Nervous <i>hoors</i> were adrift in the rustle of nylons and the fixing of garter-belts and lost in the misty valleys of their own cheap scent. (KBCB)
	<i>Hoor</i>	HOM	2	<u>Man</u> She saw a fella that was the spit of him thumbing a lift out on the Esker Line. He had the very same cap on him the <i>hoor</i> was wearing when he sauntered out of jail. (KBCB)
	<i>Townie</i>	TO	25	<u>Person from a town who has moved to a rural area and maintains urban ways</u> Mountain bleakness, the lapidary rhythms of the water, the vast schizophrenic skies [...] The truth was confronting me—I was a born <i>townie</i> , and I had made a dreadful mistake in coming here.
	<i>Knacker</i>	KNA	14	<u>A person of low social class; traveler (also <i>skanger</i>, <i>scobe</i>)</u> Smithfield? Where the hell is that? On the Northside, yeah? Are you mad? Sure, there's nothing but <i>knackers</i> up there. They'll have you robbed before you move in! (DRII)

	<i>Culchie</i>	CUL	9	<u>A person from rural Ireland</u> (Dolan 2006, 70) I wonder why he went and killed his father, anyway. A lot of those <i>culchies</i> are mad, though. They're so repressed, like. (DRYS)
	(A) <i>Spa</i>	SPA	8	<u>A fool</u> She never really looked at him or talked to him, except once she offered him a Rolo and he said okay (why did you say okay, you <i>spa</i> ?) (DRYTAD)
	<i>Slapper(s)</i>	SLA	6	<u>A women of loose morals; syn. <i>bitch</i></u> That makes you the sad sap who's left holding her while her <i>slapper</i> of a mother goes back to school and pretends she's never had a baby (PHKU)
	<i>Latchiko</i>	LAT	4	<u>'Unpleasant, disagreeable person'</u> (Dolan 2006, 137) 'He'll have the McGroartys, sure enough. McGroartys are born <i>latchiko</i> . McGroartys would hop into a Feud on account of two flies fucking. (KBCB)
	<i>Brasser</i>	BRA	3	<u>Prostitute; syn. <i>Bitch</i></u> 'I'll tell my version and then you can tell your pack of lies. Anyway, Denise, this <i>brasser</i> here was waiting for me when I got home.' (RDWD)
<i>Slang: Sex-related Terms</i>				
	<i>Mickey</i>	MIC	46	<u>Penis</u> A girl in Baldoyle had to be brought into hospital in Jervis Street after she fainted [...] after the Weirdy Fella'd jumped up in front of her [...] and had shown her his <i>mickey</i> . (RDPC)
RIDE	<i>To ride</i>	RIDV	58	<u>To have sex</u> ... You know the Aer Lingus staff, they're lovely, even if you'd only <i>ride</i> two out of every five of them. (PHDA)
	<i>A ride</i>	RID	34	<u>Attractive person; sex</u> I fancied him as well. I thought he was an absolute <i>ride</i> . (RDWD) I won't think about Lorna again after I start tapping some fine blondie wan below in Australia [...]. It's only the want of a <i>ride</i> is making me all emotional at the moment. (DRYS)
	<i>To shift</i>	SHIF	6	<u>To have sex</u>

				A girl he had <i>shifted</i> once, Audrey, sat at the end of the first pew, sobbing like a madwoman (NCCSM)
	<i>(To) Do a line</i>	DAL	10	<u>To cheat on sb.</u> The whole place has it that Bobby is <i>doing a line</i> with a little strap of a wan from town that bought one of Pokey Burke's houses. (DRYS)
	<i>(To) Do the dirt</i>	DOT	9	<u>To cheat on sb.</u> 'If I remember correctly, Ross, the reason they broke up was because she <i>did the dirt</i> on him with you' (PHKU)
	<i>Gee</i>	GE	6	<u>Dublin slang for female genitalia</u> Someone had drawn tits and a <i>gee</i> on South America (RDWD)
	<i>Fleadh</i>	FLE	4	<u>Cork slang 'flah' = sexual intercourse, synonymous to fuck</u> 'They say the missus' eyes straighten in her head when she gets <i>fleadhed</i> , Mr Mannion?' (KBCB)
TABOO WORDS				
FUCK	<i>Fucking</i>	FING	1457	
	<i>Fuck</i> (n.)	FN	476	Oh <i>for fuck's sake</i> Marian, don't be such a baby (DRTRM) 'Why don't ye hate them?' 'Because I'm not an ignorant <i>fuck</i> , like you.' (DRII)
	<i>Fucker</i>	FE	130	The gulls of Bohane are one ignorant pack of <i>fuckers</i> .
	<i>Fuck off</i>	FO	100	I gets pure wicked with her and tells her to <i>fuck off</i> (DRYS) Her and Lauren have really become bosom buds since Erika <i>fucked off</i> . (PHDA)
	<i>Fucked</i> (adj.)	FD	37	Your life is <i>fucked</i> , Rosser! (PHKU)
	<i>Fuck</i> (interj.)	FINT	137	The next morning I went to her room. <i>Fuck it</i> , I was going to be strong. (KBDLI)
	<i>Fuck</i> [(v.) to throw; to treat someone unfairly/in a humiliating manner]	FUV	33	I stort gathering up the video tapes and <i>fucking</i> them inside for the bin men to take away (PHKU) Lillis, listen to me. Don't <i>fuck</i> this up. Do you hear me? You have everything you've ever wanted now. (NCCSM)
	<i>Fuck</i> [(v.) to have sex]	FUS	25	No <i>fucking</i> , just loving, OK, Struan? (NCCSM)
S H I T	<i>Shit</i> (+variants)	SHI	626	Every week the papers are full of <i>shit</i> (DRII)

				Then Bobby went pure solid <i>apeshit</i> . That whole thing about him doing the dirt on Triona with Seanie's wan was all <i>bullshit</i> . (DRYS)
	<i>Shite</i> (+variants)	SHITE	131	I frightened the <i>shite</i> out of them (RDWD)
BLOOD	<i>Bloody</i>	BL	139	'It's the mattress,' she says. 'the <i>bloody</i> thing is weeping.' (CKF)
	<i>Bleeding</i>	BLE	137	Are you <i>bleaten</i> (bleeding) deaf or something? (PHKU)
FECK	<i>Fecking</i>	FENG	19	I can't wait till this swelling goes down and I can open my <i>feckin</i> eyes (DRYTAD)
	<i>Feck off</i>	FEO	15	'You told me to <i>feck off</i> back to Tipperary, except your didn't use the word feck' (PHDA) 'Look at you, sitting there like a clump of muck,' she says ' <i>Feck off</i> .' ' <i>Feck off</i> yourself,' she says. (NCY)
	<i>Feck</i> (interj.)	FEC	8	'Oh <i>feck!</i> ' she yelps, holding onto her arm. (NCY)
	<i>Feck</i> (n.)	FEN	7	What the <i>feck</i> do they know about anything? (DRYS) This free childcare year is going to be the making of us. And better again, I got a Montessori teacher for <i>feck-all</i> (DRYS)
	<i>Fecker</i>	FER	6	Wee <i>fecker</i> had a tongue for it, so he had (KBCB)
	<i>Fecked</i> (adj.)	FED	1	She says she's <i>fecked</i> if she's having you going around with scabby ears like half the young ones in Dublin (NCY)
		<i>Arse</i>	ARS	221
SHAG	<i>To shag</i>	SHA	14	(Synonymous with <i>fuck</i> : v. to have sex) Hillary said that it was obvious we were going to <i>shag</i> when he offered to share a taxi home with me. (DRYS)
	<i>Shagging</i>	SHING	5	How the fuck should I know? I was only here the once before like, and I

				was looking for a ride, not to send her a <i>shaggin'</i> postcard! (DRTRM)
	<i>Shag off</i> (interj.)	SHOF	4	Men like Paddy said their piece and <i>shagged off</i> and wouldn't countenance backchat (DRYTAD)
<i>Taboo: (Euphemistic) Expletives</i>				
	<i>Jesus</i> (+variants)	JE	351	" <i>Jaysus</i> , it's yer own bleedin' fault he dun it!" (DRTRM)
	<i>Bejaysus</i>	BEJ	8	Mumbly Dave woke up that morning and said <i>Bejaysus</i> , you're uglier than I thought you'd be (DRYTAD)
	<i>Begod</i>	BEG	7	Well, that's the solid Jaysus finest! Oh <i>begod</i> , don't worry at all! Sure I was only being foolish thinking Master Cunliffe would appreciate my holding his post open while he recovered from his injuries! (DRYTAD)
	<i>Gick</i>	GI	3	<u>Shit!</u> -IT'S ALL DOWN YOUR LEG <i>-GICK GICK.</i> -LA LA (RDPC)
<i>Taboo: Abuse terms</i>				
	<i>Bitch</i>	BIT	133	Unpleasant woman
	<i>Wan</i>	WA	121	Disrespectful term for 'woman'
	<i>Bastard</i>	BAS	120	Derogatory for a despicable or unpleasant person
	<i>Bollocks</i>	BOL	81	Despicable or notorious individual
	<i>Cunt</i>	CU	38	Vulgar term for a person/fellow (not necessarily a woman)
	<i>Strap</i>	STR	5	Derogatory term for a shameless/promiscuous woman
DISCOURSE PRAGMATIC MARKERS				
LIKE	<i>Initial Like</i>	IL	29	'Do you think it's possible to ever get over losing someone? <i>Like</i> , really get over it?' (NCCSM)
	<i>Medial Like</i>	ML	1077	'Oh my God, thank you so much for stopping. We are, <i>like</i> , so dead' (DRTRM)
	<i>Final Like</i>	FL	143	Okay, I better go, I'm nearly home...And make sure you say that to Lorna, won't you? Just casually, <i>like</i> . (DRII)
NOW	<i>Initial Now</i>	IN	155	'Ah, there you are, good girl. <i>Now</i> , there's a tin of Mr. Sheen and a duster in that press. You get cracking on the sitting-room.' (NCY)
	<i>Final Now</i>	FNO	29	'Does Pauline want to share with Sandra, though?' [...] 'Share with someone from Dublin? I don't know.' [...] 'Sandra mightn't want to move anyway, <i>now</i> . I think she likes Carrs and all that crowd there.' (ENDD)

YOU KNOW	<i>Initial You Know</i>	IYK	28	' <i>You know</i> , I think it's nearly time that we were making tracks,' Kinsella says, 'It's a long road home.' (CKF)
	<i>Medial You Know</i>	MYK	40	'I'm off duty, <i>you know</i> , and anyway what you need is a doctor.' (ENMF)
	<i>Final You Know</i>	FYK	85	'I need a place out here, Benni. Gather me thoughts, <i>you know?</i> ' (KBCB)
SO	Tag SO (agreement/ 'then')	SOTA	86	'[...]When can I see him?' Deegan hesitates. 'Let me think-' 'Would now suit you?' 'Now? Aye, I suppose it would.' 'Right. I'll follow you, <i>so</i> .' (CKWBF)
	<i>SO +per. pro+ vb</i> (emphasis)	SOEM	54	I'll use the awful word Daddy said to Mammy, <i>so I will</i> . (DRYS)
	Standalone SO	SAS	10	-If they blew it up they'd blow up the new one as well, said Liam. -No, they wouldn't; that's stupid. -They would <i>so</i> . (RDPC)
SURE	Initial Sure	ISU	148	' <i>Sure</i> haven't you the finest childer that ever walked through the chapel gates?' (CKWBF)
	Final Sure	FSU	1	-There's more caravans than people <i>sure</i> , Paula. (RDWD)
	Right	RIG	125	'Jesus, Billy, it's nathin to me either way. We'll give ye one minute to get inside and then we're off outa here te fuck, <i>right?</i> ' (DRTRM)
	Stressed Some	SS	51	'Ah, leds! That was <i>some</i> show you put on the lest dee!' (PHDA)
	Clause final But	FB	15	'I'm nto sure congratulations is the actual word,' I go. 'He's fourteen years of age.' She's there, 'He's nearly fifteen, <i>but</i> . And I wadn't much older when me and Kennet had hour foorst.' (PHDA)
	Yerra	YER	15	Mumbly Dave said Siobhan was a sneaky little bitch. Then he got kind of sorry and said <i>Yerra</i> lookit, women don't be in their right mind half the time, with their periods and what have you. (DRYTAD)
	Arra(h)	AR	7	'There's a Calm has held for a good stretch in Bohane,' said Ol' Boy. 'Be a hoor if it went the road, like.' 'I ain't the one been wieldin' a shkelp.' ' <i>Arra</i> , you know it's Hartnett has the Smoketown trade.' (KBCB)
	Tag What	WHA	4	'Well,' he goes, 'what a fortuitous s...s...s...set of circumstances, <i>wha?</i> ' (PHDA)

BOOSTERS				
	GenX So	GSO	47	'So eh, have you all your shopping done?' 'God, no, not a bit. I'm <i>so last minute</i> , it's crazy' (DRTRM)
	Right	RI	41	I seen Bobby, but he looked like he was <i>right</i> busy. (DRYS)
	Pure	P	24	Friesians are <i>pure</i> gentle auld crathurs. If they were Limousins, Daddy would say, they'd trample you to get to the field. (DRYTAD)
	Pure Solid	PS	18	My children never call to me, even. They're <i>pure solid</i> ashamed of me, after all I done for them. (DRYS)
	Solid	ST	16	Oh, he's <i>solid</i> gorgeous, so he is. (DRYS)
	Fierce	FI	11	Did you not hear the news? Well, I'm <i>fierce</i> sorry now to be the bearer of sorrow, but it looks like he was killed last night. (DRYTAD)
	Stone	ST	7	My father went <i>stone</i> mad on the drink when my mother died, so he did. (DRYS)
NON-STANDARD QUOTATIVES				
	Go	QGO	3174	You don't enjoy yourself. Your da doesn't listen to anything you say. You tell him that the postman lost your letter to Gwen and he <i>goes</i> , 'Oh, really? That's nice.' (NCY)
	Be Like	QBEL	1124	'So I said yeah that's totally cool, and he <i>was like</i> , yeah cool well I'll be in Renard's after the gig, so maybe I'll see you there, and then I snogged him for a bit, so we're definitely going in' (DRTRM)
	Be There	QBT	1089	<i>She's there</i> , 'Ross, you are so a snob.' (PHDA)
	Give	QGI	7	He bursts into the room, <i>giving it</i> , 'What's the stordee?' (PHKU)
	Be All	QBO	2	<i>She's all</i> , 'How dare you moralize with me!' (PHDA)
	Be There + -ing	QBTVI	2	<i>I'm there going</i> , 'I'm not going to hit him again. He's already decked.' (PHKU)
	Be Here	QBH	1	'Your daughter was expelled from her last school,' he goes. 'Is that correct?' 'Well, [...] I'd prefer to say that, educationally, it just wasn't a good fit.' <i>He's here</i> , 'Yes, expelled for bullying.' (PHDA)

APPENDIX 2

EXAMPLE OF EMAIL TEMPLATE SENT TO AUTHORS

Dear Mr. Barry,

My name is Ana María Terrazas-Calero and I am a PhD student at Mary Immaculate College (Ireland). My thesis deals with the representation of fictional Irish-English in a set of contemporary Irish novels, for which I have built a corpus of books by 8 different Irish writers. I am attaching here a publication I wrote with one of my supervisors, Dr. Carolina P. Amador-Moreno, on the language used in a different series of novels, so that you get a better idea of what my thesis is about.

That said, I am writing to inform you that I have added *City of Bohane* and *Dark Lies the Island* to my *Corpus of Contemporary Fictionalized Irish English* novels because the language used in them is everything I was hoping to find for my thesis.

I would also like to ask if you would be so kind as to give me access to pdf or .txt copies so that I would be better able to digitally analyze the linguistic features of the novels for the type of statistical analysis that is done in Corpus Stylistics. Needless to say, I would never make the originals available to anyone. I would only use them for my thesis, and I would acknowledge you and your publisher in any presentation I do at conferences and in any publications in which I quote your books. On that note, I would also like to ask if you can grant me permission to quote your books in publications?

I understand that you might need to talk to your publisher before you can grant permission to give me copies, if you decided to do so, and permission to quote the books. If you prefer me to contact them directly, please let me know.

If there is anything you would like me to clarify, I will be more than happy to do so.

I look forward to attending one of your public readings at some stage.

Thank you in advance for your help.

Looking forward to hearing from you.

Kind regards,

Ana María.

APPENDIX 3

CoROCK (books/case study)					
Publication	Title	Book Code	Novel/Play	So	Totally
2000	<i>The Miseducation Years</i>	TMY	N		x
2001	<i>The Teenage Dirtbag Years</i>	TDY	N		x
2003	<i>The Orange Mocha-Chip Frappuccino Years</i>	OMFY	N		x
2005	<i>P.S. I Scored the Bridesmaids</i>	PSSB	N		x
2005	<i>The Curious Incident of the Dog in the Nightdress</i>	CIDN	N	x	x
2007	<i>Should Have Got off at Sydney Parade</i>	SHGSP	N	x	
2008	<i>This Champagne Mojito is the Last Thing I Own</i>	CHAMP	N		x
2009	<i>Rhino What You Did Last Summer</i>	RHINO	N	x	x
2011	<i>NAMA Mia!</i>	NAMA	N	x	x
2013	<i>Downturn Abbey</i>	PHDA	N(CoFIrE)	x	x
2014	<i>Keeping up with the Kalashnikovs</i>	PHKU	N(CoFIrE)	x	
2015	<i>Seedless in Seattle</i>	SS	N		
2016	<i>A Game of Throw-ins</i>	GoT	N	x	x
2017	<i>Operation Trumpsformation</i>	OT	N		
2018	<i>Dancing with the Tsars</i>	DwT	N	x	x

APPENDIX 4

Contrastive Intensifying *Fucking* Classification¹⁰⁷

MACKENZIE'S CLASSIFICATION		COFIRE	
FEATURES	EXAMPLES	EXAMPLES	
IN/WITHIN NOUN PHRASE			
It is non-gradable	*A more <i>fucking</i> mess that I've seen...		
Preceding head N in the place of attributive/descriptive adjectives, classifiers or numerals in complex NPs	A scary <i>fucking</i> film <i>Fucking</i> tooth decay The <i>fucking</i> seven hills of Rome	NCY: It's none of your <i>fucking</i> business	
		DRYTAD: Aboy Johnsey, aboy Johnsey, you <i>fuckin</i> legend	
		KBCB: Yes a shallow <i>fucking</i> town.	
It can precede descriptive adjectives	A <i>fucking</i> scary film	PHDA: [...] just staring at Sorcha, with their big <i>focking</i> dumb mouths open	
Predicted NOT to precede added numerals or determiners	* <i>Fucking</i> three foxes (cf. (three <i>fucking</i> foxes)	PHKU: 'Theer's anutter one!' someone else shouts. ' <i>fooken</i> two mower'	
	* <i>Fucking</i> the accident (cf. the <i>fucking</i> accident)	LCIE: She looks <i>fucking</i> thirty	
IN/WITHIN VERBAL PHRASE			
Precedes lexical verbs	<i>fucking</i> (well) leave		ENDD: I'll <i>fucking</i> sue you.
		After negation	DRTRM: I don't <i>fuckin</i> ' believe dis!
Only AFTER all auxiliaries (including <i>to</i>)	*He <i>fucking</i> has disappeared (cf. he has <i>fucking</i> disappeared)	KBCB: An' she got my boy Wolfie in a love muddle 'n' all, and that ain't like Wolfie, no sir it jus' <i>fuckin</i> ' ain't, like.	
		PHKU: 'Ross, will you look after the shop for me?' And I'm like , 'Of course I <i>focking</i> will! You just go and do what you have to do!'	
CANNOT precede copulas unless there is contrastive stress on the copula	*He <i>fucking</i> is lazy vs. He <i>fucking</i> IS lazy		
IN/WITHIN ADJECTIVAL PHRASE			
Must precede head adjective (attributive or descriptive, basic,	a <i>fucking</i> scary film	DRYS: I'd look <i>fuckin</i> pretty on a building site, wouldn't I?	

¹⁰⁷ Mackenzie's (2019) summarized classification with examples, reproduced with his permission, and examples from CoFIrE.

comparative or superlative)				
Collocates with gradable and ungradable adjectives	<i>Fucking</i> dead	KBCB: The dreams is gone halfway <i>fuckin'</i> alive on me		
Follows any degree adverb	Really <i>fuckin'</i> amazing	RDWD: He was a right <i>fuckin'</i> creep in it anyway.		
IN/WITHIN ADVERBIAL PHRASE				
Precedes head adverb	He drove <i>fuckin'</i> fast	PHKU: ... I don't want to – I don't know – die <i>fuckin'</i> suddenly without having said it to you		
Only adverbs with representational meaning	I saw him only <i>fuckin'</i> yesterday			
Follows any degree adverb	Really <i>fuckin'</i> well	PHDA: very <i>fuckin'</i> quickly		
IN/WITHIN PREPOSITIONAL PHRASE				
It CANNOT precede prepositional heads	*Do it <i>fuckin'</i> afterwards	DRTRM: Shut <i>fuckin'</i> up, an' ged in will ye		
It modifies the head of the NP within the PrepP	*Put it <i>fuckin'</i> under the bed (cf. Put it under the <i>fuckin'</i> bed)			
TMESIS				
Inserted between compounds	Lovey- <i>fuckin'</i> -dovey	Non-Standard Tmesis	Compounds	PHDA: The elevator takes for- <i>fuckin'</i> -ever.
			Fixed units	DRTRM: Fine, there's a tenner. Merry <i>fuckin'</i> Christmas
			Names	PHKU: I feel like I'm in a Rick <i>fuckin'</i> Ross video
			Numbers	KBCB: I'm forty- <i>fuckin'</i> -three and I'm sat around talkin' <i>fuckin'</i> gang fights?
PHDA: The six-foot- <i>fuckin'</i> -ten dude sitting in front of me is certainly of that view				
Inserted into one word, rendering units without morpheme status	unbe- <i>fuckin'</i> -lievably	Standard Tmesis	DRYTAD: con- <i>fuckin'</i> –sortium	
			KBCB: a- <i>fuckin'</i> –broad	
			PHDA: sur- <i>fuckin'</i> -prise	
			PHKU: un- <i>fuckin'</i> -believable; lemon- <i>fuckin'</i> -ade	
WITH PRONOUNS				
<i>It CAN precede</i>				

Indefinite pronouns (and adverbs), especially containing <i>any</i>	They'll take <i>fucking</i> anyone for that job	
Negative pronouns	He said <i>fucking</i> nothing all night	DRTRM: I dun nuttin, I dun <i>fffuckin'</i> nuttin, I'm telhn' ye, dunno whad 'er bleedin' problim is, stewpid cow
<i>It CANNOT precede</i>		
Interrogative pronouns	* <i>Fucking</i> what are you talking about?	PHKU: There was something I didn't tell you – the day you called here with your old pair and <i>fucking what's</i> his face
		PHDA: 'Did you talk to um?' meaning Kennet. I'm like, 'Not yet, no' She goes, ' <i>Fooken when</i> then?'
Relative pronouns	*The guy <i>fucking</i> who said that was crazy	
Demonstrative pronouns	*I want <i>fucking</i> this	PHKU: <i>Focking</i> this again, I'm thinking
Reciprocal pronouns EXCEPTION: it appears possible with tnesis	*Can't they be nice to <i>fucking</i> each other? Can't they be nice to each <i>fucking</i> other?	
Reflexive pronouns EXCEPTION: it appears possible with tnesis	* Get <i>fucking</i> yourself out of here! Get your <i>fucking</i> self out of here!"	
Personal pronouns (whether subject or object) EXCEPTION: the pronoun has contrastive stress	*They are looking for <i>fucking</i> me/you/him/her... I want <i>fucking</i> HIM, not <i>fucking</i> YOU	PHDA: We sit down, then in walks the judge and we have to all, like, stand up again, just for <i>fucking</i> him PHKU: We morch blindly into that dork, dork jungle with me – <i>focking</i> me!
Possessive pronouns EXCEPTION: the pronoun has contrastive stress	*That suitcase is <i>fucking</i> his That's <i>fucking</i> MINE, not <i>fucking</i> YOURS	PHKU: I'm thinking to myself, Yeah, remember whose gaff this is now. Here's a hint. It's not <i>focking</i> yours.

APPENDIX 5

Intensifying *Fucking* negative emotion conveyance in CoFIrE

Negative emotive InF catalogue in CoFIrE		
Emotive meaning	Example	Source Text
Anger	I want my fuckin pension you little prick, Mickey roared and roared	DRYS
Insults	'I s'pose you know that possessed fuckin ' she-devil above in the house will put me in the ground?'	KBDLI
Dislike	'Happy? Who's happy in fuckin ' Bohane? Ya'd be a long time scoutin' for happy in this place.'	KBCB
Sarcasm	He was full of old sugary shite, that Mumbly Dave. Sure, he was a gas character. Ha ha fucking ha. How could he have a new joke or bit of smartness ready every single time the Lovely Voice came near them?	DRYTAD
Exasperation	'Would you put some fucking clothes on when you're walking around this house?'	PHDA
Disdain	Most of what you got me is lame. I've arranged everything into two piles – the things I'm going to reluctantly keep and the things you better have fucking receipts for.'	PHKU
Disbelief	She pushed back the couch and she found the laptop underneath and she brought it to the cable by the retro phone desk - a fucking cable!	KBDLI
Annoyance	'You been soundin' kinda bitter this weather, G. If you don't mind me sayin', like?' 'It's bred into me, Benni.' 'Ah, stop, will you? The fuckin ' martyrdom!'	KBCB
Mockery	Hey, you should a seen this lad the whole way to Shannon! Crying like a child! Will you give him a lend of your make-up there hey, it might fuckin cheer him up a bit! Fwahahaha!	DRYS
Self-deprecation	I was hopeless, useless, good for fuckin nothing.	RDWD

Disgust	When I hear her go , ‘Oh my God, the <u>fucking</u> smell!’ I know it’s the right place.	PHKY
---------	--	------