

“Ach, it’s a great wee day so far, isn’t it?”

A corpus stylistic study of discourse-pragmatic markers in *Derry Girls*

Alma Peris Alonso



Master’s Thesis in English Linguistics

Department of Foreign Languages

University of Bergen

September 2023

Abstract in English

This thesis aims to analyse the use of the discourse-pragmatic markers (DPMs) *ach*, *och*, final *but* and *wee* in the dialogue of the television series *Derry Girls* (Leddy et al., 2018–2022), the language of which constitutes a fictional representation of Northern Irish English. The thesis analyses those DPMs from two points of view: (1) from a dialectological point of view, comparing their use – i.e., frequency and functions – in the series dialogue to their use in the Northern Irish component of the SPICE-Ireland corpus; and (2) from a characterisation point of view, exploring the idiolects of seven main characters through said DPMs and other keywords in their dialogue.

In order to approach both sides of the thesis' objective, a corpus containing the series dialogue was created and subsequently explored employing both quantitative and qualitative methods. Among the former were keyword searches and frequency count comparisons, as well as log-likelihood statistical significance tests. The qualitative methods employed included concordance searches and the close reading of the results of these and the keyword searches.

The main findings uncovered in the analysis were that *ach* and *och* have a wider range of functions than previously reported, and that *wee* can function as a minimiser and mitigator in Northern Irish English, as suggested in previous studies. However, due to the fictional nature of the dialogue, and to the fact that SPICE-Ireland was compiled almost two decades ago, these results should be verified in future studies by analysing more up-to-date real-life spoken language. In terms of characterisation, the analysis showed that the four titular *Derry Girls* have distinct idiolects when compared to each other and when contrasted with the older main characters.

Keywords: corpus stylistics, *Derry Girls*, Irish English, television dialogue, discourse-pragmatic markers, *ach*, *och*, *wee*

Abstract in Spanish

Este trabajo de final de máster tiene como objetivo analizar el uso de los marcadores pragmáticos del discurso *ach*, *och*, final *but*, y *wee* en el diálogo de la serie de televisión *Derry Girls* (Leddy et al., 2018–2022), cuyo lenguaje es una representación ficticia del inglés de Irlanda del Norte. El estudio analiza estos marcadores desde dos puntos de vista: (1) desde el punto de vista dialectológico, comparando su uso (frecuencia y funciones) en la serie con su uso en el componente norirlandés del corpus SPICE-Ireland; y (2) desde el punto de vista de la caracterización de siete personajes principales a través de dichos marcadores y de otras palabras clave de su diálogo.

A fin de realizar ambas partes del objetivo, se creó un corpus con el diálogo de la serie y se exploró empleando tanto métodos cuantitativos como cualitativos. Entre los primeros se emplearon la búsqueda de palabras clave y la comparación de frecuencias, además de la prueba de verosimilitud logarítmica como medida de significancia estadística. Como métodos cualitativos se emplearon las concordancias y la lectura atenta de éstas.

El análisis demostró, entre otros, que las funciones de *ach* y *och* son más diversas de lo recogido en estudios anteriores. Además, se probó que *wee* realiza en efecto funciones de marcador del discurso en el inglés de Irlanda del Norte. Sin embargo, el carácter ficticio del lenguaje analizado, y el hecho de que SPICE-Ireland se creó hace dos décadas, supone que estos resultados deberán verificarse en futuros estudios que analicen el habla norirlandesa actual. Por lo que respecta a la caracterización de los personajes estudiados, el análisis mostró que cada una de las titulares *Derry Girls* tiene su propio idiolecto, y que como grupo su lenguaje se distingue del de los personajes adultos.

Palabras clave: estilística de corpus, *Derry Girls*, diálogo televisivo, inglés de Irlanda, marcadores pragmáticos del discurso, *ach*, *och*, *wee*

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Carolina Amador-Moreno, for her guidance, support, and encouragement throughout this process. Your expertise and feedback were invaluable in this endeavour. I could not have asked for a more knowledgeable and patient supervisor; without you, this thesis would not be what it is today.

I am also grateful to my fellow students for their useful feedback, both in the more structured seminars and outside of them; thank you for turning what could have been a solitary experience into a shared journey.

Last, but not least, I would like to thank my family, my partner's family, and my partner especially, for their unwavering support throughout this year. Thank you, Matthew, for believing in me whenever I doubted myself and for making sure a cup of tea was never missing on my desk.

TABLE OF CONTENTS

Abstract in English	iii
Abstract in Spanish	iv
Acknowledgements	v
Table of contents	vi
List of tables	ix
List of figures	x
Disclaimer	x
List of abbreviations	xi
1. INTRODUCTION	1
2. THEORY AND PREVIOUS STUDIES	3
2.1 Terminology regarding Irish English	3
2.2 The study of fictional dialogue	3
2.2.1 The study of IrE in fiction	4
2.2.2 Television dialogue	6
2.2.3 Corpus stylistics	8
2.3 Northern Irish English (NIrE)	9
2.4 “Sure that’s not what they meant”: The study of pragmatics in IrE	10
2.4.1 The pragmatics of IrE.....	11
2.4.2 The study of DPMs in IrE	12
2.4.2.1 DPMs: definitions and functions.....	14
2.4.2.2 Existing research on DPMs in IrE	16
2.4.2.2.1 <i>Sure</i>	16
2.4.2.2.2 <i>Like</i>	17
2.4.2.2.3 <i>Now</i>	19
2.4.2.2.4 <i>And</i>	20
2.4.2.2.5 <i>So</i>	21
2.4.2.2.6 Taboo language	22

2.4.2.3	DPMs analysed in this thesis.....	25
2.4.2.3.1	<i>Ach/Och</i>	25
2.4.2.3.2	Final <i>but</i>	26
2.4.2.3.3	<i>Wee</i>	27
2.5	Concluding remarks.....	28
3.	METHODOLOGY AND DATA.....	29
3.1	Data.....	29
3.1.1	The <i>Derry Girls</i> corpus	30
3.1.1.1	Data collection.....	30
3.1.1.2	The annotation system.....	32
3.1.1.3	Post-annotation steps.....	35
3.1.1.4	The <i>Derry Girls</i> Corpus: description.....	37
3.1.1.5	Strengths and limitations of the process.....	38
3.1.2	The corpus used to compare and contrast the DGC data	39
3.1.2.1	Comparability between the DGC and SPICE-Ireland.....	42
3.1.3	Legal and ethical considerations	43
3.2	Methods.....	44
3.2.1	Quantitative methods.....	44
3.2.1.1	Frequency counts.....	45
3.2.1.2	Significance testing	45
3.2.1.3	Keyword analysis	47
3.2.2	Qualitative methods.....	48
3.2.2.1	Concordance searches	48
3.2.2.2	Answering questions 1 and 2: DPM functions in the DGC and SPICE-Ireland..	48
3.2.2.3	Answering question 3: Linguistic characterisation in the DGC.....	49
3.3	This thesis' approach to corpus linguistics	50
3.4	Concluding remarks.....	50
4.	RESULTS AND DISCUSSION.....	52
4.1	DPMs in the DGC corpus	52
4.1.1	<i>Ach/och</i>	52

4.1.1.1	Frequencies of use	52
4.1.1.2	Functions of <i>ach</i> and <i>och</i>	56
4.1.2	Final <i>but</i>	66
4.1.3	<i>Wee</i> as a DPM	68
4.1.3.1	Frequencies of use of <i>wee</i>	69
4.1.3.2	<i>Wee</i> : patterns of use.....	71
4.2	Linguistic characterisation in the DGC	76
4.2.1	Most frequent words.....	76
4.2.2	Keyword analysis	80
4.2.2.1	Character names and other terms of address	82
4.2.2.2	Taboo language	87
4.2.2.3	Dialectalisms	93
4.2.2.4	DPMs.....	95
4.2.2.5	Keyword analysis of character groups: Derry Girls against Adults....	98
4.3	Summary of findings.....	100
5.	CONCLUSION.....	102
5.1	Limitations of the present study	103
5.2	Recommendations for future research.....	103
6.	REFERENCES	105
6.1	Primary sources.....	105
6.2	Secondary literature.....	105
APPENDIX A.....		112
APPENDIX B.....		114

List of tables

Table 2.1. Characteristics of discourse-pragmatic markers.....	15
Table 3.1. List of <i>Derry Girls</i> episodes contained in the DGC corpus	30
Table 3.2. Tags employed in the annotation of the DGC	34
Table 3.3. DGC character subcorpora	37
Table 3.4. DGC demographic partitions.....	38
Table 3.5. Partitions of the SPICE-NI data	41
Table 4.1. Frequencies for <i>ach/och</i> in DGC and SPICE-Ireland.....	53
Table 4.2. Log-likelihood tests for <i>ach/och</i>	54
Table 4.3. Raw frequencies and percentages of use for <i>ach</i> and <i>och</i> in initial and medial positions in the DGC and SPICE-NI.....	56
Table 4.4. Normalised and raw frequencies for final <i>but</i> in SPICE-Ireland	67
Table 4.5. Table 4.5. Log-likelihood tests for final <i>but</i>	67
Table 4.6. Normalised and raw frequencies for <i>wee</i> in the DGC and SPICE-Ireland.....	70
Table 4.7. Log-likelihood tests for <i>wee</i>	70
Table 4.8. DPM <i>wee</i> against Brinton’s (2017) characteristics of DPMs.....	75
Table 4.9. Word frequency lists for selected characters from <i>Derry Girls</i>	77
Table 4.10. Keywords in each of the character subcorpora.....	80
Table 4.11. One-to-one character keyword comparisons.	81
Table 4.12. Keyword comparison of character groups.....	98
Table 4.13. Raw and normalised frequencies of <i>yeah</i> and <i>aye</i>	100
Table A.1. List of tags used in the annotation of the DGC corpus	112
Table A.2. Tags for elements excluded from the analysis	113
Table B.1. Normalised and raw frequencies of <i>fuck</i> and its derived terms with and without Michelle’s dialogue	114
Table B.2. Log likelihood tests for <i>fuck</i>	114
Table B.3. Raw frequencies and normalised frequencies (in parentheses) of <i>bloody</i>	114
Table B.4. Log likelihood tests for <i>bloody</i>	115
Table B.5. Raw and normalised frequencies (in parentheses) of DPM <i>sure</i>	115
Table B.6. Log likelihood tests for <i>sure</i>	115

List of figures

Figure 4.1. Proportion of use of DPM <i>like</i> in three sentence positions in each character's dialogue.	79
Figure 4.2. Character relationship map as defined by their keywords.	83
Figure 4.3. Normalised frequencies of <i>fuck</i> (per 10,000 words) in the DGC with and without Michelle's dialogue.	90
Figure 4.4. Normalised frequencies of <i>fuck</i> per (10,000 words) in DGC and SPICE-NI.	91
Figure 4.5. Normalised frequencies (per 10,000 words) of <i>bloody</i>	92
Figure 4.6. Normalised frequencies (per 10,000 words) of DPM <i>sure</i>	97
Figure 4.7. Normalised frequencies (per 10,000 words) of <i>aye</i> and <i>yeah</i>	99

Disclaimer

Both the dialogue in the show *Derry Girls* and the data contained in the SPICE-Ireland corpus include instances of non-standard uses of English grammar and potentially offensive taboo language. In order to preserve the integrity of the data, these instances have been neither censored nor standardised, and, as a result, they appear in this thesis as part of examples and quotes where relevant for the analysis.

List of abbreviations

AF	–	Adult Female
AM	–	Adult Male
DGC	–	<i>Derry Girls</i> Corpus
DPM	–	Discourse-Pragmatic Marker
IrE	–	Irish English
LDE	–	(London)Derry English
LL	–	Log-likelihood
NI	–	Northern Ireland
NIrE	–	Northern Irish English
OED	–	Oxford English Dictionary
OF	–	Old Female
OM	–	Old Male
ROI	–	Republic of Ireland
SPICE-Ireland	–	Systems of Pragmatic annotation in the spoken component of the International Corpus of English – Ireland
YF	–	Young Female
YM	–	Young Male

1. INTRODUCTION

This thesis examines the use of discourse-pragmatic markers (DPMs) in the television series *Derry Girls* (Leddy et al., 2018–2022) as a method of dialectal representation and character construal. This comedy is comprised of three seasons, with a total of nineteen episodes. The show is set in the eponymous Northern Ireland city¹, against the backdrop of the Troubles in the early nineties, although historical events are not the focus of the plot. Instead, the story follows a catholic teenage girl, Erin Quinn, and her friends, as they navigate the challenges of adolescence.

Derry Girls is notable for featuring the local vernacular at the forefront: (London)Derry English (henceforth LDE), as McCafferty (2001) terms it. Lisa McGee, the creator of the show, based the story on her own experience of growing up in the city and used LDE in the dialogues (*Belfast Telegraph*, 2017). In addition to this, most of the characters in the show are played by local actors (McCreedy, 2019, p. 280).

The show's depiction of LDE, a variety seldom heard outside Northern Ireland, has been noticed by the – mostly British – audience, to the point that Channel 4, the network that broadcasts *Derry Girls*, released a glossary of 'Derryisms' to help viewers understand the dialogues (Little, 2017). Upon the show's success, awareness quickly turned to enthusiasm, with fans using the dialectalisms featured in the series on social media and in memes (McCreedy, 2019, p. 282). In fact, it has been argued that, without LDE, *Derry Girls* would have not reached the level of success it enjoys (McCreedy, 2019).

The saliency of language in *Derry Girls* did not go unnoticed during my first viewing of it, back in 2019, and this thesis is not the first study to approach the series from a linguistic angle. Indeed, with viewership figures on Channel 4 of over three million viewers on average ("*Derry Girls*", n.d.), plus the international availability of the series on the catalogue of the streaming platform Netflix, *Derry Girls* has exposed millions of viewers to LDE, and as such it was bound to attract scholarly attention. McCreedy (2019), for example, analyses the use of obscure expressions and cultural references in the show, central to its

¹ This city has received two names: 'Derry' and 'Londonderry'. Throughout this thesis, I will refer to this city as either 'Derry', the term used in the show, or '(London)Derry', following McCafferty (2001). The latter designation has the advantage of combining both names into one term. No ideological implications are intended by the use of these terms.

perceived authenticity, while Valleriani (2021) discusses the portrayal of Received Pronunciation as a foreign accent in the show, a role-reversal from what is usually seen in British media. For her part, Diaz-Sierra examines the authenticity of the characters' accents as perceived by Northern Irish people (2022). Beyond linguistics, Jarazo Álvarez (2023) has analysed the use and subversion of road movie tropes in the third episode of the show's second season. In addition, Henry et al. (2023) have studied the representation of queerness in the show, highlighting the potential of comedies to subvert and reimagine traditional portrayals of queer characters.

Owing to the geographical and linguistic setting of *Derry Girls*, the present thesis falls within the field of Irish English studies, where it joins an increasing number of studies that focus on the representation of this linguistic variety in fiction. In this thesis, the focus is on the DPMs that appear in the characters' dialogue. The questions that will guide my research are:

1. How frequently do these features appear in the show's dialogue and what functions do they realise?
2. How do the frequency and use of these features compare to those in the Northern Irish component of the SPICE-Ireland corpus?
3. How does the use of DPMs contribute to the characterisation of the characters in the show?

In order to answer these questions, the following chapters will provide an overview of the relevant literature regarding Irish English, its representation in fiction, and its pragmatic profile (Chapter 2), followed by a description of the methods of analysis employed (Chapter 3), the analysis of the dialogues themselves and a discussion of the findings (Chapter 4). The concluding remarks will address the relevance of the results and point out new avenues of research that may emerge from this study (Chapter 5).

2. THEORY AND PREVIOUS STUDIES

In this chapter I provide an overview of the existing research on fictional dialogue in an Irish context, and on television dialogue, as well as an introduction to the field of stylistics (section 2.2). In section 2.3 I offer a summary of the development and some characteristics of Northern Irish English, the language variety employed in the data I will analyse. After this, in section 2.4 I turn to the study of pragmatics, with a focus on Irish English (2.4.1) and the study of discourse-pragmatic markers (2.4.2). This last subsection includes a description of discourse-pragmatic markers associated with Irish English, a review of the existing research on them (2.4.2.2), and an overview of the markers that will be the focus of the present thesis (2.4.2.3).

2.1 Terminology regarding Irish English

In order to refer to the English spoken in Ireland, throughout this thesis I will be using the term *Irish English* (IrE), as it seems to be the preferred term among researchers in recent years (Amador-Moreno, 2010a, p. 8); furthermore, it has the advantage of paralleling the designations for other varieties of English such as Canadian English, Australian English, etc. (ibid.; Hickey 2005, p. 20). In the past, the terms *Hiberno-English* and *Anglo-Irish* have been used in place of *IrE*, but these terms can lead to confusion in a linguistic context.² On a related note regarding terminology, I will be following Amador-Moreno's (2010a, p. 10) use of *Ireland* to refer both to the Republic of Ireland (ROI) and Northern Ireland (NI), without intending to take any political stance by these uses.

2.2 The study of fictional dialogue

Literary and stylistic scholars have long been interested in the study of fictional representations of natural language (Terrazas-Calero, 2022, p. 1). In cases where a writer has attempted to represent a regional or social variety, we can speak of "literary dialect" (Ives, 1971, p. 146). More generally, one can speak of "fictional dialogue" (Amador-Moreno & Terrazas-Calero, 2022, p. 517).

Two approaches to the study of fictional language have been identified in the past: (1) a stylistic approach, which "considers the role and effectiveness of the dialect and non-standard

² See Amador-Moreno (2010, pp. 7-9), Hickey (2005, p. 20), or Hickey (2007, p. 5-11) for a discussion about their respective advantages and drawbacks.

within the literary work as a whole”, and (2) a dialectological approach that examines “the significance provided by the use of the dialect and nonstandard within the literary work as evidence for the dialect” (Kirk, 1997, p. 203). The validity of fictional language as a source of linguistic evidence has been contested, due to its allegedly questionable realism and consistency, and in the case of written works, the fact that it lacks several of the paralinguistic characteristics of spoken language, such as intonation or gestures (Amador-Moreno, 2010b, p. 532; Terrazas-Calero, 2022, p. 1). In addition, the English alphabet is poorly equipped to represent different phonetic realisations of the same word that might be representative of different dialects (Hodson, 2014, p. 12). Filmic and televisual dialect does not have the disadvantages that print media face with respect to the representation of orality in writing (*ibid.*), since their medium is audiovisual. That being said, on-screen representations of English varieties are not automatically more authentic than written ones, as scripts often exclude false starts or misunderstandings, common in real-life language (Bednarek, 2019, p. 14).

Despite these limitations, it has been argued that “as long as it is observed and evaluated in relation to the spoken language, the study of literary dialect can be a valuable complement to the evidence of real speech data” (Labov, 1972, p. 109). Accordingly, many of the works discussed in the following subsections adhered to this recommendation, comparing their literary data to existing corpora of oral language.

2.2.1 The study of IrE in fiction

The representation of IrE in fiction has changed since the days of Elizabethan theatre (Amador-Moreno, 2010a, p. 108). The Stage Irishman, a stock character of the era used for “parody and ridicule” (Hickey, 2007, p. 8), spoke in a stereotyped brogue that focused on phonetic features of mostly rural IrE, as perceived by non-Irish writers (*ibid.*; Amador-Moreno 2010a, pp. 91–2). In the late eighteenth century, some playwrights began to incorporate the syntactic features of IrE into their works (Amador-Moreno, 2010a, p. 95). Later, in the late nineteenth and early twentieth century, the Irish Revivalists³ continued this trend with some authors incorporating IrE’s grammatical, lexical, and phonetic features in their writing (*ibid.*, p. 100).

Despite the persistence of stereotyped representations of IrE in cinema (Walshe, 2009), there are nowadays portrayals of this variety that consider the intended audience of the text –

³ This is the term applied to the authors from the period known as the Irish Literary Revival (1890-1922), who employed IrE in their works. They attempted to portray this language variety with realism and dignity, without reducing it to clichéd phrases, with some notable exceptions (Amador-Moreno, 2010, p. 97).

Irish or from overseas –, and their ability to understand the language and the cultural references in the text (Amador-Moreno 2010a, p. 106). This awareness of the audience is present in *Derry Girls*, as creator Lisa McGee aimed to employ LDE dialectalisms that “would not be totally beyond the comprehension of outsiders [i.e., the largely British audience]” (Little, 2017). Nevertheless, the glossaries based on the show’s lexis mentioned in the introduction point to a labour of balance between understandability and authenticity in the series’ portrayal of LDE. In his study of IrE representation in film, Walshe also subscribes to this balance ideal when he points out that writers should “words which are exclusive, in so far as they are not used by non-Irish people, yet, at the same time, inclusive, in that non-Irish people are familiar with them” (2009, p. 156; see also Amador-Moreno 2010b, p.533).

A common thread in Amador-Moreno’s (2010a) chapter on the use of IrE in fiction is the use of this language variety as a characterisation tool, at times by invoking the stereotypes associated with IrE and at others as a more dignified identity marker. Indeed, in media, dialogue is “an essential tool in building the characters in a story” (Kozloff, 2000, p. 43), as it provides the audience with insights about their inner life and background (ibid.). In *Derry Girls* the language variety each character employs in their dialogues also serves to anchor them to their environment, i.e., the titular city, to make them feel like an organic part of it, or, in the case of James, Michelle’s English cousin, to foreground their otherness (Valleriani, 2021).

The representation of IrE has received a respectable amount of scholarly attention; consequently, an exhaustive list of all the available research on the matter would be a lengthy enterprise. However, one may cite a few examples, such as McCafferty (2005), who examined the use of IrE grammatical features in the works of NI writer William Carleton, using them as evidence of the historical development of said features in real-life NIrE. The above-mentioned study by Amador-Moreno, among others, dealt with the representation of north and south Dublin English in Paul Howard’s *Ross O’Carroll-Kelly* novels (2012; 2015), and how this portrayal of both Dublin English varieties was used as a characterisation tool (2016). Amador-Moreno and Terrazas-Calero also examined the *Ross O’Carroll-Kelly* novels, focusing on a range of discourse markers present in the texts (2017). More recently, this same body of work, plus the works of other contemporary Irish novelists have been explored by Terrazas-Calero (2022), again with a focus on discourse markers. Walshe, for his part, has focused on audiovisual media, researching the use of IrE dialects in both films (2009; 2016) and television (2011).

2.2.2 Television dialogue

The language in films and television, the latter being the focus of the present thesis, can be regarded as a form of literary dialect, in the same vein that theatre is (Walshe, 2011). In contrast to literary dialogue and film dialogue, televisual dialogue has been relatively understudied (Bednarek, 2010, p. 63). In the last fifteen years, however, the tide has begun to turn thanks to the work of scholars such as Paulo Quaglio and Monika Bednarek. The former conducted a corpus stylistic study (Quaglio, 2009) of the dialogue in the American sitcom *Friends* (1994–2004) using methodology adapted from Biber (1988, as cited in Quaglio, 2009, p. 30), whereas the latter has published several papers and books with television dialogue as their object of study in one way or another. Notable is Bednarek's (2010) study of the dialogue in the American dramedy *Gilmore Girls* (2000–2007), where she explored the language of the show from different perspectives, including characterisation and ideology, and employing a mix of quantitative corpus stylistics methods and qualitative analysis throughout. This same scholar also analysed characterisation in the show *The Big Bang Theory*, focusing on how the 'nerdiness' of one of the main characters is linguistically constructed (Bednarek 2012).

Bednarek (2019) also interviewed a group of screenwriters and collated their insight into the writing process of telecinematic dialogue, covering areas such as characterisation, swearing, and dialect. Swearing was also the focus of a study of the evolution of taboo language in catchphrases carried out by Beers-Fägersten and Bednarek (2022). More recently, Bednarek (2023) has turned to characterisation once more, this time examining a variety of case studies from the United States and Australia using corpus methods.

Telecinematic dialogue has also been studied from a gender perspective. For example, Bednarek (2015) analysed the use of taboo language by the female protagonists of three shows – *Weeds*, *Nurse Jackie* and *Saving Grace* –, finding that the three characters challenged stereotypical expectations of reputable behaviour by cishet white women (p. 446). Another researcher, Gregori-Signes, has studied the representation of female characters in the situation comedy *3rd Rock from the Sun*, both qualitatively (2007) and using corpus linguistic methods (2017).

As far as Irish television productions are concerned, Walshe (2011) has examined the language of the comedy *Father Ted* (1995–1998) in search of distinctive features of IrE. This show had already been studied and compared to the American situation comedy *The Simpsons* in terms of their violations of Gricean maxims in order to create humour (Hughes, 2004). *Father*

Ted was produced by the same company as *Derry Girls*, Hat Trick Productions, with both series having in common the fact that they are geared towards a British audience despite their being set in Ireland.

A peculiarity that distinguishes televisual dialogue from its filmic and literary counterparts is the fact that the former is usually the product of a group effort: television series are written by groups of screenwriters, which can affect the consistency of the characterisation, as different writers may have a slightly different approach to a character's psychology and behaviour (Bednarek, 2010, p. 15). Even though filmic scripts can also be penned by several screenwriters, the episodic and seasonal format of television series entails that each episode might be the work of a different author, and that writing teams might change from season to season as some people depart and others are hired in their place. However, this does not seem to be the case for *Derry Girls*, as the only writer credited for the episodes is the show's creator Lisa McGee (Internet Movie Database, n.d.). With only one writer at the helm, one could expect a reasonable degree of consistency in each character's unique voice.

As mentioned at the beginning of section 2.2, a problem posed by the use of dialogue as linguistic data is that of authenticity. As Amador-Moreno (2010a) writes, "fictional dialogue simulates real conversation within certain limitations, and it is in relation to real verbal interactions that fictional interactions are understood and interpreted by the audience" (p. 89). This relationship between fictional and real-life dialogue is a reason to use a real-language reference corpus when examining fictional language (Amador-Moreno 2010b, p. 535). In the case of audiovisual media, it has been argued that the distance between the fictional dialogue and the real world widens, as the number of middlemen is higher than in literary language:

[Dialogue] has been scripted, written and rewritten, censored, polished, rehearsed and performed. Even when lines are improvised on the set, they have been spoken by impersonators, judged, approved, and allowed to remain. Then all dialogue is recorded, edited, mixed, underscored and played through stereophonic speakers with Dolby sounds. The actual hesitations, repetitions, digressions, grunts, interruptions, and mutterings of everyday speech have either been pruned away, or, if not, deliberately included. (Kozloff, 2000, p. 18)

The above applies to all kinds of audiovisual dialogue; when it comes to the representation of dialects in this type of media, some writers prefer to leave the matter to the actors, provided that the cast members concerned share their dialectal background with the character they are playing (Bednarek, 2019, p. 55). However, written works can also pass through the hands of

editors, and their representation of speech is limited to the letters of the alphabet, whereas audiovisual dialogue benefits from the presence of paralinguistic features such as accent, gesture, and intonation.

A further consideration regarding authenticity is that of linguistic variation. No speaker of a variety uses dialectal forms all the time; therefore, if a character's speech contains linguistic variation, it must be seen as an attempt to infuse their speech with the characteristics of real language and not as an inconsistency in their characterisation (Amador-Moreno 2010b, p. 533).

2.2.3 Corpus stylistics

This thesis lies at the intersection between corpus pragmatics and stylistics. The latter field is commonly defined as the study of style, i.e., “the study of how the linguistic choices evident in a text contribute to the overall meanings and effects of that text” (McIntyre & Walker, 2022, p. 16). Therefore, stylistics functions as the link between linguistics and literary criticism. Within that field, the present study is concerned with the concept of characterisation, understood as “the process by which we comprehend and conceptualise characters as we read” (McIntyre & Walker 2022, p. 182), or as we watch, in the case of audiovisual media such as *Derry Girls*.

Characterisation emerges from top-down and bottom-up processing (Culpeper, 2001, as cited in McIntyre & Walker 2022, p. 183). The former concerns the use of “schematic knowledge to form an impression of character” (McIntyre & Walker 2022, p. 183), that is, drawing from real-life experiences in order to understand characters, taking into account the social categories to which they belong, such as gender, age, class, occupation, etc (ibid.).

The second mechanism of characterisation, bottom-up processing, can be defined as the procedure of taking information about character from linguistic cues in the text itself (McIntyre & Walker 2022, p. 184). These cues can be explicit, i.e., the character directly explaining “something intrinsic to their personality”; or implicit, i.e., the insights gained from the character's use of certain linguistic, paralinguistic and non-linguistic features (ibid.). Among these features that operate as characterising devices are lexis, accent, and dialect (Culpeper 2001, in McIntyre & Walker 2022, p. 185). The DPMs characteristic of NIrE that can be found in the dialogue of *Derry Girls*, as the dialectal lexical elements they constitute, are the characterisation devices this study aims to analyse.

Stylistics, like many other fields within linguistics, has benefitted from the use of corpora. McIntyre and Walker, in their 2022 publication on the subject, have discussed the

concept of *corpus-informed stylistics*, which according to them involves “using pre-existing large-scale corpora to support the stylistic analysis of single texts or textual extracts” (p. 26), with the pre-existing corpora being chosen on the basis of their “suitability to test and potentially explain particular language intuitions, or answer particular questions about the target text” (p. 64). Such is the approach taken by this thesis, as the dialogue in *Derry Girls* is compared to the SPICE-Ireland corpus (Kallen & Kirk, 2012). However, the dialogue lines have also been compiled into a corpus for easy sampling.

2.3 Northern Irish English (NIrE)

Having discussed the concepts of literary dialect and the field of corpus stylistics, we now turn to the language variety employed in *Derry Girls*, LDE, or, more broadly, NIrE. (London)Derry is the second largest city in NI and the largest in county Derry. The city lies near the border with county Donegal, part of the Republic of Ireland (ROI). Both mentioned counties form part of the Ulster province, the northernmost of Ireland’s four traditional provinces, which nowadays is largely a part of the United Kingdom (counties Donegal, Cavan, and Monaghan are part of the ROI) (Hickey, 2007, p. 85).

Ulster remained largely Irish-speaking until the 17th century (Hickey, 2007, p. 95; Corrigan, 2010, p. 19). Following the flight of the native Irish aristocracy in 1607, James I encouraged the building of plantations in the province, bringing Lowland Scottish and English settlers to work on them (Hickey, 2005, p. 21; Hickey, 2007, p. 87). After this initial wave of settlement, more English and Scottish immigrants would arrive in the region throughout the next centuries, bringing with them their language varieties, namely Scots and northern and Midlands dialects of English (Hickey, 2007, p. 87). These language varieties made inroads into the native Irish population, resulting in their becoming bilingual in Irish and English over the course of the seventeenth and eighteenth centuries (Hickey, 2005, p. 24). The further spread of English during the 1800s would eventually tip the balance in favour of English, with more and more of the population abandoning Irish in favour of the introduced language (*ibid.*).

As the groups of incomers brought by the plantations tended to concentrate in different areas of Ulster, distinct speech regions began to emerge, and have survived until the present day:

1. Ulster Scots: eastern county Donegal, eastern county Down, north-eastern county Derry and northern county Antrim.

2. South Ulster English: in the southern parts of counties Fermanagh, Monaghan and Armagh.
3. *Gaeltacht* (Irish-speaking area): western county Donegal.
4. Mid Ulster English: the remaining areas of Ulster, including the Belfast metropolitan area, most of county Derry, and the city of (London)Derry. (Hickey, 2007, p. 93)

Traditionally, the English and Scots speech areas have been distinguished from one another on the basis of phonology (Corrigan, 2010, p. 17). For instance, Ulster Scots speakers tend to observe the Scottish Vowel Length Rule⁴, whereas speakers of South Ulster English conserve “the historical phonemic vowel length system of West Germanic” (ibid.). Mid Ulster English, appropriately, exhibits a “modified Scots vowel length pattern.” (See McCafferty, 2007, p. 125, for a discussion of how the Scottish Vowel Length Rule partially applies to Mid Ulster English)

In addition to NI’s history of settlement, another factor that may have influenced the region is faith, which acts as a marker of ethnicity. The main two ethnic groups in NI are Catholics and Protestants (Corrigan, 2010, p.25). Among the urban areas in the region, Derry stands out for being the only major city with a Catholic majority (over 70% of the population identifies as such) (Hickey, 2007, p. 343). Scholars have researched if this ethnic divide has led to linguistic differences between the two populations (Corrigan, 2010, p. 27), most notably McCafferty (1999; 2001).

2.4 “Sure that’s not what they meant”: The study of pragmatics in IrE

Pragmatics, the study of “how speaker meaning is interpreted in context” (Vaughan & Clancy, 2011, p. 47), is a relatively new field of enquiry, and its application to IrE is even newer (ibid.). In consequence, the pragmatics of IrE, never mind that of the northern varieties of IrE, had been relatively understudied until relatively recently (Amador-Moreno, 2010a, p. 115; Corrigan, 2010, p. 99). This situation began to be remedied in 2005, with the publication of a collection of papers on the subject entitled *The Pragmatics of Irish English*, edited by Barron and Schneider, although few of the studies contained in the volume dealt specifically with NIrE (ibid.).

This publication emphasised the need for a sub-field termed *variational pragmatics*, at the intersection of pragmatics and modern dialectology, which sought to “explore the effects of

⁴ In varieties affected by the Scottish Vowel Length Rule, the length of a vowel is determined by its phonetic context. Thus, when a vowel is followed by certain consonants or an inflectional suffix, the vowel is long (Corrigan, 2010, pp. 31–2).

macrosocial features on language in (inter)action” (Murphy, 2015, p. 66). The articles included in Barron and Schneider’s volume examined the features mentioned in three areas: the private sphere, the public sphere, and the official sphere (Hickey, 2007, p. 370). Ultimately, Barron and Schneider’s publication is notable for having originated a perspective shift regarding the study of IrE: from then on, researchers have studied it “as a variety in its own right” and not solely as a dialect that serves as a comparison to British English (Terrazas-Calero, 2022 p. 18).

Thanks to the contributions to Barron and Schneider’s volume, and other publications that followed, like Hickey’s (2007), some insights about the pragmatic profile of IrE were brought to light. In the following subsection, I present an overview of what makes IrE distinct from other varieties of English in terms of pragmatics, before diving into the study of discourse-pragmatic markers in IrE, i.e., the linguistic items that will be analysed in this thesis.

2.4.1 The pragmatics of IrE

It has been noted that the pragmatic profile of IrE has been shaped by the historical and social circumstances in which this variety of English developed. Hickey (2007) has written that discourse interaction in Ireland is mostly grounded on practices that originate in the rural context that characterised most of Irish society until the twentieth century, where social interactions happened between friends, family, and acquaintances, and were largely consensual (p. 371). As a consequence, conversations between Irish strangers tend to include the search for common acquaintances or experiences (ibid.). Hickey also highlights the importance of cooperation and support between interlocutors, which favours frequent backchannelling (1) and reassurances (2), (2007, p. 372):

(1) Yeah, right, sure, of course

(2) Just a sec, I’ll be with you in two minutes (ibid.)

Indeed, Farr (2005) highlights that in Irish culture engagement with one’s interlocutor “must be very obvious visibly or audibly” (p. 207). In contrast, contradiction and direct criticism are disfavoured and must be accompanied by a hedge (Hickey, 2007, p. 372), i.e., a softening device that enables the speaker to distance themselves from their utterance (Brown and Levinson, 1987, p. 146):

(3) JOE: Look at this. They’re bombing the bridge, **apparently**. (*Derry Girls*, 1x01)

As we can see in (3), by ending his sentence with *apparently*, Joe indicates both that he is not completely sure about the veracity of the information he is reporting, and that he does not take responsibility for its being true or not.

In addition to this, pointing out social differences between interlocutors is discouraged, and those with higher social status than their conversation partners may soften their contributions to the exchange and “background [their] social position”, opting to employ first name terms instead of titles, and to treat each other in a familiar way (Farr, 2005, p. 239; Farr & O’Keeffe, 2002, p. 42).

Another characteristic of the pragmatic profile of IrE that is frequently mentioned in the existing literature is indirectness (Amador-Moreno, 2010a, p. 115; Clancy & Vaughan, 2012, p. 226), often reflected in the use of hedges as a softening device (Vaughan & Clancy, 2011, p. 49), and in the use of “off-the-record utterances” such as understatements, irony, or ellipses (Kallen, 2005a, p. 113). Millar (2015) notes that this indirectness, together with the search for common ground and consensuality has been linked to the national Irish character, understood as “a dislike of self-revelation; over-assertiveness and flattery” (p. 294). In this respect, the use of the pragmatic repertoire of IrE can be seen as indexing Irishness itself (Hickey & Amador-Moreno, 2020, p. 15, as cited in Terrazas-Calero, 2022, p. 106), i.e., as being a sign of Irish identity. Finally, as we will see in more detail in the following section, IrE exhibits some autochthonous uses of pragmatic markers common to other varieties of English, as well as DPMs of its own not found in other dialects.

2.4.2 The study of DPMs in IrE

In their 2011 review of the existent research on IrE pragmatics, Vaughan and Clancy noted the preponderance of corpus linguistics methods (p. 50); thus making the present thesis a contribution to this tradition, so to speak. These authors concluded their review by listing some areas where there was still little scholarly work done, among them DPMs in IrE (*ibid.*, p. 51). However, some studies on the matter had already been included in Barron and Schneider’s (2005) volume mentioned above. For instance, Kallen (2005b), besides examining the role of communicative silence and understatements in IrE, analysed the functions of two pairs of discourse markers (*I say* versus *I’d say*, and *I mean* versus *you know*), where the second member of each pair has a downgrading function. Kallen used data from the ICE corpus to compare IrE to British English and concluded that the downgrading discourse markers (*I’d say* and *you*

know) are more frequent in IrE than their non-downgrading counterparts, which are more frequent in British English.

For her part, Amador-Moreno (2005) focused on the literary representation of IrE through the use of the discourse markers *and*, *sure*, *troth*, and *arrah*, using two novels from Donegal writer Patrick McGill as data. *And* tends to precede questions in IrE and seems devoid of meaning in this particular usage. *Sure*, a “dialectal realization of the adverbial form *surely*” (Amador-Moreno, 2005, p.83), and a common sentence opener in IrE, was found to have a wide range of functions in this variety of English, that of emphatic device chief among them. *Troth* (“truth”), another sentence opener, seemed to function similarly to *sure*. Finally, *arrah*, which also appears at the beginning of utterances, was found to signal the attitude of the speaker, commonly one of impatience and displeasure (p. 92).

Finally, while O’Keeffe’s (2005) chapter in Barron and Schneider’s volume focused on question forms in an Irish radio phone-in segment, she noted the organisational role that discourse markers *well* and *and* have in interactions (p. 349), as well as the use of first name vocatives⁵ to hedge the force of utterances (p. 358).

Despite these contributions, there would not be a publication exclusively devoted to DPMs (for matters of terminology, see section 2.4.2.1 below) in IrE until 2015, when this gap was addressed by *Pragmatic Markers of Irish English*, edited by Amador-Moreno, McCafferty and Vaughan (2015). This volume contained a collection of studies that approached DPMs “in terms of pragmatic and sociolinguistic variation in a wide range of informal, intimate, professional and/or literary settings in IrE” (Terrazas-Calero, 2022, p. 19). In addition to this, in most of the studies featured in Amador et al.’s volume, the data are sourced from both purpose-built and already existing corpora of IrE, among them the ICE-Ireland, which will also be used in the present thesis (Amador-Moreno et al., 2015, p. 2). Finally, the editors adopted an inclusive approach to what constitutes a DPM, and thus the volume contains studies of more non-traditional PMs such as turn openers, turn-initial elements, tag questions, and vocatives (ibid., p. 6). This sets a precedent for the present thesis, which will also examine some unconventional DPMs, as will be explained below.

There have been other studies of DPMs in IrE in the wake of Amador-Moreno et al.’s (2015) volume. McCafferty and Amador-Moreno (2019), for instance, took a historical

⁵ Vocatives are a type of term of address, the latter being defined as “a device used to refer to the addressee(s) of an utterance” (Leech, 1999, p. 107).

approach to the study of the DPM *sure* in a corpus of letters exchanged between Irish emigrants and their relatives. Regarding the study of IrE in fictional discourse, both Terrazas-Calero (2022) and Walshe (2016) have analysed the use of DPMs in fictional media. The former examined a selection of DPMs in a corpus of contemporary Irish prose fiction, whereas the latter looked at the DPMs, including sentence tags, in a corpus composed of films set in Northern Ireland.

2.4.2.1 DPMs: definitions and functions

It would be pertinent at this point to offer a brief overview of the study of DPMs and the different definitions adopted by scholars throughout the decades. Scholarly interest in the matter began in the 1980s, and since then DPMs have been studied from a variety of perspectives, including sociolinguistics, historical linguistics, and second-language acquisition studies, among others (Amador-Moreno, 2005, p. 73).

In the field of DPMs, Deborah Schiffrin's *Discourse Markers* (1987) has been identified as one of the pioneering publications in the field (Aijmer & Simon-Vandenberg 2011, p. 224). The title of her book is also one of the most frequently used terms used to refer to this category of linguistic elements. However, the term *discourse markers* is not without its downsides, as it can be used in a rather narrow sense to refer to markers that function as “sequentially dependent elements which bracket units of talk” (Schiffrin, 1987, p. 31), such as *however* or *furthermore*. An alternative label is *pragmatic marker*, defined as a word or a short phrase that has a metalinguistic function in discourse (Brinton, 1996), which emphasises the interactional function of the particle, such as opening a turn or ceding it (Aijmer & Simon-Vandenberg, 2011, p. 227).

The most recent term employed to refer to this category of words is *discourse-pragmatic marker*, championed by Tagliamonte (2012) and Pichler (2016), and used, for instance, by Corrigan (2015) and Terrazas-Calero (2022) in their work on DPMs in IrE. As Pichler (2016) points out, the heterogeneous linguistic items subsumed under the label DPM share the following characteristics:

- a) they perform a range of interpersonal and/or textual functions in discourse;
- b) their use is motivated mainly by their functionality (p. 3).

Corrigan notes the usefulness of the term *DPM* in the sense that it encompasses a broad range of “linguistic items or expressions like interjections, pragmatic particles, discourse markers,

quotatives, intensifiers, general extenders [or] tag questions” (2015, p. 37). This follows the aforementioned trend of including non-traditional DPMs in this category. 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”. Their existence out of clause boundaries can be demonstrated by performing a deletion test, which normally shows that the sentence without the DPM is still well-formed (Hickey, 2015, p. 19).

Table 2.1. Characteristics of discourse-pragmatic markers

<p>Phonological and lexical characteristics</p> <ul style="list-style-type: none"> 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.
<p>Syntactic characteristics</p> <ul style="list-style-type: none"> 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).
<p>Semantic characteristics</p> <ul style="list-style-type: none"> g) Pragmatic markers have little or no propositional/conceptual meaning, but are procedural and non-compositional.
<p>Functional characteristics</p> <ul style="list-style-type: none"> h) Pragmatic markers are often multifunctional, having a range of pragmatic functions.
<p>Sociolinguistic and stylistic characteristics</p> <ul style="list-style-type: none"> 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 stigmatised and negatively evaluated, especially in written or formal discourse. l) Pragmatic markers may be used in different ways and with different frequencies by men and women. <p style="text-align: center;">(Brinton, 2017, p. 9)</p>

As far as the characteristics of DPMs are concerned, Brinton (2017) offers a useful summary, presented in Table 2.1 above. These characteristics can prove useful to identify potential, hitherto unstudied DPMs, such as *wee*, which will be discussed in more detail below. As an

addendum to the semantic characteristics of DPMs, Brinton (2017, p. 28) has noted that, even if they are somewhat bleached of their original meaning, DPMs tend to retain traces of it, mentioning the example of *well*, which conserves the idea of “acceptability”.

With regards to subdivisions of the DPM category, Fraser offered a classification that distinguishes between four categories. The first is basic PMs (his term), which signal the illocutionary force of the utterance. The second, commentary markers, encode a comment on the basic message. In the third, parallel markers, the message encoded is different from the basic message. Finally, we have discourse markers (Fraser 1996, 2006, as cited in Aijmer & Simon-Vandenberg, 2011, p. 227). However, it has been pointed out that many DPMs, such as *well* or *oh*, do not fit into this taxonomy (Aijmer & Simon-Vandenberg, 2011, p. 228), in addition to the fact that many DPMs are multifunctional (*ibid.*). In general terms, however, their functions can be summarised as commenting on and aiding in the interpretation of the message and signalling the speaker’s position in relation to the hearer’s (Corrigan, 2010, p.68).

DPMs have been associated with informality and spoken discourse, as they seem to appear more frequently in this medium than in written discourse (Amador-Moreno & McCafferty, 2015, p. 271); and also with women’s speech, as some researchers argue that women are more likely to use linguistic devices that signal tentativeness, a function DPMs can have. (Brinton, 1996, pp. 35). This association with female speech is controversial, however (*ibid.*). This, paired with their relative frequency in discourse (Hickey, 2007, p. 375), suggests that the dialogue of *Derry Girls*, dominated by young female characters, could contain a reasonable amount of DPMs (i.e., enough data) such as to warrant their analysis, assuming that the frequency of these items mirrors somewhat that of real-life spoken interaction.

2.4.2.2 Existing research on DPMs in IrE

This subsection will provide an overview of the available literature on DPMs prominent in IrE, including the main findings regarding their use and functions.

2.4.2.2.1 *Sure*

Owing to the flourishing of scholarly interest in the pragmatics of IrE, there is a growing body of work centred around both DPMs specific to IrE and DPMs that, despite appearing in other English varieties, have some unique usage patterns in IrE. One of the most widely used DPMs in IrE is *sure*, pronounced /ʃʌr/ when used in this capacity, as opposed to its adjectival pronunciation (/ʃʊər/) (Murphy, 2015, p. 71). Walshe (2016) notes that this is the most frequent

DPM in ICE-Ireland (p. 340). As mentioned above, Amador-Moreno (2005) studied *sure* in Pat McGill’s novels. Walshe (2009) examined its use in films set in Ireland, noting the versatility of this DPM. Murphy (2015), on her part, investigated the impact of gender and age on the use of *sure* among IrE speakers, with her results suggesting that the employment of this DPM does not vary much across these demographic categories. Among the functions of *sure* identified in her analysis are that of a downtoner to soften disagreement (4), and as an appeal for consensus (5):

(4) Ellen: So she told me this morning that they had the vet with him [dog] he gave him two injections as he has heart trouble.

Nora: And he’s so young.

Ellen: He’s not **sure**, he’s over his time for a dog of his age. (Murphy, 2015, p. 83)

(5) Marie: And **sure** isn’t Carlos above with her from Spain.

Ellen: There you are. (Ibid.)

O’Sullivan (2015), in a paper on the use of DPMs in Irish radio advertising, found evidence that *sure* can index a more rural Irish identity when contrasted with other markers such as *like* (p. 335). Finally, McCafferty & Amador-Moreno (2019) conducted a diachronic study of this DPM using data from the *Corpus of Irish Correspondence (CORIECOR)*, which contains letters spanning from the late seventeenth century to the early twentieth. The evidence these researchers found suggested that American English uses of *sure* might have developed from IrE uses of this DPM brought by Irish emigrants (p. 74).

2.4.2.2.2 *Like*

The other prominent DPM in IrE is *like*. This marker is common in other varieties of English, where it can function as a quotative (Hickey, 2007, p. 131; Romaine & Lange, 1991) i.e., to introduce reported speech (6) or as a focusing device that precedes new information (7):

(6) And I was **like** “You’re not allowed to smoke in here.” (My example)

(7) And there were **like** people blocking. (Romaine & Lange, 1991, p. 245)

This DPM can also be used as an exemplifier equivalent to “for example” (Amador-Moreno, 2012, p. 31). IrE also exhibits these uses of *like*, in addition to some unique ones. For example, in a diachronic study using data from the *CORIECOR* (see 2.4.2.2.1 above for a description of

this corpus), Amador-Moreno and McCafferty (2015), noted that *like* is used in the corpus as a hedge or mitigator (8), to signal imprecision (9), as a focusing device (10) and to highlight explanations (11):

(8) I told Robert when I received Elizabeth's letter that she said I was the best **like** but he did not agree with her on that (Amador-Moreno & McCafferty, 2015, p. 278)

(9) They put **like** wheels under them and takes them along. (Ibid.)

(10) These boats carry passengers and general cargo and **like** all the CPR boats are very popular (Ibid., p. 279)

(11) eased with their visit and delighted with the situation [...] the Farm, and they say **like** the Candidate who could not be elected for want of votes, you cannot be at home till you have society and Voters around You (Ibid.)

In contrast, Corrigan (2010) notes several uses specific to NI, where *like* is often employed in sentence-final position. In NIrE, sentence-final *like* can denote the end of old information (12) or be used as a mitigator to soften the force of an utterance (13) (pp. 100–1):

(12) You know, it was good craic **like**. (Corrigan, 2010, p. 100)

(13) It's not really rough, but it's just it's not great either **like**. (Ibid.)

Murphy (2015), however, also reported the use of *like* as a hedge in her study, which was based on data from the ROI. In a similar vein, Walshe (2016) notes that sentence-final *like* is more common in the ROI data than in the NI data in the SPICE-Ireland corpus, while also pointing out the higher frequency of this DPM in the ROI component (compared to the NI component) of the film corpus he compiled (p. 334). Outside of Ireland, final *like* has been reported in Scottish English (Miller & Weinert, 2017), where it can function to anticipate and counter assumptions. With regard to its social profile, Schweinberger (2015) noted that *like* seemed to be most frequently employed by females in their twenties and early thirties, and that in both older and younger age groups there was not a significant difference in use between genders (p. 128). *Like* has also been linked to the speech of teenagers in general (Diskin, 2017, p. 144).

This DPM has also been studied in the context of fictional discourse. For example, Amador-Moreno (2012) investigated the use of *like* in Paul Howard's novel *The Curious Incident of the Dog in the Nightdress*, set in Dublin, reporting that the majority of uses of this

DPM occurred in medial position, in contrast to the frequent use of *like* in clause-final position that is commonly associated with IrE (p. 30). The researcher suggested that this might reflect a linguistic change in progress in Dublin English, where the local use of *like* might be transitioning from more typically Irish uses to more globalised ones (ibid., p. 33). These findings were similar to those of Nestor et al. (2012), who, in a study of *like* use among L2 IrE speakers, found that their informants who resided in Dublin were more likely to use medial *like* than those in County Mayo, a more rural setting. Amador-Moreno approached the matter of *like* in fictional discourse again in 2015, examining how *like*, in addition to other DPMs and quotative devices, is used in Paul Howard's above-mentioned novel to recreate Dublin English; in particular, she examined the differences in the use of *like* between the characters from South Dublin (the 'posh' side), and those from North Dublin. The former would be more likely to use *like* as a DPM than the latter, as it can be seen as a marker of sophistication (p. 386). Finally, Terrazas-Calero (2022) studied the uses of *like* in a corpus of contemporary Irish fiction. This researcher also reported that in the data examined this DPM was most commonly found in medial position, thus offering further evidence for the shift mentioned above (p. 210). Terrazas-Calero also found *like* to be employed most frequently by females (p. 213), in line with Schweinberger's (2015) results. With regard to its functions, Terrazas-Calero noted that the most common functions of clause-medial *like* were hedging, indicating a lack of clarity, and emphasising the information that follows the DPM (i.e., a focusing device) (p. 221).

2.4.2.2.3 *Now*

Another DPM that is relatively frequent in IrE, although not exclusive to this variety is *now*. This marker is multifunctional, being able to fulfil several of its possible functions simultaneously (Clancy & Vaughan, 2012, p. 230). *Now* can be used in an utterance-initial position to signal a change in topic (Aijmer, 2002, p. 62), as we can see in example (14). Another of its uses is to convey the speaker's urge to take the floor (ibid., p. 94), as exemplified by (15).

(14) JAMES: Sister Michael, can I just quickly ask, um, since there's no male student toilets on the premises, am I permitted to use the staff ones?

SISTER MICHAEL: Absolutely not. **Now**, I think you all owe Tina an apology, don't you? (*Derry Girls*, 1x01)

(15) JENNY: **Now**, the wee girl claims that she's not bothered, but reading between the lines... (*Derry Girls*, 1x01)

As a more organisational marker, *now* can be used in utterances in which the speaker “progresses through a cumulative series of subordinate units”, such as a list (Schiffrin, 1987, p. 232):

(16) CLARE: What’s she doing now?

JAMES: Reading her book. **Now** she’s looking at the woman beside her. **Now** she’s getting up. **Now** she’s coming this way. **Now** she’s standing right in front of us. (*Derry Girls*, 2x03)

But the pragmatic function of *now* in IrE, in clause-final position, is that of a hedge, and thus it constitutes a politeness strategy that shields the user from threats to face (Clancy & Vaughan, 2012, p. 237):

(17) Speaker 1: And has there ever been a situation that you can tell me about where the executive has gone against you on something? ...

Speaker 2: No never. I don’t think so **now**. (Ibid.)

(18) GERRY: I don’t think she was talking to me, Joe. I think she was talking to the person in the full-length white frock who just managed to upstage the actual bride.

SARAH: Ach, I wouldn’t say I upstaged her **now**, Gerry. (*Derry Girls*, 2x04)

This use of *now* makes it salient in IrE discourse when compared to other varieties of English (ibid., p. 399). In this specifically IrE context *now* has been studied by Clancy and Vaughan (2012) and Millar (2015), and its acquisition by newcomers to Ireland (i.e., not native speakers of IrE) has been examined by Migge (2015).

2.4.2.2.4 *And*

The DPM *and* has also been studied in the context of IrE, specifically in literary discourse, by Amador-Moreno (2005). This scholar suggested that this marker could have its origin in the Irish interrogative particle *an*, thus making it a transfer from this language (Amador-Moreno, 2005, p. 83). Even though *and* appears as a DPM in other varieties of English, the IrE version of it seems devoid of meaning:

(19) “It’s yerself that is the decent youngster, God bless ye!” he said, and there were tears in his eyes. “**And** isn’t this a fine warm place ye are inside of this wet night?” (ibid., p. 82)

In contrast, the non-dialect-specific *and* tends to have a linking function, marking an utterance as a continuation of an interaction (Schiffrin 1987, pp. 146–7):

(20) ERIN: I cannot actually believe this. I'm missing David Donnelly's gig, Michelle.

MICHELLE: **And** all because of that miniature motherfucker. (*Derry Girls*, 1x01)

2.4.2.2.5 *So*

A further DPM that appears in IrE is *so*. At the beginning of a sentence, this DPM is a call for the audience's attention, as well as indicating that what is to follow is something that the speaker has been wanting to talk about (Amador-Moreno, 2019, pp. 92–3). This use also appears in other varieties of English (*ibid.*). A more exclusively Irish usage of *so* is its employment in utterance-final position, where it is used to indicate consent or acquiescence (Hickey, 2007, p. 371):

(21) CIARAN: This is a nice surprise. When Sarah asked me to the cinema, I thought, well, I thought she meant just the two of us.

GERRY: No, Ciaran, because that would be... What's the word now? Normal.

CIARAN: Double date it is, then, **so**. (*Derry Girls*, 2x02)

Binchy (2005) examined the use of this marker as an utterance-final mitigator that serves to soften the force of asking for money in commercial transactions, and is often reinforced with *please* or *now* (Walshe, 2016, p. 123):

(22) CUSTOMER: Pine Needle please.

SERVER: That's one eighty-eight **so please**. (Binchy, 2005, p. 330)

So can also appear as a sentence tag, with the whole tag functioning as an emphasiser. The tag tends to feature *so* followed by the verb from the main clause, or an auxiliary (Walshe, 2009, p. 127), resulting in structures like the following:

(23) SARAH: And big Shay has eyes like a hawk, **so he does**. (*Derry Girls*, 1x04).

In interactions, the tag can be a response to an interlocutor's utterance:

(24) JENNY: I don't know what you're talking about.

ERIN: Aye, **so you don't**. Supergrass! (*Derry Girls*, 2x02)

Taking a historical perspective, Amador-Moreno (2019) has studied the uses of *so* in the CORIECOR, focusing on letters from the 1800s. In the data she examined, this DPM was most frequently used to close a topic or to mark the boundaries between two different topics (p. 99):

(25) I have nothing more to say to you so I must conclude as the mail shall soon start. **So** I send my best respects to Father and Mother Brothers and Sisters and also to Uncle John.
(Ibid.)

2.4.2.2.6 Taboo language

Taboo language has also been studied in the context of pragmatics, with these expressions often being treated as DPMs (see Terrazas-Calero, 2022). Here we understand taboo language as language use that is potentially “offensive, inappropriate, objectionable, or unacceptable in any given social context” (Beers Fägersten, 2012, p. 3). In English, taboo expressions usually involve:

- (1) religion (with examples like *sweet suffering Jesus*, usually termed “swear words”),
- (2) sexuality, bodily parts and/or effluvia (like *motherfuckers*, *dicko*, *shite*, referred to as “expletives”),
- (3) words related to animals (such as *bitch*, which can be termed “insults”) (Murphy, 2010, pp. 167–8, Schweinberger, 2018, p. 2; Stapleton, 2010, p. 290). Taboo words and expressions must not be taken literally, however, as they are instead used to express “strong emotions and attitudes”, particularly anger (Andersson and Trudgill, 2007, p. 195; Stapleton, 2010, p. 294).

With regard to their syntactic functions, taboo expressions can be part of lexicalized constructions (26), used as intensifiers before a noun, adjective or adverb (27), or used as a standalone utterance (28) (Pinker, 2008, p. 350):

(26) I don't give a **fuck**!

(27) My **fucking** boss, **fucking** annoying, you've done it **fucking** wrong

(28) **Fucking** hell! (My examples)

In her review of the existing literature on taboo language, Stapleton (2010) discusses the interpersonal functions of this linguistic behaviour identified by linguists. These functions are context-dependent, affected by social norms, formality levels, or the relative status of interlocutors and their familiarity with each other (p. 291). In general, taboo language is regarded as offensive and associated with a lower socioeconomic status, thus being avoided by reasons of social image (ibid.). At the same time, these very associations turn the use of taboo language into an identity-affirming practice that carries covert prestige⁶ among teenagers and lower-class males (ibid.).

The existing literature on the matter has identified the expression of emotions as a common interpersonal function of swearing (Stapleton, 2010). Although often associated with anger and other negative emotions, taboo language is used to express a range of inner states, including happiness, sadness, enthusiasm, or fear (p. 294). Other functions are humour and verbal emphasis, the latter through the above-mentioned use of expletives as intensifiers, e.g., *fucking amazing* (ibid.). The third interpersonal function of swearing identified by researchers concerns affirming in-group relationships, building solidarity with one's peers, and, conversely, establishing boundaries with non-group members (ibid.). Lastly, as mentioned above, taboo language is employed as an identity-constructing linguistic resource, taking on different meanings across genders and age groups (ibid., p. 299).

In the same vein, previous research suggests that “swearing is socially stratified along dimensions of age, gender, and social class,” with teenagers and young adults being the most prolific swearers (Schweinberger, 2018, pp. 3–4), and people over 60 being the least prone to using taboo language (McEnery and Xiao, 2004, p. 242). In the case of IrE, it has been reported that Irish people are prone to swearing, commonly by using religious expressions, such as *Jesus!* (Walshe, 2009, p. 129). In addition, there are taboo expressions particular to IrE, such as *feck*, a euphemism for *fuck* (Murphy, 2009, p. 91). Accordingly, swearing as a sociolinguistic phenomenon has received attention from IrE scholars. For example, Murphy (2009) examined the pragmatic functions of *fuck* in IrE in terms of age and gender with the aid of corpus-based methods. This researcher found that, in the small corpus that provided the data, *fuck* was more frequently used by males than females, and that speakers of both genders in their twenties were more likely to use this swear word than older speakers (p. 94). In 2010, this same scholar

⁶ In sociolinguistics, “covert prestige” refers to a situation where non-standard features, or entire dialects are attributed high prestige in a speech community, as opposed to situations of “overt prestige”, where the standard varieties are the ones considered prestigious. (Trudgill, 1972)

provided more insight into the use of *fuck* in IrE, this time with a focus on women's language. In this study, Murphy reported that among women *fuck* was rarely used as an insult, being employed instead to indicate "camaraderie and closeness" (2010, p. 186). As a result, Murphy hypothesises that *fuck* has undergone semantic bleaching and pragmatic strengthening (ibid., p. 187). This could point to *fuck* being in the process of becoming a pragmatic marker, losing its original meaning and acquiring pragmatic uses.

A further study of swearing in IrE was published by Schweinberger in 2018. In this corpus-based study, Schweinberger reported that speakers of IrE follow the general trend of teenagers and young adults being more likely to swear. In addition, males were found to use more taboo language than females, and speakers from NI swore more often than their ROI counterparts (ibid.). Accordingly, one could expect the dialogue in *Derry Girls*, i.e., a fictional representation of NirE, in addition to the main characters being teenagers, to contain a variety of taboo expressions, but due to matters of space, this thesis will not analyse them at length unless they appear as keywords (see section 3.2.1.3 for a discussion of keywords).

In fiction, taboo language is often employed to add a sense of authenticity and realism to the language, for example by using swearwords or expletives related to a specific regional variety when attempting to portray it (Bednarek, 2019, p. 47; Walshe, 2009, p. 150). In addition to this, other functions of swearing are aiding in characterisation, expressing emotion, eliciting a reaction in the audience, and creating humour through unexpectedness (Bednarek, 2019, p. 48; Walshe, 2011, p. 139). Taboo expressions are also often present in catchphrases, i.e., salient expressions that are associated with a character in a series through their repeated use (Beers-Fägersten & Bednarek, 2022, p. 197).

In his study of the representation of IrE in the TV series *Father Ted*, Walshe (2011) reported that religious expressions, "are used predominantly as pragmatic markers to express shock, excitement, surprise, impatience, anger, or just as fillers" (p. 137):

(29) Oh, **Holy Mother of God**, he's dead!

Oh, **God**. How on earth did that all get there? (Ibid.)

Another scholar that covered swearing in fictional representations of IrE is Terrazas-Calero (2020, 2022), whose examination of DPMs in contemporary Irish fictional prose analysed the use of the expletive *fucking* as an intensifier, among other features. This researcher found that

the use of this taboo word in the novels followed the pattern noted by Murphy (2009) of males using intensifying *fucking* more frequently than females (Terrazas-Calero, 2020, p. 266; 2022, p. 174). Terrazas-Calero also compared her findings to a sample from the Limerick Corpus of Irish English (LCIE), where the gender distribution in the use of *fucking* was similar, albeit with women accounting for a higher proportion of the occurrences (ibid., p. 175). Regarding the emotions expressed by intensifying *fucking*, Terrazas-Calero reported that most uses of this expletive were either positive or neutral, a finding that was repeated in the data from the LCIE. Taken together, the findings regarding gender and functional distribution of *fucking* point to the novels analysed by Terrazas-Calero being an authentic representation of the use of this expletive in real life, as it mirrors patterns to real-life use.

2.4.2.3 DPMs analysed in this thesis

This subsection will provide a summary of the existing research and insights gleaned by it on three DPMs present in the dialogue of *Derry Girls*. The first two, *ach/och* (section 2.4.2.3.1) and final *but* (section 2.4.2.3.2), have been chosen for being prominent in NIrE, the language variety employed by most characters in *Derry Girls*. The third item, *wee* (2.4.2.3.3), has been identified as a potential DPM in development (Kallen, 2005a), and thus it has been chosen due to its novelty and lack of research about it.

2.4.2.3.1 *Ach/Och*

There has been relatively little research on DPMs that focuses exclusively on Northern Irish English (NIrE). In her 2010 book about this language variety, Corrigan provides a table of discourse particles (her term) that have been used in NI (p. 98) but cautions that most of them have fallen out of use, with the notable exception of the pair *ach/och*, which remain in frequent use across all ages in NI (ibid.). *Ach* is an exclamation that expresses displeasure (30) (ibid., p. 86), in addition to being used to “signal that new information is to follow” (31) (Corrigan, 2015, p. 46), whereas *och*, also an interjection, can be translated as “alas!” (ibid., p. 45):

(30) JENNY: Think I'll just have to mull it over.

MICHELLE: **Ach**, Jenny, mull this over [*puts middle finger up at Jenny*] (*Derry Girls*, 1x01)

(31) GERRY: Do I want to know why you're reading a book on the Russian Revolution?

SARAH: **Ach**, it's just one of my wee interests. (*Derry Girls*, 3x05)

- (32) ERIN: He's not coming.
MARY: **Och**, love. (*Derry Girls*, 2x05)

Both DPMs, which are hypothesised to originate from Irish and/or Scots Gaelic (Corrigan, 2015, p. 45), occur at the beginning of utterances and are highly salient to outsiders (Corrigan, 2010, p. 99). Despite the distinctiveness of these two markers, Corrigan points out that they have received little scholarly attention, save for her own work (2010; 2015). In her 2015 study, with data sourced from sociolinguistic interviews with Northern Irish participants (p. 39), *ach* appeared to be most frequently employed by males (although the researcher notes that there were fewer female participants, which may have skewed the data), and among females, amongst the youngest group. It will be of interest to compare these figures with those of the SPICE-Ireland corpus and with the data from the *Derry Girls* transcripts.

Besides Corrigan's work, the other available study on *ach/och* is from Walshe (2016), who examined the presence of these two DPMs in a corpus of films set in Northern Ireland. When compared to a corpus of films set in the ROI, Walshe noted that *ach/och* were much more frequent in Northern Irish films (2016, p. 334).

On a final note, Corrigan notes that the orthographic distinction between *ach/och* is not always kept consistent by NIrE speakers, especially in IrE literature (2015, p. 45). This seems to be the case in the subtitles for *Derry Girls*, where Channel 4 and Netflix occasionally differ in their transcriptions of these two interjections when they appear in the dialogue. Therefore, considering *ach* and *och* together in the analysis of the data might be a necessity due to the difficulty in distinguishing between the two.

2.4.2.3.2 Final *but*

Another DPM seemingly autochthonous to NI is utterance-final *but*, which has received more scholarly interest than *ach/och* (Corrigan, 2015, p. 47). In contrast to its regular use as a conjunction that introduces a contradiction, final *but* functions as a retrospective focuser (pp. 28–9), roughly equivalent to ending a sentence with “though” (Walshe, 2009, p. 124):

- (33) SPEAKER 1: She didn't need it, no.
SPEAKER 2: She got cured, **but**. (Corrigan, 2015, p. 48)

Once more, in Corrigan's study, it is young women who exhibit the most frequent use of this marker, with middle-aged respondents being the least likely to use it, which might be a sign

that final *but* is subject to age grading⁷ (2015, p. 55). Contrary to Corrigan's claim, Walshe (2009) ascribed final *but* to Dublin English, although he acknowledged that other authors, namely Hickey and Harris, had associated the marker with NIrE (p. 124).

2.4.2.3.3 *Wee*

A particle that may be considered a DPM is *wee*, normally an adjective meaning "little", very common in NI (Amador-Moreno, 2010a, p. 62). Kallen (2005a) suggests that it can work as a minimiser, free of its usual reference to a small size (p. 138). In (34) below, we can see an example of *wee* in its normal adjectival function, whereas (35) presents an instance of *wee* being used as a minimiser to downtone the seriousness of the situation expressed in the utterance.

(34) JOE: You're even afraid of that **wee** girl. (*Derry Girls*, 1x05)

(35) OFFICER: We've just had a **wee** security alert there. I'm afraid we're gonna have to carry out a **wee** evacuation. (*Derry Girls*, 2x02)

Minimisers like *wee* can be considered one of the linguistic expressions of the pragmatic phenomenon of mitigation, understood as a reduction in the degree of illocutionary force of an utterance with the aim of reaching a communicative goal (Albelda Marco & Estellés Arguedas, 2021, p. 72; Villalba Ibáñez 2020, p. 68). Mitigation is related to other pragmatic concepts like vague language, hedges, and politeness, although the boundaries between all these are not always clearly defined (ibid.), and the attribution of one or another pragmatic meaning to a linguistic expression is usually context-dependent (ibid., p. 73). However, there seems to be a consensus among researchers that mitigation seeks to prevent damage to the speaker's positive face or self-image (Albelda Marco & Estellés, 2021, p. 77; Caffi, 1999, p. 881; Villalba Ibáñez, 2020, p. 71).

In order to distinguish mitigation from other pragmatic phenomena, Villalba Ibáñez (2020) proposes three tests:

- 1) An absence test, where the potentially mitigating element is eliminated from the utterance in order to assess if the pragmatic meaning changes.

⁷ A linguistic feature is affected by age grading when its use changes in frequency throughout a speaker's life, without the whole of the linguistic community changing. For example, young and old speakers might use a vernacular expression, but middle-aged people avoid it because it is considered informal, and they need to express themselves in a more formal register at work. When these middle-aged people retire, they might start to use the feature again once they no longer feel the pressure to employ a formal register (Tagliamonte, 2012, p. 47).

2) A commutation test, employed with potentially mitigating elements whose elimination would render the utterance ungrammatical. Here the element in question is substituted by another one from the same grammatical category in order to observe if the pragmatic meaning of the utterance changes.

3) A solidarity test that checks if the potentially mitigating element co-occurs with other mitigators. (Villalba Ibáñez, 2020, pp. 73–5)

As a relatively common word in the dialogue in *Derry Girls*, *wee* offers an intriguing avenue for exploration, as close analysis of its usage in interaction might provide some evidence of its pragmatic meaning or meanings, including its possible mitigating function.

2.5 Concluding remarks

This chapter has offered an overview of the theory and existing literature on the two main fields this thesis is concerned with: the study of fictional language, with an emphasis on telecinematic dialogue, and the study of the pragmatics of IrE.

First, we have explored the study of fictional language, both in the context of IrE and of telecinematic dialogue, identifying the – few – existing studies that focus on the intersection between those two contexts. The problem of the authenticity of fictional language has also been discussed, and methodological approaches to address this challenge have been mentioned. This subsection has also introduced the linguistic field of stylistics and the concept of characterisation, including an explanation of the two main approaches to this process of conceptualising characters and how they can be applied to the present study.

Secondly, and after an introduction to NIrE, the main variety represented in the dialogue in *Derry Girls*, this chapter has discussed the development of the study of IrE pragmatics. An outline of the pragmatic profile of IrE, as described by the existing literature on the matter, has been presented. This has been followed by a theoretical discussion regarding DPMs, the main focus of the linguistic analysis in this thesis, and a review of the available research about DPMs present in IrE, including studies that used fictional language as data. The review was divided into two parts. The first one was concerned with DPMs that, despite appearing in other varieties of English, have functions specific to IrE. The second part has covered the DPMs that will be analysed in this thesis. This subsection has addressed three DPMs that have been associated with NIrE.

3. METHODOLOGY AND DATA

This chapter covers the data, methodology and research design of the present thesis. The study's objective is to analyse the use of DPMs in the television series *Derry Girls* and how they contribute to the characterisation of seven of its characters. To provide a frame of reference, the dialogue in the show will be compared to the real-life, spoken language contained in the Northern Irish component of the SPICE-Ireland Corpus. Section 3.1 is devoted to data description: subsection 3.1.1 and its subdivision cover the data collection and corpus creation process, whereas subsection 3.1.2 provides a description of the comparison corpus. Section 3.2 details the quantitative (3.2.1) and qualitative (3.2.2) methods employed in answering the research questions. Section 3.3 discusses the general methodological approach adopted in the present study. Finally, a summary of the chapter is offered in section 3.4.

3.1 Data

The data analysed in this thesis are sourced from the dialogue in *Derry Girls*, which has been compiled into a corpus, as I will discuss in detail below in subsection 3.1.1. A corpus can be defined as “a finite body of text, sampled to be maximally representative of a particular variety of language, and which can be manipulated using a computer” (McEnery & Wilson, 2001, p. 73). Representativeness is a key feature of corpora, as the aim of researchers who employ them is to be able to make generalisations and establish theories about language use derived from the evidence gathered from a corpus (McIntyre & Walker, 2019, p. 3). For instance, if one wanted to study the speech of teenagers in Belfast, one would have to construct a corpus that contained speech from all the neighbourhoods in the Belfast urban area, with a balanced representation of male and female speakers, of different social classes and ethnicities found in the city, etc. In the case of *Derry Girls*, maximal representativity of the show's language can be achieved by including all the dialogue spoken in the show in the corpus.

However, it has been argued that a single text (e.g., *Derry Girls*) cannot constitute a corpus, as corpora should aim to represent a language variety (such as LDE, or NIRE more broadly, the variety employed in *Derry Girls*), and a single text cannot do that on its own (McIntyre & Walker, 2019, p. 4). In consequence, one would need to compile a variety of sources that employ the variety one wants to study. That being said, in the field of corpus stylistics, i.e., the discipline that encompasses the present thesis, studies have employed corpus

methods on a single text (for examples see McIntyre & Walker, 2019, p. 5). As far as the study of television dialogue is concerned, authors like Bednarek and Quaglio (as discussed in more detail in Chapter 2, section 2.2.2) have created corpora composed of the episodes of a single television series, which arguably constitute one text.

3.1.1 The *Derry Girls* corpus

This section describes the process of building the *Derry Girls* Corpus (DGC), including the steps taken in gathering the data, formatting, and annotating it, and making the partitions required to facilitate the analysis of the language contained in the corpus.

3.1.1.1 Data collection

Table 3.1. List of Derry Girls episodes contained in the DGC corpus

	Title	Code	First aired
Season 1	“Episode 1”	1x01	4 Jan 2018
	“Episode 2”	1x02	11 Jan 2018
	“Episode 3”	1x03	18 Jan 2018
	“Episode 4”	1x04	25 Jan 2018
	“Episode 5”	1x05	1 Feb 2018
	“Episode 6”	1x06	8 Feb 2018
Season 2	“Across the Barricade”	2x01	5 Mar 2019
	“Ms de Brún and the Child of Prague”	2x02	12 Mar 2019
	“The Concert”	2x03	19 Mar 2019
	“The Curse”	2x04	26 Mar 2019
	“The Prom”	2x05	2 Apr 2019
	“The President”	2x06	9 Apr 2019
Season 3	“The Night Before”	3x01	12 Apr 2022
	“The Affair”	3x02	19 Apr 2022
	“Stranger on a Train”	3x03	26 Apr 2022
	“The Haunting”	3x04	3 May 2022
	“The Reunion”	3x05	10 May 2022
	“Halloween”	3x06	17 May 2022
	“The Agreement”	3x07	18 May 2022

Following Quaglio’s (2009) example in his corpus study of the dialogue in the sitcom *Friends*, I decided to employ transcripts, not scripts, of the nineteen episodes of *Derry Girls* as the data for the DGC. Table 3.1 presents a list of the episodes and their corresponding codes, which are

used when presenting examples from the data. The reason for avoiding scripts is that their content may differ from the finished product, as actors can be given the freedom to adapt their lines and improvise (Walshe, 2011, p. 42).

The transcripts for the DGC were obtained from the website *Forever Dreaming*, which hosts transcripts of TV series and films submitted by fans. The website advises that the transcripts may not be verbatim, (“About”, *Forever Dreaming*, n.d.), as some contributors might choose wordings that differ from the exact lines an actor spoke. Besides the characters’ dialogue, the transcripts include descriptions of characters’ actions and gestures, song titles from the show’s soundtrack and the lyrics that can be heard onscreen.

After obtaining the transcripts, and to ensure their accuracy, I rewatched all *Derry Girls* episodes and amended the transcripts where they differed from what was said in the show. Non-verbal sounds were excluded from the dialogues, as well as instances of stammering (e.g., a character repeating the pronoun *I* at the beginning of a sentence was transcribed as “I”). All mentions of the show’s soundtrack (including song titles and lyrics) were removed. As spoken language lacks explicit punctuation, the division of the dialogue in sentences and phrases required the interpretation of the original transcribers (McEnery & Wilson, 2001, p. 44); accordingly, during the rewatching process, I changed some instances of punctuation in cases where I felt that the choices made by the transcriber did not accurately reflect the intonation patterns and length of pauses present in the spoken dialogue. This process replaced the original transcriber’s interpretation with mine, and it remains true that the sentence divisions in the transcripts are but an attempt to represent features of oral language in writing.

The Northern Irish English word for *children*, frequent in the show, was transcribed using the spelling *wean*, following Hickey’s (2007, p. 110) example. The contraction of *going to* was rendered as *gonna*, as it was the form that appeared in the closed captions for the episodes available on Netflix, as opposed to the Scots spelling *gonnae*. In a similar vein, the contraction of *can* and *not* was spelt *cannot* instead of *cannae* or *canny*. The abbreviated form of *because* was spelt *cos*, and the possessive adjective *my* was transcribed as either *my* or *me* in order to reflect its variable pronunciation by the actors. Dialectal forms of *old* were transcribed as *aul*, following Kallen and Kirk (2012). Similarly, *craic* was the preferred spelling for this word denoting “chat, enjoyment, fun, gossip, way of things” (ibid., p. 26). Instances of non-standard usages were transcribed as they were (e.g., *ain’t* was not replaced with *isn’t*). The descriptions of actions, gestures and soundtrack were deleted, as they do not constitute dialogue.

Furthermore, character names were added at the beginning of lines. Finally, dialogue lines corresponding to news broadcasts, either on TV or on the radio, were deleted, as they are not spoken by characters in the show and represent a different genre.

Each episode transcript was saved as a .docx file at first, but they were also combined into a single document, which was converted to plain text format (.txt with UTF-8 encoding) to ensure it would be readable by the software I used to analyse the data, *AntConc* (Anthony, 2022). This software provides a free analysis toolkit for texts. The version I employed, *AntConc* 3.5.9, enables the user to conduct concordance searches, create concordance plots and frequency-based word lists, carry out keyword and collocation searches, as well as produce collocation and n-gram lists. The software allows the user certain flexibility in establishing their own definition of a token (i.e., word), as well as in choosing statistical parameters for the different tools (e.g., the *p*-value when calculating which words constitute relevant keywords). Matters of quantitative methods will be explained in more detail in section 3.2.

3.1.1.2 The annotation system

The next stage in the creation of the DGC was annotating the data. Annotating is the process of adding “interpretative, linguistic information to a corpus” (Leech, 1997, p. 2), which can be done automatically, automatically with manual corrections, or entirely manually (McEnery & Hardie, 2012). As most currently available annotating software does not cover pragmatic features, it was decided that manual annotation would be the most appropriate method. In addition, manual annotation has the advantage of avoiding false positives, barring human error (Smith et al., 2008, p. 165).

Smith et al. (2008) distinguish two approaches to manual corpus annotation: the bottom-up approach, and the top-down approach. The former involves a “close, manual inspection of all the texts in a corpus, coding instances of linguistic features as they are encountered” (p. 164), usually drawing from a theoretical framework. In the latter, the top-down approach, the items to be annotated are retrieved via the concordancing function of the corpus analysis software employed by the researcher (Smith et al., 2008, p. 165). Owing to the open-ended nature of some of the pragmatic categories analysed in the present thesis (namely taboo language), it was decided that the bottom-up approach would be better suited to annotate the corpus, for trying to retrieve all instances of such types of linguistic items via searches would necessarily lead to overlooking many instances of them.

Manual annotation, however, is not without its drawbacks. The main disadvantage is the fallibility of the person carrying out the annotating, as errors can always occur despite careful revision of the work done (McEnery & Hardie, 2012, p. 30). And in the case of pragmatic features their annotation requires interpretation on the part of the researcher, which can influence the results (Landert et al., 2023, p. 15). On the other hand, having only one transcriber ensures consistency regarding the systematicity of the tagging process, whereas having more than one person carry out the task might lead to incongruencies in the interpretation of the tagged features.

Following Terrazas-Calero's (2022) example, the annotation process started by reading the transcripts and highlighting the relevant pragmatic features. To ensure no tokens had been overlooked, the transcripts were read five times. The annotating system itself was also inspired by the one employed by Terrazas-Calero (2022), who created tags that combined the linguistic category (pragmatics, vocabulary, grammar, etc.), the specific item (quotative, intensifier, DPM, etc.) and the position in the clause (initial, medial, etc.) of the annotated item. This researcher employed *eXtensible Markup Language* (XML) tags, a system that consists of start and end tags that bracket the tagged item:

(1) <Sureinitial>Sure</Sureinitial>, she'd put you to shame. (*Derry Girls*, 2x06)

These tags are used across numerous fields due to their flexibility and adaptability (McIntyre & Walker, 2019, p. 90), so they were adopted for the present thesis. However, the most recent version of the corpus software I used to analyse the data does not admit XML-based tags. In order to solve this issue, an earlier version of this programme was employed (Anthony, 2020).

The annotation system consists of two main tags:

1) An opening/closing tag that encodes information about the speaker.

E.g., <E> </E> for Erin, <MI> </MI> for Michelle.

2) A tag that indicates the element being coded together with its position in the sentence, where applicable.

E.g., <AOI> for the DPMs *ach/och* in sentence-initial position, and <AOM> for *ach/och* in sentence-medial position.

When combined they appear as in the following example:

(2) <MI> Ach <AOI></MI>, some of them are rides. (1x01)

Here <MI> and </MI> are the opening and closing tags, which indicate that the speaker is Michelle, and the tag <AOI> denotes that the tagged item is *ach/och* in sentence-initial position.

Table 3.2. *Tags employed in the annotation of the DGC*

Character tag	Character name
E	Erin
MI	Michelle
C	Clare
O	Orla
Q	Mary
S	Sarah
J	Joe
Category tag	Meaning (pragmatic item)
AO	<i>Ach/Och</i>
BF	Final <i>but</i>
W	Wee
META	Metapragmatic comments made by the characters
Position tag	Meaning
I	Sentence-initial
M	Medial position
F	Final position
A	Standalone element

Table 3.2 presents the meanings of the tags. Please note that the character tag list is non-exhaustive, due to space limitations; for a complete list of character tags, please refer to Appendix A. In addition to the information provided by the table, I would like to offer some clarifications regarding the choices made in creating the tags and their applications:

- 1) *Ach* and *och* were treated jointly for reasons explained in Chapter 2, section 2.4.2.3.1. And, as they were found to occur exclusively in sentence-initial position, it was not considered necessary to include this information in the tags.
- 2) It was decided to tag the metapragmatic comments found in the dialogue, as they could provide some insight into the pragmatic profile of NIRE as represented in the show.
- 3) Due to the fact that the word *wee* had a tendency to appear within the clause structure, its occurrences were not tagged with position tags; instead, instances with pragmatic meaning were tagged with <D>, and instances without it with <A>.

The tags presented in Table 3.2 are combined, as mentioned above, to provide information about the DPMs that are the object of this thesis. Consider the following example:

(3) Now, more bacon, anyone? Not a sausage? <J> Wee <WD></J> omelette, maybe? (1x02)

Here, the opening and closing tags (<J>, </J>) indicate that it is Joe speaking; the *W* in <WD> that the item is an instance of *wee*; and the *D*, that it has pragmatic meaning.

The advantage of this system is that it enables searches at two levels of specialisation. One can search “<AO” if one wants to retrieve all instances of *ach/och*, regardless of position, or for *ach/och* in a specific position (“<AOM” for *ach/och* in medial position, for example). However, in order to maximise readability, the tags in the extracts from the DGC used as examples in Chapter 4 will be omitted, and character names will be added at the beginning of utterances. For example, (3) above would be rendered as:

(4) JOE: Now, more bacon anyone? Not a sausage? Wee omelette, maybe? (1x02)

The same readability principle will be applied to the comparison corpus, SPICE-Ireland, discussed below in section 3.1.2.

3.1.1.3 Post-annotation steps

After the annotation of the corpus, a copy of it without speaker tags was created, so that these tags would not distort the results of the different queries performed on *AntConc*. Furthermore, following Bednarek’s (2010, 2012) example, it was decided to create several subcorpora consisting of all the dialogue pertaining to a specific character, as well as comparison corpora containing all the dialogue from the other characters. Six character subcorpora were created, using the following criteria:

1. The character (not the actor) had to be from NI. This excludes James, one of the main characters, due to his being English, as well as Gerry, Erin’s dad, due to his being from the ROI.
2. The character had to be present in all the episodes of the show, barring 3x05 due to its being centred on a school reunion that concerns the mothers of the main characters.

The application of these criteria resulted in the selection of the following characters:

1. Erin Quinn: a teenage girl from a working-class, Catholic family. Arguably the protagonist of the show, Erin is an aspiring writer, confident in her literary talent. She is also prone to get in trouble alongside her friends.
2. Michelle Mallon: one of Erin's close friends, also Catholic and working-class. Not too preoccupied with her academic life, Michelle is rebellious and transgressive, often the mastermind behind the girls' mischief.
3. Clare Devlin: another of the Derry Girls' titular characters, Clare is in many respects the diametral opposite of Michelle. Clare is an academically oriented, teenager who prefers to follow rules. She tags along with Erin and Michelle's harebrained schemes rather reluctantly and is often the voice of reason.
4. Orla McCool: Erin's cousin, also part of the friendship group. Orla is unselfconscious and childlike, often seeming to live in her own world. She seems happy to join in every of her friend's ideas, although she does not always understand the whole scope of what is going on.
5. Mary Quinn: Erin's mother. Strict but caring, the mistress of her household, Mary usually behaves sensibly, but she is not without a mischievous streak. She works as a shop assistant but is more commonly shown carrying her household duties, in a home with a traditional gender-role division.
6. Sarah McCool: Orla's single mother, sister to Mary, and Erin's aunt. She's kind but rather self-centred and vain. Sarah's ditzy nature provides a comedic contrast to Mary's more serious character, and both share several plotlines that exploit this dynamic.
7. Joe McCool: Father to Mary and Sarah, and grandfather to Erin and Orla, Joe is fiercely protective of his daughters and granddaughters, and very dismissive of their male relations.

Characters 1–4 represent a gamut of teenage personalities that can be compared to one another, while characters 5 and 6 portray two types of middle-aged women. Character 7 is the sole representative of both the male and the elderly demographics. The former group can also be compared to the latter. In contrast, significant cross-gender and cross-class comparisons are not possible due to all the selected characters being working-class females, and the lone male character also being working-class. *Derry Girls* contains few middle-class or upper-class characters, and none of them fit the selection criteria. In addition, the two most prominent male

characters in the show, Gerry (Erin’s dad) and James (Michelle’s cousin, and also part of the titular gang), are from the ROI and England, respectively.

Once the characters had been chosen, separate .txt files with their respective lines of dialogue were created, in addition to files containing all the remaining dialogue to enable comparisons of each character’s language to the language of all of the others in the show. In total, fourteen subcorpora were created in this fashion.

3.1.1.4 The *Derry Girls* Corpus: description

As indicated above, the DGC consists of the dialogue from all 19 episodes of *Derry Girls*, totalling 65,498 words. DGC is only available in .txt format and is annotated for selected pragmatic features. DGC is not marked for paralinguistic features, intonation, lengths of pauses, emphatic stress or overlap of character lines. DGC is a specialised corpus of transcribed dialogue and is maximally representative of the television series *Derry Girls*, but not representative of television dialogue in general. The dialogue contained in the DGC constitutes a fictional portrayal of NIrE, and of LDE in particular, but other varieties, namely English English and English from the ROI, also appear in it.

The character subcorpora made from the DGC are summarised in Table 3.3. As mentioned above, all of these characters’ dialogue is a portrayal of NIrE.

Table 3.3. DGC character subcorpora

Character subcorpora		Remaining dialogue subcorpora	
Erin	10,307 words	No Erin	55,191 words
Michelle	8,135 words	No Michelle	57,363 words
Clare	5,414 words	No Clare	60,084 words
Orla	2,184 words	No Orla	63,314 words
Mary	6,582 words	No Mary	58,916 words
Sarah	4,272 words	No Sarah	61,226 words
Joe	3,458 words	No Joe	62,040 words

In addition to these subcorpora, partitions of the data along gender and age lines were created, in order to be able to compare different demographics beyond these seven characters, not just within the DGC, but also with the demographics represented in SPICE-Ireland, which are described below. Table 3.4 below presents said partitions. The reason for creating partitions

with and without the non-NI characters was to enable comparisons between these two groups when analysing features that are heavily associated with NIrE, such as *ach/och*.

Table 3.4. *DGC demographic partitions*

Partition name		Description
Female	47,186 words	Dialogue from all the female characters
Male	18,194 words	Dialogue from all the male characters
DGC-NI	53,106 words	Dialogue from all the Northern Irish characters
Female NI	43,498 words	Dialogue from all the female characters from NI
Male NI	9,608 words	Dialogue from all the male characters from NI
Female non-NI	3,688 words	Dialogue from all the female characters who are not from NI
Male non-NI	8,586 words	Dialogue from all the male characters who are not from NI
Young female (YF)	28,266 words	Dialogue from all the female characters under 25 years old from NI
Adult female (AF)	14,516 words	Dialogue from all the female characters aged 25–50 from NI
Old female (OF)	670 words	Dialogue from all the female characters aged 51 and over from NI
Young male (YM)	693 words	Dialogue from all the male characters under 25 years old from NI
Adult male (AM)	3780 words	Dialogue from all the male characters aged 25–50 from NI
Old male (OM)	5186 words	Dialogue from all the male characters aged 51 and over from NI

Note: The ages are approximate, as most characters' ages are not mentioned in the show.

3.1.1.5 Strengths and limitations of the process

The advantages of creating a corpus, and its subcorpora, tailored to the research questions that guide this thesis are twofold. First, the corpus-creating process enabled me to become closely acquainted with the data, as it involved reading all of the dialogue lines repeatedly. Secondly, the creation of the DGC, with its bespoke annotation for the DPMs I am interested in examining, ensures that I have the exact data I need to answer questions about characterisation and the use of specific pragmatic features in the show's portrayal of NIrE.

Resorting to building one's own corpus is not without its disadvantages, however. The main drawback of this process was the amount of time and work devoted to it, as most of the tasks were done manually, which also left room for human errors, as mentioned in section

3.1.1.2. Secondly, my aforementioned familiarity with the data might have created impressions about it that might unwittingly bias my analysis. Another weakness of the DGC, which affects the demographic partitions, is the fact that not all gender and age groups are equally represented. The two groups that are most affected by this are young males and old females, both of which have fewer than 700 words of dialogue to represent them, which risks making comparisons between these and other groups meaningless. Lastly, the fact that this corpus is comprised of copyrighted data imposes a limitation on the reproducibility of the results, as the DGC cannot be made available to other researchers to use in their own projects or to verify my results. In addition, the results cannot be replicated⁹ reliably either: if another researcher were to attempt to recreate the DGC with the system of annotation that I have described in section 3.1.1.2, there is no guarantee that they would apply said system in the exact same way I did, which could potentially lead to different results.

Finally, and as touched upon in Chapter 2, section 2.2, there remains the matter of the corpus's representativeness and authenticity with regard to LDE, and NlRE more broadly. It must be remembered that the speech contained in the DGC is a fictional construction of LDE, not real-life spoken language, hence the need to contrast it with a corpus comprised of real speech.

Having described in detail the fictional data that will be examined in the present thesis, as well as the collection and preparation process that was undertaken to be able to work with them, we will now turn to the real-life corpus that will be used for comparison purposes.

3.1.2 The corpus used to compare and contrast the DGC data

As noted by Amador-Moreno and Terrazas-Calero (2022), in order to make comparisons between fictional data and real-life spoken data, one needs to use “a baseline corpus of contemporary spoken language” (p. 520). Other scholars who have worked with literary dialect have recognised this need and, consequently, employed a variety of real-life corpora in their work, such as (example). With regard to researchers who have examined the representation of IrE in fiction, one can cite the example of Walshe (2009), who used the ICE-Ireland corpus and Terrazas-Calero (2022), who employed the Limerick Corpus of Irish English and the British National Corpus.

⁹ As explained by McEnery and Hardie (2012), “a result is replicable if a reapplication of the methods that led to it consistently produces the same result” (p. 16).

The corpus chosen as a comparison for the present thesis is SPICE-Ireland (Kallen & Kirk, 2012), which stands for “Systems of Pragmatic annotation in the spoken component of the International Corpus of English – Ireland.” This corpus contains 626,597 words from a variety of spoken genres such as broadcast news, business transactions, scripted speeches, parliamentary debates, face-to-face conversations or classroom discussions (Kallen & Kirk, 2012, p. 9). The texts that make up the data were collected between 1990 and 2005. More importantly, SPICE-Ireland is divided into two components: one for the ROI, which encompasses 314,309 words worth of spoken data; and one for NI, which totals 312,288 words (ibid.). This division is relevant for the present thesis, for it allows me to focus only on the NI data to draw comparisons with the *Derry Girls* dialogue.

As can be deduced from the meaning of its acronym, SPICE-Ireland is annotated for pragmatic features, including DPMs. In addition, the data are annotated for non-linguistic behaviours, such as yawns, and overlaps and pauses are also marked. Among the DPMs that have been annotated in the SPICE-Ireland data are *sure*, *like*, *ach/och* and religious expressions such as *God*, *oh God*, or *oh my God*. The only features examined in this thesis that have been overlooked in the SPICE-Ireland annotation are the adjective *wee* and the taboo language that is not religious in nature.

Besides the fact that the NI data were clearly identified in the corpus files and that the corpus has been annotated for pragmatic features, my reason for choosing SPICE-Ireland as my comparison corpus is that the texts that constitute it have been tagged according to their genre. Given that most character interactions in *Derry Girls* portray informal conversations between friends and family members, I have decided to focus on the data from texts tagged as face-to-face conversations in SPICE-Ireland, which for the NI component of this corpus total 91,929 words. In addition, more formal discourse situations could potentially lack the pragmatic features that are of interest to the present study (Terrazas-Calero, 2022, p. 95), as they are associated with more informal registers.

Owing to the fact that three of the DPMs which constitute the focus of this thesis have been associated specifically with NIrE in the existing scholarly literature, the ROI component of the face-to-face conversation files in SPICE-Ireland will also be used when discussing those linguistic items (namely, *ach/och*, final *but* and *wee*). The purpose is to verify that these indeed occur more frequently in the NI component than in the ROI one. In order to distinguish them, the NI component will be referred to as *SPICE-NI* and the ROI one as *SPICE-ROI*.

To enable comparisons between different demographic groups, copies were made of the face-to-face conversation files. The data contained in these were then divided according to gender and age group. This was facilitated by the comprehensive speaker data provided by Kallen and Kirk (2012, p. 63). The result was eight partitions of the face-to-face conversation files sourced from NI speakers. The partitions are listed in Table 3.5:

Table 3.5. *Partitions of the SPICE-NI data*

Partition name	Number of words
Female	65,735
Male	24,344
Female young	17,342
Female adult	22,670
Female old	16,641
Male young	13,026
Male adult	4,206
Male old	1,685

It must be noted that the word counts of these partitions, and any measurements derived from them (e.g., normalised frequencies of particular words), are not comparable to those in the DGC. This is due to the format differences in the tagging systems, for *AntConc* overlooks some of the tagged words in the SPICE-Ireland data. The only word counts that are comparable to those in the DGC are the totals for the NI and ROI components of the face-to-face conversation data, as they are provided by the corpus creators (Kallen & Kirk, 2012, p. 9). That being said, the partition word counts remain comparable to each other, as they are all tagged in the same way and thus parsed equally by *AntConc*.

Table 3.6. *Age groups in SPICE-Ireland and their condensed versions*

Age groups in SPICE-Ireland	Condensed age groups
Group 0 (0–18 yo)	Young
Group 1 (19–25 yo)	
Group 2 (26–33 yo)	Adult
Group 3 (34–41 yo)	
Group 4 (42–49 yo)	
Group 5 (50+ yo)	Old

The age groups created for the partitions are a condensed version of the ones employed in SPICE-Ireland (Kallen & Kirk, 2012, p. 60). This simplification was motivated by the need for

the SPICE-NI data to be divided along the same lines as the data in the DGC. As not all the speakers whose language is represented in SPICE-Ireland shared their details with the corpus creators, not all the data could be included in the partitions.

3.1.2.1 Comparability between the DGC and SPICE-Ireland

With regards to the comparability of the section of SPICE-Ireland that I have chosen and the DGC, first one must mention that the informants who provided the data for the former were required to have finished second-level education (Amador-Moreno, 2010a, p. 140). This contrasts with the main characters whose speech is represented in the *DGC*, i.e., the titular *Derry Girls*, who have yet to complete their second-level studies. Secondly, the data in SPICE-Ireland are more than a decade older than that in DGC; even though the show is set in the 1990s, it was broadcast between 2018 and 2022 and one cannot assume that the dialogue represents faithfully how teenagers spoke in 1990s Derry.

Turning to more linguistic matters of comparability, the biggest drawback posed by using SPICE-Ireland as a comparison corpus is the differences between its annotation scheme and the one I devised for the DGC. For instance, we must consider the fact that the annotators have used a different definition of what constitutes a DPM than the one I have used. They considered DPMs as elements that “comment on the speaker’s attitude towards the core illocution” (Kirk, 2015, p. 92). In addition, items in a clause-peripheral position are not tagged as DPMs in this corpus; only when they appear attached to a longer utterance have they been considered as DPMs (Kallen & Kirk, 2012, p. 42), whereas I have tagged them as such in the DGC. Finally, the prosodic annotation system applied to the SPICE-Ireland data (Kallen & Kirk, 2012, pp. 35–41), which involved adding numbers to the beginning of words (e.g., *1FIErce*, *2MEAdow*), makes concordance searches more difficult, as searching for “ach” in *AntConc*, for example, would not produce any occurrences of the item that are prefixed with a number. As a result, any searches conducted on the SPICE-Ireland data require attaching a wildcard marker (*) at the beginning of the search term, and the subsequent sifting of the results to discard false positives (e.g., “each” appearing in searches of “*ach”).

As far as the corpus software used to access the corpora is concerned, the fact that SPICE-Ireland is provided in .txt format enables me to open it in *AntConc*. Much like with the DGC, the angular brackets used in the SPICE-Ireland annotation system entail that I will have to work with an earlier version of the software – *AntConc 3.5.9* (Anthony, 2020) –, instead of the most recent one.

Despite the present limitations, SPICE-Ireland remains the best choice of a comparison corpus for the present thesis, and it is no easy feat to locate alternatives that cover such a specific variety of English as NIrE, so one must do the best with the available resources.

3.1.3 Legal and ethical considerations

One of the main ethical problems posed by the building of a corpus is the obtention of permission from copyright holders (Kennedy, 1998, p. 76, as cited in Baker, 2006 p. 37). That being said, the UK's Intellectual Property Office (2014) establishes an exception to copyright that circumvents the need for the owner's permission, as it "allows researchers to make copies of any copyrighted material for the purpose of computational analysis if they already have the right to read the work (that is, they have 'lawful access' to the work)", always for non-commercial purposes. As all seasons of *Derry Girls* are available on Netflix as of October 2022, and I have a subscription to this streaming service, I have lawful access to the material, so it can be argued that this study falls under the exception to copyright above mentioned and thus permission would not be necessary in this case. As far as access to SPICE-Ireland is concerned, I obtained permission to download said corpus by applying for an academic licence through the University of Zurich's website.

Another ethical challenge regarding spoken language data collection concerns the privacy of the respondents whose speech is recorded. This does not affect the DGC, owing to the fact that the speakers represented in it are fictional characters, but it does apply to SPICE-Ireland. Indeed, McEnery and Hardie (2012, p. 62) stress the importance of protecting the privacy not only of the respondents but also of those spoken *about* in the data. Appropriately, the personal details of both groups have been anonymised in SPICE-Ireland, including personal names, addresses and workplaces (Kallen & Kirk, 2012, p. 14).

The last ethical consideration that I would like to address in this section involves the potential public repercussions of this study. Since this thesis focuses on NIrE, an underrepresented variety of English, the discussion about this dialect presented here might be one of the few occasions a potential reader encounters this variety in any form, thus having the potential to shape their perception of it. The fact that I am focusing on dialectal features might predispose non-linguist readers towards a negative view of NIrE. This concern is founded on the idea that, outside linguistic circles, dialectal variation can be regarded as a mere patois devoid of intrinsic value (Cameron, 1995, p. 41). As McEnery and Hardie (2012) note, a researcher might have limited power to correct the misinterpretations of one's academic work,

but they ought to try “as a part of the more general responsibility of academia to the society that supports it” (p. 67).

3.2 Methods

Having described the corpus that will be the object of analysis, the DGC, as well as the real-life spoken corpus that will be used as a reference, SPICE-Ireland, we now turn to the methods of analysis that will be deployed in this study in order to answer the research questions. To this end, it might be useful to reiterate said questions here:

1. How frequently do the DPMs *ach/och*, final *but* and *wee*, appear in the show’s dialogue and what functions do they realise?
2. How do the frequency and use of these features compare to those in the ICE-NI corpus?
3. How does the use of these DPMs contribute to the characterisation of the characters in the show?

Matters of frequency, as mentioned in research questions 1 and 2, will necessitate a mostly quantitative approach, whereas examining the use of the DPMs and how they affect the characterisation of the characters in the show will be a more qualitative endeavour; however, all questions will require a mix of methods, in order to both exploit their strengths and compensate their weaknesses. The following two subsections, accordingly, contain descriptions of the quantitative (subsection 3.2.1) and qualitative (subsection 3.2.2) techniques employed in the analysis.

3.2.1 Quantitative methods

Let us begin with the quantitative methods, which, as their name implies, rely on tallying data and performing statistical tests on them in order to explain the observed phenomena (McEnery & Wilson, 2001, p. 76). A benefit of these methods of analysis is that the findings obtained through them can be generalised to larger samples. In addition, they allow for comparisons to be drawn between datasets “as long as valid sampling and significance techniques have been employed” (ibid.). These significance techniques enable the researchers to distinguish between – statistically – relevant and irrelevant results, as will be explained below, although perhaps at the expense of sacrificing nuance in the data and sidelining interesting, but rare, cases (ibid., p. 77).

3.2.1.1 Frequency counts

Frequency counts are a basic statistical measure that consists of “a simple tallying of the number of instances of something that occur in a corpus” (McEnery & Hardie, 2012, p. 49). This measure constitutes an example of descriptive statistics, i.e., statistics that simply describe the data instead of testing for their significance. Frequency counts can be expressed as a simple number or as a percentage of the whole corpus (ibid.), but the most common way of conveying the frequency of an item is with the *normalised frequency*, which expresses how many instances of said item can be found in every x words (ibid.). The value of x can be adjusted in proportion to the size of the corpus: in large corpora, one may express normalised frequencies in instances per million words, for instance. In this thesis, since both the DGC and the portion of SPICE-Ireland that I will use as a comparison corpus are under 100,000 words, I will express normalised frequencies in occurrences per 10,000 words. The main advantage of using normalised frequencies is that it enables comparisons between corpora of different sizes, which is relevant for the present thesis.

In order to answer research questions 1 and 2, frequency counts were employed extensively, and comparisons across demographic and character groups were carried out where there was sufficient evidence. Frequency counts were also employed to aid in answering research question 3. Lists of the twenty most frequent words in each of the seven studied characters’ dialogue were created, and also for the DGC as a whole, the latter acting as a baseline. The most common words in any corpus tend to be grammatical words such as *the* or *to*, so this test aimed to check if any DPMs, taboo expressions or character names were among each character’s most used words. Character names (or even terms of address) could aid in mapping the relationships between characters, while the presence of taboo language could bring insight into a character’s attitude towards social norms. Finally, the study of a character’s most frequently used DPMs could help gauge their adherence to typical NIrE usages of these linguistic items, including in terms of gender and age groups.

3.2.1.2 Significance testing

Much as frequency counts can be a useful metric, it is not easy to assess if the difference in frequency of use of a linguistic item across corpora or subcorpora is a product of chance or if it is meaningful. In order to address this difficulty, researchers employ significance tests, i.e., statistical tools devised to test the level of confidence one can have that the observed difference is dependable (Gabrielatos, 2018, p. 230). There are several significance tests available to

scholars, although an exhaustive explanation of their workings is beyond the scope of this thesis. A very common test is the chi-square test, but, as noted by McEnery and Wilson (2001) and McEnery and Hardie (2012), this test is unreliable with very small frequencies, and, given the small size of the DGC, it seems unsuitable for the present thesis. The log-likelihood (LL) test (Dunning, 1993) has been proposed as a better alternative for smaller datasets. This test also has the advantage of not assuming that the data are in a normal distribution. A normal distribution involves the data clustering around a mean value, resulting in a bell curve graph when represented visually (*ibid.*, p. 51). Language data do not normally follow this distribution, hence why choosing a test that does not require it is crucial (McEnery & Hardie, 2012, p. 52).

Once the appropriate significance test has been selected – in the case of this thesis, the log-likelihood test – there remains the question of choosing a cut-off point, known in statistics as an *alpha level* (McIntyre & Walker, 2022, p. 145), for what one considers statistically significant. In statistics, the standard alpha level is 0.05 (McEnery & Hardie, 2012, p. 52). This cut-off point means that there is a 5% likelihood that a researcher would get the same or a larger frequency difference when no such difference exists in reality (Gabrielatos, 2018, p. 231). This likelihood is customarily expressed via the *p*-value (*p* stands for “probability”). For example, the *p*-value for the standard alpha level of 0.05 is $p < 0.05$ (McIntyre & Walker, 2022, p. 144). However, in linguistic research, this threshold is usually lowered to at least $p < 0.01$ (*ibid.*, p. 158), meaning that results are considered significant when they have at least a 99% probability of not being due to chance. This was the cut-off point selected for the present study. In order to interpret the result of a log-likelihood test, the researcher needs to consult a table of critical values, which relates log-likelihood test values to different *p*-values (see McIntyre & Walker, 2022, 157-8 for a more detailed explanation of the process). For my chosen *p*-value of 0.01, the critical value is 6.63, which means that if the result of the log-likelihood test is higher than 6.63, the result can be considered statistically significant (*ibid.*).

The LL significance test will be applied to pairs of observed frequencies whose comparison is relevant to the analysis. For example, the frequency of each studied DPM in both DGC-NI and SPICE-NI will be compared in order to ascertain if the usage of these linguistic items in *Derry Girls* conforms to real-life speech uses. The test results will be reported in section 4.1 below. The tool used to perform the tests is an online calculator developed by Rayson (2008), which enables word frequency comparisons between two (sub)corpora.

3.2.1.3 Keyword analysis

Keyword – or keyness – analysis is a quantitative method that can bring insight into the linguistic characterisation of each of the *Derry Girls* characters covered in the present thesis. A keyword is a term that appears more frequently in a corpus (named *target corpus*) when compared to a reference corpus, which is usually larger than the target corpus (McEnery & Hardie, 2012, p. 245). Keyword analyses can also yield *negative keywords*, understood as words that appear less frequently in the target corpus than in the reference corpus. For a word to be considered a keyword, the difference in frequency between the target and the reference corpus must be statistically significant (*ibid.*), i.e., the difference in frequency must not be due to chance. The corpus-parsing software *AntConc 3.5.9* offers both the chi-square and the LL tests in the parameters of its keyword search tool. Owing to the reasons explained in the previous section, the LL test was selected, with the alpha level set at 0.01.

Something a researcher must bear in mind when conducting keyword searches is that, even though the results are arrived at via objective mathematical equations, a number of subjective decisions have been made in the process, regarding alpha levels, the keywords that are considered of interest (such as DPMs in the case of this thesis) and the very corpus employed as a comparison (Gabrielatos, 2018, p. 253). Therefore, one cannot say that a quantitative analysis is an objective process, and all the subjective decisions must be adequately explained (*ibid.*).

For each of the seven analysed characters, a keyword search was conducted in *AntConc*, using each character's dialogue as a target corpus and the remainder of the series' dialogue as a comparison corpus. Then the target corpora and their respective comparison corpora were swapped, and the keyword search was repeated to detect any other statistically relevant terms. The results were grouped according to their category, with special attention paid to terms of address, as they can index relationships between characters, taboo expressions, and DPMs. Each of these groups was then analysed qualitatively, as will be explained in section 3.2.2.3.

Following these keyword searches, it was decided to compare the characters against each other, in order to uncover individual differences that could provide insight on the dynamics of specific character pairs. The results were cross-tabulated for ease of comparison and the new keywords that appeared were analysed qualitatively. That being said, the more specific the keyword searches were, and the smaller the comparison corpora were, the less useful results were obtained. In particular, using the Orla subcorpus, the smallest of the character subcorpora,

as a comparison corpus, yielded the lowest number of keywords, and none of them constituted new information.

3.2.2 Qualitative methods

In contrast to the more numerical approach of quantitative methods, in qualitative ones, the data are used as a source of instances of language phenomena, in order to identify and describe them (McEnery & Wilson, 2001, p. 76). Qualitative approaches have the advantages of affording the same attention to all examined items regardless of frequency and allowing for fine distinctions to be made as it is not necessary to impose a limited number of classifications on the data (ibid.). In the subsections below, I will proceed to explain the main *AntConc* tool employed in the qualitative analyses (3.2.2.1), followed by a description of the analyses undertaken to answer research questions 1 and 2 (3.2.2.2), and research question 3 (3.2.2.3).

3.2.2.1 Concordance searches

As far as qualitative analyses of corpus data are concerned, concordancers are probably the most useful tool to visualise the occurrences of the phenomenon one is studying. A concordancer enables a researcher to “search a corpus and retrieve from it a specific sequence of characters from any length” (McEnery & Hardie, 2012, p. 35), such as a word or phrase, which is then displayed onscreen “as an output where the context before and after each example can be clearly seen” (ibid.). This is where the utility of a concordancer lies for the analyst, as seeing the linguistic item they are interested in studying in their context makes the close reading and interpretation of the data possible. In the case of discourse-level features, such as the DPMs that constitute the focus of this thesis, attempting to discern the functions they perform would be impossible without examining them in the context of the utterance they accompany, and often, without the wider context of the interaction of which the utterance forms part.

3.2.2.2 Answering questions 1 and 2: DPM functions in the DGC and SPICE-Ireland

To answer questions 1 and 2, I referred to the existing literature on DPMs present in NIRE, discussed in the previous chapter, in order to make a classification of the DPM occurrences in the DGC according to their pragmatic function. Three challenges made themselves apparent almost immediately: (1) in order to classify the function of a DPM it was not always enough to rely on the written transcription of the dialogue, as non-verbal communication and the scene context provided essential information that aided interpretation; (2) the functions of the studied

DPMS reported in the existing literature did not constitute an exhaustive classification, and thus I encountered occurrences that did not fit into the premade categories, thereby requiring the creation of new categories; and (3) many occurrences were found to be multifunctional, which, while enriching the qualitative interpretation of the data, posed a difficulty in terms of quantifying the percentage of each function within the total uses of a given DPM.

These challenges only increased when analysing the DPM occurrences in the comparison corpus, SPICE-Ireland: unlike with the DGC, I did not have access to the original recordings, and thus my close analysis relied solely on the transcription and annotation of the conversations. Therefore, determining the emotions conveyed by the respondents' words was difficult, often impossible, especially if one considers the possibility of the speakers being ironic. In order to avoid misrepresenting the data, I noted where I could not ascertain the evaluative stance of the speaker.

These challenges, however, had a paradoxical benefit, as the realisation that the functions exhibited by the DPMs in the data were more varied than what the reviewed literature had primed me to expect prevented me from trying to shoehorn the DPM occurrences into categories into which they did not fit, as well as pay closer attention to the expressive nuances present in the dialogue.

3.2.2.3 Answering question 3: Linguistic characterisation in the DGC

Let us now turn to matters of linguistic characterisation. As discussed in section 3.2.1.3, in order to retrieve relevant data that would enable a qualitative analysis of the linguistic characterisation of the seven studied characters, frequency counts and keyword searches were conducted and the results that were relevant to the present thesis, namely DPMs and taboo language, were selected for analysis. Terms of address, such as names and terms of kinship and endearment (e.g., *aunt*, *love*, *son*) were also considered on the basis that they can index interpersonal relationships (McCarthy & O'Keeffe, 2003, p. 154).

Using the concordancing tool, the occurrences of the key terms were retrieved and examined together in order to uncover any patterns of use that may inform the characterisation of each character and their relationships to each other. As with research questions 1 and 2, I referred back to the existing literature to interpret the characters' usage of the DPMs that surfaced in the keyword searches, particularly with regards to which usages are characteristic of which demographic groups, in order to ascertain whether the characters conform to them or

not. Close reading was applied to the occurrences yielded by the concordancing tool, particularly when interpreting characters' attitudes to each other.

3.3 This thesis' approach to corpus linguistics

Before concluding this chapter, I would like to discuss the general methodological approach adopted in this thesis. Corpus linguistics is not a unified field. Tognini-Bonelli (2001) argued that one can identify two main approaches to this discipline: a corpus-driven approach, and a corpus-based one. Studies that adopt the latter approach tend to employ corpus data in order to answer research questions or to test hypotheses that have usually been arrived at by examining the available literature on a particular issue (McEnery & Hardie, 2012, p. 6). In contrast, practitioners of corpus-driven linguistics posit that any theories and hypotheses that the researcher wishes to validate should emerge from the corpus itself, instead of being formulated *a priori* (ibid.). An approach to corpus linguistics particular to the field of stylistics is *corpus-informed stylistics* (McIntyre & Walker, 2022, p. 26) already discussed in Chapter 2, section 2.2.3, which is the approach adopted by this thesis.

Despite the open-ended nature of the three research questions, my corpus-informed stylistic approach is closer to a corpus-based approach than to a corpus-driven one, as I referred to the literature that I reviewed in Chapter 2 in order to classify the DPM occurrences according to their pragmatic function, as explained in more detail in subsection 3.2.2.2 above. Moreover, in analysing the linguistic characterisation of the characters I linked their speech patterns to those reported in prior studies as pertaining to certain demographic groups, as discussed in subsection 3.2.2.3. However, the existing classifications were at points slightly constraining and not fully appropriate, as unexpected patterns of usage emerged in the analysis, which necessitated adopting a more hybrid standpoint, somewhere between corpus-driven and corpus-based, if arguably closer to the latter.

3.4 Concluding remarks

This chapter has described the steps taken in building the *Derry Girls* Corpus, which provides the data for the present thesis. The description included the data collection process, the system employed to annotate the data and the different corpus partitions made to enable cross-character comparisons. The strengths and limitations of the process were also considered. Subsequently, the chapter discussed the comparison corpus, SPICE-Ireland, which constitutes the real-life

speech baseline against which the language in the DGC can be considered. Next, the legal and ethical issues related to the use of these two corpora were examined.

The second section of the chapter focused on the methods employed in the analysis of the data. Quantitative methods were discussed first, namely frequency counts and keyword searches, followed by qualitative methods. The advantages and disadvantages of the methods were also considered, as well as the challenges encountered in their application to the data. Finally, I discussed the general methodological approach this thesis takes. The next chapter will now present the results obtained by employing the methods described above, as well as their analysis and interpretation.

4. RESULTS AND DISCUSSION

This chapter is divided into two main sections. The first one (4.1) is devoted to answering research questions 1 and 2, and thus in it I examine the frequency of the studied DPMs in the DGC corpus, their functions, compare them with real-life, spoken NIrE as represented in the SPICE-NI, and the SPICE-ROI corpus where applicable. The second section (4.2) is devoted to addressing research question 3, that is, to analysing how these pragmatic elements contribute to the characterisation of the seven characters from the show that were described in the previous chapter. In addition to the study of their pragmatic profiles, a keyword analysis of their dialogue will be conducted. Finally, the chapter concludes with a summary of the main findings (4.3).

4.1 DPMs in the DGC corpus

This subsection discusses the use, i.e., frequency and functions, of the DPMs *ach/och*, *final but*, and *wee* in the DGC, and compares them to those in SPICE-NI.

4.1.1 *Ach/och*

The first two DPMs to be analysed in the present chapter are *ach* and *och*. The following two subsections explore their frequency of use in the DGC and SPICE-NI, drawing comparisons between characters and between real-life speakers (4.1.1.1), as well as analysing the functions these two DPMs adopt in both corpora (4.1.1.2).

4.1.1.1 Frequencies of use

Table 4.1 below presents the raw and normalised frequencies of *och* and *ach* in the DGC and SPICE-Ireland for all the main characters, and by gender and age groups, whereas Table 4.2 shows the results of the *ach/och* frequency comparisons between selected pairs of (sub)corpora.

Taking *ach* and *och* together, the character whose dialogue contained the most occurrences per 10,000 words of these DPMs was Sarah, followed by Mary, although the frequency difference between these two characters was not found to be statistically significant, according to the results presented in Table 4.2 below. Individually, *ach* was employed most frequently by Sarah, and *och* by Orla. Among the titular Derry Girls (Erin, Michelle, Orla and Clare), Michelle employed *ach* the most frequently.

Table 4.1. Frequencies for *ach/och* in *DGC* and *SPICE-Ireland*
RF= raw frequency, *NF*= normalised frequency (per 10,000 words).

DGC						
Character/speaker	Ach		Och		Ach and Och	
	RF	NF	RF	NF	RF	NF
All	60	9.16	12	1.83	72	10.99
Erin	1	0.97	4	3.87	5	4.83
Michelle	10	12.29	1	1.23	11	13.52
Clare	0	0	1	1.86	1	1.86
Orla	1	4.58	1	4.58	2	9.16
Mary	18	27.35	3	4.56	21	31.91
Sarah	18	42.13	1	2.34	19	44.48
Joe	2	5.78	1	2.89	3	8.68
NI	60	11.30	12	2.26	72	12.66
NI Female	53	12.18	11	2.53	64	14.71
NI Male	7	7.29	1	1.04	8	8.33
Non-NI female	0	0	0	0	0	0
Non-NI male	0	0	0	0	0	0
YF	12	4.24	7	2.47	19	6.71
AF	41	28.24	4	9.64	45	31.12
OF	0	0	0	0	0	0
YM	0	0	0	0	0	0
AM	2	5.31	0	0	2	5.31
OM	5	9.64	1	1.93	6	11.57
SPICE-NI						
All	4	0.43	70	7.61	74	8.05
Female	3	0.48	57	8.70	60	9.15
Male	1	0.41	13	5.36	14	5.77
YF	3	1.40	10	4.67	13	6.08
AF	0	0	27	8.07	27	8.07
OF	0	0	13	5.34	13	5.34
YM	0	0	8	3.96	8	3.96
AM	0	0	1	1.66	1	1.66
OM	0	0	0	0	0	0
SPICE-ROI						
All	2	0.21	3	0.31	5	0.52

Broadening the scope to all the characters in the show, it can be seen that only Northern Irish characters used these two terms, with characters from other regions never uttering them. This is consistent with the literature that describes *ach* and *och* as typically NIrE DPMs (Corrigan, 2010; Walshe, 2016). In examining the data for each gender, it becomes apparent that female NI characters used these two DPMs with a higher frequency than their male counterparts,

although the difference was not deemed statistically significant according to the LL test result (see Table 4.2 below).

By gender and age group, it was adult female NI characters that led the use of *ach* and *och* in terms of their frequency per 10,000 words, and the difference between this group and all the other ones in DGC-NI was found to be significant.

Table 4.2. Log-likelihood tests for *ach/och*
NF = normalised frequency (per 10,000 words)

Target corpus	Comparison corpus	NF ¹⁰ in TC	NF in CC	LL
DGC-NI	SPICE-NI	12.66	8.05	9.78
AF (DGC)	AM (DGC)	31.12	5.31	10.59
AF (DGC)	All other groups (DGC)	31.12	7.01	38.65
SARAH	NO SARAH	44.48	8.66	27.79
MARY	SARAH	31.91	44.48	1.09
NI female (DGC)	NI male (DGC)	14.71	8.33	2.67
Female (SPICE-NI)	Male (SPICE-NI)	9.15	5.77	2.66
SPICE-NI	SPICE-ROI	8.05	0.52	74.04
DGC-NI (<i>ach</i>)	SPICE-NI (<i>ach</i>)	9.16	0.43	94.28
DGC-NI (<i>och</i>)	SPICE-NI (<i>och</i>)	1.83	7.61	19.67

p < 0.1; critical value for LL = 6.63, where LL > 6.63 means that the result is significant

However, it must be considered that some age groups are underrepresented in *Derry Girls*, particularly young male characters and older female characters, which do not count among their ranks any main characters, which may explain why these two groups account for zero occurrences of *ach* and *och*. Consequently, the only age group where cross-gender comparisons are possible is the adult age range, where female characters led the use of these DPMs by a wide margin, a difference in usage that was deemed statistically significant (Table 4.2). These results are not consistent with the reported frequencies of use in the existing literature, where it was found that males employed *ach* more frequently than females, and that among the latter it was the youngest speakers that used *ach* more often (Corrigan, 2015).

A constant in the DGC data was that every character and group, except Erin, Clare and Orla, employed *ach* more frequently than *och*. Erin and Clare uttered *och* more frequently,

¹⁰ Although the LL tests were performed using the raw frequencies for each item, it was decided to show the normalised frequencies next to the LL test results, as they provide a clearer picture of the frequency difference between the pair of compared (sub)corpora. Had there been enough space, both the raw and the normalised frequencies would have been shown. This applies to the tables throughout the chapter, except where space allowed for the inclusion of both types of frequency measurements.

whereas Orla used each DPM only once. The functions of each of these DPMs will be examined in more detail below.

Compared to the DGC-NI data, the relative frequency of *och* in contrast to *ach* in SPICE-NI was notable. As these two DPMs can be difficult to distinguish, this difference might be due to transcriber preference. Regardless, the resulting normalised frequencies for *ach* were much lower in SPICE-NI than in DGC-NI, while the inverse was true for *och*, and the LL tests for both DPMs yielded significant results when these differences were compared between both corpora (see Table 4.2). However, if we consider both DPMs at once, it can be seen that the frequency of use was higher for both genders in DGC-NI, and once more the difference was deemed significant by the LL test. In SPICE-Ireland, adult females were also the most prolific users of *ach* and *och*, with females, in general, employing these DPMs more frequently than males, although, once again, the frequency difference between genders was not found statistically significant by the LL test (see Table 4.2). For males, the use of these DPMs seemed to decline with age, as the younger males used *ach* and *och* more frequently than their older counterparts. The inverse was true for DGC-NI, for it was the older characters who employed these DPMs most often among the males. As was the case with the use of *ach* in the DGC by age and gender, the results obtained from SPICE-NI did not align with those reported by Corrigan (2015).

For comparison's sake, a search for *och* and *ach* was run in the ROI face-to-face-conversation section of SPICE-Ireland. Its results are presented in Table 4.1 above. Unlike in the DGC, these two DPMs appeared in the language of non-NI speakers, but at a much lower frequency than among NI speakers, almost fifteen times lower. This difference was found to be highly significant according to the result of the LL test (see Table 4.2). These results lend support to the view that *ach* and *och* are typically NIRE DPMs and are consistent with the representation of their use in the dialogue in *Derry Girls*, where only Northern Irish characters employ these DPMs.

Let us now consider the positions *ach* and *och* can occupy in utterances. Table 4.3 shows the raw frequencies and percentages of use for *ach* and *och* in initial and medial positions. No instances of these DPMs in final position were found in the DGC. Instead, these DPMs were overwhelmingly used in initial position. The two occurrences of *ach* in medial position were found in the dialogue of Uncle Colm, Joe's brother, and they will be examined below, in subsection 4.1.1.2.

Similarly, the position within utterances of *ach* and *och* in the DGC and SPICE-NI coincided almost entirely. As can be seen in Table 4.3, in SPICE-NI *ach* was exclusively used in initial position, and the same can be said of *och* in 96% of instances.

Table 4.3. Raw frequencies and percentages of use for *ach* and *och* in initial and medial positions in the DGC and SPICE-NI

	Ach		Och	
	Initial	Medial	Initial	Medial
DGC	58 (97%)	2 (3%)	12 (100%)	0 (0%)
SPICE-NI	4 (100%)	0	69 (96%)	3 (4%)

4.1.1.2 Functions of *ach* and *och*

Having examined the use of *ach* and *och* by different demographic groups among the characters and after comparing these to the results found in SPICE-NI, the next point in this analysis concerns the functions these two DPMs realise in their occurrences. To this end, I first turned to Corrigan's (2010) description of *ach* and *och*'s usage in NIrE, which was discussed in section 2.4.2.3.1 of Chapter 2. Starting with *ach*, this researcher reported that its two functions in NIrE discourse are expressing displeasure and signalling that new information is to follow. Out of the 60 occurrences of *ach* in the DGC, 20 seem to fit the expression of displeasure, as in the following instances:

(1) JAMES: Guys, all we can do is try our best.

MICHELLE: **Ach**, don't be such a fruit, James! (1x03)

(2) GERRY: I've got the map.

JOE: Aye. [*Snatches the map from Gerry*] Well, now I have the map.

MARY: **Ach**, Da... (1x05)

In both examples, *ach* signals the disapproval of either the words (as in 1) or the actions (2) of the addressee of this DPM.

The signalling of new information also seems to be represented in the DGC instances of *ach*, and 21 out of the 60 occurrences of this DPM can be said to fulfil this function, such as in the following instances:

(3) ERIN: **Ach**, I just love Halloween. It's the one night of the year that Protestants and Catholics set aside their political and religious differences and just come together... (3x06)

(4) MARY: [*commenting on the train journey the family is on*] **Ach**, isn't this lovely? (3x03)

(5) MARY: [*at a wake, contemplating the dead person*] **Ach**, when all was said and done, she wasn't a bad old soul. (2x04)

In these three examples, the new information that follows *ach* consists of unprompted opinions; one could infer, then, that *ach* acts as an opener of sorts. However, *ach* also appears at the beginning of responses:

(6) ERIN: Bomb! He has a bomb! Listen, Mammy! Listen!

MARY: **Ach**, Erin, that's just the big clock. (1x05)

(7) JOE: You could have bagged yourself a man like [Gabriel], love.

MARY: Wise up, Daddy. Hardly.

JOE: It's true. Never understand why you settled for thon thing. [*Referring to Gerry*].

MARY: **Ach**, I wouldn't be interested in a man like Gabriel, Daddy. (3x02)

(8) MARY: I think I will go for a smoky eye, you know.

SARAH: **Ach**, I'm still not convinced, you know, Mary. (3x06)

In all these examples, *ach* introduces a response that, besides including new information, constitutes a disagreement with the previous utterance. This seems to be the case for all the instances of *ach* as a marker of new information, with the following exception:

(9) GERRY: Do I want to know why you're reading a book on the Russian Revolution?

SARAH: [*nonchalantly*] **Ach**, it's just one of my wee interests. (3x05)

Here *ach* merely introduces Sarah's reply to Gerry's question, which she utters without any aggressiveness in her tone.

Besides the discussed examples, 20 occurrences of *ach* in the DGC could not easily be said to fulfil either of the two functions mentioned by Corrigan (2010). Among these are seven instances where *ach* appears in conjunction with the entrance of a character onscreen, as can be seen in the following four examples:

(10) [*SARAH enters the living room and sees that Fionnula, a neighbour, has called by*]

- SARAH: **Ach**, Fionnula, what about you? (1x02)
- (11) [*GERALDINE enters the Quinns' living room*]
 MARY: **Ach**, Geraldine. (2x03)
 GERALDINE: I hope you don't mind; your front door was open.
- (12) [*SISTER MICHAEL enters the video club where Dennis works*]
 DENNIS: **Ach**, Sister, it's yourself. (3x01)
- (13) [*CATHY enters the living room*]
 SARAH: **Ach**, Cathy, it's yourself.
 CATHY: Long time, girls. (2x06)

In each of these examples, we find *ach* at the beginning of the utterance, followed by a vocative addressed at the character that either has just entered (11) – (13), or that the speaker is surprised to see (10). In a way, it could be argued that these examples also constitute instances of *ach* introducing new information, in the sense that the sentences they precede refer to the appearance of another character. However, *ach*'s role as a character introducer seems specific enough to warrant separate discussion. In this usage, *ach* need not necessarily be followed by a vocative, as exemplified in (14):

- (14) [*AIDEEN enters the train carriage and notices Mary*]
 AIDEEN: Mary? Ha. I don't believe it.
 MARY: **Ach**, would you look who it is.
 AIDEEN: **Ach**, Joe, Gerry. You keeping well?
 GERRY: Not so bad, not so bad. (3x03)

Here, Mary does not recognise Aideen, as she has not seen her in over a decade, so she is unable to use her name after *ach*, but she still begins her utterance with this DPM. Aideen's subsequent usage of *ach* is more in line with examples (10) – (13), for she addresses Joe and Gerry by their first names. Later in this same episode, Aideen is 'introduced' once more with *ach*, this time by Sarah, once the latter realises who she has been talking to:

- (15) SARAH: [Aideen]'s in prison.
 AIDEEN: She's not anymore. But a ten-year stretch does wonders for the old figure.
 SARAH: [*finally recognising AIDEEN*] **Ach**, Aideen, what about you? (3x03)

Here, Aileen's 'appearance' is not literal, in the sense that she has been onscreen with Mary and Sarah for most of the episode at this point, but Sarah's recognition is still marked by *ach*.

There are still other minor uses of *ach* that do not fall into Corrigan's categorisation. In seven instances, this DPM is employed at the beginning of a response to a compliment:

(16) SARAH: Orla really thrives in the wild, you know. That time we went camping in Portsalon, sure she was like Mowgli running around them woods. She was that happy, Mary, I honestly thought about just leaving her there.

ORLA: [*fondly*] **Ach**, Mammy. (2x01)

(17) BRIDIE: [*sarcastically*] You should be proud of yourself.

SARAH: [*failing to perceive the sarcasm*] **Ach**, thanks very much. (2x04)

(18) SARAH: And fair play to you, Cathy, you've kept them eyebrows shipshape, so you have. Credit where credit's due.

CATHY: **Ach**, you're too kind, Sarah. (1x06)

(19) AIDEEN: My God. Sarah McCool. Have you got a picture in the attic?

SARAH: **Ach**, away of that. (3x03)

(20) GERRY: I don't think she was talking to me, Joe. I think she was talking to the person in the full-length white frock who just managed to upstage the actual bride.

SARAH: **Ach**, I wouldn't say I upstaged her now, Gerry. (2x04)

(21) COLM: He was telling me there, John was, that every being in the place is talking about Sarah's frock.

SARAH: **Ach**, really? (2x04)

Examples (17), (20) and (21) have been considered examples of responses to compliments because that is how they are interpreted by Sarah, despite Gerry's line in (20) being intended as a criticism of her wardrobe choices, Colm's line in (21) being a neutral, factual statement, and Bridie's sentence in (17) being sarcastic. Sarah's inappropriate reactions contribute to the humour of the scene.

Out of these examples of responding to compliments, only (16)–(17) express unreserved acceptance (despite the compliment in [16] including abandoning Orla at a camping site, thus evidencing both Sarah and Orla's quirkiness). Examples (17), (18) and (20) portray a mild rejection of the compliments, which has been described as a common response to flattery in Irish society (Kallen, 2005a). It could be argued that the rejection present in these uses of *ach* could be interpreted as the characters expressing displeasure at receiving a compliment, but in

(17)–(18), the utterances headed by *ach* are said in a friendly tone by the character, which to me signals more modesty than displeasure on the part of the speaker.

In two instances *ach* is used to express something akin to fondness and approval:

(22) ERIN: You asked Granda to the prom?

ORLA: Well, everyone kept saying you have to ask a fella you really like. And he's the fella I like the most.

SARAH: [*fondly*] **Ach**, Orla, love. (2x05)

(23) MARY: Even so, Sarah, wouldn't you like to know that Mammy's happy, that she's at peace now?

JOE: I'd give anything to speak to her again.

MARY: [*fondly*] **Ach**, Daddy. (3x04)

In these instances, a non-dialectal synonym for *ach* could be *aww*, which can express endearment and thus be uttered in response to something the speaker finds touching (Tinkerbell, 2004). The expression of fondness is reinforced using vocatives, such as the term of endearment *love* in (22), which accompanies the first-name vocative *Orla*, and the kinship term *Daddy* in (23). One can compare Mary's use of *Daddy* here with her use of *Da* in example (2), where *ach* is employed with its more common function of expressing displeasure. This contrast will be discussed in more detail in subsection 4.2.2.1. Using *ach* to express endearment seems the opposite of its above-discussed expression of displeasure, which evidences the multifunctionality of this DPM in the dialogue in *Derry Girls*.

The last two occurrences of *ach* that I would like to discuss are the only ones that occur in medial position:

(24) COLM: There was a knock at the door, this must have been, **ach**, we're talking eight, half eight, for I was halfway through me dinner. (1x02)

(25) COLM: I knew a fella once, Tommy Duddy, he spoiled his vote. Now this would have been back in, **ach**, we're talking '88, '89. Or was it '90? (3x07)

Uncle Colm's dialogue is characterised by long, winding monologues that bore everyone around him. Here, his use of *ach* mid-anecdote right before trying to recall a fact suggests that this DPM is functioning as an imprecision marker in the vein of medial *like*. Perhaps *ach* adds

a connotation of frustration at the inability to remember the exact timing or date, owing to its other use as a marker of displeasure, but confirming this would require more evidence than is available in the DGC.

Turning now to *och*, its reported function in Corrigan's (2010) work is that of expressing sorrow. Out of the 20 occurrences of this DPM in the DGC, only three can be said to express this feeling unambiguously:

(26) ERIN: [*about her prom date*] He's not coming.

MARY: **Och**, love. (2x05)

(27) CLARE: But it's history! She's making us miss history!

ORLA: **Och**, but I really wanted to find out what that Cromwell fella got up to next. I really enjoy him. (2x06)

(28) CLARE: **Och**, I'm not even sure I trust Chelsea anymore. (2x06)

In these three cases, the speakers express their sadness and disappointment with their current situation with utterances headed by *och*. The remaining 9 occurrences of this DPM seem to align with two of the functions of *ach* that have been identified in the DGC: expressing displeasure (29)–(33) and preceding new information (34)–(35):

(29) MARY: I said you could invite one friend to the reception! One!

ERIN: **Och**, Mammy, they don't come separately. (2x04)

(30) JOE: Exactly, love. Everybody knows you didn't mean to kill the old boot. God rest her soul.

MARY: I didn't kill her!

JOE: **Och**, you know what I mean, not kill... hex. (2x04)

(31) MARY: Well, if he really likes you, it won't matter what you wear.

ERIN: **Och**, come off it! (2x05)

(32) GERRY: That's not a healthy eating plan, That's a no eating plan!

MARY: **Och**, don't be so dramatic, Gerry. We can have oranges. (3x05)

(33) MICHELLE: **Och**, fuck off, Dennis. (3x07)

(34) MARY: **Och**, this is lovely, Gerry. (2x02)

(35) SARAH: [*about Orla trying on her prom dress*] **Och**, Mary, you'd think the wean's been dropped into it. (2x05)

As was the case with *ach*, *och* is also used to express disagreement, as seen in (29)–(31). This apparent overlap in the functions of these two DPMs lends support to Corrigan’s observation that these two linguistic items can be hard to distinguish, even by speakers of NIrE.

The last function of *och* that appears in the DGC is that of replying to offers, present in the following instances:

(36) ERIN: [*upon receiving a present*] **Och**, lads, you really shouldn’t have. (2x01)

(37) MAN: Let me give you something for your trouble.

ERIN: **Och**, not at all. (3x01)

In both examples, the speaker verbally rejects the offers, despite accepting the presents she is being given. In doing so, she manages to protect the givers’ positive face, as she is accepting their gifts, while also preventing damage to their negative face by downplaying the (perceived) obligation of giving her something.

It is interesting to note that in examples (34)–(35), where *och* precedes new information, the sentences that follow *och* are invariably unprompted positive opinions, which was also the case for examples (3)–(5), concerning *ach*. This contrasts with its association in the existing literature (Corrigan, 2010) with the expression of negative emotions, and highlights the multifunctionality of these DPMs.

Attributing functions to the occurrences of *ach* and *och* in the DGC has been aided by the knowledge of the audio-visual material the dialogue is sourced from. Without the non-verbal communicative elements present in the scenes, it would be difficult to assign emotions to each example of these DPMs, as gestures and tone of voice are invaluable in interpreting the character’s inner state. Therefore, it must be taken into account as we now turn to *ach/och* in SPICE-NI that this contextual information is not available to the researcher when working with this corpus, which makes the functional analysis of these DPMs more challenging. Indeed, I have been unable to categorise some of the occurrences of *och*, as, lacking information about the speaker’s tone of voice, it is nigh impossible to discern the emotions and attitudes expressed by them. And, as far as distinguishing between displeasure and sorrow is concerned, the verbal context of the DPM was considered in order to decide between both feelings, but these distinctions are tentative.

After these methodological caveats, let us examine the instances of *ach* in SPICE-NI. By looking at their respective contexts, it appears that all four of them would fall under the “expression of displeasure” category:

- (38) A: It’s not worth it.
B: **Ach**, it is. (P1A-022)
- (39) E: No, I hate Christmas.
A: **Ach**, Sean. (P1A-023)
- (40) C: What was that, Mark, I didn’t hear you.
D: **Ach**, no. (P1A-030)
- (41) D: I supposed it’s alright.
A: **Ach** well. (P1A-022)

In (38) the displeasure is inferred from the fact that B disagrees with A. In (39), A seems to disapprove of E’s opinion on Christmas. In (40), D’s displeasure arguably stems from C not having heard what they said. Lastly, (41) could be interpreted as an expression of resignation.

Let us move now onto *och*, which was much more frequent in SPICE-NI than *ach*, as discussed above. Out of 70 instances of this DPM in said corpus, 9 could be interpreted as expressions of displeasure, much like *ach*. Consider the following examples:

- (42) B: I know, I don’t even feel comfortable talking about him on tape.
E: **Och**, don’t be stupid. (P1A-007)
- (43) A: I couldn’t speak. I just went “He’s not my type.” [Name] thinks Justin’s completely ugly.
Och for fuck’s sake. Don’t listen to her. Justin, Justin’s a big screw. (P1A-017)

In (42) the use of a derogatory term tipped the scales in favour of displeasure, as opposed to sorrow; the same logic was applied to the use of taboo language in (43). That being said, the expression of sorrow through the use of *och* also appeared in SPICE-NI:

- (44) D: Edward, did you see *Prisoner* Friday night?
A: No.

D: **Och** ‘ve you not seen it? It’s on tonight. Did you not see the one about Sandy? (P1A-034)

(45) B: I failed my exam as a protest.

A: **Och** Michael. (P1A-016)

(44) and (45) are two out of the 12 occurrences that could be interpreted as expressing sorrow, but I must acknowledge that (45) could very well express displeasure too.

In a similar vein to *ach/och* in the DGC, *och* in SPICE-NI appears preceding new information, including positive, unprompted opinions (46)–(47), responses to an interlocutor (48)–(49) and utterances that express disagreement (50)–(51):

(46) D: **Och** well I’m glad Benjamin’s got a job now. He’s not the worst, you know. (P1A-003)

(47) B: **Och** your house is just so gorgeous now. You must be just so happy here. (P1A-006)

(48) A: Is he the one who writes the books?

D: **Och**, he does write. (P1A-008)

(49) A: But what about the rest of the staff: Roger and Millie et cetera?

C: **Och** Roger’s just Roger, like. You give him a bit of shit and you’ll get a bit of shit back. (P1A-024)

(50) B: And that was the end of that.

A: **Och** it wasn’t really, like. (P1A-044)

(51) C: It must be really awkward, like, you know, I mean, if it’s all in school and she’s...

A: **Och** she gets on well with people. **Och**, I don’t know. (P1A-026)

Interestingly enough, in (51) *och* is used initially to express disagreement, and then again when the speaker backtracks. In this example, together with (50), *och* also carries a connotation of frustration or displeasure. Another function of *och* that can be found in SPICE-NI is that of responding to a compliment:

(52) B: Oh that’s very nice now. Doreen this is really great.

A: **Och**, thanks a million. (P1A-007)

(53) B: You are successful.

D: But I end up like over-compensating saying “**Och**, I’m just working away.” Instead of saying “Yeah, work’s going well” or whatever. But yes, they would-

B: But you are always better to underplay it.

D: Yes, oh absolutely. (P1A-020)

In (52), the receiver of the compliment accepts it unreservedly, but (53) presents a different situation. In this example, D is discussing their professional success, and how, instead of accepting the compliments they receive about it, they downplay their achievements. The example of a reply they quote begins with *och*, but what is most interesting is how their interlocutor, B, opines that downplaying one’s success is the better course of action, on which D agrees. This exchange constitutes an example of the Irish dislike for self-aggrandisement (Millar, 2015, p. 294) discussed in section 2.4.1.

A use of *ach/och* that is underrepresented in SPICE-NI compared to the DGC is that of acknowledging the entrance of a person. However, this can be explained by the fact that the conversations transcribed in the first corpus do not include the beginning of the exchanges, so the way the speakers greet each other is not reflected in the data. As a result, only one instance of this type could be identified:

(54) D: Now I must make the point three or four weeks ago I was coming back from lunch, you know, up Queen Street so I was at those traffic lights opposite the Centra, and I saw him and I thought “I can’t be arsed talking to you”. But he was waiting to cross the fucking lights as well. So I was standing with the mobile phone and all out. And I could see him going, you know, looking like this just to make sure it was me “Sarah” “**Och**, Roger how are you doing?” (P1A-020)

As can be seen, in (54) the speaker is not even greeting any interlocutor; instead, their use of *och* happens within an anecdote she is relating in which she feels forced to greet someone to whom she does not want to talk.

Finally, and in the vein of Uncle Colm’s dialogue in examples (24)–(25) above, in SPICE-NI there is one instance of *och* being uttered as a marker of imprecision:

(55) C: Ah I just have an awful lot of revision to do over the Christmas period. Now whether I get that put in or not is another story. But, when I think about it in the long term I don’t

particularly want to do repeats in August, so I'll try, **och**, I'd say if you do a couple of hours here and a couple of hours there... (P1A-039)

The fact that this speaker employs *och* instead of *ach*, the latter being employed by Uncle Colm in the DGC, once more points to the fuzzy boundary between these two DPMs.

In general, it appears that the dialogue in *Derry Girls* reflects the uses of *ach/och* in NIrE as they are represented in SPICE-NI with a reasonable degree of accuracy.

4.1.2 Final *but*

Having analysed the use of *ach* and *och* in the DGC and compared it to that in SPICE-NI, we now turn to final *but*. As mentioned in Chapter 2, section 2.4.2.3.2, this is another DPM that has been associated with NIrE in the available literature on DPMs in IrE. In DGC-NI, there was just one instance of final *but*, which resulted in a normalised frequency of 0.19 tokens per 10,000 words:

- (56) EMMETT: Look, I'm sure you have questions. I just don't really want to get into it all.
MARY: Yeah, well, we didn't really want you to get into our boot, **but**. (1x05)

In contrast, there were 23 clear instances of final *but* in SPICE-NI, yielding a normalised frequency of 2.50 tokens per 10,000 words. In addition, this frequency difference between DGC-NI and SPICE-NI was deemed statistically significant by a LL test (see Table 4.5 below). It could be argued, therefore, that final *but* is underrepresented in the dialogue in *Derry Girls*. As far as the function of this DPM is concerned, it can be seen in example (56) above that, as mentioned in the literature discussed in Chapter 2, final *but* works in the same way as *though* at the end of a sentence. The same can be observed in the following examples, randomly sampled from SPICE-NI:

- (57) B: But whenever Granda died he changed and he, he joined just to sort of respect him or whatever. [...]
A: He's not a team player, **but**. (P1A-005)
- (58) C: And I said that there to this one one day I said to her "you know there's very few ma-very little manners no breeding". She says "We're not a lot of horses", because I said "breeding".

[...]

C: I was going to say, no you're well-shoed **but**. (P1A-037)

(59) B: Yeah it would be cool in a bathroom actually. I was thinking on a hearth it would be nice, or you know on a mantelpiece. [...]

B: But it was quite hideous **but**. (P1A-035)

In (57) and (58), the person uttering the final *but* is replying to another speaker, but, as is the case in (59), this DPM can be used in a response to the speaker's own words. Unfortunately, due to the scarcity of data in the DGC, the usage demographics of this DPM cannot be investigated and compared to those in SPICE-NI to the same degree that other DPMs allow. However, we can compare the data from SPICE-NI to the results obtained by Corrigan (2015), which were discussed in section 2.3.2.3.2 above. In her study, this researcher noted that final *but* might be subjected to age grading, a phenomenon that manifests as adults using a non-standard feature less than younger and older speakers (p. 55). As we can see in Table 4.4 below, this pattern does not appear in the SPICE-NI data.

Table 4.4. Normalised and raw frequencies for final *but* in SPICE-Ireland

SPICE-NI		
Speaker	Raw frequency	Normalised freq. (per 10,000 w)
All	23	2.50
Female	15	2.28
Male	8	3.29
Female young	3	2.18
Female adult	6	2.65
Female old	2	1.20
Male young	6	4.61
Male adult	2	4.76
Male old	1	5.93
SPICE-ROI		
All	18	1.91

Table 4.5. Table 4.5. Log-likelihood tests for final *but*
NF = normalised frequency (per 10,000 words)

Target corpus	Comparison corpus	NF in TC	NF in CC	LL
DGC-NI	SPICE-NI	0.19	2.50	14.67
Female (SPICE-NI)	Male (SPICE-NI)	2.28	3.29	0.67
SPICE-NI	SPICE-ROI	2.50	1.91	0.75

p < 0.1; critical value for LL = 6.63, where LL > 6.63 means that the result is significant

Since the literature review showed a certain disagreement among scholars regarding the regional prevalence of final *but*, with Corrigan (2015) and Hickey and Harris (cited in Walshe, 2009, p. 124) attributing it to NIrE, whereas Walshe claimed that it was characteristic of Dublin English (ibid.), a search was run in SPICE-ROI, which yielded 18 occurrences of final *but*, as shown in Table 4.4 above. Subsequently, a log-likelihood test was performed to compare this frequency to that of final *but* in SPICE-NI, the result of which is shown in Table 4.5. The test deemed the frequency difference between both subcorpora as not statistically significant. This casts doubt on final *but* being exclusively NIrE, but its presence in SPICE-NI does not allow one to attribute it to Dublin English either. In order to conclude either way, more research, conducted on more recent, spoken, real-life data would be required.

In *Derry Girls*, the only character who employs final *but* is a middle-aged woman, Mary, and a look at the normalised frequencies of this DPM's usage in SPICE-NI, presented in Table 4.4, shows that adult females were also the most likely among their gender to use it, although this difference between genders is not statistically significant (see Table 4.5 above). At any rate, Mary's usage of final *but* reinforces her status as a speaker of IrE, which had already been portrayed through her frequent use of *ach/och* (second only to Sarah, also a middle-aged female character), another DPM closely associated with (Northern) IrE. Taking together *ach/och* and final *but*, it appears that Mary contrasts with the remaining main characters in terms of how distinctively (Northern) Irish her speech is, in terms of DPM use. The remaining sections of this analysis will reveal if this pattern stays consistent.

4.1.3 *Wee* as a DPM

We now turn to the last of the pragmatic elements on which section 4.1 focuses: *wee*. Studying the use of this potential DPM in the DGC posed some difficulties, the main one being that its position in the sentence could not be used as a sign that the word carried pragmatic meaning. *Wee* always appears as a noun modifier, firmly embedded in the sentence structure, and not in peripheral positions like other DPMs, and thus the coding process for this linguistic item was difficult. Consequently, it was decided to put categorical distinctions between adjectival and discourse *wee* aside and instead look for its patterns of use in the data, as discussed below. The relatively low frequency of *wee* (with only 96 occurrences in DGC-NI) made this manual analysis possible. The next subsection, 4.1.3.1, will discuss the frequency of all occurrences of this element in the DGC and SPICE-NI, as well as their partitions, while subsection 4.1.3.2 will be devoted to the functions that *wee* exhibited in the data contained in both corpora.

4.1.3.1 Frequencies of use of *wee*

To enable comparisons between characters, the normalised frequencies of *wee* in their dialogue were calculated. The results are presented in Table 4.6. Among the individual characters analysed, Orla was the one who employed *wee* the most, with a normalised frequency which almost doubled that of the second most prolific user, Sarah. The frequency difference between these two characters' use of *wee* was not deemed significant according to the results of the LL tests, shown in Table 4.7. However, the difference in use between Orla and the rest of the characters was found significant. At the other end of the scale was Erin, whose normalised frequency for *wee* was almost the same as that of non-NI male characters, a similarity corroborated by a log-likelihood test that deemed the difference between the two subcorpora not significant (see Table 4.7 above). By gender, male characters were more frequent users of *wee* than female characters, although the difference was not found statistically significant in the corresponding log-likelihood test. Once age groups are considered, it can be seen that it was the adult males that led the usage of *wee*, followed by young males. However, this last group is severely underrepresented in the DGC, which means that every instance of use becomes more than 14 occurrences per 10,000 words when the normalised frequency is calculated, which significantly skews the results. In a similar vein, the fact that old female characters never utter *wee* is not particularly significant if their underrepresentation in the DGC is taken into account.

Furthermore, the frequencies for adult male and adult female characters might have been skewed by a plot line in episode 1x06, which involves the main adult characters looking for a missing "wee docket". As a result, 3/15 occurrences of *wee* for adult female characters corresponded to "wee docket", and the same was true for 2/15 occurrences of *wee* for adult male characters. Disregarding the instances of "wee docket", the normalised frequencies for *wee* were 16.53 for adult females and 34.39 for adult males. With these revised figures, adult males still led the use of *wee*, but adult females fell from third to fourth place, surpassed by old males. With regards to the significance tests, they were performed twice when comparing adult males to the remaining character groups, once including the "wee docket" occurrences in the frequency counts and once without them. The results showed that, once the "wee docket" occurrences were disregarded, the difference in frequency between the adult male subcorpus and the remaining subcorpora was not statistically significant, even though the normalised frequency for the first group almost doubled that of the second (see Table 4.7 below).

Table 4.6. Normalised and raw frequencies for *wee* in the DGC and SPICE-Ireland. RF = raw frequency, NF = normalised frequency (per 10,000 words).

Character/speaker	DGC		SPICE-NI	
	RF	NF	RF	NF
All	100	15.27	121	13.16
Erin	6	5.80	–	–
Michelle	14	17.21	–	–
Clare	4	7.39	–	–
Orla	13	59.52	–	–
Mary	12	18.23	–	–
Sarah	13	30.47	–	–
Joe	6	17.35	–	–
DGC-NI	96	18.08	–	–
NI Female	72	16.55	86	13.08
NI Male	24	24.98	34	13.97
Non-NI female	0	0	–	–
Non-NI male	4	4.66	–	–
Young female	41	14.48	14	10.19
Adult female	30	19.34	32	14.12
Old female	0	0	15	9.01
Young male	2	28.37	18	13.82
Adult male	15	39.81	7	16.64
Old male	7	17.50	2	11.87
SPICE-ROI				
Speaker	RF		NF	
All	10		1.06	

Table 4.7. Log-likelihood tests for *wee*. NF = normalised frequency (per 10,000 w.)

Target corpus	Comparison corpus	NF in TC	NF in CC	LL
DGC	SPICE-NI	15.27	13.16	1.20
DGC-NI	SPICE-NI	18.08	13.16	4.57
ORLA	NO ORLA	59.52	13.74	17.05
ORLA	SARAH	59.62	30.47	4.72
FEMALE (DGC)	MALE (DGC)	16.55	24.98	2.84
AM (DGC)	All other groups (DGC)	39.81	16.40	8.04
AM (DGC) without “wee docket”	All others (DGC) without “wee docket”	34.39	15.19	6.10
ERIN	MALE non-NI (DGC)	5.80	4.66	0.12
SPICE-NI	SPICE-ROI	13.16	1.06	113.83
AM+AF (SPICE-NI)	All other groups (SPICE-NI)	14.51	10.87	1.79
AM (SPICE-NI)	All other groups (SPICE-NI)	16.64	11.95	0.64

$p < 0.1$; critical value for LL = 6.63, where LL > 6.63 means that the result is significant

When examining the normalised frequencies of *wee* in SPICE-Ireland, *wee* was more common in the speech of NI speakers than in the speech of ROI speakers, as was expected from the fact that this word is associated with NIrE (in an Irish context; in Great Britain, it is instead associated with Scottish English). By comparing the normalised frequency of *wee* in SPICE-NI to that in the DGC-NI corpus, it could be seen that the difference was not statistically significant, and therefore it could be argued that *wee* is proportionately represented in *Derry Girls* when compared to real-life speech.

By gender, both males and females exhibited similar frequencies of use in the SPICE-NI data. When age was taken into account, it was once more adults who revealed themselves as the most frequent users of *wee* in each gender, with adult males being the most constant users overall, as was the case in the DGC. However, the differences between groups were less stark in SPICE-NI, and neither of the significance tests that included an adult SPICE-NI partition yielded significant results (see Table 4.7).

4.1.3.2 *Wee*: patterns of use

Having looked at the frequencies of use of *wee* among the characters represented in the DGC, and in the SPICE-NI data, we now turn to the examination of how *wee* is used.

(60) TEACHER: Dig deep, something you hate, something you despise. No holding back.
Come on! Get it out!

JAMES: The fact that people here use the word *wee* to describe things that aren't even actually that small! (2x02)

To non-NIrE speakers like James, and most of *Derry Girls*' audience (Irish and Scottish viewers excepted), the word *wee* can be highly salient. What James is expressing in the quote above is his intuition that this adjective does not just function as a regional stand-in for *little*, but it has additional uses that can be more difficult to parse for those unfamiliar with NIrE. This quote, furthermore, serves to signal those in the audience who may have not yet realised this multifunctionality of *wee* that they should pay attention to the show's language if they want to fully engage with it.

Let us now explore the occurrences of *wee* in the DGC to confirm or refute James' intuition that it means more than just *little*. However, we first need to consider that a common use of *wee* in the DGC is indeed describing small things:

- (61) ORLA: When we were **wee**, we always went to parties dressed the same. (2x05)
- (62) JOE: They try to catch him but their boat's too **wee**. (3x03)
- (63) GERRY: You're absolutely sure we need the big clock, love? Definitely don't wanna bring the **wee** clock? (1x05)

In these examples, *wee* simply modifies the noun it describes, although in (63) the use of this word adds an element of affection to the utterance. In a similar vein to these instances, *wee* can be found preceding *bit* to form a quantifying or intensifying expression, equivalent to *a little bit*:

- (64) DAVID: In this game, you need a **wee** bit of integrity. (1x04)

The cluster *wee bit* occurs 8 times in the DGC. In contrast, *little bit* appears only twice. These uses of *wee* as just an adjective meaning “small” are also frequent in SPICE-NI, where at least 65/121 (i.e., 54%) instances of *wee* fulfil that function, as seen in examples (65)–(66):

- (65) B: Then she was expecting a baby and she was in the hospital and she had she'd this **wee** boy. (P1A-033)
- (66) A: We decided she wants to do a **wee** bit of praying and we decided to push her. (P1A-043)

But beyond these uses of *wee*, there are other environments where this adjective appears relatively frequently. First, there are in the DGC 11 instances of the structure *adj + wee + derogatory term*, exemplified by the following quotes:

- (67) DENNIS: You **cheeky wee rat!** (3x07)
- (68) FIFI: **Lying wee shites!** (3x06)
- (69) MARY: She always was one **cold, self-serving wee madam!** (2x06)

In (67), *wee* functions as a mitigator, softening the insults it precedes. If one performs the deletion test proposed by Villalba Ibáñez (2020), which was explained in section 2.4.2.3.4, one can see that the resulting utterance is more forceful than the original one containing *wee*. In addition, the speaker (Dennis) is addressing the insult to Michelle, his employee, so it could be argued that the mitigating use of *wee* in (67) also seeks to protect their relationship by

diminishing the threat to Michelle's positive face. Example (68) also seems to pass the deletion test. However, it must be noted that the addressees of the utterance are out of earshot in the scene, and as a result, one cannot say there is facework present in this example.

In contrast to (67)–(68), in (69) it could be argued that *wee* acquires the derogatory connotations of the linguistic items that precede it, as opposed to attenuating their meaning. Consequently, here *wee* would serve as a further insult that metaphorically diminishes the addressee of the offensive phrase. This function of *wee* also appears in SPICE-NI, as exemplified in (70)–(71):

(70) B: Simon Brown's a **wee** frigging **wee** poof. (P1A-018)

(71) A: Well in my day it was um Martin's Wine Barrel.

D: It was a grotty **wee** hole in your day. (P1A-024)

The adjective *wee* can also be found in the DGC in the context of offers (72)–(74) and requests (75)–(77). In contrast, in SPICE-NI *wee* does not appear in offers, only in requests (78)–(79):

(72) JENNY: [*to the Derry Girls*] Oh, and FYI, the prom queen vote closes today.

AISLING: [*to the Derry Girls*] Here's the **wee** ballot. (2x05)

(73) MICHELLE: Sit yourselves down and have a **wee** drink. (1x02)

(74) JOE: Now, more bacon, anyone? Not a sausage? **Wee** omelette, maybe? (1x03)

(75) CIARAN: Look, why don't you just go home and have a **wee** look for your wee docket?
(1x06)

(76) FIFI: Can I see your **wee** passes there? (3x06)

(77) POLICE OFFICER: We've just had a **wee** security alert there. I'm afraid we're gonna have to carry out a **wee** evacuation. (2x02)

(78) C: Could I have a **wee** glass of red, please? (P1A-019)

(79) B: Oh yes I always said, "Please now wait a **wee** minute." (P1A-008)

It is in these cases that it could be said that *wee* is acting as a minimiser, as Kallen (2005a) suggested. In these examples, *wee* serves to diminish the illocutionary force of the offer or the request, therefore lessening the force of the imposition, of the potential face threat that these types of speech acts entail. (Brown & Levinson, 1987, p. 66) For instance, in (73), without *wee* the offer would sound less friendly to the hearers.

In addition, *wee* adds various nuances to the examples. For instance, in (74) *wee* contributes to the action of offering, making it more attractive from the point of view of the speaker, who is trying to sound accommodating. In (75), *wee* attaches a notion of collaboration to the speaker's suggestion, who by using this word involves himself in the task he is suggesting. In (76), the use of *wee* serves to make the request sound less intimidating. Finally, in example (77), *wee* also accompanies the noun phrase *security alert*, which arguably serves to minimise the importance of the threat such an alert could represent, especially if one takes into account the historical context of the show, where terrorist attacks were not as infrequent as one would desire.

There are similar occurrences of *wee*, where this adjective arguably functions as a minimiser, but not in the context of offers or requests:

(80) SARAH: We heard about your **wee** divorce. (2x06)

(81) SARAH: So, heads, the **wee** IRA man comes with us. Tails, we leave him. (1x05)

(82) MAE: Really. It'll be a laugh. And I'm definitely an upgrade [compared to Erin]. Just a **wee** joke, there. (2x05)

In (80), Sarah's use of *wee* might help soften the act of mentioning a potentially painful memory, a divorce. In (81), *wee* is not physically describing the man, who is of average height, and not much younger than the character who says the line; instead, it could be argued that *wee* serves to minimise the fact that the man belongs to the IRA, which puts the Quinn family in danger of becoming their accomplices if they decide to help him cross the ROI border. (82) is excerpted from a scene where Mae has just offered to be Clare's prom date after Erin spurns the latter. Mae's use of *wee*, together with the hedge *just*, is intended to minimise the positive face threat that her previous sentence is.

This usage of *wee* as a softening device is consistent with the literature (discussed in Chapter 2, section 2.4.1) that describes the preference of IrE speakers for indirectness, which is expressed, among others, through hedges and minimisers. *Wee* would, therefore, be another tool in the repertoire of NIrE speakers to follow the pragmatic conventions of their linguistic community.

Table 4.8. *DPM wee* against Brinton's (2017) characteristics of DPMs

Characteristic	Does <i>wee</i> fulfil it?
Phonological and lexical	
a) Pragmatic markers are often 'small' items.	Yes.
b) Pragmatic markers may form a separate tone group, but they may also form a prosodic unit with preceding or following material.	<i>Wee</i> does not form a separate tone group, as it is embedded in clauses.
c) Pragmatic markers do not constitute a traditional word class, but are most closely aligned to adverbs, conjunctions, or interjections.	No, as <i>wee</i> is an adjective.
Syntactic	
d) Pragmatic markers occur either outside the syntactic structure or loosely attached to it.	No, <i>wee</i> is embedded within the syntactic structure.
e) Pragmatic markers occur preferentially at clause boundaries (initial/final) but are generally movable and may occur in sentence-medial position as well.	<i>Wee</i> always occurs in sentence-medial position.
f) Pragmatic markers are grammatically optional but at the same time serve important pragmatic functions (and are, in a sense, pragmatically non-optional).	Yes.
Semantic	
g) Pragmatic markers have little or no propositional/conceptual meaning, but are procedural and non-compositional.	Yes (in the instances where <i>wee</i> is not used in its literal, descriptive sense).
h) Pragmatic markers are often multifunctional, having a range of pragmatic functions.	Yes.
Sociolinguistic and stylistic	
i) Pragmatic markers are predominantly a feature of oral rather than written discourse; spoken and written pragmatic markers may differ in form and function.	No data, as this study focuses on (fictional) oral discourse.
j) Pragmatic markers are frequent and salient in oral discourse.	Yes.
k) Pragmatic markers are stylistically stigmatised and negatively evaluated, especially in written or formal discourse.	No data.
l) Pragmatic markers may be used in different ways and with different frequencies by men and women.	With regard to frequency, true for DGC data, and not true for SPICE-Ireland data.

Besides its use in derogatory expressions and face-threatening acts, *wee* also appears in the context of characters expressing positive opinions, both in the DGC (83)–(84) and in SPICE-NI (85)–(86):

- (83) ORLA: Their **wee** beekeeper costumes are so **cracker**. (2x03)
- (84) SARAH: Ach, it's a **great wee day** so far, isn't it? (2x04)
- (85) B: And here's me I said "Now, wasn't that a lovely **wee** film?" (P1A-033)
- (86) A: And that's a brilliant **wee** machine that wee one there (P1A-041)

Examples like these could be linked to the tendency towards understatement that is present among IrE speakers (Amador-Moreno, 2010a, p. 116). In (83)–(84) the minimising effect of *wee* helps avoid a potentially hyperbolic appraisal.

Having explored the various functions of *wee* in both the DGC and SPICE-NI, it would be useful to compare the functions exhibited by this linguistic item against the characteristics of DPMs outlined by Brinton (2017), mentioned above in section 2.4.2.1, in order to see if we can include *wee* in the DPM category with more confidence. The comparison is presented in Table 4.8 above.

According to the analysis carried out in this section, *wee* appears to agree with 6/12 characteristics, and disagree with 2. For the remaining ones, there is either no data or *wee* fulfils one of the options mentioned in the characteristic. As a result, one could regard *wee* (when it is not used in its literal sense of “small”) as a DPM in NIrE, albeit a non-prototypical one.

4.2 Linguistic characterisation in the DGC

Having addressed research questions 1 and 2, we now turn to the third and last research question, concerning the linguistic characterisation, mainly via pragmatic elements, in the dialogue of *Derry Girls*. The first source of these elements will be each character's most frequently used words, examined in subsection 4.2.1 below, and the second will be the keyword lists obtained from each character subcorpus through *AntConc*, analysed in subsection 4.2.2.

4.2.1 Most frequent words

Table 4.9 presents the thirty most frequent words in the Derry Girls dialogue and for each of the seven characters studied in this thesis. As could be expected, the majority of the most frequent words were grammatical words like pronouns, articles and prepositions. That being said, there were some interesting items on the list. First names appeared on several of the characters' top thirty words, the most notable example being *Mary*'s sixth position in Sarah's column, the highest-ranked occurrence of a non-grammatical word on the table. By looking at

the concordance lines of Mary in the Sarah subcorpus, one can see that out of 62 occurrences of this item, 59 constituted vocatives, so they signalled interactions between Sarah and Mary, indicating that these two characters share a considerable portion of their screentime on the show.

Table 4.9. Word frequency lists for selected characters from Derry Girls

Key: the *green* keywords are terms of address; the *red*, taboo expressions; the *blue*, dialectalisms; and the *purple*, DPMs.

	All	Erin	Michelle	Mary	Clare	Sarah	Joe	Orla
1	You	you	The	You	I	You	you	I
2	The	the	You	The	The	I	the	A
3	I	I	A	I	You	A	A	The
4	A	A	I	To	To	To	I	You
5	To	To	To	A	A	The	To	It
6	It	It	And	It	It	Mary	It	And
7	Of	we	That	Of	We	It	what	To
8	And	that	It	What	Is	That	Of	Is
9	That	what	Of	Is	And	Not	that	Of
10	Is	Of	We	And	Of	Of	No	Like
11	What	Is	On	We	I'm	And	In	We
12	We	and	Is	That	Not	On	is	No
13	Not	not	Fuck	No	What	I'm	and	Have
14	This	this	Fucking	Do	That	it's	don't	Not
15	Me	just	Just	Not	This	Me	not	On
16	I'm	it's	Me	In	Are	In	be	Oh
17	Just	have	Not	This	it's	For	with	So
18	In	I'm	This	For	Just	Have	have	In
19	On	On	Are	Have	Be	He	this	She
20	Have	Be	In	On	Me	But	well	But
21	it's	No	What	All	Do	Do	on	That
22	No	So	James	Be	In	Is	me	Aye
23	Do	well	Do	I'm	Like	Her	it's	Just
24	Be	my	it's	Are	My	This	know	Me
25	Are	are	So	Da	Have	Well	that's	Are
26	For	In	For	Gerry	So	What	love	Do
27	So	Do	Be	Know	No	Be	I'll	don't
28	My	For	Have	Was	don't	So	do	Erin
29	Well	don't	Up	Well	Erin	We	he's	My
30	don't	Me	About	Just	Think	With	for	I'm

Also notable was the presence of two taboo terms, *fuck* and *fucking* one after another in Michelle's column, them being the only instances of this type of language in Table 4.9, which set Michelle apart from the other characters, giving her dialogue a potentially more vulgar profile, which will be discussed in more depth in section 4.2.2.2 below. Finally, both Clare and Orla had *like* in their respective columns. Seeing that one of the functions of this linguistic item is that of a DPM, a search was run in the Orla and Clare subcorpora to assess if its use aligned with more traditionally Irish patterns or with more recent, globalised uses of this DPM, as discussed in Chapter 2, section 2.4.2.2.2. The results from the Orla subcorpus showed that this item was mostly used as a verb, and the only instance where it was used as a DPM did not occur in the typically Northern Irish sentence-final position:

- (87) JOE: What even is brunch?
ERIN: It's a sort of half breakfast, half lunch.
ORLA: **Like** a cornflake sandwich. (3x05)

In (87) *like* functions as an exemplifier, one of the uses of this DPM reported by Amador-Moreno (2012). In the Clare subcorpus DPM *like* occurred more frequently: of 39 instances of *like*, 9 could be said to function as a DPM, with none of them occurring in sentence-final position either. The exemplifying function was also present in Clare's dialogue:

- (88) JAMES: Yes, I do! We have a lot in common, Clare.
CLARE: **Like** what? I mean, apart from the fact you appear to be wearing his clothes!
(1x02)

In the Clare subcorpus, DPM *like* appeared in initial position only in 2 instances, whereas medial *like* occurred 7 times, mostly as a focuser preceding new information (89)–(90), but also as an imprecision marker (91):

- (89) CLARE: I met this guy that went to her school, he said she had to leave for, **like**, being a bully. (2x05)
(90) CLARE: So... now that I think about it, weren't those computers, **like**, relatively new?
(3x01)
(91) CLARE: Yeah! I mean, obviously, I totally agree, I'd die for her as well. But I'm also conscious of the fact we've only known her, **like**, two days. (2x02)

Clare’s preference for more globalised uses of *like* can be interpreted as a fictional representation of the shift away from the more traditional uses of this DPM in IrE among young, urban IrE speakers that was reported by Nestor et al. (2012) among real-life IrE users, and by Terrazas-Calero (2022) in fictional representations of Dublin English. In order to assess if Clare and Orla were outliers or if the other characters would exhibit a similar pattern in their usage of *like*, a concordance search was conducted in the seven character subcorpora. Figure 4.1 below presents the proportion of use of *like* in initial, medial, and final positions for each of the characters.

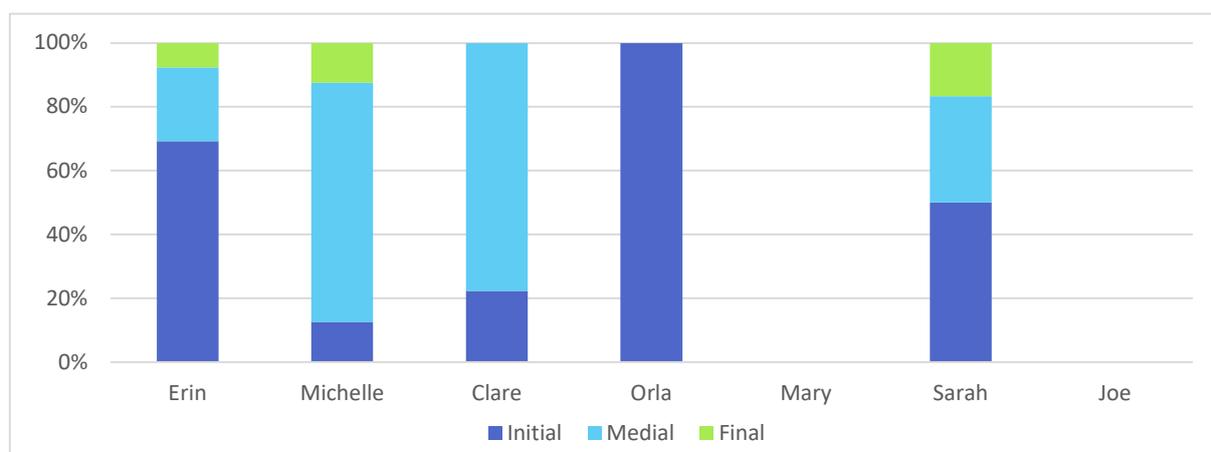


Figure 4.1. Proportion of use of DPM *like* in three sentence positions in each character’s dialogue

As can be seen, even though Erin and Michelle do employ final *like*, they still show a preference towards the less dialectal uses of *like*. Interestingly enough, Sarah exhibits a similar pattern, despite her being in her late thirties, which contrasts with the other two adult characters, Mary and Joe, who do not use DPM *like* at all in their dialogue, not even the more traditional final *like*. This could be interpreted as Sarah being more sensitive to language innovation than Mary and Joe. Sarah’s usage of initial and medial *like*, juxtaposed with her frequent use of *ach*, portrays her as someone whose idiolect is recognisably regional and, at the same time, attuned with changes that transcend dialectal borders.

On their part, the Derry Girls (Erin, Michelle, Clare and Orla) show a pattern that is coherent with adolescent speech, thus making them seem more realistic by grounding their dialogue in real-life linguistic phenomena. If we consider *like* in conjunction with their DPM usage frequencies as described in section 4.1, we can see that the speech of each Derry Girl begins to acquire distinctive traits. For instance, Clare’s total absence of final *like* in her speech

is accompanied by very low occurrences of *ach*, *och* and *wee* in her dialogue when compared to the remaining characters. Erin’s dialogue exhibits a similar pattern, although she does employ final *like*. In contrast, Orla employs *wee* and *och* frequently, despite never uttering *like* in final position, and therefore one could say that her idiolect is more recognisably NIrE from a DPM point of view. Finally, Michelle, in addition to her dialogue having the highest proportion of final *like* among the Derry Girls, employs *ach* most frequently in her friendship group and uses *wee* to a degree on par with Sarah, Mary and Joe. Therefore, like Orla’s, Michelle’s dialogue has a more NIrE profile than that of Erin or Clare’s.

4.2.2 Keyword analysis

Table 4.10. Keywords in each of the character subcorpora.

Key: the *green* keywords are terms of address; the *red*, taboo expressions; the *blue*, dialectalisms; and the *purple*, DPMs.

Erin	more	<i>Orla, mammy, Clare, Toto, Michelle, we</i>
	less	<i>Erin, fucking</i>
Michelle	more	<i>fuck, fucking, James, ride, Clare, dickhead, fanny, Macca, hole, Clive</i>
	less	<i>Michelle, God</i>
Clare	more	<i>Michelle, Erin, we</i>
	less	n/a
Orla	more	<i>Cracker¹², aunt, Jon, lion</i>
	less	n/a
Mary	more	<i>Da, Janette, Sarah, Gerry, punt¹³, cold, purse, O’Shea</i>
	less	<i>James</i>
Sarah	more	<i>Mary, eyebrows, Gerry, Ciaran, Brad, daddy</i>
	less	n/a
Joe	more	<i>son, bloody, Jim, ah, thon, Seamus, love</i>
	less	n/a

Note: the keywords in italics represent words that are only used in one episode by the character in question, often in short stretches of dialogue where the word is repeated several times in the space of a few conversational turns, so they have been mostly disregarded in the discussion that follows.

Having discussed the DPMs present in the character’s most frequent words, let us now turn to the keyword analysis of each character’s subcorpus, conducted as detailed in Chapter 3, section 3.2.1.3.

¹² In NIrE, *cracker* means “really good”. (MacQuillan, 2020)

¹³ The punt was the currency of the Republic of Ireland until the euro was adopted in January 1999. (Oxford English Dictionary [OED], 2023d)

Table 4.11. One-to-one character keyword comparisons.

Key: the *green* keywords are terms of address; the *red*, taboo expressions; the *blue*, dialectalisms; and the *purple*, DPMs.

		Comparison corpora						
		Erin	Michelle	Mary	Clare	Sarah	Joe	Orla
Target corpora	Erin		+Michelle +mammy +Orla - fucking - Erin	+ Clare + Orla + James - Sarah - Erin - ach	+ Clare + Orla - I	+ Clare + Michelle + mammy - ach	+ Clare + we + Michelle + mammy + James + Orla - son - love - ah	- Erin
	Michelle	+ fuck + fucking + Erin - Michelle - mammy - Orla		+ fuck + fucking + James + Clare - da - Sarah - God	+ fuck + fucking + Clare + him - Michelle - I - God - mean	+ fuck + fucking + James + Clare - Mary - sure - God	+ fuck + fucking + James + Clare + Erin - love - I'll - son - ah	+ fuck + fucking - cracker
	Mary	+ da + Gerry + Sarah + Erin + ach - Clare - Orla - James	+ Gerry + da + Sarah + God + daddy - James - Clare		+ da + Sarah + Gerry - Michelle	+ Sarah - Mary	+ da + Erin - son	+ da
	Clare	+ Erin + I - Orla	+ Michelle + I + God + mean - fucking - him	+ Michelle		+ Michelle - Mary	+ Erin + Michelle + we - love	n/a
	Sarah	+ Mary + Gerry + ach + Ciaran - Michelle - mammy	+ Mary + Gerry + daddy + sure + God - James	+ Mary - Sarah	+ Mary + Gerry - Michelle		+ Mary - son	+ Mary
	Joe	+ son + Gerry + bloody + Mary + love + ah - we - Orla	+ love + Gerry + I'll + bloody + son + Jim + ah + boy	+ son	+ son + love + Gerry + bloody - we	+ ah + son + bloody - Mary		n/a
	Orla	+ Erin + Mary + cracker	+ cracker	+ cracker	+ cracker	- Mary	+ Erin	

The first set of results can be seen in Table 4.10 above, which presents the words that every character is more and less statistically likely to say when compared to the rest of the show's dialogue. The resulting keywords were grouped and, if they were considered for analysis, colour-coded according to the type of linguistic item they represented; each type will be discussed below. After comparing the character subcorpora against the remaining dialogue in the DGC, I decided to compare all of these characters against each other to further uncover any interesting phenomena. Keyword analyses for each character pairing were carried out twice so that each character could be both target and comparison in the test. Table 4.11 above presents the results, with positive keywords preceded by a plus sign (+) and negative keywords preceded by a minus sign (-).

Before proceeding to analyse the keyword categories, we must mention that the differences in subcorpora size act as a distorting factor for the keyness analysis results. This affects Orla the most: for instance, despite *Erin* being her most commonly used term of address (as can be seen in Table 4.9), the fact that characters with a higher amount of dialogue also mention her a lot entails that Erin does not show up among her keywords.

4.2.2.1 Character names and other terms of address

As can be seen, character names, frequently used, but not exclusively, as vocatives, are common in Tables 4.10 and 4.11 and we can use them to map the relationships between the characters. Figure 4.2 presents a diagram of these character connections as represented by each character's keywords.

Visualising the keyword relationships between characters like this enables us to identify two groupings of characters as defined by their connections:

1. The teenage Derry Girls: Erin, Michelle, Clare, Orla and James, with the latter being excluded from analysis due to his being English.
2. The adults in the Quinn-McCool family: Mary, Sarah, Joe, and Gerry, with the latter being excluded from analysis due to his being from the ROI.

Each of these groups has a character that is well connected within their group while also acting as a node point between their group and the other: Erin and Mary. Consequently, they could be considered as the centre of this character network. Let us now examine each of the character groups, beginning with the Derry Girls. This group seems to be divided along family dynamics via the cousin pairs Erin-Orla and Michelle-James. However, we can also see that Erin,

Michelle and Clare had each other's names among their keywords, which highlights the higher frequency with which they address or discuss each other compared to other characters, so they could be regarded as the core members of their friendship group. Orla, in contrast, only appeared on Erin's keywords, suggesting that she is a more peripheral member of the group, whose main tie to them is being Erin's cousin. The same happens to James, Michelle's cousin, who is only present in Erin and Michelle's keywords.

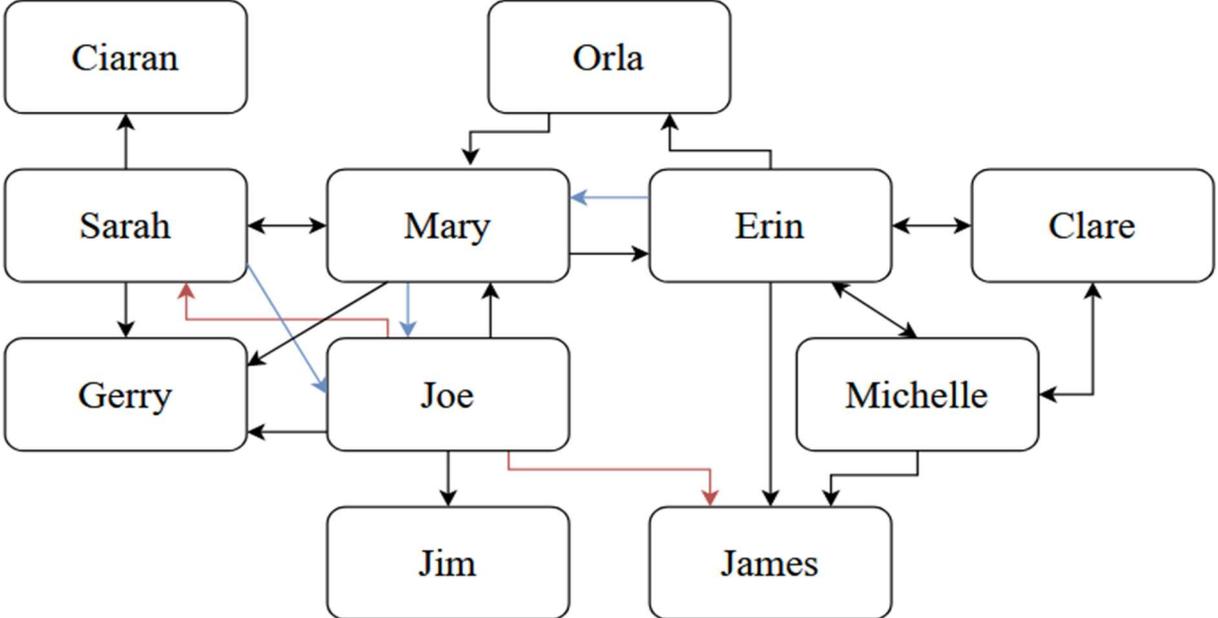


Figure 4.2. Character relationship map as defined by their keywords
 Note: The arrows point towards the characters that appear in the other character's keywords. The black arrows represent first-name keywords. The blue arrows represent non-first-name keywords that refer exclusively to one character. The red arrows represent more indirect mentions, i.e., via vocatives that are not first names and that do not always refer to the same character.

Beyond their friendship group, it can be seen that Erin's relationship with her mother, Mary, is prominent in the dialogue, as *mammy* was a keyword for Erin, and *Erin* for Mary. Of the 43 times she employs this word throughout the corpus, 40 refer to her mother. On her part, Orla notably had *aunt* among her keywords, a term that in her dialogue always refers to her aunt Mary. This makes Orla part of the four characters whose keywords cross the adults versus teens divide, the other three being Mary, Erin, and Joe. Regarding their negative keywords, the Derry Girls in general were found to be less likely to employ *Sara* and *Mary*, highlighting once more the generational divide. In addition, *Orla* appeared in Michelle and Clare's negative keywords when compared to Erin's dialogue, thus reinforcing Orla's standing on the fringes of her friendship group.

Turning now to the second group of characters (Mary, Sarah, and Joe) one can see that the Derry Girls were absent from their keywords list, even though Erin and Orla are Mary and Sarah's daughters, respectively, and Joe's granddaughters. In Mary and Sarah's lists, each other's names were found, as well as, *Gerry*, Mary's husband's name. In addition, Mary's list contained *da*, and Sarah's, *daddy*, each character's preferred term of address for Joe. Finally, *Ciaran*, Sarah's boyfriend's name, appeared on her list. The similarities in Mary and Sarah's lists regarding terms of address highlight their close-knit relationship and the high overlap between their connections to other characters, barring Ciaran. If we consider Joe's keyword *love*, which he uses to address both his daughters, as I will discuss in more detail below, we can see that Mary, Sarah and Joe also form a trio of closely connected characters, in the same vein as Erin, Michelle and Clare. The fact that Mary and Sarah had keywords that refer to Gerry points to his also being part of the adult core group of characters, but the study of his dialogue is beyond the scope of this thesis.

In contrast to most of the other characters', Joe's keyword list is conspicuous for its almost total absence of first names, except for *Jim*, the name of a neighbour who is also an elderly man like him. But once we examine the vocatives present in it, we find that Joe's connections cross the generational borders. Instead of names, what we find in Joe's keywords are the terms of endearment *love*, *son* and *boy* (this last one only appearing in Table 4.11, when Joe's dialogue is contrasted with Michelle's), whose usage I believe merits a closer analysis. Starting with *boy*, I must note that it is employed by Joe almost exclusively (in 7 out of 9 occurrences) to address Gerry, his son-in-law; the remaining two instances are directed to Ciaran, his daughter's boyfriend. In contrast, Joe never employs *son* to refer to these two male characters: out of the 18 occurrences of this word in Joe's dialogue, a third of them refer to James, with the rest addressing minor male characters with whom Joe has little relationship. If we delve further by looking at the contexts in which each of these two terms of address appear, a pattern begins to emerge:

- (92) JOE: Hey **boy**, it's dinner she invited you for, not an orgy. (1x06)
- (93) JOE: Don't you take that tone with me, **boy**. (1x05)
- (94) JOE: Just what the hell is your problem, **boy**? (3x01)

The language employed by Joe in the examples above ranges from vaguely hostile (92)–(93) to openly antagonistic (94). Watched in the context of the episodes, none of these lines is said by Joe in jest, which suggests that Joe’s feelings as denoted by the dialogue are genuine.

(95) JOE: [James]’s still a fella. There’s still a good chance that he’s a rapist. (*To James*) I mean no offence, **son**. (1x02)

(96) GABRIEL: Yeah, that’s right. I’m from Kildare originally.
JOE: Very sorry to hear that, **son**. (3x02)

(97) JOE: Listen, listen, have you killed anyone, **son**? (1x05)

Joe’s use of *son*, as shown in examples (95)–(97) presents a different picture. In contrast to *boy*’s hostile environments, *son* appears with politeness formulae in four instances, as exemplified in (95)–(96). In (95), the “no offence, son”, addressed to James, follows an insult from Joe (“bastards” and “rapist”), so it constitutes an apology. The only time where Joe uses *son* in a line that could be considered offensive, shown above in (97) makes sense in the context of its episode, 1x05, where a member of the IRA hides in the boot of the family car to get across the Irish border and is discovered by the main characters. Joe is sympathetic to the IRA member and wants to drive him across the border; therefore, in his quest to aid the fugitive, Joe asks the man if he has killed someone with the aim of convincing the others that the IRA member is not a dangerous man.,

It can be seen that *boy* is employed in a more antagonistic way by Joe than *son*. What can we make then of the fact that he uses the former term exclusively to refer to his daughters’ male partners, who are never addressed as *son*? To shed further light on Joe’s relationship with Gerry and Ciaran, we can search for their names in the Joe subcorpus. The results show that Joe never utters Ciaran’s name onscreen; Gerry, however, appears 14 times in Joe’s dialogue.

(98) JOE: There’s nothing wrong with the Kimberleys, Carlos. **Gerry** here is just a bit of an arsehole. (3x04)

(99) JOE: You know what is wrong with you, **Gerry**? You’re an awful wuss. (1x05)

(100) GERRY: You’re looking well, Joe.

JOE: It’s not all shite you talk, **Gerry**. (2x05)

In his dialogue, Joe refers to his son-in-law Gerry with a variety of abusive terms, as shown in (98)–(99). Even Joe’s compliment to Gerry (100) is backhanded. This language suggests that Joe’s opinion of Gerry is quite negative, which reinforces the negative associations of the vocative *boy* as employed by Joe, given that it’s almost always used to refer to Gerry. From this, we can extrapolate that Joe’s opinion of Ciaran is not better, for he is always addressed as *boy* by Joe. The fact that Ciaran is never mentioned by name by Joe is easily explained by Ciaran’s more peripheral role both in the show and Joe’s life: Gerry is the protagonist’s dad, appears in every episode, and lives in the same house as Joe; in contrast, Ciaran is introduced at the end of season 1 and appears in only four episodes.

The final term of address that appears on Joe’s keyword list is *love*. This vocative, which Joe employs exclusively to address female characters, appears 19 times in the Joe subcorpus. Of these, 9 instances refer to Mary, 4 to Sarah, 1 to Erin and none to Orla. In contrast, Joe addresses Mary by name 11 times, but Sarah only twice, and Orla only once. Erin, Joe’s other granddaughter, does not appear by name at all in Joe’s dialogue (Joe appears under *granda* and *grandad* in Erin’s dialogue 9 times, with 7 of them being vocatives). This suggests that Joe’s character interacts more with his daughters than with his teenage granddaughters, at least verbally. Of the 19 instances of *love* in Joe’s dialogue, only two are more negatively coloured (101)–(102), both being addressed to Mary. The remaining occurrences appear with politeness formulae, as in (103), or in the same utterance as adjectives that suggest a positive evaluation such as *nice* (104) or *grand* (105):

- (101) JOE: (*to Mary*) No chance, **love**. (1x02)
- (102) JOE: (*to Mary*) There’s no excuse for it, **love**. (3x04)
- (103) JOE: (*to Mary*) Would you not reconsider? Please, **love**. Just the once. (1x06)
- (104) JOE: (*to Sarah*) Oh, that’s very nice, **love**. (1x06)
- (105) JOE: (*to Mary*) It’ll be grand, **love**. (2x04)

Thus far, Joe’s interactions with female characters, and with his daughters Mary and Sarah in particular, seem rather positive. For comparison’s sake, let us consider the use of Mary and Sarah as vocatives in the Joe subcorpus. *Mary* appears 12 times in Joe’s dialogue, while *Sarah* occurs only twice. Some occurrences are shown here:

- (106) MARY: That's enough, **Da**.
 JOE: The tight bastard is trying to starve us all, **Mary**. (1x02)
- (107) JOE: **Mary**, come on. Would you not reconsider?
 MARY: No, **Da**.
 JOE: Please, love. Just the once.
 MARY: I'm not letting you hit him, **Da**. (1x06)
- (108) MARY: Ach, **Da**, not the bingo thing, still!
 JOE: She's a cheating old bitch!
 JOE: Her nephew brought her this pen back from New Jersey, it changes the numbers. I'm telling ya! It's witchcraft, **Mary!** (1x03)
- (109) MARY: I'm not interested in your coloureds, Sarah. It's darks I need. Oh, listen, **Da**.
 JOE: Do not be starting at me about darks, **Mary**. I've given you whatever darks I have. Stick on a half load and be done with it, woman.
 MARY: A half load goes against everything I stand for. You know that, **Da**. (1x06)
- (110) SARAH: **Daddy**, how many times? You can't strim the grave of another man's wife. It's inappropriate.
 JOE: You could rear cattle in it, **Sarah**. (3x04)

As can be seen, these exchanges contain some abusive language on Joe's part (*bastard* in [106], *bitch* in [108]), but none of it is directed towards his Mary or Sarah. However, in all these interactions Joe disagrees with his daughters. There seems to be a correlation, therefore, between Joe's use of his daughters' first names as vocatives and their exchanges being more confrontational than when he addresses his daughters with *love*.

4.2.2.2 Taboo language

After first names and other terms of address, the next prominent category of keywords in the seven characters' dialogue is taboo language. With regard to positive keywords, Michelle and Joe have taboo expressions in their respective lists, as seen in Table 4.10. The latter counts the intensifier *bloody* among his keywords, whereas the former has six: *fuck*, *fucking*, *ride*, *dickhead*, *fanny* and *hole*. *Ride* is an Irish slang term for sexual intercourse (Terrazas Calero, 2022), and by extension a noun that designates sexually attractive people, as depicted in example (111) below. On its part, *hole* is a synonym for *arse*, as can be seen in examples (112)–(113), whereas *fanny* is slang for female genitalia (114):

- (111) MICHELLE: A French fella. That's what I want to do. Nation of **rides**. My fanny is going funny just thinking about it. (1x02)
- (112) MICHELLE: She can stick her party up her **hole**. (3x07)
- (113) MICHELLE: Kiss my actual **hole**, Erin. (3x02)
- (114) MICHELLE: Tell me you haven't just invited girls from school. There'd better be fellas at this thing, otherwise it's gonna be a total fucking festival of the **fanny**. (3x07)

Michelle's relatively frequent use of *ride* and *fanny*, terms with sexual associations, characterises her as more sexually forward than her friends, something that can be corroborated in Michelle's interactions with them:

- (115) ERIN: I'm not sure about [partying with the boys we have met].
 MICHELLE: I knew you'd fanny out!
 ERIN: **These lads have moves. You said so yourself, Michelle. And I haven't got any moves!**
 MICHELLE: Look, let's just get in there, have a few drinks and just loosen up a bit. (2x01)
- (116) CLARE: Stop saying lezzie! Just because I'm gay doesn't mean I see every other gay woman as a potential suitor.
 MICHELLE: Aye, keep saying things like "suitor", Clare. **That'll help you offload the old virginity.** (3x06)

As we can see, Michelle considers herself more knowledgeable and worldly than Erin and Clare in terms of sexual relations; therefore, it is understandable that sex-related terms appear in her keywords. Interestingly, Orla is always left outside these discussions; throughout the show, she does not have any romantic interests whatsoever. In episode 2x05, where the girls go to prom, she takes her grandfather, Joe, as her date. In contrast, the process of finding a date is a major plotline for the other girls.

Let us now turn to other taboo keywords. In Table 4.11, *fuck* and *fucking* appeared in Michelle's positive keywords when her dialogue was contrasted against the other six characters, and the same can be said for Joe and *bloody*. In addition, the taboo term *God* appeared in Mary, Sarah, and Clare's positive keywords, and in Michelle's negative ones when contrasted to these other three characters. Erin and Clare also counted one taboo term among their negative keywords, *fucking* in both cases, which appeared when their dialogue was compared to

Michelle's. This suggests that there might be a dichotomy between using religious taboo expressions and non-religious taboo expressions. The dialogue in *Derry Girls* provides some metapragmatic commentary on the matter:

(117) MICHELLE: My **fanny** is going funny just thinking about it.

ERIN: Could you not use that word, Michelle?

MICHELLE: What, **fanny**?

CLARE: Why do you always have to be so coarse? (1x02)

(118) SEAN: Does every **bastard** in Derry own this thing?!

GERRY: It would seem so.

GERALDINE: [*to SEAN*] **Language**, you! **Jesus**, girls, I'm ready to drop here. I haven't had a bite since 11. Any sign of Janette? (3x05)

In these two examples, characters are reprimanded for their use of 'inappropriate' language. Example (117) concerns a conversation between teenage girls, whereas (118) is an excerpt of a scene with adult characters. Therefore, it can be seen that the distaste for taboo language is not solely confined to older characters, an age group that has been found to favour swearing less in studies conducted on real-life speakers (Murphy, 2009). In addition, Geraldine in (118) exemplifies the different attitudes regarding the two kinds of taboo language mentioned at the beginning of this subsection: she admonishes SEAN for his use of *bastard*, followed immediately by her using *Jesus* to start a sentence. This suggests that, for Geraldine, religious expressions are not censurable, whereas other kinds of taboo language are.

This double standard regarding taboo language suggests that a character who employs non-religious taboo language frequently, like Michelle, is comfortable defying social norms, thus aiding in her characterisation as a more rebellious teenager when compared to her friends Erin and Clare, who have *fucking* among their negative keywords. Alternatively, it could be argued that by swearing frequently, Michelle conforms to typical adolescent linguistic behaviour, as this age group has been found to swear more than older speakers, barring young adults (Schweinberger, 2018, pp. 3–4). In contrast, the fact that Clare counts God among her positive keywords aligns her with Mary and Sarah, two older characters, suggesting that Clare has more conservative or adult views regarding swearing, thus portraying her as a more atypical adolescent. This would leave Erin as the mid-point between Michelle's linguistic teenage rebelliousness and Clare's propriety.

Let us now delve more deeply into the demographics of *fuck* and *fucking*, in order to assess the extent of Michelle’s uniqueness among the *Derry Girls* characters. A concordance search in the Michelle subcorpus revealed that *fuck* occurs 74 times in this character’s dialogue, or 90.96 times per 10,000 words. *Fucking* was only slightly less frequent in the Michelle subcorpus, occurring 72 times overall, or 88,51 times per 10,000 words. For reference, *fuck* occurred 89 times overall in the DGC, and *fucking* appeared 104 times, meaning that 83% of *fuck* instances and 69% of *fucking* instances corresponded to Michelle’s dialogue alone. Michelle’s outlier status as far as these two swearwords are concerned is even more evident when we compare the frequencies of *fuck*, *fucking* and other related terms (*fucked*, *fucko*, etc.) in the subcorpora that include Michelle’s dialogue (Female, Female NIrE, and Young female), to versions of these subcorpora with her dialogue removed, created *ad hoc* for this comparison. The results are presented in Figure 4.3 below (for the raw frequencies see Table B.1 in Appendix B):

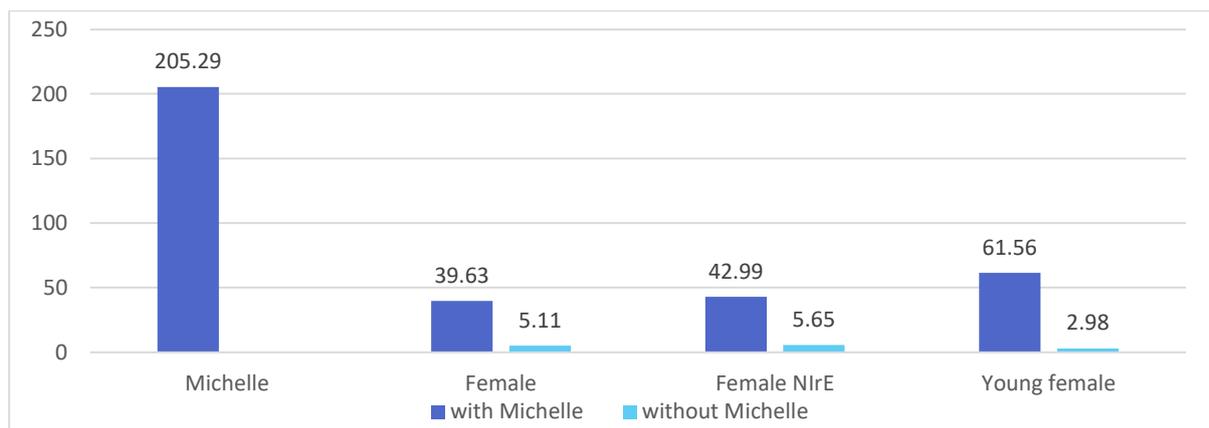


Figure 4.3. Normalised frequencies of *fuck* (per 10,000 words) in the DGC with and without Michelle’s dialogue

As can be seen, removing Michelle’s dialogue from the subcorpora decreases drastically the frequency of *fuck* and its derived terms, reinforcing the idea that Michelle is an outlier in matters of swearing. In addition, the significance tests found all these differences significant (see Appendix B, Table B.2). For further comparison, Figure 4.4 below presents the frequencies of these terms in the remaining DGC subcorpora and the partitions of SPICE-NI:

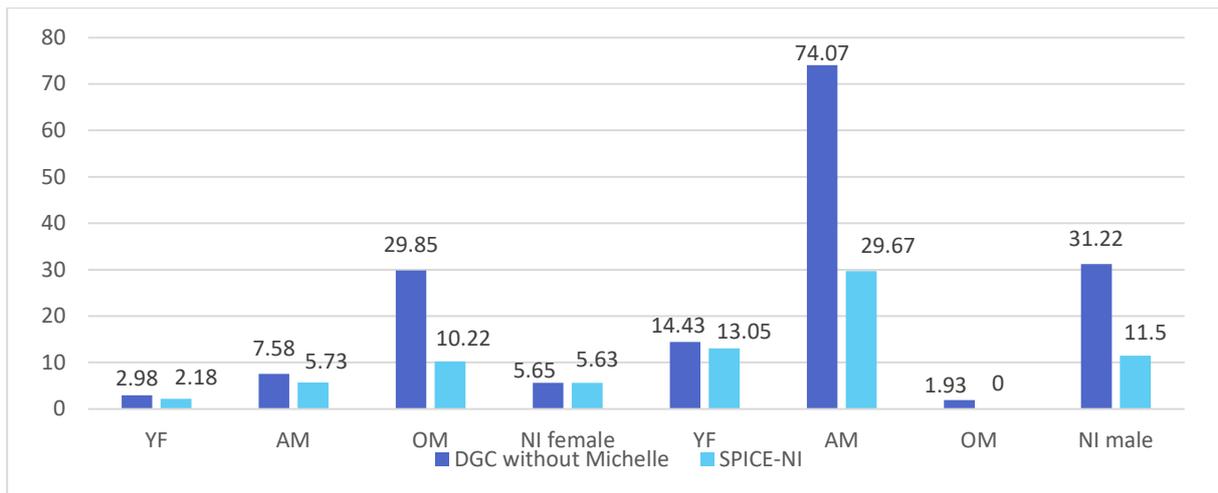


Figure 4.4. Normalised frequencies of fuck per (10,000 words) in DGC and SPICE-NI

As the table shows, without Michelle’s dialogue skewing the data, the frequency of *fuck* and its derived terms followed a similar pattern in both corpora, with these swearwords being more common among males than among females, a gendered difference that was found to be statistically significant by the LL tests (see Appendix B, Table B.2). In addition, in both corpora females used *fuck* more the older they were, while for males the usage was highest for adults and lowest for elderly speakers. Michelle’s disproportionate usage of these swearwords, consequently, contributes to her characterisation as a rejecter of standard social norms, given her penchant for using language that is disapproved of and more in line with adult male patterns of speech. At the same time, we must remember that swearing has been found to carry covert prestige among teenagers and members of lower socioeconomic classes (Stapleton, 2010), both of which apply to Michelle; therefore, her frequent swearing also portrays her as a socially aware teenager with street smarts, as mentioned above. In contrast, Erin and Clare’s reluctance to employ non-religious taboo language could suggest that these two characters want to distance themselves from linguistic uses that do not carry overt prestige, considering that both of them have academic aspirations that could take them outside of their region:

(119) CLARE: Passing those exams was our only chance. We’re girls, we’re poor, we’re from Northern Ireland and we’re Catholic, for Christ’s sake!

ERIN: Oh, my God, she’s right! What type of future will we have?! (3x01)

The next taboo keyword, *dickhead*, also comes from Michelle’s dialogue. This term is mainly employed by this character to express her disapproval of others:

(120) JENNY: I'd quite like some feedback on the poetry assignment, actually.

MICHELLE: [*whispering*] **Dickhead**. (2x02)

(121) RITA: And the mouthy curly one thinks we shouldn't [go back to find James] because..?

MICHELLE: Because he's a **dickhead**, Rita. (2x03)

Of 9 occurrences of this term in the Michelle subcorpus, 4 refer to James, Michelle's cousin, which points to their strained relationship:

(122) MICHELLE: You're so lucky, Erin. I begged me ma to let me have [an exchange student], but she said we've no room now that my **dickhead ballbag English prick of a cousin** has moved in. Talking about you, James, in case you're wondering. (1x04)

In (122), Michelle explicitly addresses her frustration at her having to live with James. The string of abuse she directs at him, with *dickhead* at the forefront, is on par with her usual treatment of him.

Let us now consider the remaining taboo keyword, *bloody*, whose usage is monopolised by Joe in a similar vein to Michelle's dominion over *fuck* and *fucking*. Indeed, in the DGC, Joe is responsible for 13/26 instances of *bloody*, or 50%. In order to ascertain if this speech pattern is in line with Joe's gender and age or if he is also an outlier, a search was conducted in the demographic partitions of both the DGC and SPICE-NI. The results are presented in Figure 4.5 below (for the raw frequencies, see Table B.3 in Appendix B):

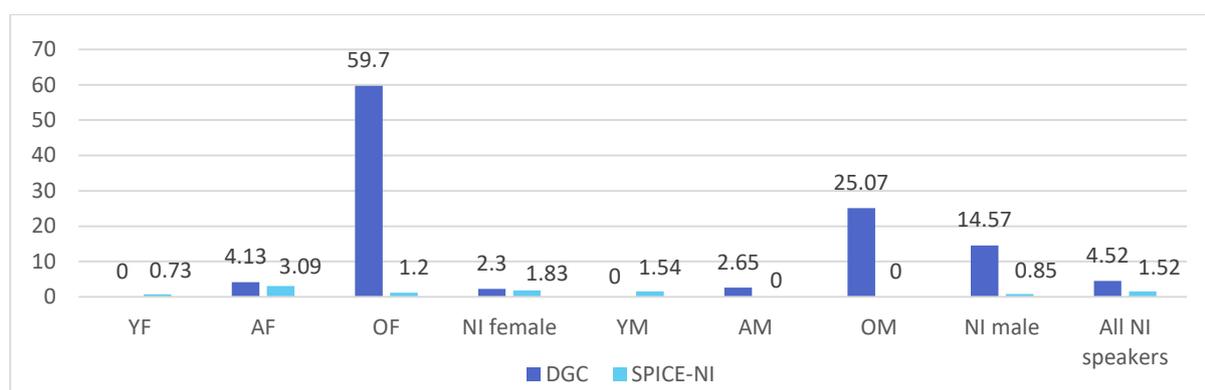


Figure 4.5. Normalised frequencies (per 10,000 words) of *bloody*.

Looking at age groups, no patterns could be discerned in SPICE-NI, whereas in the DGC it seems that the older a speaker was, the more likely they were to use *bloody*. However, we must

bear in mind that the small size of the OF partition in the DGC entails that every instance of a word turns into 14.93 occurrences per 10,000 words, and Joe, i.e., only one character, is responsible for all the occurrences of *bloody* among old male speakers in the DGC. Therefore, there is not enough evidence to assess whether Joe is an outlier or not. Although the literature on swearing has found that elderly speakers are less prone to swearing (McEnery & Xiao, 2004, p. 242; Schweinberger, 2018), taboo language is more common among working-class male speakers, such as Joe. Therefore, it appears that for him, his gender and socioeconomic class might trump his age in matters of swearing. Interestingly enough, however, Joe does not approve of the use of taboo language by young female speakers (123), which suggests that he might hold onto more traditional values regarding gender and swearing.

(123) EXCHANGE STUDENT: I am finished talking. She is racist, patronising **bitch**.

ERIN: I am *a* racist, patronising **bitch**.

JOE: Now, come on, girls. **Language**, please. (1x04)

Looking at Figure 4.5 again, if we consider gender groups, it seems that in SPICE-NI, *bloody* was slightly more frequent among female speakers, although the difference was not found statistically significant (see Appendix B, Table B.4 for the LL results). In contrast, in the DGC male speakers led the usage of this taboo intensifier. However, the vast majority of occurrences in the male dialogue are attributed to Joe, which skewed the figures considerably. Without his contribution, it was females that employed *bloody* most commonly, although the small size of the OF subcorpus, which was the DGC partition where *bloody* was most frequent, also distorts the normalised frequencies of this element. In any case, Joe's prominent use of *bloody* also points to him as a character who disregards, to an extent, the societal disapproval of swearing, although it is somewhat prestigious within his socioeconomic class. It could be argued that Joe in some respects is the Michelle of the adult character group, although he would disapprove of her swearing habit, judging by his dialogue in (123) above.

4.2.2.3 Dialectalisms

We now turn to a less crude group of keywords: words associated with NIrE. Only three examples appeared in Tables 4.10 and 4.11: *thon*, among Joe's keywords, *cracker*, in Orla's and *ride*, in Michelle's. The latter has already been discussed in the previous subsection. *Thon* is a demonstrative adjective and pronoun which points to something more remote than *that* (OED, 2023e), used in Scotland and NI, seemingly with both singular (124) and plural (125)

referents. In the DGC, this demonstrative occurred 13 times, of which 7 corresponded to Joe's dialogue.

(124) JOE: I've read through **thon** introduction thirty bloody times. (3x07)

(125) JOE: Aye, **thon** boys would get in where draughts wouldn't. (2x03)

The remaining 6 occurrences were uttered by his brother, Colm, and Joe's sister-in-law. These three characters are in their old age, which suggests a possible link between *thon* and older demographics. A search in SPICE-NI yields 4 occurrences of *thon*: of these, 3 correspond to old speakers, while the remaining one is said by a young female. This last occurrence seemed unusual, and thus I examined in more detail the corpus file that contained it (P1A-040). The file represented part of a conversation between two old speakers and two young ones, where one of the older speakers also employed *thon*, albeit later in the conversation. It could be the case that the younger speaker borrowed the idiolect of their older interlocutors in this conversation, as all the other occurrences of *thon* corresponded to elderly speakers. Returning to Joe, the fact that one of his keywords is a dialectalism that is potentially associated with older speakers helps characterise him as a traditional older Northern Irish man.

Regarding *cracker*, an adjective meaning "really good" (MacQuillan, 2020), Orla was responsible for 10 of 13 occurrences of the term in the DGC, the remaining 3 corresponding to Michelle, Sarah and Erin. In SPICE-NI, *cracker* appears only once, nominalised:

(126) B: The [story] in between Fred and Gerry. [*Laughter*] It's a **cracker**. (P1A-044)

In this example above, *cracker* is uttered by an adult speaker. From this and the evidence from the DGC, one cannot say that *cracker* is associated with any particular age group. However, the fact that this term is a keyword for Orla might point to her having an optimistic and cheerful disposition. In addition, as discussed in section 4.1.3.1, Orla employs *wee* the most among all studied characters, even if this dialectalism did not appear among her keywords. Her frequent use of two terms associated with NIrE helps characterise Orla as a member of this speech community, and the fact that she employs them more often than other characters contributes to shaping her idiolect.

4.2.2.4 DPMs

Let us now examine the last group of keywords, DPMs. Although section 4.1 of this chapter already examined the uses of *ach*, final *but*, and *wee* in the DGC and SPICE-NI, the presence of DPMs in the keyword search results merits a closer analysis at the individual character level, not the least because most of the DPMs in the results were not the ones analysed above.

Among the DPMs present in Tables 4.10 and 4.11, *ach* is the only one that had been touched upon in section 4.1. In Table 4.11, this DPM appears in Mary and Sarah's positive keywords when their dialogue is contrasted with Erin's. As discussed in section 4.1.1.1 above, Mary and Sarah are the most prolific users of this DPM, and in SPICE-NI it also features most frequently in the speech of adult women. Mary and Sarah's conspicuous use of a DPM characteristic of NIRE aids in grounding them as typical Northern Irish women. In addition, Erin's statistical avoidance of *ach* helps contrast her speech with her mother and aunt's and it could signal that her identity as a Northern Irish girl is still not fully developed or that she has a more cosmopolitan outlook than her elder relatives, as discussed in section 4.2.1 above. For comparison, and as shown in Table 4.1 in section 4.1.1.1, *ach* is totally absent from Clare's dialogue, whereas Michelle is its most prolific user among the Derry Girls. Clare is characterised as academically inclined and more rule-abiding than her friends, as shown in examples (127)–(128):

(127) CLARE: Of course it's your fault! You've dragged me down to your level, your stupidity has finally rubbed off on me. I was a scholar when I met you, Erin, a scholar!
(3x01)

(128) ERIN: Oh my God, I passed, I passed!

JAMES: Me too!

MICHELLE: Me three!

MICHELLE'S MUM: Aye, by the skin of your teeth, Michelle.

CLARE: Straight As, straight As! (3x01)

Consequently, Clare's avoidance of dialectalisms might stem from her educational aspirations and having internalised the stigmatisation of English dialects that deviate from the standard. In contrast, Michelle shows lower academic expectations in her dialogue (128), and her frequent swearing, already discussed, reveals her as a character sensitive to the covert prestige of stigmatised language forms, which can explain her more frequent use of *ach*. Once more, Erin

aligns more with Clare than with Michelle, as the occurrences of *ach* in Erin’s dialogue was just 0.97 per 10,000 words, whereas Michelle’s was 12.29. Perhaps not coincidentally, this was the same pattern regarding the use of final *like* among these teenage characters, as discussed in section 4.2.1 above.

Continuing with Clare’s linguistic characterisation, we now turn to one of her keywords present in Table 4.11 when her dialogue was compared to Michelle’s: *mean*. A concordance search in Clare subcorpora showed that, in 21 of 29 occurrences *mean* was used as the DPM *I mean*. Not being associated with any particular dialect of English, the uses of *I mean* are rather varied. Kallen (2005b) notes its use as an upgrader, i.e., an emphatic device, whereas the *OED* classes it as a filler (2023c), and the Merriam-Webster dictionary points to its use as a reformulation marker and as a marker of being “unsure of what to say or how to say it” (n.d.). Within Clare’s dialogue, I could identify 10 instances of *I mean* used as an emphatic device, as exemplified in (124), one as a reformulation marker (125) and five signalling Clare’s being unsure (126):

(129) CLARE: I want no part of this! **I mean**, we’re stealing a body! (1x02)

(130) CLARE: Well, that’s a start. **I mean**, not a great start. (3x07)

(131) CLARE: [*to the girl she likes*] Is that a joke? **I mean**, cos it’s fine. If it is a joke. **I mean**, it’s hilarious, but I just wanted to check in case it isn’t. (3x07)

As stated before in this section, the fact that Clare seems to prefer non-dialectal and non-disapproved-of linguistic forms might point to a desire to distance herself from what she could see as a unfavourable background, or simply to her tendency to follow the rules. In addition, her signature DPM being a marker of hesitation or uncertainty helps characterise her as someone prone to doubt and worry, and aware of it, as shown below:

(132) ERIN: Everything makes you nervous, Clare. You’re a walking cack attack.

CLARE: I can’t help that. It’s the way God made me. (1x06)

We now turn to Sarah, among whose keywords in Table 4.11 is *sure*. A concordance search in her subcorpus showed that in 17 of 21 occurrences, this item was functioning as a DPM. As discussed in Chapter 2, section 2.4.2.2.1 *sure* has been identified as the most common DPM in IrE, and among its functions are softening disagreement and appealing for consensus. In Sarah’s

dialogue, *sure* functions as a softener in 2 instances, as exemplified in (133), and as an appeal for consensus in 12, as represented in (134). *Sure* also seemed to be used as an emphasiser on 3 occasions (135):

- (133) SARAH: [*of a hair dye*] The label said mahogany. **Sure**, that’s barely oak. (2x03)
- (134) SARAH: **Sure**, didn’t I predict this? (1x05)
- (135) SARAH: The last time I went to see a psychic, **sure** I was tortured. The spirits were coming at me left, right and centre. (3x04)

A search for *sure* in the DGC and SPICE-NI partitions was performed, which yielded the following normalised frequencies (the raw frequencies are available in Appendix B, Table B.5):

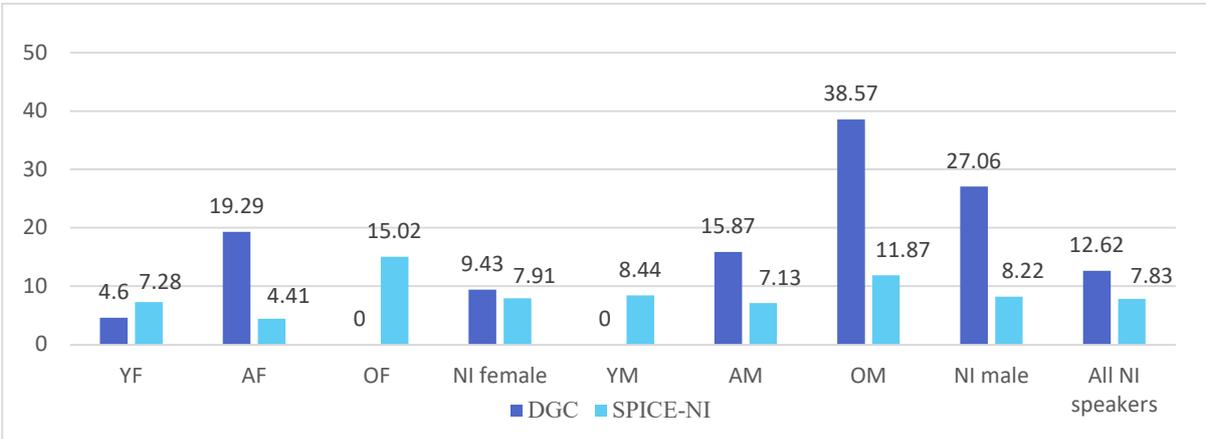


Figure 4.6. Normalised frequencies (per 10,000 words) of DPM *sure*

The absence of *sure* among old females and young males in the DGC might be due to the relative lack of dialogue these two groups have in the show. Apart from these anomalies, elderly speakers seemed to be the more prolific users of *sure* across both corpora, and indeed when the frequency difference between old speakers and the rest of the speakers was tested for significance the results were positive for both DGC and SPICE-NI (see Appendix B, Table B.6 for the data). Sarah being a prominent user of this DPM, then, cannot be attributed to it being most common among adult females. Perhaps her use of this characteristically IrE DPM, together with her frequent use of *ach* indexes her (Northern) Irish identity, as was argued in section 4.1.1.2. In addition, her frequent usage of *sure* as an appeal for consensus conforms to the pragmatic profile of IrE, which favours agreement and social cohesion (Hickey, 2007, p. 372).

The last DPM present in the results of the keyword searches comes from the Joe subcorpus: *ah*. Of 50 occurrences of this interjection in the DGC, 14 corresponded to Joe’s dialogue. The multifunctionality of *ah* is evident in its entry in the *OED* (2023a), where it is reported that this interjection can express a wide array of emotions, both positive and negative, as well as function as a filler, and signal disagreement. In Joe’s dialogue, *ah* does not seem to express positive emotions, ranging instead from neutral utterances (136), through more negative ones (137), to rather confrontational dialogue (138):

- (136) JOE: [*giving money to a train employee*] I’ve got it. Here, son. **Ah**. Here, hold on, son. You’ve done yourself. That’s 50p too much. (3x03)
- (137) AIDEEN: How’s life been treating you?
JOE: **Ah**, not great, I’m afraid. I caught dementia. (3x03)
- (138) GERRY: It’s not my fault.
JOE: **Ah**, sure, nothing ever is. (2x02)

That a DPM with an expressive range as wide as *ah* is never used by Joe to express positive emotions helps portray him as a rather grumpy character. Taken together with his use of *boy* to address his daughters’ partners, one can get the sense that Joe can be painted as a rather unpleasant character, depending on whom he interacts with.

4.2.2.5 Keyword analysis of character groups: Derry Girls against Adults

Having compared the seven characters against the rest of the DGC and each other, one last keyword search was run, this time based on the two groups of characters that emerged from analysing the connections in their vocative keywords as mapped in Figure 4.1 above. Once more, the test was run twice, with each group as both target and comparison. Table 4.12 presents the results:

Table 4.12. *Keyword comparison of character groups.*

Key: the green keywords are terms of address; the red, taboo expressions; and the purple, DPMs.

Derry Girls	James, Clare, fucking, fuck, Michelle, yeah, ma, mammy, we
Adults	Gerry, Mary, Da, <i>Janette</i> , love, Sarah, bloody , Deirdre, son, Ciaran, <i>Maeve</i> , <i>ach</i> , <i>O’Shea</i>

Note: The keywords in italics are confined to just one episode, so they have been disregarded.

We can see that only two new keywords surfaced in this test: *Deirdre* and *yeah*. *Deirdre* is Michelle’s mum, neighbour and friend to Mary and Sarah, hence her name being relatively more frequent in the Adults’ dialogue than in the Derry Girls’. The other keyword, *yeah*, is a form of *yes* that originates from US English (OED, 2023f). *Yeah* can be contrasted to another word used to express agreement, *aye*, which is associated with IrE, among other regional varieties of English spoken in the British Isles (OED, 2023b). Let us compare the frequencies of *aye* and *yeah* between both character groups:

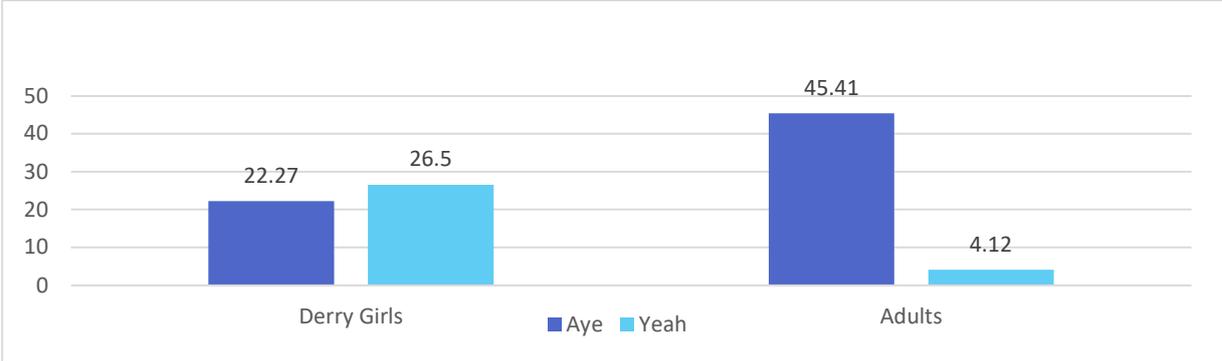


Figure 4.7. Normalised frequencies (per 10,000 words) of *aye* and *yeah*

As can be seen in Figure 4.7, while the Derry Girls employed both terms with a similar frequency, the Adults showed a clear preference towards *aye*. This could mean that the Adult characters were too old to adopt *yeah* when it expanded outside of the US, their linguistic patterns being more established and conservative, more strictly Northern Irish. In contrast, while the Derry Girls retain the use of *aye*, which indexes their regional identity, they seem more open to linguistic innovation as younger characters, with perhaps a more global outlook.

Let us examine the frequencies of both *aye* and *yeah* for all seven characters, presented in Table 4.13 below. The differences that surface once the normalised frequencies of *yeah* and *aye* are analysed at the individual level seem to follow the pattern that had been observed with respect to other elements: Michelle and Orla employ *aye*, the dialectal marker of assent, more frequently than *yeah*, the more globalised marker, while the inverse is true for Erin and Clare. None of the frequency differences between each Derry Girl’s use of these two words was found to be statistically significant, however, despite the difference sizes. In contrast, for the Adult characters the frequency differences between *yeah* and *aye* in their respective subcorpora were all found to be significant. In addition, all the Adult characters employed *aye* more commonly than *yeah* at the individual level as well as a group, the only noticeable difference being Joe’s

total avoidance of *yeah* compared to Mary and Sarah’s sparse use of it. This can be attributed to Joe being a generation older than Mary and Sarah, which potentially makes his speech more conservative than his daughters’.

Table 4.13. Raw and normalised frequencies of *yeah* and *aye*
RF = Raw Frequency, *NF* = Normalised Frequency, *LL* = Log-likelihood

Character	Yeah		Aye		LL
	RF	NF	RF	NF	
Erin	33	32.02	18	17.43	4.48
Michelle	10	12.29	18	22.13	2.32
Clare	17	31.40	7	12.93	4.30
Orla	9	41.21	15	68.68	1.52
Mary	4	6.08	28	42.54	20.25
Sarah	2	4.68	24	56.18	21.94
Joe	0	0	13	37.59	18.02

p < 0.1; critical value for *LL* = 6.63, where *LL* > 6.63 means that the result is significant

Thus concludes the analysis of keywords in the DGC character subcorpora. The next, and final, section of this chapter will present a summary of the main points discussed in the pages above.

4.3 Summary of findings

This chapter aimed to answer the research questions that guide the present thesis. Section 4.1 addressed the usage of *ach/och*, final *but* and DPM *wee* across the corpora DGC-NI and SPICE-NI (research questions 1 and 2, as stated in the introductory chapter). In the case of *ach/och* it was found that, while the frequency of use of these related DPMs in both corpora, in addition to their functions, was similar, the results did not fully correspond with prior research for two reasons: (1) males had been previously reported as the most frequent users of these DPMs, whereas in the corpora examined in this thesis it was females; and (2) the range of functions exhibited by *ach/och* in DGC-NI and SPICE-NI exceeded what had been discussed in earlier studies. It was also found that, although there is an overlap in the functions of *ach* and *och*, each DPM has unique uses.

As far as final *but* was concerned, the results showed that it was underused in the DGC and that its uses coincided with those noted in previous studies, but the presence of this DPM in both SPICE-NI and SPICE-ROI brought into question the claims that associated it with NIrE. Regarding *wee*, the results showed that this term has both pragmatic and non-pragmatic uses, the latter including those of minimiser and mitigator. These uses were found in both DGC-NI

and SPICE-NI, and the frequency of use of *wee* was similar in both corpora, in addition to being much higher than in SPICE-ROI, thereby lending support to the notion that this word is more associated with NIrE.

Section 4.2 sought to address research question 3, concerning the linguistic characterisation of seven main characters in *Derry Girls* through an analysis of the DPMs, terms of address, taboo language and dialectalisms present in their most frequent words and their keyword lists. The analysis showed that the characters formed two groups, the teenage Derry Girls and the Adults, with Erin and Mary functioning as the main link between both. The Derry Girls in general showed less regional and more globalised patterns of use of terms such as *like* and *yeah*, with Clare and Erin being the two characters with the least regionally marked speech, and Michelle and Orla the two Derry Girls that appear more genuinely Northern Irish through their dialogue. Michelle, in particular, deviated from the rest of the group in terms of her use of taboo language. Within the Derry Girls group there also seemed to be a negative correlation between academic ambition – exemplified by Clare, and, to a lesser extent, by Erin – and the use of NIrE dialectalisms, pragmatic or not, which could point to Clare’s awareness of the stigmatisation of dialectal forms, and/or to her adherence to overtly prestigious speech uses, which avoid both said dialectalisms and taboo language. On the other end of the scale, Michelle seems to align with covertly prestigious linguistic uses. The Adult characters, on their part, exhibited more typically NIrE speech patterns than their younger counterparts, although Sarah in her use of *like* aligned more with the Derry Girls. Furthermore, Mary and Sarah’s idiolects resembled one another more closely than either of them did Joe’s, who, in a similar way to Michelle, stood out due to his use of taboo expressions and terms that are characteristic of his dialogue. Joe also diverged from Mary and Sarah by employing keywords that in both the DGC and SPICE-NI were associated with elderly speakers.

5. CONCLUSION

This thesis sought to examine the use of the DPMs *ach*, *och*, final *but* and *wee* in the television series *Derry Girls* in terms of their role in the representation of NIrE and the construction of seven main characters in the show. The dialogue of the show was compiled into a corpus, the DGC and annotated for the relevant pragmatic features. In order to establish a baseline for NIrE speech usage, the DGC was compared to the Northern Irish component of the real-life spoken corpus SPICE-Ireland, which I termed SPICE-NI.

The results showed that *ach* and *och* were, respectively, over- and underrepresented in *Derry Girls*, although, if considered jointly, their frequency in both corpora did not exhibit significant differences. Importantly, their functions were shown to be more varied than the existing literature (Corrigan 2010) had described, with data from both corpora providing evidence for this. As far as final *but* was concerned, its use in *Derry Girls* was much less common than one could have expected according to the data from SPICE-NI. In addition, this DPM was found in the ROI partition of SPICE-Ireland, and the frequencies in both partitions did not show any statistically significant differences, which suggests that final *but* is not exclusive to any particular variety of IrE, being instead a feature of this dialect in general, contrary to previous claims (Corrigan, 2015; Walshe, 2009). The examination of the pragmatic uses of *wee* in the *Derry Girls* dialogue also produced some novel results. This term had only been briefly discussed by Kallen (2005a) in order to mention its possible function as a minimiser, something that seems to have been corroborated in the present study. In addition, it has been found that *wee* can also function as a mitigator and that it can carry a range of interpersonal nuances. *Wee* was, furthermore, compared to Brinton's list of DPM features (2017, p. 9), and it was shown to fulfil more than half of them, which constitutes a foundation for considering *wee* as a DPM, albeit a non-prototypical one.

With regard to the linguistic characterisation of the seven *Derry Girls* characters chosen for analysis, the keyword searches and subsequent qualitative analyses showed a divide between the teenage characters and the adult ones. The first group was in general less likely to employ DPMs associated with NIrE and to use other DPMs such as *like* in ways connected to IrE usages. For the second group, the inverse was true. Furthermore, each group was shown to have a member whose idiolect constituted a departure from the other characters: Michelle among the teenagers, and Joe among the adults. Finally, among the four adolescent characters

there seemed to be a cline of linguistic ‘NIrE-ness’ that correlated negatively with the character’s academic aspirations and their willingness to follow rules: the more academically accomplished and rule-abiding the character was, the less dialectal features their dialogue exhibited.

5.1 Limitations of the present study

As discussed in Chapter 3, section 3.1.1.5, the replicability and reproducibility of this thesis are severely limited due to the copyright-protected nature of the data contained in the DGC. In addition, one must bear in mind that the results of the analysis regarding the frequency of linguistic features in the DGC cannot be generalised to the NIrE speech community directly, for the language in this corpus is a fictional representation of said dialect. That being said, the analysis showed that the functions of *ach*, *och*, and final *but* which appeared in the DGC were also present in SPICE-Ireland. In addition, while it might be true that the results derived from the analysis of the latter corpus due to its data having been compiled between 1990 and 2005, the results still provide new insights into relatively understudied DPMs.

As far as the study design and methods are concerned, what characterised the development of this thesis was the ever-narrowing scope of its aims. A wider range of DPMs associated with IrE was considered for analysis until relatively late in the research process, to the extent that the DGC is tagged for them as well (see Appendix A, Table A.2 for a list of the tagged features that were eventually disregarded). As a result, the annotation of the DGC was much more time-consuming than the final list of studied features might suggest. The reasons that ultimately led to the exclusion of some of the linguistic items were space constraints and the decision to focus on features associated with NIrE specifically. However, there are remnants of the initial scope of the thesis. Terms of address and taboo language, in addition to a wider range of DPMs, were included in the discussion of linguistic characterisation presented in section 4.2 because all these items had been initially included in the general analysis, and their presence in keyword lists, together with the insights they revealed about the characters, led to their being relevant enough to not be excluded from the discussion.

5.2 Recommendations for future research

This thesis is but a contribution to the study of NIrE and its fictional representations. Future research is limited by the present lack of up-to-date corpora that focus on this English variety. Therefore, an essential step in order to progress the synchronic study of this dialect would be

the collection of more real-life, spoken NIrE data. This would be no easy task, much beyond the scope of a simple master's thesis, and likely not an individual effort, but I believe it would open fruitful avenues of study. For example, a researcher could endeavour to replicate the findings regarding the various functions of *och* and *ach*. Our knowledge of the pragmatic roles performed by *wee* would also benefit from a similar approach. Recent data would also be essential in searching for more evidence concerning the geographic boundaries of final *but*, although addressing this matter would require an up-to-date corpus that covers both NI and the ROI.

Even if one limits the scope to studies focusing on fictional representations of NIrE, and to *Derry Girls* in particular, there are still many directions towards which a researcher could take their enquiries. For instance, it was beyond the scope of this study to conduct a thorough analysis of the use of terms of address, including vocatives, in *Derry Girls*, as well as a study of the whole range of taboo language employed by the characters and the functions it performs. In addition, this thesis has only examined at length four DPMs due to their being associated with NIrE, with some less detailed explorations of DPMs that appeared in keyword lists. Consequently, more conscientious analyses of the DPMs discussed in section 4.2 could be performed on the same data, as well as of DPMs present in the dialogue of *Derry Girls* that have not been touched upon in this thesis, such as *so* or *now*. Beyond the field of pragmatics, *Derry Girls* could still be studied in terms of its representation of (N)IrE lexical and morphosyntactic features. Furthermore, the dialectal features present in the dialogue of the show could be compared to those of other television shows based in Ireland, and Northern Ireland more specifically.

6. REFERENCES

6.1 Primary sources

Leddy, C., Lewin, L., McGee, L., Mulville, J. (Executive producers). (2018–2022). *Derry Girls* [TV series]. Hat Trick Productions.

6.2 Secondary literature

About us. (n.d.) Forever Dreaming Transcripts. Retrieved February 26, 2023, from <https://transcripts.foreverdreaming.org/app.php/aboutus>

Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. John Benjamins.

Aijmer, K., & Simon-Vandenberg, A. (2011). Pragmatic markers. In J. Zienkowski, J. O. Östman, & J. Verschueren (Eds.), *Discursive pragmatics* (pp. 223–247). John Benjamins.

Albelda Marco, M., & Estellés Arguedas, M. (2021). Mitigation revisited. An operative and integrated definition of the pragmatic concept, its strategic values, and its linguistic expression. *Journal of Pragmatics*, 183, 71–86. <https://doi.org/10.1016/j.pragma.2021.07.002>

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). Mouton de Gruyter.

Amador-Moreno, C. P. (2010a). *An introduction to Irish English*. Equinox.

Amador-Moreno, C. P. (2010b). How can corpora be used to explore literary speech representation?. In Anne O’Keeffe & Michael McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 531–544). Routledge.

Amador-Moreno, C. P. (2012). 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. Special Issue A New Approach to Literature: Corpus Linguistics*. 12(2), 19–38.

Amador-Moreno, C.P. (2015). “There’s, like, total silence again, roysh, and no one says anything”: Fictional representations of “new” pragmatic markers and quotatives in Irish English. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 37–64). John Benjamins.

Amador-Moreno, C. P. (2016). The language of Irish writing in English. In R. Hickey (Ed.), *Sociolinguistics in Ireland* (pp. 299–319). Palgrave Macmillan.

Amador-Moreno, C.P. (2019). Discourse-pragmatic variation. In *Orality in written texts* (pp. 91–134). Routledge.

Amador-Moreno, C. P., McCafferty, K., & Vaughan, E. (Eds.). (2015). *Pragmatic markers in Irish English*. John Benjamins.

Amador-Moreno, C. P., & McCafferty, K. (2015). “Sure this is a great country for drink and rowing at elections”: Discourse markers in the Corpus of Irish English Correspondence, 1750–1940. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 37–64). John Benjamins.

Amador-Moreno, C. P., & Terrazas-Calero, A. M. (2017). Encapsulating Irish English in literature. *World Englishes*, 36(2), 254–268. <https://doi.org/10.1111/weng.12257>

- Amador-Moreno, C. P., & Terrazas-Calero, A. M. (2022). Using corpus linguistics to explore literary speech representation: non-standard language in fiction. In Anne O’Keeffe & Michael McCarthy (Eds.) *The Routledge Handbook of Corpus Linguistics* (2nd edition) (pp. 517–531). Routledge.
- Andersson, L., & Trudgill, P. (2007). Swearing. In L. Monaghan & J. E. Goodman (Eds.), *A cultural approach to interpersonal communication* (pp. 195–199). Blackwell.
- Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Anthony, L. (2022). AntConc (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.
- Barron, A., & Schneider, K. P. (Eds.). (2005). *The pragmatics of Irish English*. Mouton de Gruyter.
- Bednarek, M. (2010). *The language of fictional television: Drama and identity*. London: Continuum.
- Bednarek, M. (2012). Constructing ‘nerdiness’: Characterisation in *The Big Bang Theory*. *Multilingua* 31, 199–229.
- Bednarek, M. (2015). ‘Wicked’ women in contemporary pop culture: ‘Bad’ language and gender in *Weeds*, *Nurse Jackie* and *Saving Grace*. *Text&Talk*, 35(4), 431–451.
- Bednarek, M. (2019). *Creating dialogue for TV: Screenwriters talk television*. Routledge.
- Bednarek, M. (2023). *Language and characterisation in television series: A corpus-informed approach to the construction of social identity in the media*. John Benjamins.
- Beers-Fägersten, K. (2012). *Who’s swearing now? The social aspects of conversational swearing*. Cambridge Scholars Publishing.
- Beers-Fägersten, K., & Bednarek, M. (2022). The evolution of swearing in television catchphrases. *Language and Literature* 31(2), 196–226.
- Binchy, J. (2005). Three forty two so please: Politeness for sale in Southern-Irish service encounters. In A. Barron & K. P. Schneider (Eds.), *The pragmatics of Irish English* (pp. 313–335). Mouton de Gruyter.
- Brinton, L. J. (1996). *Pragmatic markers in English: Grammaticalization and discourse functions*. Mouton de Gruyter.
- Brinton, L. J. (2017). *The evolution of pragmatic markers in English: Pathways of change*. Cambridge University Press.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Caffi, C. (1999). On mitigation. *Journal of Pragmatics* 31, 881–909.
- Cameron, D. (1995). *Verbal Hygiene*. Routledge.
- Carter, R. & McCarthy, M. (2006). *The Cambridge grammar of English*. Cambridge University Press.

- Clancy, B. & Vaughan, E. (2012). 'It's lunacy now': A corpus-based pragmatic analysis of the use of *now* in contemporary Irish English. In Migge, B. & Ní Chiosáin, M. (Eds.), *New perspectives on Irish English* (pp. 225–246). John Benjamins.
- Corrigan, K. P. (2010). *Irish English volume 1 – Northern Ireland*. Edinburgh University Press.
- Corrigan, K. P. (2015). "I always think of people here, you know, saying 'like' after every sentence": The dynamics of discourse-pragmatic markers in Northern Irish English. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 37–64). John Benjamins.
- Derry Girls*. (n.d.). Internet Movie Database. Retrieved February 10, 2023, from <https://www.imdb.com/title/tt7120662/>
- Derry Girls*. (n.d.). Wikipedia. Retrieved February 5, 2023, from https://en.wikipedia.org/wiki/Derry_Girls
- Derry Girls Transcripts*. (n.d.). Forever Dreaming Transcripts. Retrieved September 20, 2022, from <https://transcripts.foreverdreaming.org/viewforum.php?f=1222>
- Diskin, C. (2017). The use of the discourse-pragmatic marker 'like' by native and non-native speakers of English in Ireland. *Journal of Pragmatics* 120, 144–157. <https://doi.org/10.1016/j.pragma.2017.08.004>
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61–74.
- Farr, F. (2005). Relational strategies in the discourse of professional performance review in an Irish academic environment: The case of language teacher education. In A. Barron & K. P. Schneider (Eds.), *The pragmatics of Irish English* (pp. 203–234). Mouton de Gruyter.
- Farr, F., & O'Keeffe, A. (2002). *Would* as a hedging device in an Irish context: An intra-varietal comparison of institutionalised spoken interaction. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 25–48). John Benjamins.
- Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 225–258). Routledge.
- Gregori Signes, C. (2007). What do we laugh at? Gender representations in 3rd Rock from the Sun. In J. Santaemilia, P. Bou, S. Maruenda & G. Zaragoza (Eds.), *International perspectives on gender and language* (pp. 726–750). Universitat de València.
- Gregori Signes, C. (2017). "Apparently, women don't know how to operate doors": A corpus-based analysis of women stereotypes in the TV series *3rd Rock from the Sun*. *International Journal of English Studies*, 17(2), 21–43.
- Henry, S., Bryan, A., & Neary, A. (2023). "Laughing ourselves out of the closet": Comedy as a queer pedagogical form. *Ethics and Education*, DOI: 10.1080/17449642.2023.2188744
- Hickey, R. (2005). Irish English in the context of previous research. In A. Barron & K. P. Schneider (Eds.), *The pragmatics of Irish English* (pp. 17–43). Mouton de Gruyter.
- Hickey, R. (2007). *Irish English: History and present-day forms*. Cambridge University Press.

- Hickey, R. (2015). The pragmatics of Irish English and Irish. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 17–36). John Benjamins.
- Hodson, J. (2014). *Dialect in film and literature*. Bloomsbury Publishing.
- Hugues, B. M. (2004). Violations of Grice’s conversational conventions as humour in Irish and American television comedies. *TEANGA, the Journal of the Irish Association for Applied Linguistics*, 20, 94–109. <https://doi.org/10.35903/teanga.v20i.507>
- Intellectual Property Office. (2014). Exceptions to copyright. *Gov.uk* <https://www.gov.uk/guidance/exceptions-to-copyright#text-and-data-mining-for-non-commercial-research>
- Ives, S. (1971). A theory of literary dialect. In J. V. Williamson & V. M. Burke (Eds.), *A various language: Perspectives on American dialects* (pp. 145–177). Holt, Rinehart, & Winston.
- Jarazo Álvarez, R. (2023). “It’s fucking tiny.” Road movies, youth and (im)mobility in *Derry Girls* (2018–). *Atlánticas. Revista Internacional de Estudios Feministas*, 8(1), 192–213. doi: <https://dx.doi.org/10.17979/arief.2023.8.1.873>
- Kallen, J. L. (2005a). Politeness in Ireland: ‘In Ireland, it’s done without being said’. In L. Hickey & M. Stewart (Eds.), *Politeness in Europe* (pp. 130–144). Multilingual Matters.
- Kallen, J. L. (2005b). Silence and mitigation in Irish English discourse. In A. Barron & K. P. Schneider (Eds.), *The pragmatics of Irish English* (pp. 47–72). Mouton de Gruyter.
- Kallen, J. L., & Kirk, J. M. (2012). *SPICE-Ireland: A user’s guide*. Cló Ollscoil na Banríona.
- Kirk, J. M. (1997). Irish English and contemporary literary writing. In J. Kallen (Ed.), *Focus on Ireland* (pp. 189–205). John Benjamins.
- Kozloff, S. (2000). *Overhearing film dialogue*. University of California Press.
- Labov, W. (1972). Some principles of linguistic methodology. *Language in society* 1(1), 97–120.
- Landert, D., Dayter, D., Messerli, T., & Locher, M. (2023). *Corpus pragmatics (Elements in pragmatics)*. Cambridge University Press. Doi:10.1017/9781009091107
- Leech, G. (1997). Introducing Corpus Annotation. In R. Garside, G. Leech & T. McEnery (Eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Routledge.
- Leech, G. (1999). The distribution and function of vocatives. In H. Hasselgård & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 107–120). Rodopi.
- Little, I. (2017, December 20). *Derry Girls* could become TV hit – If viewers can understand them. *Belfast Telegraph*. <https://www.belfasttelegraph.co.uk/entertainment/news/derry-girls-could-become-tv-hit-if-viewers-can-understand-them-36423218.html>.
- MacQuillan, S. (2020, February 4). *The Derry Girls Dictionary: 10 mad Derry Girls phrases explained*. Ireland Before You Die. <https://www.irelandbeforeyoudie.com/the-derry-girls-dictionary-10-mad-derry-girls-phrases-explained>
- McCafferty, K. (1999). (London)Derry: Between Ulster and local speech – Class, ethnicity and language change. In P. Foulkes & G. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 246–264). Arnold.

- McCafferty, K. (2001). *Ethnicity and language change: English in (London)Derry, Northern Ireland*. John Benjamins.
- 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.
- McCafferty, K. (2007). Northern Irish English. In David Britain (Ed.), *Language in the British Isles* (pp. 122–134). Cambridge University Press.
- McCafferty, K., & Amador-Moreno, C. P. (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). John Benjamins.
- McCarthy, M. J., & O’Keeffe, A. (2003). “What’s in a name?”: Vocatives in casual conversations and radio phone-in calls. In P. Leistyna & C. F. Meyer (Eds.), *Corpus analysis* (pp. 153–185). Brill.
- McEney, A., & Xiao, Z. (2004). Swearing in modern British English: The case of *fuck* in the BNC. *Language and Literature* 13(3), 235–268.
- McEney, T., & Hardie, A. (2012). *Corpus linguistics*. Cambridge University Press.
- McEney, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh University Press.
- McIntyre, D., & Walker, B. (2022). *Corpus stylistics: Theory and practice*. Edinburgh University Press.
- Merriam-Webster. (n.d.). I mean. In *Merriam-Webster.com dictionary*. Retrieved July 20, 2023, from <https://www.merriam-webster.com/dictionary/I%20mean>
- Migge, B. (2015). Now in the speech of newcomers to Ireland. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 292–317). John Benjamins.
- Millar, S. (2015). Blathering Beauties: The use of pragmatic markers on an Irish beauty blog. In C. P. Amador-Moreno, K. McCafferty & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 390–407). John Benjamins.
- Miller, J., & Weinert, R. (1995). The function of LIKE in dialogue. *Journal of Pragmatics* 23, 365–393. [https://doi.org/10.1016/0378-2166\(94\)00044-F](https://doi.org/10.1016/0378-2166(94)00044-F)
- Murphy, B. (2009). ‘She’s a fucking ticket’: The pragmatics of *fuck* in Irish English—an age and gender perspective. *Corpora*, 4(1), 85–106. DOI: 10.3366/E1749503209000239
- Murphy, B. (2010). *Corpus and sociolinguistics. Investigating age and gender in female talk*. John Benjamins.
- Murphy, B. (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). John Benjamins.
- Nestor, N., Ní Chasaide C., & Regan, V. (2012). Discourse ‘like’ and social identity: A case study of Poles in Ireland. In Migge, B. & Ní Chiosáin, M. (Eds.), *New perspectives on Irish English* (pp. 327–353). John Benjamins.
- O’Keeffe, A. (2005). You’ve a daughter yourself? A corpus-based look at question forms in an Irish radio phone-in. In A. Barron & K. P. Schneider (Eds.), *The pragmatics of Irish English* (pp. 339–366). Mouton de Gruyter.

- O’Sullivan, J. (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). John Benjamins.
- Oxford English Dictionary. (2023a). Ah, int. & n. In *Oxford English Dictionary*. https://www.oed.com/dictionary/ah_int
- Oxford English Dictionary. (2023b). Aye, adv. & n. In *Oxford English Dictionary*. https://www.oed.com/dictionary/aye_adv
- Oxford English Dictionary. (2023c). Mean, v.¹ In *Oxford English Dictionary*. https://www.oed.com/dictionary/mean_v1?tab=meaning_and_use#37459701
- Oxford English Dictionary. (2023d). Punt, n.⁷ In *Oxford English Dictionary*. https://www.oed.com/dictionary/punt_n7
- Oxford English Dictionary. (2023e). Thon, pron.² & adj. In *Oxford English Dictionary*. https://www.oed.com/dictionary/thon_pron2
- Oxford English Dictionary. (2023f). Yeah, adv. In *Oxford English Dictionary*. https://www.oed.com/dictionary/yeah_adv
- Pichler, H. (Ed.). (2016). *Discourse-pragmatic variation and change in English: New methods and insights*. Cambridge University Press.
- Quaglio, P. (2009). *Television dialogue: the sitcom Friends against natural conversation*. John Benjamins.
- Rayson, P. (2008). *Log-likelihood and effect size calculator*. UCREL web server. <http://ucrel.lancs.ac.uk/llwizard.html>
- 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. <https://doi.org/10.2307/455799>
- Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press.
- Schweinberger, M. (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). John Benjamins.
- Schweinberger, M. (2018). Swearing in Irish English: A corpus-based quantitative analysis of the sociolinguistics of swearing. *Lingua* 209, 1–20. <https://doi.org/10.1016/j.lingua.2018.03.008>
- Smith, N., Hoffman, S., & Rayson, P. (2008). Corpus tools and methods, today and tomorrow: Incorporating linguists’ manual annotations. *Literary and linguistic computing* 23(2), 163–180. Doi:10.1093/llc/fqn004
- Stapleton, K. (2010). Swearing. In M. A. Locher & S. L. Graham (Eds.), *Interpersonal pragmatics* (pp. 289–306). De Gruyter.
- Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell.
- Terrazas-Calero, A. M. (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).

- Terrazas-Calero, A. M. (2022). 'Jaysus, keep talking like that and you'll fit right in': An investigation of oral Irish English in contemporary Irish fiction [Doctoral dissertation, Mary Immaculate College, University of Limerick].
- Tinkerbell. (2004, July 4). *Aww*. Urban Dictionary. <https://www.urbandictionary.com/define.php?term=aww>
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2), 179–195. doi:10.1017/S0047404500000488
- Valleriani, L. (2021). Why is he making a funny noise?: The RP speaker as an outcast. In D. Montini & I. Ranzato (Eds.), *The dialects of British English in fictional texts* (pp. 194–210). Routledge.
- Vaughan, E., & Clancy, B. (2011). The pragmatics of Irish English: The use of English in Ireland shows specific features which contribute to its unique profile. *English Today* 27(2), 47–52. <https://doi.org/10.1017/S0266078411000204>
- Villalba Ibáñez, C. (2020). Recognising mitigation: Three tests for its identification. *Journal of Pragmatics* 167, 68–79. <https://doi.org/10.1016/j.pragma.2020.06.015>
- Walshe, S. (2009). *Irish English as represented in film*. Peter Lang.
- Walshe, S. (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. doi: 10.1558/sols.v5i1.127
- Walshe, S. (2016). Irish society as portrayed in Irish films. In R. Hickey (Ed.), *Sociolinguistics in Ireland* (pp. 320–343). Palgrave Macmillan.

APPENDIX A

Table A.1. List of tags used in the annotation of the DGC corpus

Main character tag	Character name
E	Erin
MI	Michelle
C	Clare
O	Orla
Q	Mary
S	Sarah
J	Joe
Other character tag	Character name
G	Gerry
JA	James
DEN	Dennis
SM	Sister Michael
CIA	Ciaran
JEN	Jenny
DEI	Deirdre
GD	Geraldine
SEA	Sean
CO	Uncle Colm
PET	Father Peter
BOY	Minor young male characters
GIRL	Minor young female characters
MAN	Minor adult male characters
WOMAN	Minor adult female characters
OM	Minor old male characters
OW	Minor old female characters
Category tag	Meaning (pragmatic item)
AO	<i>Ach/Och</i>
BF	Final <i>but</i>
W	Wee
META	Metapragmatic comments made by the characters
Position tag	Meaning
I	Sentence-initial
M	Medial position
F	Final position
A	Standalone element

Table A.2. Tags for elements excluded from the analysis

Category tag	Meaning (pragmatic item)	
V	Vocative	
T	Taboo language	
SU	Sure	
L	Like	
Vocative subcategory tags	Meaning	Example
E	Term of endearment	<i>Love, son</i>
K	Family/kinship term	<i>Mammy, ma, da</i>
F	Familiariser	<i>Girls, boy, folks</i>
FN	First name	<i>Mary, Sarah</i>
KFN	Family term and first name	<i>Aunt Sarah</i>
T	Title, followed by name or not	<i>Mrs Quinn, Sister Michael</i>
H	Honorific	<i>Sir, ma'am</i>
SW	Taboo language as vocative	<i>Wank features, Dicko</i>
NN	Nickname	<i>Sugar tits</i>
P	Pronoun as vocative	<i>You, you guys</i>
Y	Phrase headed by <i>you</i>	<i>You poor critter</i>
Taboo language semantic subcategory tags	Meaning	Example
R	Religious expression	<i>Hell, God, Christ, Jesus</i>
NR	Non-religious taboo language	<i>Fuck, cock, motherfucker</i>
Taboo language syntactic subcategory tags	Meaning	Example
Q	Taboo language following question word	<i>What the hell How the fuck What in under Christ</i>
I	Intensifier	<i>The bastarding rosary Fucking savages</i>
D	Taboo expressions acting as DPMs	<i>God, mammy, you're up early.</i>
X	Standalone expressions	<i>Oh my God!</i>
L	Lexicalised expression	<i>I don't really give a flying fuck</i>
V	Verb	<i>Fuck off</i>
T	Term that describes another character (but is not a vocative)	<i>That miniature motherfucker</i>
A	Adjective	<i>Fucked up</i>
NM	Noun between a verb and its preposition	<i>Chill the fuck out</i>

APPENDIX B

Table B.1. Normalised and raw frequencies of fuck and its derived terms with and without Michelle's dialogue

Group	Raw frequency (normalised frequency per 10,000w)	
	With Michelle	Without Michelle
Michelle	167 (205.29)	–
Female	187 (39.63)	20 (5.11)
Female NI	187 (42.99)	20 (5.65)
Young female	174 (61.56)	6 (2.98)

Table B.2. Log likelihood (LL) tests for fuck.

Target corpus	Comparison corpus	NF in TC	NF in CC	LL
DGC Female	DGC Female without Michelle	39.63	5.11	126.06
DGC-NI Female	DGC-NI Female without Michelle	42.99	5.65	123.20
DGC YF	DGC YF without Michelle	61.56	2.98	145.29
DGC-NI Male	DGC-NI Female without Michelle	31.22	5.65	34.94
SPICE-NI male	SPICE-NI female	11.50	5.63	7.73
MICHELLE	DGC AM	205.29	74.07	129.16

p < 0.1; critical value for LL = 6.63, where LL > 6.63 means that the result is significant

Table B.3. Raw frequencies and normalised frequencies (in parentheses) of bloody

Group	DGC	SPICE-Ireland-NI
Young female	0	1 (0.73)
Adult female	6 (4.13)	7 (3.09)
Old female	4 (59.70)	2 (1.20)
Female NI	10 (2.30)	12 (1.83)
Young male	0	2 (1.54)
Adult male	1 (2.65)	0
Old male	13 (25.07)	0
Male NI	14 (14.57)	2 (0.85)
Totals	24 (4.52)	14 (1.52)

Table B.4. Log likelihood (LL) tests for bloody

Target corpus	Comparison corpus	NF in TC	NF in CC	LL
DGC-NI FEMALE	DGC-NI MALE	2.30	14.57	19.26
DGC-NI FEMALE	DGC-NI MALE without JOE	2.30	1.02	0.71
SPICE-NI FEMALE	SPICE-NI MALE	1.83	1.52	1.31

p < 0.1; critical value for LL = 6.63, where LL > 6.63 means that the result is significant

Table B.5. Raw and normalised frequencies (in parentheses) of DPM sure

Group	DGC	SPICE-Ireland-NI
Young female	13 (4.60)	10 (7.28)
Adult female	28 (19.29)	10 (4.41)
Old female	0	25 (15.02)
Female NIrE	41 (9.43)	52 (7.91)
Young male	0	11 (8.44)
Adult male	6 (15.87)	3 (7.13)
Old male	20 (38.57)	2 (11.87)
Male NIrE	26 (27.06)	20 (8.22)
Totals	67 (12.62)	72 (7.83)

Table B.6. Log likelihood (LL) tests for sure

Target corpus	Comparison corpus	NF in TC	NF in CC	LL
DGC-NI OLD	DGC-NI ADULT AND YOUNG	34.15	9.95	17.49
SPICE-NI OLD	SPICE-NI ADULT AND YOUNG	14.73	6.34	10.09

p < 0.1; critical value for LL = 6.63, where LL > 6.63 means that the result is significant