A Corpus-Based Comparative Pragmatic Analysis of Irish English and Canadian English

By

Yasir Sulaiman Abdurahman Almuways

A thesis submitted in fulfillment of the requirement for the Degree of Doctor of Philosophy

Supervisors: Dr. Anne O'Keeffe and Dr. Joan O'Sullivan

Submitted to the University of Limerick

January 2022

ABSTRACT

This PhD thesis is a comparative study of the spoken grammar of Irish and Canadian Englishes within the framework of Variational Pragmatics at the formal level, used to study the pragmatic variation (the intra-varietal differences) in terms of forms and pragmatic functions. It is a study of spoken grammar as a whole (in a comparative and representative way between and across two varieties of English). Corpus linguistics is used as a methodological tool in order to conduct this research, exploring the nature of spoken grammar usage in both varieties comparatively in relation to their pragmatic functions and forms. The study illustrates an iterative approach in which top-down and bottom-up processes are used to establish pragmatic markers and their pragmatic functions in spoken grammar in the two varieties. Top-down analysis employs a framework for spoken grammar based on existing literature while the bottom-up process is based on micro-analysis of the data. The corpora used in the study are the spoken components of two International Corpus of English (ICE) corpora, namely ICE-Ireland and ICE-Canada comprising 600,000 words each (approximately). Methodologically, this study is not purely corpus-based nor corpus-driven but employs both methods. This iterative approach aligns with the notions of corpus-based versus corpus-driven linguistics and perspectives. Corpus tools are used to generate wordlists of the top 100 most frequent word and cluster lists. These are then analysed through qualitative analysis in order to identify whether or not they are a part of the spoken grammar. This process results in a candidate list that can then be functionally categorised and compared across varieties in terms of forms and functions. Specifically, the study offers insights on pragmatic markers: discourse markers, response tokens, questions, hedges and stance markers in Irish and Canadian English. The results offer a baseline description of the commonalities and differences in terms of spoken grammar and pragmatics across the two varieties of English which may have application to the study of other varieties of English. Also, the prominent forms of spoken grammar across these two varieties can be further explored from a macro-social perspective (e.g. age, gender, or social class) and a micro-social perspective (e.g. social distance or social dominance) and how these interplay with pragmatic choices.

DECLARATION

I hereby declare that this thesis is entirely my own work and has not been submitted to any other university or higher education institution. Where use has been made of the work of other people, it has been fully acknowledged and referenced.

Signed:

Yasir Sulaiman Abdurahman Almuways

Date:

January 1 2022

DEDICATION

On this beautiful occasion, as I am submitting my PhD thesis, I would like to thank and dedicate this accomplishment to my parents who were the reason for me to be in this life.

*To my father **Sulaiman Abdurahman Abdulaziz Almuways**, who has supported me at all levels (especially financially) from the day I was born until this day.*

*To my mother **Bidriyah Humoud Ahmad Almuzayni**, for her supplications day and night for me to be successful in both worlds (in this life and in the hereafter).*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS USED

CL	Corpus Linguistics
ICE	International Corpus of English
ROI	Republic of Ireland
NI	Northern Ireland
DM	Discourse Marker
RT	Response Token
SM	Stance Marker
TQ/QT	Tag Questions/Question Tags
CA	Conversation Analyses
CP	Corpus Pragmatics
VP	Variational Pragmatics
PM	Pragmatic Markers
KWIK	Key Word in Context
CANCODE	Cambridge and Nottingham Corpus of Discourse in English
BNC	British National Corpus
LCIE	Limerick Corpus of Irish English

CHAPTER 1

INTRODUCTION

1.1 Background and Introduction to the Study

Irish English and Canadian English are considered to be amongst the major varieties of English today. Canadian English is spoken by approximately twenty million people whereas Irish English is spoken by approximately five million people (Barber 2004; Dolan 2013). Both of these English varieties have been shaped linguistically by many factors that have had a huge impact on them. Irish English has been largely influenced by British English, mainly due to its geographical location, political situation, and historical events. Another main influence on Irish English is the Irish language (see Filpulla 2002; Hickey 2011). Canadian English, on the other hand, has been affected linguistically by the settlers who had migrated to Canada. One of those settler groups comprised Irish people who migrated to Canada after the American War of 1812 (Boberg 2004; Clarke 2012; Barbuto 2013). Since then, this connection between Ireland and Canada has been documented historically in terms of culture and language (Clarke 2012). However, even though there is an established connection between the two varieties of English, there has not been a lot of comparative research done on them in general (linguistically). This is especially the case in terms of spoken grammar as it has not received as much attention as written grammar (Carter & McCarthy 2006). In other words, there is no single work that has compared or provided a full description of the spoken grammar of Irish and Canadian Englishes (taken as a whole variable) in comparison to each other. Thus, this current study seeks to address this research gap. A core component of spoken grammar which has been under-explored in Variational Studies (Carter & McCarthy 2006) is that of pragmatic markers. These items will be specifically explored in this study. Pragmatic markers (PMs) have been defined by Carter and McCarthy (2006) as "a class of items which operate outside the structural limits

of the clause and which encode speakers' intentions and interpersonal meanings" (p.208). The term "pragmatic markers" is both general and comprehensive and includes "discourse markers which indicate the speaker's intentions with regard to organising, structuring and monitoring the discourse; stance markers, which indicate the speaker's stance or attitude vis-à-vis the message; hedges, which enable the speakers to be less assertive in formulating their message; and interjections, items which indicate affective responses and reactions to the discourse" (Carter & McCarthy 2006, p.208).

In this study, the focus will be on providing a comparative description of spoken grammar across two varieties without considering different variables such as age, gender, setting, and so forth (which makes it a non-traditional sociolinguistics study, prioritizing the *what* and not so much the *why* although the *why* could be seen as relating to geographical region). Thus, the focus will be on providing a broad-ranging comparative and representative description of spoken grammar across two varieties: Irish English and Canadian English. Variational Pragmatics (VP) is designed to investigate both the social space (factors) (which is the main focus of VP) and geographical space; however, my study investigates only the forms and functions in geographical space rather than social space.

This primary framework can investigate intra-lingual differences (i.e. pragmatic variation within varieties of the same language) without looking at the social triggers (such as socio-economic status, ethnicity, gender, and age). Consequently, this current study falls short of being a fully-fledged VP study due to the fact that it does not take account of these variables (e.g. age, gender etc.). However, this current study can be a starting point for further research in which such variables can be studied and compared.

In addition to the above, the methodological tool used to conduct this research is Corpus Linguistics (CL). CL will be used to comparatively reveal the nature of spoken grammar usage in both varieties in relation to their pragmatic functions and pragmatic markers (features). The spoken components of two International Corpus of English (ICE) corpora, namely ICE-Ireland and ICE-Canada comprising approximately 600,000 words each, will be used in this study (see Chapter 4, section 4.4 for more details on the corpora and its comparability). The data will be analyzed using an iterative approach which uses top-down and bottom-up processes to identify pragmatic markers and their pragmatic functions in spoken grammar. The top-down analysis is based on a framework for spoken grammar from existing literature while the bottom-up process is based on micro-analysis of the data (both quantitative and qualitative). The iteration between top-down and bottom-up approaches to the data in this study aligns with the notions of corpus-based versus corpus-driven perspectives. This renders this study neither purely corpus-based nor corpus-driven; rather it is a mixture of both methodological approaches. For example, corpus tools are used to generate wordlists of the top 100 most frequent word and cluster lists (for both corpora individually). Qualitative analysis is then used to identify whether or not those word or cluster lists are a part of the spoken grammar. The result of this is a candidate list which can then be functionally categorised and compared across varieties in terms of forms and functions (see Chapter 3, section 3.5 and Chapter 4, section 4.5). This results in a very interesting overview on some similarities and differences in relation to pragmatic functions and forms in the spoken grammar of Irish and Canadian Englishes.

1.2 Rationale for the Study

Despite the historically documented connection between Ireland and Canada in terms of culture and language (Boberg 2004; Clarke 2012), there is a significant lack of comparative research done on these two varieties of English, especially on spoken grammar. This has triggered me as a researcher to choose these two varieties to compare rather than others. Therefore, the rationale for this study is to add to our understanding of the differences and similarities between Irish English and Canadian English in relation to the pragmatic functions and forms of their spoken grammar. It also, through the very nature of the research conducted, contributes to the description of both varieties individually as well. This is done by providing a full comparative and representative description of spoken grammar across these two varieties which have not been compared before. As Chapter 2 details (as well as the analysis chapters), there have been numerous studies on British English and Irish English. Many of these entail comparative analyses of many prominent pragmatic markers, see Tottie (1991); McCarthy (2002); Farr & O'Keeffe (2002); O'Keeffe & Adolphs (2008); Barron (2017a, 2019); for British English specifically see Carter & McCarthy (2006); for American English specifically see Biber *et al* (1999). Comparatively, Canadian English (in terms of pragmatics) has received very little attention (as opposed to British English, American English, and Irish English) which makes it difficult to conduct a comparative pragmatic analysis (this shows the need for providing a full description of the spoken grammar of Canadian English as starting points for further research (Tagliamonte 2005; D'Arcy 2017)). Moreover, some of these varieties mentioned above, including Canadian English, have been studied solo where some of their prominent forms have been studied (Gardner 2001; for Irish English see Farr, Murphy & O'Keeffe 2004; Murphy

2015; for Canadian English see Tagliamonte 2005; Da Mota & Herment 2016; D'Arcy 2017). However, there is very little work done on Irish and Canadian Englishes comparatively although the comparable data (ICE-Ireland and ICE-Canada) is in existence. This is why the two regional ICE corpora (ICE-Ireland and ICE-Canada) were chosen due to their comparability and design which effectively and accurately serve the goal of this study (see Chapter 4, section 4.4 for more details). Additionally, at the personal level, as someone who has lived in and studied Linguistics (BA at Simon Fraser University and MA at University College Dublin) in both countries, it is a topic of great interest to me as I have observed commonalities and differences that have not previously been acknowledged or revealed.

Lastly, the advances and progress of CL (triggered by the improvements in audio-recording and associated technology such as corpus linguistic tools) have made it possible for sufficient quantities of spoken data to be used for in-depth analysis. Carter and McCarthy (2006, 2017) acknowledge that until recently items and structures most typically found in spoken communication have not been fully covered. Thus, the predilection towards researching the written language has pushed me towards examining the spoken language as it is still under-explored.

1.3 Locating the Study

For many decades, the study of spoken interaction in everyday life has become an interesting field of study for researchers from different backgrounds such as ethnomethodology, sociolinguistics, philosophy, pragmatics, structural-functional linguistics and social semiotics (Eggins & Slade 2001). Consequently, there are a number

of disciplines involved in the analysis of discourse, spoken interaction, spoken grammar and conversational grammar. As a result, discourse analysis can be approached from different perspectives. The present study takes a comparative approach using Pragmatics and CL within the framework of VP, looking at forms and functions. As Chapter 3 details, variational pragmatic analysis is aided by CL and ultimately could be referred to as Corpus Pragmatics (CP). As discussed in Chapter 3 (section 3.4), CP is a relatively new and emerging field which has resulted from the combination of the key methodologies of CL and Pragmatics. CL refers to the study of language using corpora in which large collections of "real life" language use stored in corpora. Pragmatics has been defined as "the study of the use of context to make inferences about meaning" (Levinson 1983, p.9) and also as "the study of those relations between language and context that are grammaticalized or encoded in the structure of a language" (Fasold 1990, p.119; see also Aijmer & Rühlemann 2015; Weisser 2015; O'Keeffe 2018; O'Keeffe *et al* 2020). Thus, within the new coinage, "Corpus Pragmatics," more considerations have been made so as to optimise the use of CL for pragmatics research to provide a more accurate analysis of data through the integration of both fields. This has been increasingly recognised, albeit with some caveats (Jucker 2013; Aijmer & Rühlemann 2015; O'Keeffe 2018; O'Keeffe *et al* 2020). CL and Pragmatics have different approaches to the examination of data, yet they are not contradictory which makes it possible to utilize both of them together. CL moves from frequencies of forms to their functions (through an inductive process) known as the form-to-function approach (which has been applied in this current study, known as the bottom-up approach, see Chapters 3 and 4). Pragmatics, on the other hand, works in the opposite direction, moving from "function-to-form," starting with the specific pragmatic function

under investigation (e.g. a speech act such as apologising) to the forms which are typically used for that pragmatic function through an inductive process (see O'Keeffe 2018; O'Keeffe *et al* 2020). By combining the bottom-up CL approach and the more top-down Pragmatics approach, the approach taken in this thesis optimises both the recall of language items and the comparative analysis of both form and function in Irish and Canadian spoken grammar.

Lastly, it is worth emphasising again that this study is not a traditional sociolinguistics study and falls short of being a fully-fledged VP study due to the fact that it does not look in depth at any of the variables such as age, gender, and so forth (as mentioned earlier in section 1.1). However, it is still concerned with VP in the sense that it looks at many features and forms (with their pragmatic functions) which are pragmatic across two different varieties. In essence, it is very much a study of spoken grammar as a whole in a comparative and representative way between and across two varieties of English. As Chapter 3 will discuss, the study draws methodologically from CL and Pragmatics, within a VP framework at the formal level and is used to study pragmatic variation in geographical space.

1.4 Research Questions

The research questions of this study focus on the differences and similarities between Irish English and Canadian English in their spoken grammar in terms of form and function (note that pragmatic markers is an umbrella term for all the following features representing the spoken grammar as suggested in the opening section). This covers all the pragmatic markers of the following features: discourse markers (see section 2.4.1), response tokens

(see section 2.4.2), hedges (see section 2.4.3), stance markers (see 2.4.5), and finally questions and tags (see section 2.4.4) which all represent spoken grammar (as suggested in the opening section). Therefore, the first main question is as follows:

> How different and similar are Irish English and Canadian English in their spoken grammar in terms of form and function?

And the sub-questions are as follows:

- How different and similar are Irish English and Canadian English in their Discourse Markers in terms of form and function?

- How different and similar are Irish English and Canadian English in their Response Tokens in terms of form and function?

- How different and similar are Irish English and Canadian English in their Questions and Tags in terms of form and function?

- How different and similar are Irish English and Canadian English in their Hedges in terms of form and function?

- How different and similar are Irish English and Canadian English in their Stance Markers in terms of form and function?

1.5 Overview of the Thesis Structure

This thesis has ten chapters, starting with this introduction chapter and followed by Chapter 2 which covers the literature review and other chapters as demonstrated below:

> **Chapter 2** aims to encompass the main concepts of the spoken grammar of English. It is the foundation upon which this research has been built and is the key to its understanding. This chapter will introduce some general concepts such as definitions,

features, key concepts to understanding spoken grammar (such as context, form, and function), and standard spoken English grammar. Following that, this chapter will survey some of the prevailing studies on spoken grammar (prominent pragmatic markers) of English generally and Irish and Canadian Englishes specifically if found. In particular, this chapter will focus on the research done on discourse features and functions such as discourse markers, response tokens, stance markers, hedges, interjections (which are all known to be pragmatic markers) and questions and tags (Carter & McCarthy 2006). It may also be worthwhile to note here that Canadian English has a lot less literature on spoken grammar compared to Irish English.

Chapter 3 aims to establish a comparative framework for spoken grammar analysis across two varieties which shows how the studies in relation to spoken grammar/discourse analysis have been approached and studied in the past, and how they will be approached and studied in this current study.

Chapter 4 explains the methodological concepts, tools, and approaches undertaken in this study by discussing the main concepts and tools of corpus linguistics which have been used in the present study. This is done in order to reveal the nature of spoken grammar usage in both varieties comparatively in relation to their pragmatic functions and forms and detailing how the iterative approach (in which top-down and bottom-up processes are used to identify pragmatic markers and their pragmatic functions in spoken grammar) has been applied in this study. This is achieved by detailing the top-down analysis which was based on a framework for spoken grammar from existing literature and the bottom-up process which was based on micro-analysis of the data (both quantitative and qualitative). In addition, this chapter will comprehensively describe the corpora used in the present study in

terms of their design, limitations, and the rationale behind why they have been chosen for this present study. This chapter will also discuss how this current study is neither purely corpus-based nor corpus-driven but both due to the application of the iterative approach between top-down and bottom-up approaches to the data in this study.

Chapters 5- 9 are the analysis chapters in which Discourse Markers (DMs) (Chapter 5), Response Tokens (RTs) (Chapter 6), Questions and Tags (QTs/TQs) (Chapter 7), Hedges (Chapter 8), and Stance Markers (SMs) (Chapter 9) of Irish and Canadian Englishes are investigated comparatively from two perspectives: form and function. This is done by detailing the main quantitative findings from the datasets of ICE-Ireland and ICE-Canada corpora through the theoretical and methodological framework taken in this study (as explained in Chapters 3 and 4) in order to unpack and analyse these findings qualitatively in comparison with each other.

Chapter 10 is the concluding chapter which will revisit and reassess the research questions to ensure that the aim of this present study has been met. This chapter will also propose some suggestions for further research and some recommendations or shortcomings for researchers to either follow or avoid while conducting such a research.

1.6 Conclusion

In conclusion, this research attempts to provide a full description of the spoken grammar of Irish and Canadian Englishes in comparison with each other, which has been taken as the research gap of this current study and will be addressed and investigated. Therefore, the primary framework of Variational Pragmatics at the formal level is used to study pragmatic variation in geographical space rather than social space. Thus making it possible to

investigate intra-lingual differences (i.e. pragmatic variation within the varieties of the same language) without looking at the social triggers (such as socio-economic status, ethnicity, gender, and age); rather, it looks at all of spoken grammar of both varieties as the variable.

CHAPTER 2

SPOKEN GRAMMAR

2.1 Introduction

As established in Chapter 1, this research aims to comparatively investigate the nature of

spoken grammar usage in Irish and Canadian Englishes in terms of its forms (pragmatic

markers) and pragmatic functions within the framework of variational pragmatics (VP). VP

(as an analytical framework) has been used with corpus linguistics (CL) by many linguists

such as O'Keeffe and Adolphs (2008) and Barron (2015) (amongst others) to comparatively

explore the differences between spoken grammar features such as backchannels (or

response tokens) and TQs across different varieties of English while taking some social

variables, such as gender, into consideration (see Chapter 3 for more information on VP as

well as Barron & Schneider 2005, 2009; Schneider & Barron 2008). However, in this

current study, the focus will be on providing a comparative description of spoken grammar

(as a whole variable) across two varieties without taking account of different variables such

as age, gender, setting, and so forth (as discussed in Chapter 1). The rationale for this

approach is that the spoken grammar of Irish and Canadian Englishes has not been

compared before; thus, this current study can be a starting point for further research in

which variables can be studied and compared. Now, we will turn to the scope of this current

chapter reviewing the literature on spoken grammar.

2.2 The Scope of this Chapter

This chapter initially aims to encompass the main concepts of the spoken grammar of

English as it is a foundational component of this research. This will be achieved by

introducing some general concepts such as definitions, features, key concepts to

understanding spoken grammar (such as context, form and function), and standard spoken

English grammar. Following that, this chapter will survey some of the prevailing studies on spoken grammar (prominent pragmatic markers), of English generally, and Irish and Canadian Englishes specifically. In particular, this chapter will focus on the research on discourse features and functions such as discourse markers, response tokens, stance markers, hedges, interjections (which are all known to be pragmatic markers) and questions and tags (Carter & McCarthy 2006, 2017). It may also be worthwhile to note here that Canadian English has a lot less literature on spoken grammar compared to Irish English and this is reflected in the number of studies discussed below.

2.3 Spoken Grammar

2.3.1 What is spoken grammar?

Spoken grammar has been defined simply as "the grammar we find in regular and repeated use by the majority of native- and expert-speakers of a language in the majority of their spoken interactions" (McCarthy & O'Keeffe 2014, p.5). Because most people spend the majority of their speech time chatting in an everyday, regular manner (as opposed to the more formal type of speaking found in speeches, sermons, presentations, or interviews etc.), "it is in everyday conversations that we are likely to find the most basic and widespread forms of spoken grammar" (McCarthy & O'Keeffe 2014, p.5; Carter & McCarthy 2017). Spoken grammar has been investigated widely. For example, Biber *et al* (1999) have explored the distribution and function of the grammatical differences between written registers such as fiction writing, news writing, academic writing, and conversation in their corpus-based work *Longman Grammar of Spoken and Written English.* This work highlights the spoken grammar of American English introducing the same concepts and aspects later done by Carter and McCarthy (2006) regarding spoken grammar. Spoken

grammar is discussed in depth, with a focus on the features of spoken grammar, the importance of shared context, pragmatic markers, pragmatic functions, and questions and tags (Biber *et al* 1999). Also, Carter and McCarthy (1995) have investigated and analysed the conversational or spoken grammar of British English taken from their spoken corpus which has been predominantly collected from face-to-face conversational data. This data resulted in listing common grammatical features in spoken conversations that were rare, or functioned differently, in writing (Carter & McCarthy 2006, 2017; Carter *et al* 2011). The term "spoken grammar" has been widely used across many studies in the field. However, according to Carter and McCarthy (2017), spoken grammar can refer to another term which is "conversational grammar." They base this on Rühlemann's work which "suggests that much description of grammar in spoken corpora should be better termed conversational grammar, since it is there that outstanding differences have been brought to light" (Rühlemann 2007, p.11; Carter & McCarthy 2017).

2.3.2 Characteristics of spoken grammar

Exploring the features or characteristics of spoken grammar must involve parallel grammatical features to which they can be compared. Therefore, it is important to highlight the questions that arise regarding spoken grammar versus written grammar in terms of features and functions or meanings. These have all been answered in detail by McCarthy and O'Keeffe (2014) and Carter and McCarthy (2017). Among those pertinent questions are: Is there an actual distinction between spoken and written grammar, and if so, do the features and forms of grammar differ in speaking and in writing; and also, do the meanings of grammatical structures differ in speaking and in writing?

McCarthy and O'Keeffe (2014) collected pieces of evidence from several corpora in order to prove that there is a distinction between how people use grammar in writing and speaking. The distinction in grammar usage (between written and spoken grammar) appeared, not only with native speakers, but also with learners of English. Thus making the evidence for a noteworthy difference between spoken grammar and written grammar even more reliable in that each has distinctive forms (O'Keeffe, McCarthy & Carter 2007; O'Keeffe & McCarthy 2012). This analysis and conclusion of the distinction between spoken and written grammar is manifested in the way speakers and writers construct clauses and turns/sentences. In spoken grammar for instance, the feature that often arises, due to the nature of face-to-face talk or conversation in context and in real time, is incomplete clauses and sentences; whereas in written grammar this is not the case (McCarthy & O'Keeffe 2014). Also, the main and subordinate clauses in speaking have a special pragmatic and discourse function that is not found in written grammar. For example, *if*-clauses can have a unique pragmatic function in spoken language (such as giving polite suggestions or instructions as in *If you'd like to come this way, please*) (Schleppegrell 1992; McCarthy & O'Keeffe 2014). Other subordinating conjunctions also tend to have functions that are more prominent in spoken contexts. For instance, the meaning and the function of the word *because* has been frequently shown to have an *I'm justifying what I'm saying to you* function rather than a cause-effect function in conversation (Schleppegrell 1992; McCarthy & O'Keeffe 2014). In addition, in spoken grammar, it is common to find the repetition of the same clause: subjects, objects, and so forth. For example, the speaker doubles and repeats the subject with a noun phrase and a pronoun using a noun phrase, then repeating it as a subject pronoun and so forth (i.e. *Well*

*you know **all the old neighbourhood customers they'**ve died. And **the young people they***

go to the big supermarkets or the…) (McCarthy & O'Keeffe 2014, p.9). This pattern is very

rarely found in written texts (for examples of these features mentioned above, see Carter &

McCarthy 1995, 2006, 2017; Carter, McCarthy, Mark, & O'Keeffe 2011; McCarthy &

O'Keeffe 2014).

The main feature that makes spoken grammar more distinctive than written

grammar is the role between the speaker and the listener (Bublitz 1998; McCarthy &

O'Keeffe 2014). Since face-to-face conversations include listeners, the role of listener and

speaker alternates, and listeners are rarely solely passive participants. As a result, there are

three roles that listeners regularly play which are of interest to anyone wanting to

understand the special nature of spoken grammar (McCarthy & O'Keeffe 2014; Carter &

McCarthy 2017). The first is what is so-called collaborative completions or sentence-

completions whereby one speaker finishes a prior speaker's utterance (Gardner 2001). The

second is that listeners routinely insert extra clauses to the "sentences" formed by another

speaker. An example of this could be the *which-* comment-clause by which joint production

of a sentence has been made with its main clause at the end of the previous speaker's turn.

Another example is the sentential *which*-clause starting the new speaker's turn resulting in

a pragmatic function enabling conversational participants to evaluate events and situations,

whether used by the same speaker or by another speaker (Tao & McCarthy 2001; McCarthy

& O'Keeffe 2014). The third role is when listeners are more active and engaging in their

responses to the message, demonstrating their attitudes and stance and providing more than

yes and *no* responses (McCarthy & O'Keeffe 2014). Thus, we may conclude that the space

between two speakers' turns in speaking is not like a period or a full stop as in a written

text; "in face-to-face dialogue, the grammar can flow across turns, contributing to a sense of 'confluence,' when the whole conversation flows seamlessly" (McCarthy & O'Keeffe 2014, p.14; Carter & McCarthy 2006, 2017; McCarthy 2010).

Moreover, another aspect where spoken grammar might differ from the written grammar is the inventory of ready-made chunks (which contain between two to five words and have different terms and labels such as lexical bundles, multi-word units, or clusters); these play important interactive or discourse functions that are not found in written grammar due to the state of knowledge between speakers and listeners (such as *I think so, I know, you know what I mean,* and *I mean*) (see Greaves & Warren 2010). Some of these discourse functions are organizing and monitoring the talk, indicating stance and attitude towards the prior turn, and enabling speakers to be less assertive in formulating their message (Carter & McCarthy 2006; McCarthy & O'Keeffe 2014; also see the many chapters in Schmitt 2004 and Greaves & Warren 2010). Also, one of the distinctive features often found in spoken grammar is ellipsis (i.e. the deletion of some elements) which are often intentionally omitted by the speaker due to the fact that real-time conversation characteristically takes place between speakers who are in the same time and place and have deeply interconnecting lives which means that everything need not be said explicitly. On the other hand, written text might be written and then read in an environment in which both the writer and the reader need to have the constraints of written grammar to stand on its own two feet and persist in time and be read in a different context.

As a result, in everyday informal conversations, speakers do not tend to include items which would normally be considered obligatory by the conventions of written grammar. Take this example illustrated by McCarthy and O'Keeffe (2014):

Example 2.1

Teacher:	**Everybody finished?**
Student:	Yeah.
Teacher:	So overall do you think tourism is of a benefit to a country or is the, is the negative side too much? (CLAS)

In example 2.1, the teacher omits the word *has* before *everybody* because when in the middle of an ongoing conversation, the fact that this phrase is a question is clearly implied. This is partly because of the intonation used by the speaker as well as the (expected) repetitive use of such a question/phrasing in a classroom setting. This phenomenon is known as *situational ellipsis* which is "the non-necessity of stating explicitly everything in the situation, because one can assume that the listener(s) will simply understand" (McCarthy & O'Keeffe 2014, p.12). This kind of situational ellipses is a crucial part of spoken grammar that differentiates it from the conventions of written grammar. Auxiliary verbs (*do*, *be*, *have*) and their subject pronouns are usually dropped in these situations. So, instead of the standard written form: *Do you* want a coffee? one would simply say *Want a coffee?* in everyday speech and the listener would naturally fill in the omitted *do you*. (McCarthy & O'Keeffe 2014; also see Carter & McCarthy 2006).

Furthermore, word order has been presented in spoken grammar with unique patterns that are rare or are not found in written grammar. For example, the slots which common adverbs typically occupy in the clause can change depending on whether they are written or spoken. Conventionally, adverbs like *probably*, *maybe*, *still*, and *almost* are usually placed between the subject and main verb or after a first modal or auxiliary verb in written grammar: *She probably knows the answer*. However, because conventional spoken grammar allows other word-order patterns, adverbs can move from their traditional positions and they may occur before the subject (clause-initial position: *Probably, she*

knows the answer) or, as is more likely found in informal speech, they can occupy the final position: *She knows the answer, probably* (McCarthy & O'Keeffe 2014).

2.3.3 Standard spoken English grammar

According to Carter and McCarthy (2006, 2017), standard English grammar is most characteristically associated with written language and not spoken language. By their definition, standard grammar is "considered to be characteristic of the recurrent usage of adult, educated native speakers of a language" (Carter & McCarthy 2006, p.167). They point out however that standard spoken English grammar is different from standard written English grammar, and that "what may be considered 'non-standard' in writing may well be 'standard' in speech" (Carter & McCarthy 2006, p.168). Furthermore, according to Biber *et al* (1999), a standard English exists in spoken English just as it exists in written English, and it is wrong to assume that standard English is fixed "with little or no variability" (pp.16-17).

As a result, recognising the dialectal varieties (regional and social dialects) and describing the patterns of variation that exist within standard English, and accounting for those patterns in terms of contextual factors leads us to applying a similar approach used in the analysis of written English to spoken English. In this case, standard spoken English is defined as including grammatical characteristics shared widely across dialects, excluding those variants restricted to local or limited social regional varieties (Mac Mathúna 2004). This approach recognizes that conversation has special grammatical characteristics not typically found in writing, and so we do not impose a written standard on our analyses of conversation (Carter & McCarthy 2006; for more reading on standard spoken English

grammar see also Biber *et al* 1999).

Standards in Irish and Canadian Englishes

As Biber *et al* (1999) argue, standard English cannot be fixed (in general); similarly, Kirk (2011) argues this point in relation to Irish English by acknowledging that there is no specifically codified Irish standard English. However, there are some recognized linguistic features associated with traditional dialects and vernacular Irish English found in what he loosely refers to as "standard English in Ireland" (Kirk 2011, p.35). He elaborates by recognizing that these features are "muted [quantitatively and qualitatively] relative to the material found in a dialect or sociolinguistic study" (Kirk 2011, p.35). In addition, Hickey (2003) states that:

> The standard of the south of Ireland is a typically fuzzy, non-binary phenomenon, it is not orthographically encoded and does not seem to figure high in the consciousness of speakers. But that should not deter one from trying to come to grips with it. Speakers can move up and down on a stylistic cline for whatever reason, e.g. vernacularising their speech in, say, a family context or a familiar Irish environment. But there are a large number of speakers from the Republic of Ireland whose default speech style consists of the employment of non-regional phonological, syntactic and lexical features on which there is much unconscious agreement (p.15).

He also argues that Irish English does not have one, single recognised standard English; rather it has more than that due to considerable exposure to forms of British which resulted in the existence of supraregional varieties in Ireland (Hickey 2003). Supraregionalisation is described by Hickey as "an historical process whereby varieties of a language lose specifically local features and become less regionally bound" (Hickey 2007a, p.309). However in the Irish context, supraregionalisation is associated more with accent than with grammar (Hickey 2007a, pp.309-315).

In relation to the ICE corpora, Kirk and Kallen (2006, p.88) consider the concepts of "standardisation" and "Celticity," based on a study of the British and Irish components of ICE. In this context, "Celticity" refers to the features of lexis, grammar, and discourse, which appear in the ICE-corpora and which show influence from the Irish language. Kirk and Kallen refer to the "dual nature of Irish Standard English," which shows "both the effects of the standardisation process common to all standard Englishes and the effects of Celticisation arising from a variety of circumstances" (2006, p.88) They argue that, despite evidence of "Celticisation," the spoken and written texts of ICE-Ireland are "essentially standard" and demonstrate few features of what are associated with traditional dialects. Despite this, they argue that "standardisation is never quite fully achieved" and Irish English continues to show elements of variation in standard contexts (2006, p.109).

On the other hand, in Canadian English also, as noted by Lougheed (1985), there is no consensus on what is the standard in Canadian English and what is not. On compiling papers from a conference held for that purpose he states, "I cannot write that the Conference came to a definitive conclusion as to what is standard and not standard in Canadian English; indeed, had it done so, it would have been suspect and much less invigorating" (Lougheed 1985, p.ix). However, there is still a widely recognized standard English which has been codified in dictionaries, grammars, and usage handbooks (see sections 4.4 and 4.4.1 for the definition of what is considered standard English in the ICE Corpora).

2.3.4 Key concepts of spoken grammar

2.3.4.1 Context

Context is the core element shared by the speaker and the listener by which they can both understand utterances. Utterance is the "production of speech uttered by a speaker and perceived by a listener in a certain context" (O'Keeffe 2014, p.18). Thus, speech is produced for listeners who are usually physically present (in some capacity) at the same time of the speech production (i.e. virtual live conversations, physical presence in the same space, telephone conversations or other modes of communication where a live interaction can take place). Social convention thus necessitates that listeners should actively engage with speakers, unlike the case of written texts, which are commonly created in one time and place and consumed in another. Also, speech occurs in real-time, usually without the luxury of planning, correcting, revising, and polishing that writing allows. Thus, the complete and well-formed sentences, in the conventional sense, are often not present, nor do they need to be for communicative efficiency in everyday speech (Brazil 1995; McCarthy & O'Keeffe 2014; Carter & McCarthy 2017). According to Thomas (1995), the utterance cannot be understood without assigning the sense (the semantic/core meaning taken from the dictionary) and the reference (the meaning taken from the context in which the word occurs). In linguistics the distinction is made between the subfield of semantics, which is the study of the meanings of words without reference to their context, and the subfield of pragmatics, which is the study of the meanings of words with reference to their context which may shape the linguistic meaning of the word (Cruse 2002; Szabo 2005). Thus, the main key that should always be taken into consideration in order to understand the grammar of spoken language is its context of use since spoken grammar is the study of grammatical forms and their functions in immediate and shared contexts between the speaker and the

listener (Biber *et al* 1999; McCarthy & O'Keeffe 2014). One of the methodological tools that has been increasingly applied to contexts where the use of language is being analysed in a given context is Corpus Linguistics (CL). It has become the main tool for discourse analysis and conversation analysis at different given contexts such as courtrooms (including forensic linguistics) (Cotterill 2004), workplace discourse (Koester 2006), classroom and educational contexts (Farr 2002, 2003; Walsh 2002; O'Keeffe & Farr 2003), political discourse (Carter & McCarthy 2002), advertising and the media (O'Keeffe 2002, 2006; Charteris-Black 2004; O'Keeffe & Breen 1997) and healthcare discourse (Adolphs *et al* 2004).

However, the notion of context has proven difficult to define (see Lyons 1981; Levinson 1983; Cook 1990; Duranti & Goodwin 1992; Janney 2002). Ochs (1979) claims that "the scope of context is difficult to define... one must consider the social and psychological world in which the language user operates at any given time" (p.1). Nonetheless, the agreed upon fact about context is that it is constructed by factors both *outside* and *inside* the text or utterance. Context *inside* text has been referred to as co-text which contains grammatical and lexical cohesion within texts. In comparison, the context *outside* the text can be subdivided into situational context and background knowledge context. Thus, within the context *inside* text, the interlocutors have no need to move *outside* the utterance in order to determine the meaning, whereas, within the context *outside* the text, the interlocutors need to move *outside* the utterance in order to determine the meaning (Clancy 2010).

2.3.4.2 Form and function

The notions of form and function in any study of grammar are central to our understanding of spoken or written grammar. It is clear that there are unique items (forms) or structures found only in either the spoken or written grammar (discussed earlier in section 2.3.2). Therefore, spoken and written grammar employ grammatical resources in different ways, reflecting their different purposes and different contexts of use (see Leech 2000; McCarthy & O'Keeffe 2014; Carter & McCarthy 2017). This can be seen in the way spoken grammar exploits the deictic system to create interpersonal meanings. For example, the pronoun *we* can be used for solidarity and a collective sense. As McCarthy and Handford (2004) show, authoritative individuals in corporate settings can construe *we* to take on different meanings to reflect their more covert intentions. For example, they can use *we* to take on a more collective, inclusive meaning when used to represent the whole corporation; or, they can insert it into the highly occurring chunk *we need to* in order to lessen the authoritative command that is truly meant—*you must*, thereby creating a manufactured solidarity between the chains of command in the workplace (McCarthy & O'Keeffe 2014; also see Drew & Heritage 1992). Therefore, the face-to-face context of use and the grammar (the form) we choose intentionally or unintentionally can have a huge impact on shaping the relationship between the speakers due to its pragmatic function. Also, the nature of personal deixis can often make it difficult to interpret conversational transcripts unless they are in clusters such as *you know* and *I mean* which are analysed all together as one unit (Carter & McCarthy 2017).

2.4 Key Components of Spoken Grammar

Both Biber *et al* (1999) and Carter and McCarthy's (2006) works are considered seminal works in the realm of spoken grammar. However, the intensity of corpus linguistic and pragmatic evidence presented in the *Longman Grammar of Spoken and Written English* is far less than what is presented in *Cambridge Grammar of English* (see Biber *et al* 1999). As such, the key components of spoken grammar presented and elaborated upon in Carter and McCarthy's (2006) comprehensive work have become an essential guide to spoken and written English grammar and usage today. This text also forms the basis for this present study in terms of its definition and framework for spoken grammar. Through CL, Carter and McCarthy (2006) evidence the different forms and functions of grammar within and across turns in spoken language. For a more extensive look into this, refer to section 2.3.2 (also see Biber *et al* 1999; Carter & McCarthy 2006, specifically pp.164-240; for an exploration on features of informal spoken grammar with corpus-based evidence see Carter & McCarthy 2006, 2017).

Carter and McCarthy (2006, pp.92-122), have shown how grammatical features in spoken utterance are used in terms of forms and functions, reflecting on the creation of discourse rather than just the internal construction of phrases, clauses, and sentences. They give a detailed explanation for deixis, situational ellipsis, response tokens, headers, tails, questions and tags, and so forth. They also comprehensively explain pragmatic markers and their pragmatic functions, directly connecting to the study presented in this thesis. Carter and McCarthy (2006, 2017) have made the term pragmatic markers an inclusive term that includes discourse markers, response tokens, stance markers, hedges, and interjections. Yet all the features mentioned above are a part of spoken grammar (that we are investigating in

this study). It is important to note, however, that different studies and researchers sometimes differ in their classification and terminology related to pragmatic markers (see section 2.4.2 for the terminology related to response tokens as an example). Thus, awareness of a difference in terminology regarding pragmatic markers in terms of categorization and inclusion helps in determining the terminology which will be used consistently throughout the thesis (see Chapter 1, Chapter 4, and Appendix A of this thesis).

The key feature of spoken grammar is pragmatic markers which are known to be the core component of spoken grammar and are defined as "a class of items which operate outside the structural limits of the clause and which encode speakers' intentions and interpersonal meanings" (Carter & McCarthy 2006, p.208). For the purposes of this study, this includes discourse markers (Chapter 5), response tokens (Chapter 6), questions and tags (Chapter 7), hedges (Chapter 8), and stance markers (Chapter 9) whose definitions are provided in the following sections (note that very few interjections were found in the datasets based on the top 100, see Chapter 6 for some interjections, such as *oh*, found in the data as response tokens).

2.4.1 Discourse markers

Discourse markers (DMs) are "words and phrases which function to link segments of the discourse to one another in ways which reflected choices of monitoring, organisation, and management exercised by the speaker" (Carter & McCarthy 2006, p.208). The most common DMs in everyday informal spoken language (which contain single words and phrasal and clausal items) include *anyway*, *cos*, *fine*, *good*, *great*, *like*, *now*, *oh*, *okay*, *right*, *so*, *well*, *you know*, *I mean*, *as I say*, and *for a start* (Carter & McCarthy 2006, 2017). These

DMs serve different pragmatic functions. For more information, see Chapter 5, section 5.2, which further covers existing literature on DMs by demonstrating different perspectives on DMs, definition and function of DMs, and general characteristics of DMs.

Irish English, in regard to spoken grammar (generally in terms of pragmatic markers and specifically discourse markers) has been given more attention lately by many scholars and researchers (Amador-Moreno, McCafferty & Vaughan 2015). One of the core works on DMs of Irish English is that of Kallen and Kirk (2012). According to Kallen and Kirk (2012), DMs "do not contribute to the predication or other core function of an utterance, but express the speaker's attitude towards this core illocution within the context of emerging discourse" (p.42). They have categorised DMs under different categorizations such as syntactic DMs which are in syntactic constituents (i.e. *I know, I see, I mean, what do you think,* and *I don't think so)*, lexical DMs which are lexical items (or single words (i.e. *actually, right, well, okay,* and *yeah)*, and phonological DMs which are more associated with sounds (they are not words) (i.e. *oh, ah,* and *aw)* for SPICE-Ireland (Systems of Pragmatic Annotation in the Spoken Component of ICE-Ireland) annotation and design (for more examples see Kallen & Kirk 2012, pp.104-109). Nevertheless, these types of DMs are dealt with as one category in this present work under the category of "discourse markers" based on the definition provided earlier in this section by McCarthy. Moreover, Kallen (2006) demonstrated the dynamics of discourse marking in ICE-Ireland by exploring the distribution and functions of some DMs such as *like, you know,* and *I mean* in ICE-Corpora and also showed how particular DMs are associated with traditional dialect.

One of the DMs that has been frequently studied across varieties of English is *like* (D'Arcy 2017). The DM *like* has been studied comparatively in Irish English and south-eastern British English (SE-BrE) in terms of its use. Schweinberger's (2015) study resulted in significant differences being found between how *like* is used in terms of overall frequency, social meaning, and positioning in these two varieties of English. In addition, there is a good collection of work on pragmatic markers of Irish English compiled in *Pragmatic Markers in Irish English* (2015) where Irish English has been analysed, in some cases comparatively with other languages such as the Irish language or varieties of English such as British English (Amador-Moreno *et al* 2015). In this volume, Hickey (2015) examines the pragmatics of two structurally different languages sharing a single geographical area showing that common pragmatic features of Irish English have been shared by Irish due to the shared cultural norms of both speech communities in Ireland (Amador-Moreno *et al* 2015; Hickey 2012, 2015). Also in the same volume, Corrigan (2015) looks into the dynamics of eight pragmatic markers in Northern Irish English (not Southern Irish English which is the one examined in this thesis) based on a corpus of Northern Irish English. It was designed from sociolinguistic interviews conducted in 2008 averaging ninety minutes in length in order to conduct their pragmatic functions within the index of social categories of age and gender (Amador-Moreno *et al* 2015; Schweinberger 2015). Additionally, Murphy (2015) explores the pragmatic functions of *like* and *sure* from an age and gender perspective based on a small sociolinguistic oriented corpus containing 103,000 words. The study shows that both of these pragmatic markers are used differently and present distinctive pragmatic functions based on age and gender (Amador-Moreno *et al* 2015).

On the other hand, in Canadian English, Tagliamonte (2005) is one of the early studies that highlighted DMs in Canadian English. In her study she provides an empirical account of a number of DMs in the spoken conversation of young Canadians (from age 10 to 20, 14 male and 12 female), from the same community (sharing the same geographical area) in Toronto. This study employed the quantitative method within the tradition of sociolinguistic variation theory (e.g. Labov 1972; Sankoff 1980, 1988) using a corpus that comprises 26 speakers between the ages of 10–20 and amounts to about 200,000 machine-readable words. The aim of the study is to demonstrate the most frequent DMs among young Canadians and how they are distributed across age groups. The results show that *like*, *just*, and *so* are all salient features of Toronto Youth English, concentrated amongst 15- to 16-year olds generally and found in use with female speakers in particular. However, "only *like* declines by age 20, while *just* and *so* prevail" (Tagliamonte 2005, p.1911). D'Arcy (2008) conducted an apparent-time analysis of *like* across the full age spectrum by providing a detailed examination of its rise within a well-defined community which represents contemporary, urban Canadian English. This is done by using corpus data from a large archive of informal, spoken English collected in Toronto, Canada in the period between 2002 and 2004. The full collection comprises over 350 hours of casual conversational data of quota-based random sampling and social networking from speakers who were born and raised in the city of Toronto and whose ages are between 9 and 92 years old. This shows that "*like* has been developing systematically in the vernacular for as long as can be ascertained using apparent-time data, giving us a view of how it has come to its current state of use… [which] we can use to extrapolate to other varieties and locales" (D'Arcy 2008, p.125).

In terms of pragmatic function, Da Mota and Herment (2016) investigated the usage of the final particle *eh* in spoken Canadian English and showed the link between the use of the final particle *eh* when used as a DM and the use of high rising terminals (HRTs). This study was done based on two different sources, 1) recordings of spontaneous conversations by Canadian speakers, where no occurrences of *eh* were found; and 2) transcriptions from DVDs of comedy shows, where occurrences of *eh* were present. Their findings include that the discourse marker *eh* has the following pragmatic functions: to show the speaker's uncertainty and to leave one's statement open for a room for argument, agreement, or disagreement (Da Mota & Herment 2016; for more reading on *eh* in Canadian English, see Avis 1972).

2.4.2 Response tokens

Response tokens (RTs) are one of the ambiguous discourse features that have been examined from a variety of perspectives and occur in the form of short utterances and non-verbal surrogates (e.g. head nods) (see Fries 1952; Kendon 1967; Yngve 1970; Maynard, D. 1989, 1997; Maynard S. 1986, 1997; Tottie 1991; Drummond & Hopper 1993a, 1993b; McCarthy 2002; Gardner 2001; O'Keeffe & Adolphs 2008). RTs are also one of the discourse features that have been categorized differently in terms of terminology and inclusion. The term "response tokens" was given its own categorization by Carter and McCarthy (2006) following Gardner (2001), where it was classed as a type of discourse marker under the umbrella term, pragmatic markers (Gardner 2001; Carter & McCarthy 2006). Conversely, in *Longman Grammar of Spoken and Written English*, the term "response tokens" has been given different terms based on their pragmatic functions such as "response forms" which is more general and "backchannels" which is more specific, and

"response elicitors" which includes fixed tags (Biber *et al* 1999; also see Chapter 6 for more details). Thus, sometimes they are included within the discourse markers categorisation and sometimes they are not. Nevertheless, these signals are produced by the listener, according to Kendon (1967), as an accompaniment to a speaker. In this thesis, they are treated as a separate category under the name of "response tokens." See Chapter 6, section 6.2 for a more in-depth explanation of RTs and to view more literature on RTs in terms of forms and functions and different perspectives.

One of the most comprehensive comparative works on RTs of Irish English is O'Keeffe and Adolphs (2008) who use a variational pragmatic framework and corpus linguistic methods in order to investigate the discourse feature of RTs in terms of forms and functions across the two English language varieties (Irish and British) (see McCarthy (2015) on turn openers across Irish and British Englishes as it overlaps with O'Keeffe and Adolphs). They do this by comparing and contrasting the distribution of their forms and functions, using data from two language corpora which were assembled with the study of spoken discourse in mind, namely the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) (which contains five million words in total) and the Limerick Corpus of Irish English (LCIE) (which has just over one million words collected in the Republic of Ireland, but does not include Northern Irish English). In O'Keeffe and Adolphs' study only the top 5000 most frequent items were used to generate word and cluster lists. These were then analysed through qualitative analysis in order to identify whether or not they are a part of the response forms. This process resulted in a candidate list that could then be functionally categorised and compared across varieties in terms of forms and functions (O'Keeffe & Adolphs 2008). A similar methodological method has been employed in this

research, as explained in Chapter 4. In O'Keeffe and Adolphs' research, they use the term

listener response as the general term to refer to the activity involving vocal, verbal, and

non-verbal non-floor-holding responses when a listener responds to the floor-holding

message in a conversation. They use response token as a specific term to refer to the

specific items being analysed in the research (O'Keeffe & Adolphs 2008). Also, in their

selection process, they define and limit the response token to the item that only fills a

response slot and does not take over the speaker turn. In other words, RTs are seen as turn

yielding. In our analysis, RTs that appear as a part of a turn are not included as RTs. For

example, the word *really* in the following example 2.2 was not counted as a RT (O'Keeffe

& Adolphs 2008, p.10).

Example 2.2

A: ... basically I think I shouldn't have gone at all because the prescription he gave me
        I think I could have gotten over the counter.
B: **Really?** What did he give you?

Alternatively, *really* in example 2.3 below is counted as a RT because it does not

take over the speaker's turn (O'Keeffe & Adolphs 2008, p.10):

Example 2.3

A: And I don't think her insurance is even that cheap.
B: **Really**.

Before presenting the ultimate results (forms and functions) of their research, there

are some conclusions O'Keeffe and Adolphs (2008) reached in regard to RTs in Irish and

British Englishes based on the distribution of their forms and functions. 1) A broader range

of RTs is used by British English speakers at the single and two-word level (O'Keeffe &

Adolphs 2008). 2) It has been noticed that a broader range of RTs as single-word items has

been found in American English than in British English such as *right*, *absolutely*, *sure*,

good, *lovely*, *exactly*, *great*, *definitely*, *true*, and *really* (as noted by McCarthy 2002).

However, the Irish English single word forms only have *really*, *sure*, and *right* in common with McCarthy's findings for single word non-minimal responses in American English (O'Keeffe & Adolphs 2008). 3) It was noted that *yes* and *quite* as RTs in British English have no corresponding occurrence in the Irish English data, which suggests a higher level of formality in British English (O'Keeffe & Adolphs 2008). 4) Religious references and swear words both appear in British and Irish Englishes. However, their appearance and usage in British English is limited to *God* and *oh God* while Irish English forms comprise *God*, *oh God*, and *oh my God* and the swear words *Jesus* and *Jesus Christ*. This points to a higher level of informality in Irish English conversation (O'Keeffe & Adolphs 2008). 5) A difference was noted in the reduplication of forms in both varieties. For example, Irish English displays more reduplication of the single word form such as *yeah yeah*, *no no*, *yeah yeah yeah*, *no no no*, *oh yeah yeah yeah*, *no, no, no, no* and *yeah yeah yeah yeah* while British English displays less reduplication but contains more clusters with the vocalisation *oh* such as *oh yeah*, *oh right*, *oh yes*, *oh God*, *oh dear*, *oh I see*, *oh well*, *oh my God,* and *oh I don't know* (O'Keeffe & Adolphs 2008). Finally, some RTs, they noted, could potentially lead to cross-cultural pragmatic failure as in the Irish response form *Are you serious*? which is used in Irish English as a non-minimal response token and not found in the more dominant variety of British English. This, argue O'Keeffe and Adolphs, may cause misunderstanding in terms of how the listener orients towards the propositional content, potentially resulting in pragmatic confusion or even face threat (2008).

O'Keeffe and Adolphs' (2008) results show that while there are clear differences between British and Irish English in terms of the forms, frequencies, and sociocultural subtleties, the pragmatics of the discourse function of RTs appear to be constant. Below are

the pragmatic functions of the RTs across the two varieties of English which have been summarised into four broad functions extracted from the data (corpora) as a whole:

1. **Continuer response tokens**
2. **Convergence tokens**
3. **Engagement tokens**
4. **Information receipt tokens** (O'Keeffe & Adolphs 2008; more details on this study and findings are covered in Chapter 6, section 6.4.2).

On the other hand, Canadian English has received far less attention than Irish English in terms of spoken grammar in general (as mentioned earlier in section 2.2). Thus, fewer discourse features have been explored. Nevertheless, under the category of RTs, Farenkia (2012) conducted a study on face-saving strategies in responding to gratitude expressions in Canadian English. This study is modelled on Schneider (2005) (but with different Discourse Completion Tasks (DCTs)) in which a pragmatic phenomenon has been investigated across different varieties of English within the variational pragmatic framework (which matches the framework used in this present study). Farenkia (2012) examines politeness strategies used by a group of English-speaking Canadian university students when responding to gratitude expressions (following Aijmer's (1996) five strategies to respond to *thanks* and to minimize the indebtedness of the thanker). Similarly Schneider (2005) tested and compared the same strategies for responses directed to gratitude expressions in Irish English, English (British) English and American English (following the seven parameters: the interactional patterns in which the responses are employed; the realization of head moves and supportive moves; the types of response realizations and their variants; the frequencies of tokens belonging to these types; the

speaker strategies; the modification patterns and the situational distribution of tokens). The data from the Farenkia study used was elicited through DCTs (for more information see Schneider 2005, p.128 and Farenkia 2012, p.3). The quantitative and qualitative study addresses functional, formal, and situational aspects of RTs. The study reveals that "the participants mostly preferred "minimizing the favour" and "expressing appreciation" in their responses and native speakers of Canadian English seem to make the same choice as speakers of Irish English, American English and British English (Farenkia 2012, p.1). In terms of the realization forms of responses to thanks, the results indicate that the participants mostly employed constructions with "*no problem*, *welcome* and *pleasure"* (Vaughan & Clancy 2011, p.4; Farenkia 2012).

2.4.3 Hedges

Hedges are a range of expressions and markers used in everyday spoken language for speakers to downtone the assertiveness of a segment of discourse. In other words, speakers are usually careful to not sound blunt and assertive, so they use these expressions. Examples of these hedges include: *apparently*, *kind of*, *by any chance*, *I think*, *just* (about), *like*, *maybe*, *perhaps*, *probably*, *sort of*, and *surely* (Biber *et al* 1999; Carter & McCarthy 2006). Hedges have been studied from perspectives such as: theoretical, empirical and applied perspectives (Clemen 1997; Schröder & Zimmer 1997; Hyland 1996a, 1996b; Crompton 1997, 1998; Lindemann & Mauranen 2001; Mauranen 2004). See Chapter 8, section 8.2, in which more existing literature has been covered in terms of hedging; definition and origins and hedging within the scope of Pragmatics and Discourse.

There have been a number of studies in which hedges in Irish English have been studied and analysed with different approaches and methodologies. One of the relevant

studies on hedges which blends pragmatics and corpus linguistics is the work of Farr and O'Keeffe (2002). Along with other studies, it adds a quantitative dimension to the characterisation of Irish English, moving from frequency of form (via corpus analyses) to functional categorisation (O'Keeffe 2018; Vaughan & Clancy 2011). Farr and O'Keeffe (2002) study the occurrences of the hedges *I would say* and *I'd say* based on three one-million-word spoken corpora which are the Limerick Corpus of Irish English (LCIE), samples from the Cambridge and Nottingham Corpus of Discourse in English (CANCODE), and a corpus of American spoken data from the Cambridge International Corpus (CIC). Initially, based on the quantitative findings, they found that these two hedges are used twice as frequently by Irish speakers as by their American counterparts in the data, which led them to do further investigation on this finding in regards to Irish English. Ultimately, they present a qualitative case study analysis (which is more nuanced) of two specific contexts (radio phone-in and post-observation teacher training interaction) in which these hedges occur. The results showed that these hedges are used by Irish speakers in order to soften face threatening acts such as disagreement or giving advice. They also noted that Irish speakers have a greater tendency to use these hedges in order to downtone the assertiveness of a segment of discourse even with undisputed propositional content. As a result, they proposed that the broad pragmatic functions of these hedges are linked to a tendency by Irish English speakers (who may feel an obligation to hedge in situations where British or American speakers do not) to avoid directness. They speculate that this may relate to the Irish socio-cultural context, saying that "in Irish society... 'forwardness,' which ranges from being direct to being self-promoting, is not valued" (Farr & O'Keeffe 2002, p.42; see also Vaughan & Clancy 2011).

In addition to that, Farr, Murphy and O'Keeffe (2004) examined the occurrences of hedging across different contexts in spoken Irish English such as family discourse, radio phone-in, teacher training feedback, service encounters and female friends chatting by using the Limerick Corpus of Irish English (LCIE) in order to show the impact of context on the phenomenon of hedging. They came to the conclusion that the analysis of intra-varietal variation at the level of context has shown that the more institutionalised the discourse is, the higher the frequency of hedging items. Also, they provided a list of the most frequent hedges used in LCIE from single-word and two-word clusters which revealed that *like* and *you know* are among the most frequent hedges. Moreover, further investigation has been done on the occurrences of these most frequent hedges in relation to Irish family discourse by Clancy (2005). He found that *like* was the most frequent hedging item in his data. This hedge has also been studied extensively by many researchers and scholars including Schweinberger (2015) and Alexandra D'Arcy (2017) who compare the use of *like* across different varieties of English (Vaughan & Clancy 2011; Barron 2014; for more information on *like* in Canadian English see D'Arcy (2017) and Gabrys (2016)).

2.4.4 Questions and tags

Questions and tags have been viewed and categorised separately from pragmatic markers. However, they still belong to spoken grammar in that they bear specific pragmatic functions and forms. They occur frequently in spoken English in order to engage the listener and invite convergence with the speaker (Carter & McCarthy 2006; see Chapter 7, section 7.2, in which more existing literature on questions and tags has been covered in terms of definition and their forms and functions in spoken grammar). By way of example, question tags (QTs; also known as tag questions (TQs)) can generate several types of

meaning based on different intonations (falling tones or falling tone plus rising tone) used by the speaker in the main clause which can give an expectation of either a *yes* or *no* answer or the potential for both a *yes* or *no* answer) (Biber *et al* 1999; Carter & McCarthy 2006). This type of question is usually formed through a tag after a declarative clause. They are "highly interactive in that they may constrain the range of possible or desired responses from the addressee" (Carter & McCarthy 2006, p.725). See the examples in Table 2.1 for more illustration.

Table 2.1 *Types of question tags*

Type	Polarity	Falling tone	Falling or rising tone	Constrained or desired answer
1	affirm. + neg.	They'**ve been** affected by it,	**haven't** they? [falling tone]	agreement with *yes (Yes, they have.)*
2	affirm. + affirm.	He'**s gone** back,	**has** he? [rising tone]	agreement with *yes (Yes, he has.)*
3	neg. + affirm.	She **never talked** to anybody,	**did** she? [falling tone]	agreement with *no (No, she didn't.)*
4	affirm. + neg.	You'**ve worked** hard,	**haven't** you? [rising tone]	anticipated agreement with *yes (Yes, I have.)* but open to challenge with *no (No, I haven't.)*
5	neg. + affirm.	He **didn't get** it,	**did** he? [rising tone]	anticipated agreement with *no (No, he didn't.)* but open to challenge with *yes (Yes, he did.)*

- Types 1 and 2 contain an affirmative statement by the speaker in the main clause, and an expectation of a *yes*-answer as confirmation in the tag.

- Type 3 contains a negative statement by the speaker in the main clause, and an expectation of a *no*-answer as confirmation in the tag.

40

- Type 4 contains an affirmative statement by the speaker in the main clause, and a more neutral possibility (i.e. of a *yes-* or a *no-* answer) in the tag.

- Type 5 contains a negative statement by the speaker in the main clause, and a more neutral possibility (i.e. of a *yes-* or a *no-* answer) in the tag.

For example, the pragmatic function representing requests can also be expressed and created by QTs as interrogatives. These can be formed in the pattern of negative clause + affirmative tag with the fall and rise intonation pattern, see the examples in Table 2.2 for further illustration (Carter & McCarthy 2006, p.198; also see Figure 7.1 for more pragmatic functions generated by QTs).

Table 2.2 *Interrogatives as requests*

Polarity	Falling tone	Rising tone
neg. + affirm.	You couldn't carry **this** for me,	**could** you?
neg. + affirm.	You haven't got any chocolate **bi**scuits,	**have** you?

QTs are known to have several pragmatic functions classified under information-oriented QTs (which are mainly about the exchange of information between speaker and addressee, covering such pragmatic functions as: establishing common ground, topic-initiating, surprised reactions, stating a fact or opinion (including self-monitors), acknowledging responses, and challenging) and action-oriented QTs (which are used to "give or demand goods and services and include requests, offers and suggestions") (Barron *et al* 2015, p.6; Axelsson 2011; Kimps *et al* 2014, pp.81-82; also see Chapter 7, section 7.2 for more information).

QTs usually take place at the end of the clause (as demonstrated above); however in informal speech, QTs can interrupt the clause (see examples 2.4 and 2.5 below).

Example 2.4   **That's** odd, *isn't it*, from a tutor?

Example 2.5   **It was** perhaps your team, *was it,* that was round there?

In reporting structures, it has been noticed that QTs may appear before the reported clause, especially if the reported clause seems to be unusually long:

Example 2.6 [commenting on the recipes of a famous cookery-book writer]

> **You** always **know**, *don't you*, that what you make will be suitable, and light, and that it will taste all right too. (For more information on reporting structures and question tags, see Carter & McCarthy 2006, p.198)

Also, it has been found that anticipatory *it* clauses may also be interrupted by a QT:

Example 2.7   **It's** true, *isn't it*, what they said about him?

Additionally, there is a very prominent type or formation of questions called fixed tags (also known as invariant TQs) which appear frequently in spoken language and are considered to be one of the significant forms of spoken grammar. Fixed tags primarily perform the pragmatic function of checking and making sure that something (a segment of the topic in the discourse) has been understood, or to confirm that an action has been agreed upon. Fixed tags can also be used for shared knowledge and facilitative uses as in example 2.11 below. This type of question is served in the discourse, and spoken grammar in general, with a wide range of forms such as: *(all) right*, *okay*, *yeah*, *eh*, *don't you think?* (Biber *et al* 1999; Carter & McCarthy 2006; also see Chapter 7, section 7.4.2.1 for more on fixed tags). Below are some examples (2.8- 2.11) illustrating fixed tags in the discourse:

Example 2.8   So we're meeting at 7 outside the pizza place, **okay**?

Example 2.9   Let's stop talking in circles, **right**?

Example 2.10   Don't tell anyone about this, **yeah**?

Example 2.11   Oh well, what on earth can we do about it, **eh**?"

(Carter & McCarthy 2006, p.198).

This type of fixed-tag question formation has been captured quite often in the datasets (ICE-Ireland and ICE-Canada) along with **echo questions**. Echo questions, are very common in spoken language. They occur in the form of declarative word order and with a wh- word such as: *where? which? what? the what? a what? who? The what stuff?* and *how?*. Their typical pragmatic function in the discourse is to request further clarification about a segment in the discourse, or the noun phrases, or parts of them, which may not have been heard by the listener correctly (Biber *et al* 1999; Carter & McCarthy 2006). However, echo questions, according to Biber *et al* (1999), are the questions that "request confirmation of what has already been said, by repeating part of its content. Some echo questions repeat the structure of what was said earlier using interrogative intonation, or they make the purpose of the echo question clear by the use of the words *did you say*" (Biber *et al* 1999, p.1101). See the examples 2.12 to 2.15 below for further illustration.

Example 2.12
A: Big day tomorrow. Got to go to the Phoenix.
B: Got to go **where?**
A: Got to go to this very formal meeting of all these academic people.

Example 2.13 [Talking about problems with European bureaucracy]
A: What was it? The European Commission?
B: Mm. *Translation service*.
A: **The what**, sorry?
B: *Translation service*. And they were just so badly organized. It was just
      unbelievable.

Example 2.14

A: That looks like a *dinosaur*.
B: Like **a what**?
A: A *dinosaur*.

Example 2.15

A: But apparently the president of the guild, he's a really nice bloke Alex says, cos, you know, she does all *the party stuff*.
B: Does **the what stuff**?
A: She does all *the politics* of the department stuff.
B: Politics.

(Carter & McCarthy, p. 199)

Having discussed echo questions, we will now move on to follow-up questions (for more on echo questions, see Chapter 7, section 7.4.2.2).

Follow-up questions are commonly used in the discourse with the typical pragmatic function of serving as a signal of engagement and attention shown by the listener to keep the conversation going by inviting further responses or to expand the discourse by requesting further specification (Carter & McCarthy 2006). Follow-up questions include a variety of different types and occur frequently in speech; some of these types are reduced questions with wh- words that occur either alone, or as a wh- word + a substitute word, or with stranded prepositions as in the following examples 2.16- 2.18 below:

Example 2.16

A: You're not staying in tonight, you're going out for dinner.
B: Oh. **Where to**?
A: I'm not telling you, a surprise.

Or alone or with a substitute word as in the following examples 7.14 and 7.15:

Example 2.17

A: Is he warden of the whole thing?
B: **Who**?
A: Doctor Thornton.
B: I don't know.

Example 2.18 [In a restaurant; A is the customer; B is the waiter]

A: I'll have that one.
B: **Which one**?
A: The king prawn in lemon sauce.

(Carter & McCarthy 2006, pp.199-200)

In addition to that, according to Biber *et al* (1999) and Carter and McCarthy (2006), follow-up questions can adopt a pragmatic function in which they are used as a signal of engagement and attention shown by the listener. They are more like the response tokens (RTs) *uhum*, *yeah*, and *really* indicating supportive responses as in the following examples 2.19 and 2.20:

Example 2.19

A: I went to school with her.
B: **Did you**?
A: Mm.

Example 2.20

A: And on mama's tree she's got some raspberries and tomatoes.
B: **Does she**? That's great."

(Biber *et al* 1999, pp.190-191; Carter & McCarthy 2006, p.200)

Sometimes, the follow-up and tag questions appear in a doubling type of structure with a typical intonation pattern for the sake of marking emphasis. In this case they typically occur with a negative clause followed by an affirmative tag (usually with two affirmatives) (see Table 2.1 and the forms below for more illustration). Nevertheless, follow-up and tag questions usually occur for the sake of keeping the conversation going by inviting further responses from the listener. This can be seen clearly with formulaic follow-up questions which occur in order to expand the discourse or to request further specification. Here are some of the forms by which question tags are seen: *what about? where to? who? which one? did you? does she? oh they haven't, have they?* (with falling and rising intonation), *oh he does, does he?* (with falling and rising intonation), *like what?*

how come? so what? what for? and like what? (Biber *et al* 1999; Carter & McCarthy 2006).

Furthermore, we have what are called the **two-step questions** and responses which are one of the question constructions formed and used in speech for politeness purposes (and others mentioned below). The two-step questions generally involve a two-step process occurring in a way that the first question acts like "a preface" for the upcoming question. This type of question usually happens in order for the speaker to avoid being rude, too direct, or too general (Carter & McCarthy 2006). Also noteworthy is that yes-no questions are only rarely self-contained. While they may function to elicit specific information, such questions are normally asked as a preface to further questions. For example, the question *Are you going to the match tonight?* anticipates an answer which may then be followed by a further more personal or specific question:

Example 2.21

A:	**Are you going to the match tonight?**
B:	Yeah, I am.
A:	**Do you mind if I tag along?**
B:	Sure. We're leaving around seven.

Example 2.22

A:	**Are you in this Sunday afternoon?**
B:	I expect so. I think we might be going out later.
A:	**Okay, do you mind if I pop round to pick up the drill?**
B:	Of course not.

(Carter & McCarthy 2006, p.201)

In addition to that, we have another type of two-step question called the pre-question in which the speaker asks a question in order to have permission for the next question, or to show more respect. Politeness or formality between the speaker and the listener may trigger this type of question as in the following examples 2.23 and 2.24.

Example 2.23

A: **I wondered if I might ask you something**?
B: Sure.
A: Would you be able to write a reference for me?

Example 2.24

A: **Can you tell me something**? What time is the rubbish collection on Mondays?

(Carter & McCarthy 2006, p.201)

Lastly, we have the preface question *do you know what?* which occurs a lot in

spoken language as a four-word cluster with the pragmatic function of showing what the

speaker considers newsworthy or important information to the listener as in the following

example 2.25.

Example 2.25

A: **Do you know what?**
B: What?
A: Roger's mum's bought Rachel a jumper. Isn't that sweet?
B: Yeah.

(Carter & McCarthy 2006, p.202)

Irish English has thoroughly explored questions in spoken discourse and there have

been major works conducted related to this area. For example, we have tag questions which

have been studied in terms of form and functions. One of the more central and key works

conducted on tag questions is by Barron *et al* (2015). It provides a corpus analysis of tag

questions (TQs) use in Ireland and Great Britain using spoken data from the Irish and

British components of the International Corpus of English (ICE). Analysis is based on the

formal and functional level, investigating form- functional relationships. Findings show

many commonalties in the use of TQs across the varieties as well as some contrasts. For

example, there is a lower use of TQs in Irish English and in a range of variety-preferential

features on both the formal and functional levels. Also, the paper shows how an in-depth

analysis of form-function relations, together with a fine-tuned investigation of sub-functions, provides an insight into formal preferences. Canadian English on the other hand has been under investigation in relation to questions formed as utterance-final tags (which can also be referred to as fixed tags) in which *right?* and *you know?* and some other tags like *eh* were explored (Denis & Tagliamonte 2016). Echo questions have also been studied in Canadian English where questions like *So who? Like how? Just what?* in the conversations of young Canadians were investigated (Tagliamonte 2005; also see Chapter 7, section 7.2 for more information and examples extracted from ICE-Ireland and ICE-Canada on questions and tags).

2.4.5 Stance markers

Stance markers are the expressions that signal stances, attitudes and points of view towards segments of discourse, and they all serve the same pragmatic functions but through different forms and features. Stance and the related area of evaluation has been widely covered in written language, particularly in the context of academic writing (see for example, Jiang & Hyland 2015; Thompson & Hunston 2019) but its coverage in spoken contexts is relatively sparce in comparison. Common stance markers include: *I think*, *basically*, *to be frank*, *frankly*, *certainly*, *clearly, honestly, to be honest, hopefully, ideally, if you ask me, I'm afraid, I must admit, I must say, in fact, no doubt, obviously, of course, really, sadly, seriously, (I'm) sorry, strictly speaking, to tell you the truth*, and *unfortunately* (Biber *et al* 1999; Carter & McCarthy 2006). See Chapter 9, section 9.2, in which more existing literature on stance markers has been covered in terms of how they have been viewed, approached, and studied or investigated.

Most studies investigating stance markers have a similar focus and approach, which is usually the corpus-based approach. For example, we have the study of Precht (2003) which, like the present study, investigates two different varieties of English: British and American. Precht (2003) used the British and American conversation elements of the *Longman Corpus of Spoken and Written English* and her findings propose that American speakers tend to use more stance markers as, what she calls, "affect markers" such as *cool* and *wow* while British speakers tend to use "evidential markers" such as *a bit* in order to hedge propositions (see also Jones 2016). In general, stance has received little systematic attention in terms of how it is used across varieties of spoken language so this study will add to its description in the comparative context of Irish and Canadian spoken English.

2.4.6 Interjections

Interjections, according to Biber *et al* (1999), normally refer to exclamative utterances consisting of single words that do not easily fit into the major word classes (noun, verb, adjective, adverb) such as: *bother*, *crikey*, *damn*, *god*, *goodness* (*me*), *gosh*, (*good*) *heavens*, *hooray*, *jeez*, *ooh*, *oh no*, *oops*, *ouch*, *ow*, *ugh*, *tut-tut*, *whoops*, *wow*, *yippee*, *yuk*. All these items express positive or negative emotional reactions to what is being or has just been said or to something in the situation (Biber *et al* 1999, p.1083; Carter & McCarthy 2006). (See Chapter 6, Examples 6.17 and 6.18 for the interjection *oh* found in the datasets). We note that definitionally there is overlap between interjections and response tokens and some of the forms which manifest as interjections also function as response tokens. For example, work by O'Keeffe and Adolphs (2008) on Irish English and as a core feature of spoken language, interjections are very under-researched and so this study will add to their comparative description in the context of the two spoken varieties under examination.

2.5 Conclusion

In this chapter, we have surveyed some of the defining literature on spoken grammar. In the process, we have introduced its key concepts including definitions, key forms, features, and functions. We have also sought to answer the question: What is spoken English grammar? Moreover, this chapter has presented some of the general studies on the spoken grammar of English. Where possible, it has covered work on Irish and Canadian Englishes discourse features (forms) and functions such as discourse markers, response tokens, stance markers, hedges, and interjections. As mentioned, later chapters will provide more detailed reviews in relation to specific forms and work on Irish and Canadian Englishes. Based on the literature review provided thus far, it is clear that fewer studies have been conducted on the spoken grammar of Canadian English in comparison to Irish English. While there has been much work done on Irish English using corpus data, there is not one study that comprehensively describes the forms and functions of spoken grammar in Irish English. This points to the importance of this study in terms of providing a wide-ranging baseline description of spoken grammar of two varieties.

The next chapter (Chapter 3) will discuss how such studies of spoken grammar/discourse have been approached in the past, and how they will inform this current study. This will be done through introducing some key concepts such as: Variational Pragmatics, Corpus Pragmatics, Spoken Discourse Analysis as well as Conversation Analyses (CA) and Corpus Linguistics (CL). In addition, the next chapter will introduce the Iterative Approach that will be taken in this study in a combined top-down and bottom-up approach.

CHAPTER 3

**A COMPARATIVE FRAMEWORK FOR SPOKEN GRAMMAR
ANALYSIS ACROSS TWO VARIETIES**

3.1 Introduction

This chapter aims to demonstrate how spoken grammar/discourse analysis studies have been approached in the past, and how this current study will be approached. Before going in depth, it is important to mention that the focus on language variation at a pragmatic level is a relatively recent development because the study of language variation traditionally focused on phonological, lexical and syntactical levels (Schneider & Barron 2008; Cheshire 2016). Moreover, the focus has usually been based on historical variation and geographical variation. Thus, as the focus has shifted from phonology, lexis, and syntax, the type of variation or variables have also shifted. There is now much more emphasis placed on variation in social space (i.e. context) in addition to the historical and geographical variation (Clancy, O'Keeffe & Adolphs 2019). Yet, the pragmatic variation studies which have been done comparatively across two varieties (so far), have not provided a full description of spoken grammar (pragmatic variation) across two varieties (especially between Irish and Canadian Englishes) (see O'Keeffe forthcoming). Rather, they have looked at some prominent forms of spoken grammar across two varieties based on Variational Pragmatics (VP) from *macro-social* and *micro-social* perspectives. These allow researchers to account for the influences of different factors (of *macro* such as region, age, gender, ethnicity or social class and of *micro* such as power, social distance or register) on pragmatic choices made (Clancy 2010). In relation to Irish English, O'Keeffe (forthcoming), in her survey of corpus linguistics (CL) and Irish English, supports this point, saying that while there are many studies focusing on single items or features of lexis, form or pragmatic features (often across sociolinguistic variables of use), very few have looked systematically across at these items (for example, some discourse markers, hedges,

vagueness markers, etc. are well-researched but no one has looked at all of these features).

These types of variation: historical, geographical and social, can be examined from

different levels (originally proposed by Schneider & Barron (2008b) and developed further

by Félix-Brasdefer (2015)) in which each level is concerned with and focuses on a specific

aspect and component of pragmatic analysis as illustrated below in Table 3.1 (Schneider &

Barron 2008, 2008b, pp.20-21; Barron and Schneider 2009; Félix-Brasdefer 2015;

Schneider 2021a, 2021b).

Table 3.1 *The eight levels of pragmatic analysis (proposed by Schneider & Barron (2008)*
and developed further by Félix-Brasdefer (2015))

Level	Description
Formal	This level is concerned with the analysis of linguistic forms such as discourse markers or hedges.
Actional	This level focuses on the realisation and modification of speech acts.
Interactional	The focus here is on sequential patterns such as adjacency pairs, exchanges, or phases (for example, openings and closings).
Topic	The focus here is on how conversational topics are selected, addressed, developed etc.
Organisational	This level deals with turn-taking phenomena such as pauses, overlaps, interruptions, or backchannels.

Stylistic	The stylistic level investigates the choice of address forms/terms, including the relevant distinctions of formal – informal, serious – non-serious, transactional – interactional, and pertain.
Prosodic	The prosodic level investigates variation in intonation, pitch, stress, loudness and speech rate (which are all looked at as features of pragmatic resources).
Non-verbal	The non-verbal level looks at facial expressions, gaze, gestures and posture as pragmatic resources by which pragmatic vararion can be measured and investigated.

Thus, in this study, the primary framework of VP at the formal level is used to study pragmatic variation in geographical space rather than social space. Schneider and Barron (2008) argue that VP is "contrastive by definition" and can provide a framework through which the analysis of pragmatic similarities and differences can be conducted between and within different language varieties (p.21). Inter-varietal studies of pragmatic variation focus on comparing two or more language varieties, for instance, Plevoets *et al*'s (2008) study on Netherlandic and Belgian Dutch and O'Keeffe and Adolphs' (2008) study on response tokens (RTs). However, the present study investigates pragmatic variation intra-varietally, within the same language (but within different varieties). In other words, both of these varieties (Irish and Canadian) belong to the same language (English) as opposed to English versus French or Arabic, but within different varieties (as opposed to hedging in Irish English in family discourse versus institutional discourse) (see section 3.2 in which variational pragmatics has been explored in detail in relation to this context). Specifically,

this study looks at all of spoken grammar of both varieties as the main variable. This allows for the intra-varietal investigation of differences (i.e. pragmatic variation within the varieties of the same language in two different geographical zones) without looking at the social triggers (such as socio-economic status, ethnicity, gender, and age known as non-linguistic variables). Note that there are so many variables that we could look at in this study but because we are looking at an uncharted area in its entirety (i.e. spoken grammar), the only broad variable is a geographical one.

In addition, language variation at the pragmatic level has proposed that variational pragmatic analysis can be combined with CL resulting in Corpus Pragmatics (CP). CP has, according to Clancy and O'Keeffe (2015), "successfully combined the methodological field of conversation analysis (CA) with that of corpus linguistics in order to provide a much more fine-grained analysis of spoken language than would be possible if each were used in isolation" (p.241). As Walsh *et al* (2011) observe, "when the researcher records, transcribes, annotates and builds a small contextualised spoken corpus, a different landscape of possibilities opens up in areas beyond lexis to areas of use (especially issues of pragmatics, interaction and discourse)" (Walsh, Morton, and O'Keeffe 2011, pp.326-327). This chapter introduces some of these concepts such as: Variational Pragmatics, Corpus Pragmatics, and Spoken Discourse Analysis which details the way spoken discourse/grammar analysis has been approached in tandem with Conversation Analyses (CA) and Corpus Linguistics (CL). It will also define the processes that are required for this study, namely the Iterative Approach: a two-way process involving a top-down approach (function to form) and a bottom-up approach (form to function).

3.2 Variational Pragmatics

Variational Pragmatics is an analytical framework which was first introduced by Barron and Schneider (2005) and whose fundamental goal is to address the research gaps existing in both modern dialectology and pragmatics. As Schneider and Barron (2008) note, VP studies pragmatic variation in geographical and social space, which can investigate both inter-lingual differences (i.e. pragmatic variation across languages/ varieties) and intra-lingual differences (i.e. pragmatic variation within one variety) and the geographical and social triggers (such as region, socio- economic status, ethnicity, gender, and age) that can affect the choice of one pragmatic strategy over another (Schneider & Barron 2008a; Barron & Schneider 2009). Additionally, VP is not limited to the above-mentioned five types of macro-social factors that may have an impact on pragmatic variation, it is also open to other macro-social factors (such as education and religion) and micro-social factors (such as power and social distance or register) as well. These should all be considered as a part of this research framework due to their influence on pragmatic variation (Schneider & Barron 2008a, 2008b; Clancy 2011a, 2011b; Félix-Brasdefer 2015; Schneider 2021a, 2021b). Yet, in terms of a practical research agenda, Schneider and Barron (2008b) have suggested that, "currently, variational pragmatics concentrates primarily on macro-social variation. It aims at determining the influence of each macro-social factor on language use individually... At a later stage it will be necessary to systematically include micro-social variation and to investigate the interaction between micro-social and macro-social factors" (p.16; see also Clancy 2011a, 2011b). Nevertheless, it cannot be denied that both macro and micro-social factors have a clear impact on pragmatic choices which are essential to our

understanding of language-use differences (Félix-Brasdefer 2015; Schneider 2021a, 2021b).

Schneider and Barron (2008) criticized the degree of representativeness in studies where it assumes that language communities of native speakers are homogenous wholes when language variation is considered, thus, in a sense, negating the impact of social variables on language communities. They also claim that many researchers in this area employ participants from student communities, often from their own courses, thereby further compromising representativeness. While such studies are undoubtedly insightful, this lack of representativeness makes it difficult to formulate reliable generalisations about typical language use (Barron & Schneider 2005; Schneider & Barron 2008a, 2008b; Clancy 2011a, 2001b). As a result, VP was depicted by Schneider and Barron (2008a) as the way out of this lack of homogeneity (Schneider 2021a).

It is also noted that VP is flexible on a methodological level. For instance, it has been applied in several corpus-based studies and, as illustrated by many of the early VP studies, others have used DCTs as a means of eliciting data (see Norrby *et al* 2012). CL is particularly suited to applying the VP framework because the concept of representativeness is a defining principle in corpus design. As Clancy (2011a) notes, one of the strengths of CL is that representativeness has long been a core concern. O'Keeffe (forthcoming), in her review of work on Irish English using CL, notes that since the turn of the century, and especially influenced by Barron and Schneider (2005), there has been a growing mass of studies on Irish English that use CL as their methodology (see also Vaughan & Clancy 2016).

Barron and Schneider have cited two studies concentrating on regional language variation in English in which the issue of lack of representativeness does not appear due to the useful tools offered by CL, thereby mitigating the initial criticism of pragmatic research which was made before VP (Barron & Schneider 2005; Schneider & Barron 2008a; Clancy 2011a). These studies (which came before VP) are corpus-based studies in which Tottie (1991) and McCarthy (2002) investigate the differences between backchannels (or response tokens) in British and American English. Tottie employs the London Lund Corpus (LLC) and the Santa Barbara Corpus (CSAE) and McCarthy utilizes the Cambridge and Nottinghan Corpus of Discourse in English (CANCODE) in addition to a similar-sized sample of the Cambridge North American Spoken Corpus (CNASC). McCarthy (2002) maintains that cross-corpora comparisons of different varieties of the same language are useful for the study of variational pragmatics (Félix-Brasdefer 2015; Schneider 2021, 2021b). Thus, they provide safer ground for generalisations. All four corpora employed by Tottie and McCarthy have been specifically designed to represent standard British (LLC and CANCODE) and American (CSAE and CNASC) English, and work to alleviate some of the criticisms aimed at cross-cultural pragmatic research. We will return to the notion of representativeness in Chapter 4 (section 4.3.2), especially in regards to the corpora used for this present study, as discussed in section 4.4 (Schneider & Barron 2008; Clancy 2011a).

Therefore, having established the analytical synergy of Variational Pragmatics and Corpus Linguistics, it is argued that this study can investigate spoken grammar intra-varietally in Irish and Canadian English using representative corpus data (in this case spoken data from the International Corpus of English). According to Barron (2015), systematic contrastive and comparative analyses including Irish English and further Inner

or Outer Circle varieties are needed. In Variational Pragmatics (across the varieties of English), many of the contrastive and comparative analyses conducted thus far, in relation to spoken grammar generally and pragmatic markers specifically, are with British English and American English. The fact that British and American Englishes have been a primary focus of cross-varietal research is not surprising (although lately that focus has started slowly shifting to other Inner and Outer Circle varieties) for the following reasons. Firstly, in the 19th century and in the first half of the 20th century, language use and varietal differences (which pre-dated Pragmatics) were studied as a part of dialectology (dialect geography) which focused exclusively on investigating regional variation (Schneider & Barron 2008). During that time, the term *dialect* was used in a narrow perspective sense, describing regional varieties of a language. The two major Inner Circle varieties at the time were British English and American English (due to many factors such as their political and military power) (Schneider & Barron 2008b). Their prevalence in the literature is noted by Cheshire (1991), who states in her survey of the literature on the differences between British English and American English, that the "phonetic, phonological, lexical and syntactic differences between the two national varieties have long been recognised and described" (p.13).

With the initiation of sociolinguistics in the 1960s, the research focus shifted radically from regional variation to the social variation and factors which are now the main focus and goal of Pragmatic Variation. However, even with this shift, it was easier to compare British English and American English in relation to social factors at the pragmatic level rather than other English varieties due to the greater availability of different types of linguistic data and meta-data at different levels of linguistics (cf., e.g., Walters 1988;

Chambers & Trudgill 1998; Schneider & Barron 2008). Another significant reason for the focus on BrE and AmE is that there are simply more resources and corpora available for BrE and AmE (designed for variational pragmatic studies) compared to other Englishes largely because there is a greater critical mass of researchers interested in these inner circle varieties. This study, by focusing on varieties other than British and American English, can offer a basis for more comparative research across both the Irish and Canadian varieties of English in the future, involving the more in depth study of macro and micro-social variables.

Although this study is not a fully traditional sociolinguistics study (because the only broad variable taken to study spoken grammar here is the geographical one), it is still concerned with variational pragmatics in the sense that it looks at many forms (and their pragmatic functions) across both varieties. Thus, it is very much a study of spoken grammar as a whole (in a comparative and representative way between and across two varieties of English located in two different geographical spaces). In other words, it is beyond the scope of this study to consider extra-linguistic variables at an in-depth level, given the broad-ranging nature of this comparative and representative study of spoken grammar across two varieties. In short, because no one has yet compared spoken grammar as a variable across Irish and Canadian English, the scale of this task alone mitigates against going into the important variables such as age or gender, among other similar variables, but it paves the way for this level of research to be undertaken intra-varietally.

3.2.1 Core principles in Variational Pragmatics

VP analyses are led by the principles of empiricity, contrastivity, and comparability (Schneider 2010, Schneider 2014, pp.362-365):

Empiricity

VP research is based on empirical data but it is not prescriptive on the nature of the data. Data used in VP studies can range from collections of material observed (e.g. field notes on interaction), elicited discourse completion task (DCT) or interview data or corpus data. VP studies are not based on impressionistic and episodical (i.e., intuitive and fabricated) data) (Barron *et al* 2015; Schneider 2010; Schneider 2014).

Contrastivity

VP studies also work within the principle of contrastivity. This is essential when looking at features or patterns of a variety. As Barron *et al* (2015, p.3) note, claims of distinctiveness alone are not warranted unless they are based on empirical comparison(s) (see also Schneider 2014). In other words, simply describing a feature as frequent in a variety does not bring into relief whether this is specific to this variety. It is only by contrasting it that this can be substantiated. As Barron *et al* (2015, p.3), point out:

> linguistic features are only considered variety-specific or variety-preferential if the variety under study is contrasted with at least one other variety of the same type and language, since a mere recording of a feature or pattern in a particular variety in the absence of comparative research on further varieties does not warrant statements as to the distinctiveness nor relative saliency of that feature in the variety at hand (p.3).

As a result, it is impossible to establish any variety-exclusive and variety preferential features of any (regional, socioeconomic, ethnic, etc.) dialect, if the dialect itself is not explicitly compared to a dialect of the same kind of the same language. (Schneider & Barron 2008b; Barron & Schneider 2009).

Comparability

The third core principle is comparability. Data are considered comparable when the social variables under investigation are controlled. In other words, there is a need to compare like with like. An example of this is O'Keeffe and Adolph's (2008) response token study using British CANCODE and Irish LCIE spoken corpora. Because both corpora were designed using the same data collection and categorisation matrix (see McCarthy 1998; Farr *et al* 2004) this allowed for the controlled comparison of response tokens among (circa) 20 year old female participants, who were friends, talking in a socialising context, and so forth. This meant that the researchers were not comparing, for example, response tokens in institutional settings across age groups in Irish English with the same tokens in conversations in British family settings.

This study is based on data from two corpora (ICE-Ireland and ICE-Canada) which were designed using the same sample frame (see Kallen & Kirk 2008; *International Corpus of English* 2016). Thus, the ICE corpus suite has been designed for variational research and its metadata is replicated across all collections, making it very comparable. While the first principle of *empiricity* is generally adhered to in the examination of variation, what VP brings is a rigour through the importance it places on *contrastivity* and *comparability* so as to lead to robust findings about distinctiveness between (and across) controlled data (Schneider & Baroon 2008). This present study follows these three principles:

- *Empiricity:* the data is naturally occurring Irish and Canadian English drawn from the ICE-Ireland and ICE-Canada collections.

- *Contrastivity:* results from ICE-Ireland and ICE-Canada will be contrasted in terms of frequency and use.

- *Comparability:* because the sample frame for both corpora is the same, the data is highly comparable and for this reason, these corpora were chosen over other far less comparable options (e.g. LCIE or The Strathy Corpus).

3.2.2 Variational Pragmatics and social structure

As mentioned earlier (in section 3.2), deviations in social structure, whether macro-social factors (such as region, socio- economic status, ethnicity, gender, and age) or micro-social factors (such as power, social distance, or register), are recognised as influencing language use (see Kallen 2006, O'Keeffe & Adolphs 2008, Elwood 2011, Lucek 2011, Barron *et al* 2015). An example of the variational pragmatic framework being employed is Clancy's (2011a, b) study on two Irish families which considers the socio-economic class/variation: one is a settled family from the Limerick city area and the second is a traveller family (with the same gender profile). His research is on hedging, examining the use of kinship terms in naturally occurring data. The results showed that the settled family uses more hedging than the traveller family. Contrastively, the traveller family uses more kin titles rather than first names. These findings reflect on the notion of individuality found in the settled family and the collectivity and sense of family found in the traveller community (as suggested by Clancy). Yet, further research is required before any conclusive claims are stated due to the fact that the data underlying the analysis are not comparable on all levels. This leads to a possible influence of the social structure variation of age, socio-economic status, level of education, as well as ethnicity. (Barron & Pandarova 2016).

In addition to that we have the influence of age and gender which has been the focus of research by Murphy (2010). She investigates a range of hedging devices, taboo language, amplifiers, boosters and vague category markers in a female language corpus

consisting of three sub-corpora of 20-29 year-olds, 40-49 year-olds, and 70-80 year-olds. Murphy (2010) also created a male corpus to use for comparative purposes. Some of the significant findings of her research are that more hedges are used by women in their twenties and forties than they are with women in the 70-80 year-old range. Additionally, women in their twenties prefer using the forms *like* and *actually*, whereas those in their forties prefer using *you know* and *I think.* In regards to gender, she found that males use less hedges as they age (see more work done on age and gender how they can have an impact on pragmatic choices and variations: Farr & Murphy 2009; Schweinberger *et al* 2009; Murphy 2012; Schweinberger 2012).

Finally, Barron and Pandarova (2016) provide a corpus analysis of tag questions (TQs) across region and gender in Northern of Ireland (NI) and Republic of Ireland (ROI). They do this within the variational pragmatic framework, focusing on using contrastive comparable empirical data. The investigation is limited to the ICE-Ireland face-to-face and telephone conversations text types. Thus, a total of 241 TQs were identified in the speech of NI and ROI males and females taken together. This research is important because of the separation between the NI sub-corpus and the ROI sub-corpus which allowed the research to go further than other researches have gone. For example, Barron *et al* (2015) found that both ICE-Ireland (ROI and NI) speakers use significantly fewer TQs than ICE-GB speakers, this research sheds further light on this fact by showing that ROI speakers use significantly more TQs than NI speakers. Also, in terms of function, speakers in ICE-GB used less question TQs, compared to both NI and the ROI speakers who used questions to a similar extent and both groups used more question TQs. In addition to that, NI speakers were also found to employ more statements than ROI speakers, whereas ROI speakers

employed more statement question blends, showing more similarities in this respect to the ICE-GB data than to the NI data.

In relation to gender analysis, there are facts highlighted on this divergence. While males and females across the two datasets do not differ significantly in their use of statements and statement-question blends, there is a significant clear variation specifically in terms of women's preferences. For example, NI women use more statements than statement-question blends, while ROI women exhibit the reverse tendency. As a result, the observed regional differences can be attributed to significant variation in NI and ROI women's TQ function preference.

These research examples clearly show the interplay of social factors which demonstrates the importance of considering social distributions in future analyses of regional corpora (Barron & Pandarova 2016).

3.3 Spoken Discourse Analysis

When it comes to spoken discourse analysis, we always encounter two key methodological approaches: Conversation Analyses (CA) and Corpus Linguistics (CL) which have both been used on their own to study spoken discourse (Walsh *et al* 2011). CA is a research tradition that grew out of ethnomethodology which has some unique methodological features used for the investigation of utterances and spoken interactions by analysing data on a qualitative level (O'Keeffe 2003; Wooffitt 2005; Palma 2013). CA investigates the social organisation of conversation and what is called "talk-in-interaction." This is done by providing a detailed bottom-up review of materials collected from naturally occurring occasions of everyday interaction including tape recordings, transcriptions based on a micro

analysis approach to data in order to capture not only what has been said in the data but also how it has been said (Atkinson & Heritage 1984; Wooffitt 2005).

Walsh *et al* (2011) distinguish between 1) "pure" CL research in which "the description of the language of the corpus is an end in itself (descriptive corpus research)" (p.326), and 2) CL research which applies the methodology as a means to an end to help bring about an understanding of language in use in its wider interactional context. With this latter approach in mind, CL and CA can come together as a complementary means of analysing language because of their micro analysis approach to data. As O'Keeffe and McCarthy (2012) point out, in the beginning of CL development, the goal was to create large written corpora with the focus being on semantic and lexical patterning in order to cater to the needs of lexicographers. Because of this, these "large corpora were lexically rich but contextually poor. That is, when a researcher looks at a lexical item in a mostly written corpus of 100 million words or more, it is detached from its context" as the original focus was not on discourse context (Walsh *et al* 2011, p.327). But, when a researcher takes the time to record, transcribe, annotate, and build smaller more context oriented spoken corpora, it opens the door to more enriched areas of study "beyond lexis to areas of use (especially issues of pragmatics, interaction and discourse" (Walsh *et al* 2011, p.327).

Thus, the shared point of interest in conversation analysis that plays a crucial role is context, which makes Drew and Heritage (1992) define CA as "combining a concern with the contextual sensitivity of language use with a focus on talk as a vehicle for social interaction" (p.16). Also, in highlighting the importance of context in CA, Drew and Heritage (1992) state that, "the CA perspective embodies a dynamic approach in which the 'context' is treated as both the project and product of the participants' own actions and

therefore as inherently locally produced and transformable at any moment" (1992, p.19).

Walsh *et al* (2011) conducted a study where they combined CL and CA as an iterative process (from CL to CA, back to CL and so on) to provide enhanced descriptions of spoken interaction in the context of small group teaching in higher education. They began with CL by focusing largely on words and combinations of words, and then used CA by highlighting pertinent interactional features. In terms of the analysis steps, the first step of analysis (using CL) allowed them to extract and quantify recurring linguistic features. After linking these recurring features to their context in the corpus, they came up with contextual "patterns." The second step of analysis (using CA) drew upon these contextual patterns in the quantitative analysis and investigated them more closely. For instance, in the corpus exploration, there were interesting findings around the frequency and use of certain discourse markers around specific contexts. This led them to the CA investigation which, in turn, produced interesting findings above the level of turn and in relation to specific interactional features. As a result, the CL and CA analysis provided detailed descriptions of the interaction from three perspectives:

1. The linguistic perspective which describes the use of high frequency items, keywords, multi-word units (MWUs), discourse markers, question forms and so on.

2. Interactional perspective which focuses on turn-taking and turn design, sequential organization and so on.

3. Pedagogic perspective which looks at specific pedagogic functions at a given moment to include eliciting, explaining, instructing and so on.

Walsh *et al* drew a noteworthy conclusion of their findings in this particular research when they observed,

Had we used CL on its own we would have achieved interesting lists of high frequency items which we could have explained functionally but it would not have brought us anywhere near the depth of understanding compared with what a CA framework could explain. Had we looked at the data purely from a CA perspective, we would have possibly identified the four main speech exchange systems but we would not have been able to back up the fact that the words and patterns they contain were actually high frequency items (that is, keywords, high frequency words and multi-word units). In addition, by drawing on quantitative methods within CL, we were able to reference our findings against another dataset (in this case LCIE). We can therefore safely assert that CL and CA are "well met."

(Walsh *et al* 2011, p.344).

This meeting of CL and CA is further developed in O'Keeffe and Walsh (2012) and represented as follows:
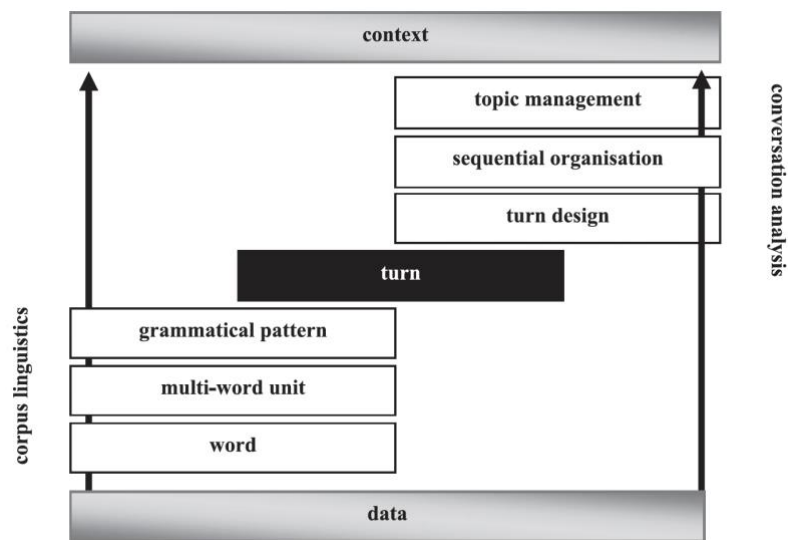
Figure 3.1 CA and CL as a combined framework for analysis of spoken language
(O'Keeffe & Walsh 2012, p.164).

As Figure 3.1 shows, the strength of the corpus lies below the turn and the strength of CA is from the turn level upwards. Of note, in Figure 3.1, we see mono-directional movement from the data to the context which moves from the bottom up. Important to this study is to engage with an iterative approach which is bi-directional, moving between bottom-up and top-down processes of analysis. Carter and McCarthy (2002) offer an interesting example of this process. In their study, they looked specifically at one BBC radio interview with the British Prime Minister, Tony Blair and used CA and CL as an

iterative process where they firstly applied the framework of CA in their analysis and then conducted corpus-based analysis on the same data. The CA analysis demonstrates that the interviewer and interviewee both adhere to and exploit the generic conventions of the interview in terms of turn-taking, topic management, and participant relationships. The interviewer reinforces an agenda of getting the interviewee to commit to action; the interviewee, in turn, responds cohesively and coherently and yet avoids direct commitment to action and maintains his topical agenda without losing face (and with useful soundbites delivered along the way, which are likely to be extracted and quoted in the following national news reports). As a result, Carter and McCarthy test how CA and CL can supplement each other and offer a more integrated way of understanding how conversational agendas are achieved when the two methods are used in combination than either of them can aspire to alone. In other words, CA and CL are both methodological means used to achieve the aim of the research questions (under investigation) (O'Keeffe 2003). The current study will build upon this approach but will also engage with the more recently coined "Corpus Pragmatics" in which CA and CL can be incorporated in a complementary way (as we discuss further below).

3.4 Corpus Pragmatics

Corpus Pragmatics (CP) is a relatively newly emerging field which resulted from the combination of the key methodologies of both Corpus Linguistics and Pragmatics which has been defined as "the study of the use of context to make inferences about meaning" and also as "the study of those relations between language and context that are grammaticalized or encoded in the structure of a language" (Levinson 1983, p.9; Fasold 1990, p.119; Aijmer

69

& Rühlemann 2015; Weisser 2015; O'Keeffe 2018; O'Keeffe *et al* in press). CP "has also successfully combined the methodological field of conversation analysis (CA) with that of corpus linguistics in order to provide a much more fine-grained analysis of spoken language than would be possible if each were used in isolation" (Clancy & O'Keeffe 2015, p.241). Thus it is evident that CL and Pragmatics have had an interactional relationship for quite some time. CL studies have previously drawn on pragmatics in order to have more accurate data analysis and pragmatic research questions have been addressed by using corpus data. However, the term Corpus Pragmatics has only been established recently. Within this new coinage, considerations have been made to best utilize CL in providing more accurate data analysis for pragmatics research; thereby increasing its recognition (albeit with some limitations) (Jucker 2013; Rühlemann & Aijmer 2015; O'Keeffe 2018). Although it is known that CL and pragmatics have different approaches to examining data, they are not contradictory, making it possible to utilize both of them together. CL moves from frequencies of forms to their functions (through an inductive process) known as the form-to-function approach (which is known as the bottom-up approach in this study, see sections 3.5, 4.5.1.2, and Appendices B-M). On the other hand, pragmatics takes the opposite route, starting with the specific pragmatic function to the forms which are typically used for that pragmatic function (through an inductive process) (known as the top-down approach in this study, see sections 3.5, 4.5.1.1, and Appendix A). This is typically used in the analysis of speech acts and related phenomena. This type of approach is known as the function-to-form approach (Vaughan & Clancy 2011; Aijmer 2018; O'Keeffe 2018). Using one of these approaches on its own can be criticized. For example, although the form-to-function analysis is effective at identifying forms in terms of frequency,

distribution, positions and collocations with different functions, it is still weak at identifying all of the occurrences of a particular function as it is form-driven (Vaughan & Clancy 2011; Rühlemann & Aijmer 2015; O'Keeffe 2018). Therefore, according to O'Keeffe *et al* (2020), CP employs both approaches. It does so as follows: First, usually by applying a form-to-function analysis in which the investigation begins with the forms under examination, using CL tools such as concordances and frequency and cluster lists. This leads the analysis towards results of the forms. Second, by applying the traditional approach associated with pragmatics in which "function-to-form" analysis is directed towards the function as previously mentioned (i.e. if one is looking at hedging, one can search for the well-documented list of possible exponents of this pragmatic phenomenon in the data). As a result, this mixture of methodologies allows linguistic items to be examined at both a structural (syntactic) and interpersonal (pragmatic) level which will enable us to understand how "words, utterances and text combine in the co-construction of meaning" (Walsh 2013, p.37; Clancy & O'Keeffe 2015, p.241; also see Chapter 4 for more information on the application of such approaches and how CP has been employed in this present research).

3.5 The Iterative Approach

As mentioned earlier (in section 1.1), the iterative approach used in this study has been formed from two different sub-fields of Linguistics (Corpus Linguistics and Pragmatics), resulting in Corpus Pragmatics. By employing top-down and bottom-up processes, the forms and pragmatic functions of spoken grammar of a language/variety (in our case, Irish and Canadian varieties of English) can be compiled and analysed. Combining these

approaches ensures that they work together to make up for any shortcomings they may have individually (as discussed in section 4.7).

Consequently, the top-down approach will be initially applied to gather the forms and pragmatic functions of spoken grammar based on the existing literature in order to conduct a function-to-form corpus analysis. These functions then provide important facts about some of the prominent forms of spoken grammar, such as pragmatic markers, in terms of frequency, functions, and so forth (Amador-Morena *et al* 2015; O'Keeffe 2018). Therefore, the top-down analysis encompasses the literature review of the following forms of spoken grammar such as: Pragmatic Markers, Deixis, Questions and Tags, Ellipsis, Headers, Tails, Social Routines, Swearing and Taboo Expressions, and Vocatives (of which some have been covered in Chapter 2, section 2.4 and in further detail in Chapter 4, section 4.5.1.1). Additionally, refer to Appendix A for the analysis of the forms and pragmatic functions of spoken grammar. It also refers to the terminology used in this study where items have various nomenclatures in the literature.

After applying the top-down approach and having an idea about what to expect from the datasets/corpora under investigation, the bottom-up approach will be applied in which corpus linguistic tools are involved (see section 4.5.1.2 for a more detailed explanation of the bottom-up process). The bottom-up approach with the application of the word frequency tool can be applied with the AntConc software (used in this present study) or any software which has the same functions (like Wordsmith Tools) (Anthony 2020; Scott 2021). The word frequency tool allows the researcher to count and present all the words in a corpus in an ordered list, which then helps to quickly identify the top 100 most frequent word-forms. For a full explanation of the use of word-frequency tools and how

they have been used in this study, refer to section 4.5.1.2. (also see Table 4.6 in Chapter 4 illustrating a sample of the top 20, and see Appendices B and C for the full list of the top 100 most frequent word-forms for ICE-Ireland and ICE-Canada).

As a next step, the clusters tool/the N-grams tool will be applied. The clusters tool is extremely important to such a study as it functions as a scanner, searching the entire corpus for 'N' length clusters (i.e. one word, two-word, three-word clusters) (refer to section 4.5.1.2 to see the functions of the clusters tool and how it has been used in this study). Use of this tool is essential in gaining a full picture of the clusters found in the datasets and avoiding any resulting amibiguity while concordancing. (For the full clusters lists of ICE-Ireland and ICE-Canada, see Appendices D- M).

After the application of the word frequency tool and the clusters tool, we move from the quantitative analysis to the qualitative analysis by using the concordance tool. The concordance tool is employed in order to determine whether or not the items mentioned in the wordlists and cluster lists are actual forms of spoken grammar. In other words, concordancing these items manually helps to eliminate the items that are not being used as forms of spoken grammar. Concordance tools can be challenging to navigate and thus Sinclair (2003) came up with a seven-step procedure to facilitate their use. These steps are discussed in Chapter 4, section 4.5.1.2 (refer to this section for a closer look at the uses of concordance tools in this research).

3.6 Conclusion

In this chapter, the primary framework of Variational Pragmatics used to study pragmatic variation in two different geographical zones has been introduced in order to investigate

intra-varietal differences (i.e. looking at differences (pragmatic variation) between two varieties of the same language). As mentioned earlier in section 3.1, VP is "contrastive by definition" and can provide a framework through which the analysis of pragmatic similarities and differences can be conducted between and within different language varieties (Schneider & Barron 2008b, p.21). Also, VP does not impose any theoretical or methodological orientation but puts pragmatics on the map of dialectology and variational linguistics (Barron & Schneider 2005, p.12). However, this chapter has discussed the analytical synergy of Variational Pragmatics and Corpus Linguistics (applied in this current study) in which it shows how VP analysis has been supported by CL in a number of ways which has led to Corpus Pragmatics through the iterative approach adopted in this study. Yet, (as mentioned before), this current study is not a fully-fledged Variational Pragmatics study (nor a corpus-based study, which will be discussed in Chapter 4, section 4.2.4, due to the methodology applied in this study). Therefore, Chapter 4 will explain the methodological concepts, tools, and approaches undertaken in this study. Doing so will reveal the nature of spoken grammar usage in both varieties comparatively in relation to their pragmatic functions and forms. Chapter 4 will also detail how the iterative approach has been applied in this study. It will also describe the corpora (data) used in the present study in terms of their design, limitations, and the rationale behind why they have been chosen. Lastly, it will demonstrate the application of the iterative approach on ICE-Ireland and ICE-Canada which will lead to quantitative and qualitative insights aided by the framework of Variational Pragmatics.

CHAPTER 4

DATA AND METHODOLOGY

4.1 Introduction

As previously stated in earlier chapters, this study seeks to address the outstanding research gap in the comparative study of Irish English and Canadian English in relation to spoken grammar; specifically, the pragmatic markers which are known to be the core components of spoken grammar (Carter & McCarthy 2006). Thus, this chapter will explain the methodological concepts, tools, and approaches undertaken in this study. For example, in this chapter we will discuss the main concepts and tools of CL which have been used as a methodological tool in order to reveal the nature of spoken grammar usage in both varieties comparatively in relation to their pragmatic functions and forms. Furthermore, this chapter will detail how the iterative approach (in which top-down and bottom-up processes are used to identify pragmatic markers and their pragmatic functions in spoken grammar) has been applied in this study. This is done by detailing the top-down analysis which was based on a framework for spoken grammar from existing literature and the bottom-up process which was based on micro-analysis of the data (both quantitative and qualitative) using corpus tools. In addition, the corpora used in this study will be described in detail in terms of their design, limitations, and the rationale behind why they were chosen. The two corpora used are the spoken components of two International Corpus of English (ICE) corpora, namely ICE-Ireland and ICE-Canada comprising approximately 600,000 words each. Methodologically, this study is neither purely corpus-based nor corpus-driven but utilises both approaches (see section 4.2.4). For example, the iteration between top-down and bottom-up approaches to the data aligns with the notions of corpus-based versus corpus-driven perspectives (which will all be elaborated and discussed in details below).

4.2 Corpus Linguistics

4.2.1 What is corpus linguistics?

Indeed, it is difficult to provide a clear-cut or precise definition for corpus linguistics (CL) that has been agreed upon by all or most linguists (McEnery & Hardie 2012; Stefanowitsch 2018). This is due to the fact that some schools of corpus linguists believe that CL is only a method of carrying out linguistic analysis, and some disagree with the characterisation of CL as a methodology (Nesselhauf 2005; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Stefanowitsch 2018). Addditionally, this issue of whether CL is a theory or a methodology has been debated from different standpoints (Tognini-Bonelli 2001; O'Keeffe & McCarthy 2012; Stefanowitsch 2018). My research position in relation to these differences will be explained in detail in section 4.2.4. CL is quite different from many other fields and sub-fields in linguistics because it does not focus directly on the study of any particular aspect of language; rather, it is an area which concentrates on a set of procedures and methods by which language research can be conducted (Bennett 2010; McEnery & Hardie 2012). Other scholars of linguistics define CL as "the study of language based on examples of 'real life' language use" (McEnery & Wilson 2001, p.1; Stefanowitsch 2018, p.20). CL is one of the fastest growing sub-fields or methodologies in contemporary linguistics which focuses on the linguistic performance and description rather than the Chomskyan model which only considers linguistic competence and universals. It also makes use of both qualitative and quantitative models of language (McEnery & Hardie 2012).

The main component of CL is corpora (the plural form of corpus) which provides a large amount of textual data used in many fields of linguistics (Hunston 2002; O'Keeffe & McCarthy 2012; Bijeikiene & Tamosiunaite 2013; Baker & McEnery 2015; Stefanowitsch

2018). Corpora are a principled collection of texts, spoken or written, stored on a computer and available for quantitative and qualitative analysis using computational tools (Biber *et al* 1998, p.4). Corpus analyses, and corpora in general, are very useful tools for many reasons: 1) They enable researchers to test hypothesis about language use in a very accurate way; 2) They provide the possibility of raising new questions and theories; 3) They enable researchers to measure the frequencies of any linguistic pattern, form, or words; 4) They help researchers to quantify linguistic patterns; 5) They highlight the most frequent variables in the language (this is usually done with large corpora); and 6) They enable researchers to find "evidence on rare or unusual cases of language" (Baker 2010, p.94; Hunston 2002; Bijeikiene & Tamosiunaite 2013; Baker & McEnery 2015). Accordingly, these affordances of using corpus analysis will be directly beneficial to me as a researcher in answering my primary research questions while keeping in mind the limitations of CL which have been dealt with and overcome in relation to this study (as mentioned in section 4.3).

A corpus has been defined in several ways; however, all these definitions essentially state the same thing using different wordings (Bijeikiene & Tamosiunaite 2013). A corpus can be defined as "any text or collection of texts" (Baker 2010, p.95), and others have defined it as a systematic collection of naturally occurring texts (of both written and spoken language) (Hunston 2002; Nesselhauf 2005; Bijeikiene & Tamosiunaite 2013; Baker & McEnery 2015). In order for the corpus to be defined as a corpus from a linguistics viewpoint, it should follow the characteristics of corpora that have been set by corpus linguists; in other words: the structure and the contents of the corpus should follow certain extra linguistic principles (Nesselhauf 2005; Bijeikiene & Tamosiunaite 2013). For

example, it should have a finite size; it should be able to represent a variety under examination; and, it should be readable (Baker 2010, p.95; Nesselhauf 2005; Bijeikiene & Tamosiunaite 2013). Thus, corpora are easily built through interviews, observations and the like. Having balanced corpora is also very important for reliable specific analysis and results, not least of all those that have a sociolinguistic dimension. For example, different age groups should be represented equally (Bijeikiene & Tamosiunaite 2013). The representativeness of the corpora is linked to the research questions. In some research, just one variety is the focus while, in others, several varieties are under scrutiny (the details of the corpus design issues are covered in section 4.3) (Hunston 2002; Bijeikiene & Tamosiunaite 2013; Baker & McEnery 2015).

4.2.2 Types of corpora and their different purposes

There are many types of corpora that can be used for different purposes and analyses (Tognini-Bonelli 2001; Hunston 2002; O'Keeffe & McCarthy 2012; Stefanowitsch 2018). For instance, we have the **general/reference corpora**; the aim of these types of corpora is to represent and present the language or the variety of the language as a whole (Hunston 2002; Nesselhauf 2005; Wynne & Berglund 2012). In comparison, we have the **specialized corpora** which include specific types of texts; this type of corpora aims to represent the sub-language (specific domains or genres within the whole language) rather than the whole language itself (Hunston 2002; Nesselhauf 2005). Another type of corpus is the **diachronic/historical corpus** which includes texts for the same language or variety gathered from different time periods, and are used mainly to track changes in language evolution (Hunston 2002; Megyesi 2009; Wynne & Berglund 2012). Furthermore, there are **synchronic/comparable corpora** which are not regional, and they represent similar texts

in two, or more languages or language varieties such as the *International Corpus of English (ICE)* (Greenbaum 1988) which includes different varieties of English like British, American, Irish, Canadian, and East African (Hunston 2002; Kallen 2006; Megyesi 2009; O'Keeffe & McCarthy 2012; Wynne & Berglund 2012; Stefanowitsch 2018). However, within the ICE corpus, we have **regional corpora,** which represent one regional variety of a language like *ICE- New Zealand, ICE- Singapore, ICE- Canada, ICE-Ireland*, and many others (as shown in section 4.4) (Hunston 2002; Megyesi 2009; Wynne & Berglund 2012; International Corpus of English 2016; Stefanowitsch 2018). The two regional ICE corpora used in this study are ICE-Ireland and ICE-Canada which have been chosen due to their comparability and the way they have been designed which can effectively and accurately serve the purpose of this study (see section 4.4 for more details). We also have the **learner corpora** which represent the language produced by the learners of any particular language as opposed to the language produced by its native speakers (Hunston 2002; O'Keeffe & McCarthy 2012; Stefanowitsch 2018). In addition to the above, there is also the **multilingual/parallel (translation) corpora** which represent and present the same texts in and from two or more languages for the sake of contrastive and comparative analysis (Megyesi 2009; Wynne & Berglund 2012). There are also **monitor corpora** which keep increasing in size with new material and texts added over time in order to monitor language change (Hunston 2002; Megyesi 2009; O'Keeffe & McCarthy 2012; Wynne & Berglund 2012; Stefanowitsch 2018).

All the types of corpora that have been mentioned above are either **spoken, written**, or **mixed** corpora (Hunston 2002; Megyesi 2009; McCarthy & O'Keeffe 2012; Wynne & Berglund 2012; Stefanowitsch 2018). Some of the examples of the spoken corpora are the

Corpus of London Teenage Language (COLT) (which has transcribed spontaneous spoken language of London teenagers), *London-Lund Corpus of Spoken English (LLC)* (which has transcribed spoken language with various text types), *Lancaster/IBM spoken English Corpus (SEC)* (which has primarily transcribed spoken language of radio broadcasts), *Santa Barbara Corpus of Spoken American English* (which has transcribed spoken interaction from all over the U.S.), and finally the *Michigan Corpus of Academic Spoken English (MICASE)* (which has transcribed the academic speech of native and non-native speakers) (Hunston 2002; Megyesi 2009; O'Keeffe & McCarthy 2012; Wynne & Berglund 2012; Stefanowitsch 2018). It should be noted that all these types of corpora and their different purposes mentioned above are distinctive from each other only in relation to the types of the texts included in the corpora, not in relation to how the texts have been treated and analysed (Tognini-Bonelli 2001; Hunston 2002; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Stefanowitsch 2018). This fact will lead us to the term **corpus annotation** which will demonstrate the distinction between the two types of corpora (annotated verses unannotated corpora) in regards to how corpora are treated and analysed (Tognini-Bonelli 2001; Hunston 2002; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Baker & McEnery 2015; Stefanowitsch 2018).

Corpus annotation has been defined as "the practice of adding interpretative linguistic information to a corpus" (Leech 2004, para 1; McEnery & Hardie 2012). In another definition, corpus annotation is defined as "the linguistic analysis encoded in the corpus data itself" (McEnery & Hardie 2012, p.13; Hunston 2002; O'Keeffe & McCarthy 2012; Baker & McEnery 2015; Stefanowitsch 2018). One of the most common types of annotation is the addition of tags or labels which indicate the word class and which words

in the text it belongs to (Hunston 2002; Leech 2004; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Baker & McEnery 2015; Stefanowitsch 2018). This type of annotation is known as part-of-speech tagging or POS tagging (Hunston 2002; Leech 2004; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Baker & McEnery 2015; Stefanowitsch 2018). This is very useful, especially with words that share the same spellings, but have different meanings or pronunciations; for example, the word *present* which can be a noun and a verb with two different meanings, and the word *record*. Both of these examples have two different pronunciations that indicate the word classes and meanings based on the stress of the word, shaping the way this word should be pronounced (Hunston 2002; Leech 2004; O'Keeffe & McCarthy 2012; Stefanowitsch 2018).

However, some scholars, like the late John Sinclair, preferred not to use or engage in corpus annotation because they believe that unannotated corpora are more authentic and pure to work with and investigate (Leech 2004). In corpus annotation, there are many types and kinds of annotation: 1) **Phonetic/prosodic annotation** which is the practice of adding information about prosodic features like stress and intonation and pauses. This type of annotation is mainly about how the word is pronounced in spoken corpora. 2) **Syntactic annotation**, by far the most common type of annotation, which is the practice of adding information about how any given sentence in a corpus is parsed in terms of the syntactic rules in general, usually referred to as Part-of-Speech (PoS) tagging. 3) **Semantic annotation** which is the practice of adding information about the semantic category of words. For example, the word *cricket* can refer to either a type of sport or an insect, and it belongs to different semantic categories even though there is no difference in spelling or pronunciation. 4) **Pragmatic annotation** which is the practice of adding information about

the kinds of speech act that occur in a spoken dialogue. 5) **Discourse annotation** which is the practice of adding information about anaphoric links in a text; for example, connecting any of the pronouns which occur in the sentence to their referents and antecedents. 6) **Stylistic annotation** which is the practice of adding information about speech and thought presentation such as direct speech or indirect speech for instance. 7) **Lexical annotation** which is the practice of adding the identity of the lemma of each word form in a text; for example, the word *lying* has the lemma *lie* (Hunston 2002; Leech 2004; McEnery & Hardie 2012; O'Keeffe & McCarthy 2012; Baker & McEnery 2015; Stefanowitsch 2018). So, having identified the types of corpora, their purposes, and how they can be treated, the corpora used for this research, the ICE-Ireland and ICE-Canada corpora, are known to be amongst the regional comparable corpora. They have been annotated at different levels (syntactic parsing, word class tagging, and textual markup) and have been determined to be most suitable for this study even though there are some limitations (as discussed in Chapter 10, section 10.3) (Kallen 2006; Schneider 2013; International Corpus of English 2016).

4.2.3 Spoken versus written corpora

Spoken and written corpora both form the core of CL which is, in turn, used to search and analyse these databases (corpora) for real language use (O'Keeffe & McCarthy 2012; Vaughan & O'Keeffe 2015). Spoken and written corpora usually present any principled collection of written or transcribed spoken language that have been stored electronically in computer databases (McCarthy & O'Keeffe 2012; Carter & McCarthy 2017).

On one hand, we have spoken corpora which are obtained and stored by transcribing actual spoken language through collections of transcripts of real speech and speech events

recorded on analogue or digital recording equipment (McCarthy & O'Keeffe 2012; Carter & McCarthy 2017). In other words, their transcripts can be taken from different sources such as ordinary conversations recorded in people's homes, workplaces, or in public (which can sometimes be either by very general demographic sampling or sometimes in more specialized contexts) (McCarthy & O'Keeffe 2012; Carter & McCarthy 2017). Before going in depth about the types of spoken corpora, it is useful to note the distinction made between spoken corpora and speech corpora which both serve different purposes. Speech corpora are "usually collections of speech (which could be anything from transcripts of spontaneous speech to recordings of people reading out loud prepared lists of single words) that are compiled for purposes such as creating automatic voice-to-text applications, telephone technology or the analysis of the phonetic substance of speech" (McCarthy & O'Keeffe 2012, p. 1; Newman 2008; Harrington, 2010). Thus, speech corpora are used for the analysis of how the speech has been produced, i.e. the speech signal itself; whereas, spoken corpora are typically used for the investigation and the analysis of what people say, why they say it, and how they use spoken language to communicate their messages and to interact with one another (McCarthy & O'Keeffe 2012). Before the existence of spoken corpora, tape recorders were common and were used for the same purposes spoken corpora are used for today, but with some limitations (Tognini-Bonelli 2010; McCarthy & O'Keeffe 2012). So, the only way for people's natural spoken interactions to be analysed before was through observation (or eavesdropping) and attempting to write down what was said in the recordings (Timmis 2009; Tognini-Bonelli 2010; McCarthy & O'Keeffe 2012). Spoken corpora can be divided into two types: 1) **Read speech** (non- spontaneous) such as excerpts from books, news broadcasts, word lists, and number sequences (Edwards & Lampert

1992; Leech *et al* 1995; Newman 2008; Baker & McEnery 2015; Love, Dembry, Hardie, Brezina & McEnery 2017). 2) **Spontaneous speech** such as dialogues, meetings, narratives (in which a story is being told) and many others (Edwards & Lampert 1992; Leech *et al* 1995; Newman 2008; Baker & McEnery 2015; Love *et al* 2017).

On the other hand, we have written corpora which are comprised of texts taken from books, newspapers, magazines, or written discourse in general which are then stored by scanning or downloading them electronically (Carter & McCarthy 2017). Written corpora are more plentiful in terms of building them because they are easier, quicker, and cheaper to be created in comparison to the long-drawn-out and costly process of building spoken corpora, which involves making recordings and painstakingly transcribing them in order for them to be in computer-readable forms.

Spoken and written corpora can be very similar or very different, depending on the way both corpora have been built and designed (Hunston 2002, 2013; Newman 2008; Dickinson 2015). The similarities or differences usually can be seen in relation to content selection, representativeness, size, structural criteria, annotation, documentation, balance, topic, and homogeneity (see section 4.3). Yet, both of these types of corpora are very useful for language studies, and they can both be used in unique ways to serve either the same, or different, goals being undertaken in the research under investigation. Also, both types of corpora have the ability to move from what is more formal to what is less formal and vice versa (Hunston 2002; Tognini-Bonelli 2010; McCarthy & O'Keeffe 2012; McCarthy & Carter 2017). However, despite their shared usefulness in present language studies, in relation to the study of grammar, the focus was mainly on written language as the baseline for grammar rules until right before the end of the 19th century. It was only after this time

period that the shift and the attention turned systematically to the spoken language which

has shown the importance of the inclusion of spoken language in the study of grammar.

According to Carter and McCarthy,

> until recently, items and structures most typically found in spoken communication have not
> been fully described. Most grammars of English have had a bias towards the written language. It is
> only recently that advances in audio-recording and associated technology have made it possible for
> sufficient quantities of spoken language to be used for analysis.
>
> (Carter & McCarthy 2006, p.164)

Also, according to O'Keeffe and Mark, "Sweet's influential work at the turn of the

twentieth century asserted the principle that spoken language was not a corrupt form of

written language and that it should be the starting point of any language description" (2018,

p.141). Sweet's work on English grammar has provided a great overview on the core

features of what we now refer to as spoken grammar such as parataxis (the placing of two

phrases and clauses side by side without a connecting word), phrases (which are now

referred to as clusters) and ellipsis (Sweet 1899, 1900; Carter & McCarthy 2017).

Furthermore, the work of Palmer (1924) has "displayed an in-depth understanding of the

grammar of speech and discourse, with insights like the greater use of coordination of

clauses rather than subordination in informal spoken contexts" (O'Keeffe & Mark 2018,

p.139).

However, depending on whether spoken corpora or written corpora are used, there

are two major impacts on grammar. One is that the analysis of corpora (the empirical data)

will shift us from prescription (about what should be said or written) to description (about

what is found) in the data (O'Keeffe & Mark 2018). The second is that the availability and

the advance of spoken corpora can be useful in allowing us to make the distinction between

spoken grammar and written grammar since the gathering momentum is increasingly there

to work on spoken grammar (O'Keeffe & Mark 2018). Therefore, in this present research,

the focus will be only on the spoken components of ICE-Canada and ICE-Ireland (in other words, only the spoken data gathered in the Republic of Ireland will be used in this research, see section 4.4.1 for the rationale behind this decision) since the aim of this research is to comparatively investigate and explore the nature of spoken grammar (and not written grammar) in Irish and Canadian Englishes in relation to their pragmatic markers and pragmatic functions (as explained in the previous chapters and discussed below in section 4.4).

4.2.4 Corpus-based versus corpus-driven linguistics

Before demonstrating the differences and the distinction between corpus-based and corpus-driven language study or linguistics, we should be aware that these two terms were created and introduced by Tognini-Bonelli in order to present two different views and perspectives about CL (Tognini-Bonelli 2001, 2010; McEnery & Gabrielatos 2006; McEnery & Hardie 2012; Hunston 2013; Baker & McEnery 2015). Thus, we know from this fact that there are points of controversy about these two terms and their core beliefs. A corpus-based perspective accepts the characterization of corpus linguistics as a *method* which supports the approach of corpus-based studies that typically use corpus data in linguistics in order to explore a theory or hypothesis (Tognini-Bonelli 2001; McEnery & Gabrielatos 2006; McEnery & Hardie 2012; Hunston 2013; Stefanowitsch 2018). On the other hand, corpus-driven linguistics rejects the characterization of CL as a method and believes instead that the corpus itself should be the single and sole source of our hypotheses about language (Tognini-Bonelli 2001; McEnery & Gabrielatos 2006; McEnery & Hardie 2012; Stefanowitsch 2018). Also, it claims that the corpus itself shapes its own theory of language and has theoretical status (Tognini-Bonelli 2001; McEnery & Gabrielatos 2006; McEnery

& Hardie 2012; Stefanowitsch 2018). As a result, most of the work of CL research is sorted and divided into one or the other category in the view of those who believe in the corpus-based versus corpus-driven dichotomy which has created this basic and binary distinction (McEnery & Hardie 2012). Alternatively, there are some linguistics scholars who reject that distinction and believe that all CL can justly be described as corpus-based (McEnery & Gabrielatos 2006; Tognini-Bonelli 2010; McEnery & Hardie 2012; Baker *et al* 2013; Stefanowitsch 2018).

Regardless of this, it is important to look in depth into corpus-based and corpus-driven analyses of language variation and use because both demonstrate how corpora can be used in different ways in order to validate, exemplify or build up a language theory (Tognini-Bonelli 2001; Hunston 2002, 2013). The primary aim of corpus-based research is to analyse the systematic patterns of variation and use for those pre-defined linguistic features with the belief that the validity of linguistic forms and structures are derived from linguistic theory (Tognini-Bonelli 2001; McEnery & Gabrielatos 2006; Baker *et al* 2013; Gatto 2015). However, corpus-driven research is more inductive in the way that "it aims to derive linguistic categories systematically from the recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli 2001, p.87; McEnery & Gabrielatos 2006; Baker *et al* 2013). This results in the emergence of the linguistic constructs themselves from analysis of a corpus as a whole (Tognini-Bonelli 2001; Gatto 2015). This aim goes along with the holistic approach to language (Tognini-Bonelli 2001).

This research aligns itself with the linguistic scholars who reject the distinction between corpus-based and corpus-driven analyses of language variation. However, with that distinction in mind, it was crucial for me as a researcher to look into both analyses of

language variation and use in depth because both demonstrate different ways corpora can be used in order to validate, exemplify, or build up a language theory. This reflects on the fact that CL is now a progressive methodology which carries out linguistic analysis in a more effective and accurate manner. Therefore, in this present study, CL is used as a methodological tool (as elaborated in section 4.5) within the theoretical framework of VP (as mentioned in Chapter 3) to explore the nature of spoken grammar usage in both Irish and Canadian English varieties comparatively in relation to their pragmatic functions and forms. This is done by employing the iterative approach in which top-down and bottom-up processes are used to compile pragmatic markers and their pragmatic functions in spoken grammar. Top-down analysis was based on a framework for spoken grammar based on existing literature while the bottom-up process was based on micro-analysis of the data. As a result, methodologically, this study is not purely corpus-based nor corpus-driven but involves both, due to the iteration between top-down and bottom-up approaches to the data in this study which aligns with the notions of corpus-based versus corpus-driven linguistics and perspectives as mentioned above (see section 4.5 and 4.7 for more information regarding the application and the rationale).

4.3 Corpus Design Issues

4.3.1 Authenticity

Authenticity is one of the primary requirements for the data used in CL in order to provide accurate evidence for what has been investigated (Tognini-Bonelli 2001; Stefanowitsch 2018). The expectation of data collected for the purpose of linguistic research, especially those that will be input into corpora, is that "all the material included in a corpus, whether spoken, written, or gathered along any intermediate dimension is assumed to be taken from

genuine communication of people going about their normal business" (Tognini-Bonelli 2001, p.55; Biber 1988; Love *et al* 2017). This is because data, or corpora, in general are used as reservoirs of evidence (Tognini-Bonelli 2001; Stefanowitsch 2018). However, if the data is described as "special purpose" or the corpora themselves are specialized corpora, then there are some special restrictions on the choice of texts that can be included in such a collection. The fact that it is not a "general purpose" corpora should be stated clearly in the documentation, and the responsibility for labelling these specialized corpora is on the originator. This is because the default understanding of the purpose of a corpus is seen as "general purpose" or "reference" (Tognini-Bonelli 2001; Hunston 2002; Baker & McEnery 2015). For example, "if it is a corpus of bible translations or lyric poetry or interviews with psychiatric patients this should be made clear, and even more critically if the language has been elicited in an artificial or experimental setting" (Tognini-Bonelli 2001, p.55).

4.3.2 Representativeness

Representativeness is closely related to authenticity in the way that a corpus cannot be considered representative if it is not authentic in the first place (Tognini-Bonelli 2001; Stefanowitsch 2018). According to Leech, "a corpus is representative when the findings based on its contents can be generalized to a larger hypothetical corpus" (Leech 1991, p.27; Tognini-Bonelli 2001). In relation to this, Atkinson and Biber mention that "most of the uses of a corpus in language work rest on the reliability of the corpus as representative of the language grammars, dictionaries, and analytical software [which] all make such assumptions" (1994, p.377; Tognini-Bonelli 2001). So, there is no point in choosing a corpus or data that is not representative. It is generally agreed that a corpus "should be representative of a certain population and that the statements derived from the analysis of

the corpus will be largely applicable to a larger sample or to the language as a whole"
(Tognini-Bonelli 2001, p.57; Stefanowitsch 2018). One of the elements that will affect the
representativeness of a corpus is size, as it can be difficult to measure and ascertain whether
or not the size of a corpus is large enough to fulfil any particular researcher's purpose
(Tognini-Bonelli 2001; Stefanowitsch 2018). Fillmore points out that native speakers
would come to this conclusion if they failed to find something that was intuitively expected
(Fillmore 1992, p.38; Tognini-Bonelli 2001). However, the presence or the absence (of
something in a corpus) is one thing, and the frequency (of it) is another thing. Therefore,
the need for a frequency-based model in order to account for the cumulative effect of data
is required (Tognini-Bonelli 2001).

According to Tognini-Bonelli (2001), CL is considered to be "the natural
correlation of the language model upheld by Firth, Halliday, and Sinclair" (p.57), a model
in which repeated events noticed in language samples at the level of individual performance
are an essential element in the formulation of generalization about language (Gatto 2015).
This notion that individual performance could lead to insights about language has received
a lot of criticism by Chomsky who rejected the idea that any corpus could be representative
of innate language knowledge. However, Chomsky's rejection of CL goes back to the
1950s when corpora were really small entities in a way that their size itself was a guarantee
of their lack of representativeness; now, corpora have become larger so the initial issue of
size is no longer as relevant (Bianchi 2012; Gatto 2015; Stefanowitsch 2018). As a result, it
is reasonable to think that "there could be much more representative corpora than Chomsky
could dream of when he first criticized corpus-based linguistics" (McEnery & Wilson 2001,
p.78; Gatto 2015).

4.3.3 Balance and size

Balance

Balance and size are two important elements in designing corpora because they can heavily impact the representativeness of a corpus, especially general corpora which depend primarily on these two elements (McEnery, Xiao & Tono 2005; Stefanowitsch 2018). What is meant by balance is the range of text categories included in the corpus (McEnery *et al* 2005; Bianchi 2012; Stefanowitsch 2018). As for the size, it is all about the type of questions that are going to be asked in the research, which sometimes do not require larger data as the production of a word frequency list can be produced from a corpus of almost any size (Hunston 2013). However, large data are usually required in order for the research to be more accurate and useful, particularly in descriptive studies which need larger data to cover the whole phenomenon under investigation. Larger databases are generally considered better because they offer more flexibility in terms of filtering out some of the output (McEnery *et al* 2005). The acceptable balance and size of a corpus are determined by its intended uses (McEnery *et al* 2005). A balanced corpus is usually large enough and is supposed to be representative of the language or language variety under investigation (McEnery *et al* 2005). This cannot be achieved unless the corpus covers a wide range of text categories with clear-cut, classified, characterized, typologized texts, and textual variation (McEnery *et al* 2005).

Size and comparability

As Pérez-Paredes (2021) notes, comparability is of key importance in the process of analysing corpus results and indeed, as discussed in Chapter 3 (section 3.2.1), it is also a

core principle of VP. According to Biber (1990) the aspect of a full range of variation is very important to the point that small corpora can be adequate for overall analyses if they can represent the range of variation (O'Keeffe & McCarthy 2012; Stefanowitsch 2018). However, despite the common ICE sample frame, the sizes of both corpora used in this study differ. ICE-Ireland has a lower word count than ICE-Canada because the spoken component of Northern Ireland (NI) has been excluded in this study because the focus has been only on the spoken component of the Republic of Ireland (ROI) (see section 4.4.1 for the rationale and the discussion). To make results from ICE-Ireland and ICE-Canada comparable, there was need to *normalise* them. This meant bringing the results to a common base (because of the relatively small size of the coropora, a base of 10,000 was used. As McEnery and Hardie (2012, p.49) note, normalisation answers the question, "how often might we assume we will see the word per X words of running text?". In the case of this study, this would equate to asking, "how often might we assume we will see the word (or multi-word unit) X per 10,000 words of running text?". Therefore, normalisation was calculated as follows (based on McEnery & Hardie 2012, pp.49-50; Vaughan & O'Keeffe 2015):

A normalized frequency (*nf*) = (number of examples of the word in the whole corpus ÷ size of the corpus) X (base of normalization) (Vaughan & O'Keeffe 2015).

4.3.4 Sampling

For a corpus to be representative and used as the basis for generalisation about the language, we have to define the target population which the corpus is aiming to represent (Tognini-Bonelli 2001; Stefanowitsch 2018). This is the most important issue in corpus design because it provides the rationale for the collection and the classification of the

corpus; such as, which types of texts are being selected, the number of texts, the selection

of particular texts, the selection of text samples from within texts, and the length of text

samples all have to be addressed, and as Biber again points out, all of the aforementioned

are sampling decisions (Atkinson & Biber 1994; Stefanowitsch 2018). Thus, the criteria for

the assembly of a corpus should be explicit and accessible to corpus users in order for them

to be capable of accurately evaluating the corpus using the criteria and relating the

statements derived from the analysis to the typology of the texts included in the corpus

(Tognini-Bonelli 2001). Also, Biber, in his article about representativeness in corpus

design, defines representativeness as "the extent to which a sample includes the full range

of variability in a population" (Atkinson & Biber 1994, p.378; Tognini-Bonelli 2001). He

proposes that variability should be considered both from situational and linguistic

perspectives, and gives examples of genre or register, which are situationally defined text

categories, and text types, which are linguistically defined text categories (Atkinson &

Biber 1994, p.380; Tognini-Bonelli 2001). Biber forms the basis of addressing the issue of

optimal text length on the distribution of linguistic features within texts, particularly on the

number of continuous words required in text samples (Biber 1993; Tognini-Bonelli 2001).

He concluded that if two samples from the same text show similar choices from the feature

list and two samples from different texts show different choices, then it could be

determined that the sample size is large enough. Alternatively, if the results are not as

apparent, then the sample size would need to be enlarged until it did become clear

(Tognini-Bonelli 2001).

   The corpora used in this study contain samples of speech by both males and

females, and include a wide range of age groups. However, the proportions of the

population used in ICE-Corpora as a whole are not representative as women are not equally represented in professions such as politics and law, and thus do not produce equal amounts of discourse in these fields. Similarly, data taken from students or academic authors are not representative of a wide range of age groups (International Corpus of English 2016) (see Chapter 10, section 10.3 for more details about ICE limitations and how they have been dealt with in this research).

4.4 The International Corpus of English (ICE)

The International Corpus of English (ICE) has been in existence for almost thirty years (since 1990) involving twenty-six research teams from different countries across the world and various organizations as well such as WHSPR and New Spirit Services (International Corpus of English 2016). The ICE project is a collection of electronic corpora of national or regional varieties of English comprising 500 texts (300 spoken and 200 written) of approximately 2,000 words each (60% spoken texts and 40% written). Consequently, each ICE corpus contains one million words of spoken and written English produced after 1989. The primary aim of the different corpora of ICE has been to provide (common) material and resources for comparative studies of English worldwide which includes countries in which English is the first language (such as Canada and Australia) or countries in which English is an official additional language (such as India and Nigeria). This involves the same corpus design (i.e. sample frame) and scheme for grammatical annotation to be followed in order to ensure compatibility among the component corpora (Greenbaum 1991; Kallen 2006; Schneider 2013).

Yet, the ICE project is not viewed or defined as "a variationist project or a sociolinguistic enterprise in the usual sense" because it is not concerned with demonstrating the relationship of speaker variables such as age, gender, ethnicity, or social class with linguistic variability, rather is it designed to examine questions of linguistic change in progress (Kallen 2006, p.3). The ICE project is in fact dedicated to the examination of "what is avowedly standard English" (Kallen 2006, p.3). The ICE definition of what is "standard English" is not to be confused with Milroy and Milroy's (1999) definition of "standard ideology" that there is only one correct spoken form of the language, modelled on a single correct written form (Milroy 1999). ICE corpora are not about making linguistic decisions for what is "standard" nor enforcing uniformity in English usage (Kallen 2006). On the other hand, as Nelson (1996) has explained, the definition of what goes into an ICE corpus is based on two non-linguistic quantifiable criteria which are that the speakers chosen were adults of eighteen years of age or older and they should have had a formal education through the medium of English to at least a secondary school level, which somehow gives an empirical definition of "standard" in ICE corpora (Nelson 1996; Kallen 2006). Additionally, the design of the spoken component of ICE contains 60% dialogic and 40% monologic material and these are divided into public and private dialogues and into scripted and unscripted monologues, as demonstrated in Table 4.1 below (Meyer 2002; International Corpus of English 2016).

Table 4.1 *The design of the spoken component of ICE corpora*

SPOKEN (300)	**Dialogues** (180)	**Private** (100)	Face-to-face conversations (90)
			Phone calls (10)
		Public (80)	Classroom lessons (20)
			Broadcast discussions (20)
			Broadcast interviews (10)

Monologues (120)	**Unscripted** (70)	Parliamentary debates (10)
		Legal cross-examinations (10)
		Business transactions (10)
		Spontaneous commentaries (20)
		Unscripted speeches (30)
		Demonstrations (10)
		Legal presentations (10)
	Scripted (50)	Broadcast News (20)
		Broadcast talks (20)
		Non-broadcast talks (10)

At the time of writing, there are 15 ICE corpora available, representing 15 different Englishes from around the world as shown in Table 4.2 (Kallen 2006; International Corpus of English 2016).

Table 4.2: *Other representative English ICE-corpora*

ICE-Great Britain	ICE-Australia	ICE-Philippines
ICE-Ireland	ICE-New Zealand	ICE-East Africa
ICE-Canada	ICE-Hong Kong	ICE-Jamaica
ICE-United States	ICE-India	ICE-Malaysia
ICE-Singapore	ICE-South Africa	ICE-Sri Lanka

While the ICE-corpora have many merits and have proven very valuable to this research there are some limitations to consider in their usage. These have been addressed in Chapter 10, section 10.3. We will now look more in depth at the sub-corpora of ICE relevant to this study in sections 4.4.1 and 4.4.2 below.

4.4.1 International Corpus of English - Ireland (ICE- Ireland)

ICE- Ireland is one of the ICE-corpora following the common ICE principles mentioned

above. It has a collection of 300 transcribed spoken texts in 15 different discourse situations covering different ranges and types of discourse situations (e.g. formal vs. informal) as illustrated in Table 4.4 It contains data from both NI and the ROI and involves 955 speakers. Each spoken and written text category is obtained equally (with the same number) from NI and ROI which makes this corpus unique because it is a cross-border data resource (Kallen 2006; Kallen & Kirk 2008; Schneider 2013). However, although ICE-Ireland sub-corpora have the same corpus design and scheme for grammatical annotation which comprise the same type of texts and present the Irish English language, both sub-corpora present two different regions in which each has different political contexts and different language origins.

This results in several facts:

1) The dialect of English spoken in Northern Ireland (presented in ICE (NI)) has been derived from Lowland Scots and different English forms spoken in the northern region which were taken to Ulster during the plantations of the seventeenth century (Hickey 2012).

2) The data drawn from different governmental, administrative, and economic environments in both sub-corpora reflects the hypothesis that governments affect the development of what is considered a norm or standard language (Kallen & Kirk 2007). This is because in lexical search results for the text categories of administrative prose, learned natural science, parliamentary debates, broadcast news, legal presentations, and face to face conversation (categories which include both informal and formal domains), some expected terminology appears in one of the sub-corpora but not the other (or it appears more often in one over the other). For example, the presentations of the official terminology in both sub-

corpora varies in terms of its linguistic origin which somehow has been triggered by political agendas.

3) The presentation of "Irishness" or Irish words in both sub-corpora is quite different. For example (Kallen 2006):

Example 4.1   *Arra* vat shall ve do full quickly and soon (*The Irishmen's Prayers* 1689).

Example 4.2   'Bud, *musha*, what's the harm o' that', he said to himself (Banim 1838).

Example 4.3   'Oh, *wurrah*, no,' said both, 'don't mention that' (Carleton 1843).

Consequently, the focus of this study will be on investigating a single language variety, in a single zone, sharing the same political and linguistic climate (the spoken component of ICE (ROI). This decision impacts the present study because pragmatic analysis depends heavily on context, which is easily influenced by different factors triggered by what has been mentioned above. Thus, Southern Irish English and Northern Irish English are in two different language zones representing a distinctive kind of cross-linguistic influence (Kallen & Kirk 2007; Hickey 2007a, 2007b; Kallen 2012). As a result, for the purpose of this study, the focus will be only on the data collected from the Republic of Ireland. See Table 4.3 below for the representation of the spoken text categories of ICE-Ireland (ROI) (Kallen & Kirk 2008).

Table 4.3: *Spoken text-categories of ICE-Ireland (ROI)*

Text category	*Number of texts*	Approx word count
Broadcast discussions	10	22,000
Broadcast interviews	5	11,000
Broadcast news	10	21,000
Broadcast talks	10	21,000
Business transactions	5	12,000

Classroom lessons	10	24,000
Demonstrations	5	11,000
Face to face conversation	45	122,000
Legal cross-examinations	5	10,000
Legal presentations	5	10,000
Parliamentary debates	5	11,000
Scripted speeches (not broadcast)	5	10,000
Spontaneous commentaries	10	21,000
Telephone conversation	5	12,000
Unscripted speeches	15	33,000
TOTAL SPOKEN TEXTS	150	351,000

Now, we will take a closer look at the ICE-Canada corpus.

4.4.2 International Corpus of English - Canada (ICE- Canada)

The International Corpus of English-Canada (ICE-Canada) is the Canadian component of ICE, collected under the direction of Nancy Belmore at Concordia University. As previously mentioned, ICE-Canada also follows the same overall design as each national corpus in ICE and thus contains 500 "texts" of approximately 2,000 words each – a total of approximately one million words. "Texts" can be based on written sources (40%) equal to 200 texts or transcribed spoken language (60%) equal to 300 texts. The spoken text categories and distributions have been demonstrated in Table 4.1 and 4.4 (UAL Dataverse 2015; International Corpus of English 2016).

Table 4.4: *Spoken text categories of ICE-Ireland (ROI and NI) and ICE-Canada*

Text category	*Number of texts*	Approx word count
Broadcast discussions	20	40,000
Broadcast interviews	10	20,000
Broadcast news	20	40,000
Broadcast talks	20	40,000
Business transactions	10	20,000

Classroom lessons	20	40,000	
Demonstrations	10	20,000	
Face to face conversation	90	180,000	
Legal cross-examinations	10	20,000	
Legal presentations	10	20,000	
Parliamentary debates	10	20,000	
Scripted speeches (not broadcast)	10	20,000	
Spontaneous commentaries	20	40,000	
Telephone conversation	10	20,000	
Unscripted speeches	30	60,000	
TOTAL SPOKEN TEXTS	300	600,000	

Moreover, both of these corpora ICE-Ireland (ROI) and ICE-Canada, bear some social information in regards to the speakers who participated in designing both of these corpora (see Table 4.5 below). This information on social factors can be very useful for conducting an empirical study which goes beyond the regional variation to the gender and age variation in Irish English and Canadian English comparatively.

Table 4.5: *Speaker biodata of spoken ICE-Ireland (ROI) speakers vs ICE-Canada speakers in relation to gender and age*

Gender	Spoken ICE-Ireland (ROI speakers)	Spoken ICE-Canada Speakers		
Female	270	466		
Male	313	752		
Unknown: no answer given	7	None		
Total	590	1218		
Age	Spoken ICE-Ireland (ROI speakers)	**Age**	Spoken ICE-Canada Speakers	
0-18	4	**10-18**	7	
19-25	156	**19-24**	43	
26-33	24	**25-30**	116	
34-41	43	**31-40**	142	

42-49	60	41-50	118
50+	94	51-60	86
Unknown: no answer given	198	61+	35
		Total	547
Total	579		

We will now look at the approaches and CL tools used in this study.

4.5 The Approaches and Corpus Linguistic Tools Used in the Study

4.5.1 The iterative approach

As mentioned earlier in section 3.5, the iterative approach is used in this research by employing top-down and bottom-up processes by which the forms and the pragmatic functions of spoken grammar in Irish and Canadian varieties of English have been compiled and analysed. The sections below will show how the iterative approach has been employed in the study and how each approach cooperatively serves the other, thus confirming the importance of using both in order to avoid any shortcomings (as discussed in section 4.7).

4.5.1.1 The top-down approach

As a starting point, the top-down approach was applied initially in order to compile the forms and the pragmatic functions of spoken grammar based on the existing literature. This top-down approach is a commonly used approach in pragmatics for function-to-form corpus analysis in which the pragmatic functions are reviewed based on existing research findings as the "seeds" or starting points (Amador-Morena *et al* 2015; O'Keeffe 2018). Appendix A presents the top-down framework for the analysis of the features and discourse

functions of spoken grammar in terms of their types, functions, forms (as well as terminology which will be used in this study). This inventory will guide the top-down analysis.

Across the literature, the following features of spoken grammar are deemed the most salient (see also Chapter 2 in section 2.3):

- **Pragmatic markers** which include discourse markers, response tokens, stance markers, hedges, and interjections.

- **Deixis** which includes personal deixis, temporal deixis, and spatial deixis.

- **Questions and tags** which include question tags, fixed tags, echo questions, follow-up questions, two-step questions (and responses), pre-question, and preface question.

- **Ellipsis** which includes situational ellipsis, textual ellipsis, and structural ellipsis.

- **Headers** commonly refer to the subject, object, object complement or prepositional complement which can be placed within the S-V-X clause structure. In other words, what conventionally comes in writing at, or towards the end of the clause (e.g. object, complement, adverbial, question-word clause) will be placed in front of it, and this is called "fronting" (Carter & McCarthy 2006). A noun phrase followed by one or more pronouns which refer to the noun phrase is the most typical type of header (Carter & McCarthy 2006).

- **Tails** are similar items (usually noun phrases) to headers in which they are placed outside the s-v-x clause structure occurring after the clause. They can be noun phrases, prepositional phrases or clauses (Carter & McCarthy 2006)

- **Social routines** are items that express greetings and leave-taking routines (Carter & McCarthy 2006).

- **Swearing and taboo expressions** are taboo naming expressions and taboo intensifiers which often take the form of interjections with the function of expressing "a variety of strong feeling, in particular, annoyance, frustration and anger" (Carter & McCarthy 2006, p.226).

- **Vocatives** are names (including abbreviated names) and titles, terms of kinship and endearment, general plural vocatives, impersonal vocatives, honorifics.

Appendix A elaborates upon this inventory, including their pragmatic functions and forms in spoken grammar.

Therefore, what has been done in relation to the top-down phase of analysis is a general review of the existing literature and a specific review of the existing literature on spoken grammar which have resulted in this inventory of forms. Some of these are general categorizations and some are more specific to Irish or Canadian Englishes and have been used as starting points (for bottom-up investigation as explained earlier).

For example, some discourse markers in Irish English are more salient and frequent than others such as *like*, *you know*, and *I mean* (Kallen 2006). On the other hand, Canadian English shows that *like*, *just*, and *so* are all salient forms of Toronto Youth English, concentrated amongst 15 to 16 year olds generally and found in use with female speakers in particular. However, "only *like* declines by age 20, while *just* and *so* prevail" (Tagliamonte 2005, p.1911). Also, based on the existing literature, it has been found that Irish English speakers use a more limited range of response tokens at the single and two-word level in comparison to British English speakers which raises the question: Is this the case in

comparison with Canadian English? Also, Irish English shows less formality in terms of response tokens as opposed to British English which raises the question: Is this the case in comparison with Canadian English? Irish English also demonstrates a wider range of religious references and swear words as well as reduplication which raises the question: How similar or different is Canadian English? (O'Keeffe & Adolphs 2008). On the other hand, response tokens such as *no problem*, *welcome* and *pleasure* in Canadian English have been expressed and used in the same way as speakers of Irish English, American English, and British English would use them (Vaughan & Clancy 2011; Farenkia 2012). Furthermore, we have the occurrences of the hedges *I would say* and *I'd say* which have been found to be used twice as frequently by Irish speakers as by their American counterparts in the data which led to further investigation on this finding in regards to Irish English. As a result, the broad pragmatic functions of these hedges are linked to a tendency by Irish English speakers (who may feel an obligation to hedge in situations where British or American speakers do not) to avoid directness; this goes back to the Irish socio-cultural context as "in Irish society... 'forwardness,' which ranges from being direct to being self-promoting, is not valued" (Farr & O'Keeffe 2002; Vaughan & Clancy 2011, p.49). However, all these "seeds" are not enough to provide a full description of the spoken grammar of both varieties of English (Irish and Canadian) because they are still seeds which only provide a glance of what may be expected and worth investigating in the corpus. For example, if this research solely searched for these items in both datasets, it would likely miss many other salient forms and features that have not been cited in the literature. Therefore, the micro-bottom-up analysis is also needed so as to reveal any forms and features not already identified in the literature and to underscore what has already been

identified in at least some of the literature but not in the comparative context of Irish and Canadian English. The combined approach will bring a quantitative and qualitative analysis of the whole spoken grammar of these two varieties.

4.5.1.2 The bottom-up approach

After the application of the top-down approach, as detailed above, the bottom-up process is based on a micro-analysis of the data (both quantitative and qualitative). According to Ädel and Reppen (2008) and O'Keeffe (2018), the bottom-up approach is known to be the purest corpus approach in which frequency-based listing of forms (in corpora) are conducted and observed in terms of their forms and meanings. Note that the quantitative presentation of the data has been unmodified and remains original. For example, the item 's' appearing in the wordlist can be the contraction of the word *is* or *has*; however, the item 'm' is definitely the contraction of the word *am*. Thus, for the accuracy of the quantitative presentation, some items have been presented as they are in the original generated form, especially after knowing that these items have no impact on the quantitative and qualitative analysis whatsoever due to the way the data has been analysed (as demonstrated below in the following sections).

The word frequency list

The bottom-up approach was applied in this study using the corpus tools provided by AntConc software. The function of the word frequency tool is to count all the words in the corpus and present them in an ordered list, which allows the researcher to quickly find which words are the most frequent in a corpus. This helps to identify the items occurring

within the top 100 most frequent word-forms which will be compared to the forms

conducted by the top-down analysis for inclusion and exclusion purposes (as shown below

in Table 4.6). Also, for comparative analysis of two or more varieties, the frequency lists

are always a good starting point within the bottom-up approach. Therefore, the word

frequency tool has been used in order to generate the wordlists of the top 100 most frequent

word-forms for both datasets in ICE-Ireland and ICE-Canada (see Table 4.6) below

illustrating a sample of the top 20, and see Appendices B and C for the full list of the top

100 most frequent word-forms):

Table 4.6: *The wordlist of the top 20 most frequent words of Spoken ICE-Ireland and ICE-Canada*

#	ICE-Ireland	Freq	ICE-Canada	Freq
1	the	15764	the	27644
2	and	8954	and	18871
3	I	7659	I	16760
4	to	7569	to	16427
5	of	7388	you	14471
6	a	6119	that	14125
7	you	6036	it	13915
8	it	6020	a	13821
9	that	5995	it's	13230
10	in	5777	of	13051
11	it's	5134	in	9742
12	was	3245	uh	7058
13	is	2973	is	6604
14	uh	2735	they	5837
15	yeah	2625	't (not)	5661
16	't (not)	2615	was	5141
17	we	2471	know	4874
18	he	2404	we	4834
19	on	2286	so	4706
20	they	2253	this	4604

Having generated the wordlist of the top 100 most frequent words, a matching

process was carried out based on the top-bottom analysis of the existing literature in which

the top 100 most frequent words were linked to their possible and potential forms and pragmatic functions that may be found in the data under investigation. So, after that the concordance process was undertaken (as discussed further below), where each word in the wordlist was examined so as to ascertain whether it functioned as a part of the spoken grammar or not. From looking at concordances of these single word items in isolation, some single word items were identified as functioning as part of spoken grammar (e.g. *because* in Canadian English (see Chapter 6, Table 6.1)). Overall, however, the word list analysis led the researcher to realize that most high frequency single-word items that functioned as part of spoken grammar were components of clusters. Therefore, it is very important to not look only at single word items but also to consider the most frequent clusters in the data in order not to miss any possible form existing at either level (as we now detail).

The clusters tool/the N-grams tool

The clusters tool is extremely important to the present study as it scans the entire corpus for 'N' length clusters (i.e. one word, two-word, three-word clusters). This allows me as a researcher to find the common clusters occurring in both datasets. Thus, in order to concordance each of the words presented in the wordlists (illustrated fully in Appendices B and C) and eliminate any that were not used in spoken grammar, clusters analysis is essential. As discussed, many frequent single word items are components of longer clusters, for example, *you know*, *by the way*, *at the same time* and so forth (see analysis chapters and Chapter 2, section 2.2.3.2).

The clusters tool/the N-grams tool has therefore been used extensively in order to gain a full picture of the clusters found in both datasets and to avoid any ambiguity while concordancing. The cluster lists of the top 100 most frequent clusters of Spoken ICE-Ireland and ICE-Canada have been generated based on two, three, four, five, and six-word clusters. Some samples extracted from both corpora are illustrated in Tables 4.5 and 4.6 below (for the full clusters lists of ICE-Ireland and ICE-Canada, see Appendices D, E, F, G, H, I, J, K, L, and M).

Table 4.7: *The word clusters list of the top 10 most frequent clusters of Spoken ICE-Ireland*

#	Two-word clusters	Three-word clusters	Four-word clusters	Five-word clusters	Six-word clusters
1	of the	I don't	I don't know	I don't know I	at the end of the day
2	it's	Don't know	I don't think	you know what I mean	in the attorney general's office
3	in the	It's a	I'm going to	I don't know what	as president of the high court
4	you know	It's not	I think it's	president of the high court	It's going to be a
5	that's	and it's	We're going to	yeah yeah yeah yeah yeah	na na na na na na
6	don't	one of the	's going to be	at the end of the	yeah yeah yeah yeah yeah yeah
7	to the	yeah yeah yeah	yeah yeah yeah yeah	It's going to be	I don't know I think
8	I'm	There's a	Don't know I	the attorney general's office	and all that kind of stuff
9	and I	I think it	at the end of	I don't know how	I don't know it's
10	on the	going to be	Don't know what	I don't know if	president of the high court and

Table 4.8: *The word clusters list of the top 10 most frequent clusters of Spoken ICE-Canada*

#	Two-word clusters	Three-word clusters	Four-word clusters	Five-word clusters	Six-word clusters
1	it's	I don't	I don't know	I'm gonna going to	mm hmm mm hmm mm hmm
2	you know	gonna going to	're gonna going to	you're gonna going to	You're gonna going to have
3	that's	Don't know	mm hmm mm hmm	I don't know if	It's gonna going to be
4	of the	a lot of	It's it's	I don't know I	this week's ecofacts and trends
5	in the	It's a	I don't think	I dunno don't know	you don't wanna want to
6	don't	It's not	two or three words	It's gonna going to	're gonna going to have to
7	I'm	going to be	's gonna going to	We're gonna going to	I don't know if I
8	going to	you don't	gonna going to be	's gonna going to be	That's it for this week
9	mm hmm	wanna want to	I'm gonna going	don't wanna want to	's it for this week's
10	You're	and it's	'm gonna going to	I don't know what	I don't wanna want to

Let us now look at a key process in the analysis of single words and clusters, namely concordancing.

The concordance tool and selection procedures

According to Sinclair, "a concordance is an index to the places in a text where particular words and phrases occur" (2003, p.173). The concordance lines are generated by software programmes or tools (like AntConc which is the one used in this study) which show search results in a "KWIC" (Key Word In Context) format and allow the researcher to see how words and phrases are commonly used in a corpus of texts. After having conducted the top-down analysis of the forms and pragmatic functions of spoken grammar by the top-down approach and having conducted the wordlists and cluster lists by the bottom-up approach,

the concordance tool has been employed in order to determine whether or not the above items mentioned in the wordlists and cluster lists are actual forms of spoken grammar. In other words, concordancing these items helps to manually eliminate the items that were not used in spoken grammar, or as forms of spoken grammar, by determining the pragmatic meanings or functions of only those items which are a part of spoken grammar. This has resulted in the candidates lists for both corpora as shown below (in section 4.6). However, upon first glance, concordance lines can be difficult to deal with in terms of interpretation because they are generated in the order in which they occur in the corpus. Sinclair (2003) recommends a seven-step procedure for "uncovering the mysteries of most concordances" (2003, p.xvi). These steps are:

1. Initiate: Look at the words that occur directly to the left and right of the node. Sinclair recommends working with no more than a single screen of concordances at any one time. Note any that are repeated. Employ the strongest pattern you find as a starting point.

2. Interpret: Look at the repeated word and formulate a hypothesis that may link them or most of them. For example, they may all have similar meanings.

3. Consolidate: Look for other evidence from adjoining words, for example, to support your hypothesis. Be prepared to "loosen" your hypothesis based on this.

4. Report: When you have exhausted the patterns you can observe and formulate an explicit, testable hypotheses.

5. Recycle: Following from the initial step, employ the next strongest pattern in the vicinity of the node and repeat steps two – four. Continue until all repeated patterns have been exhausted.

6. Result: Make a final list of hypotheses based on the node.

7. Repeat: Gather a new selection of concordances of your node word from the corpus. Repeat the steps and confirm, extend, or revise your hypotheses as you progress. (Sinclair 2003).

Note that the concordance process, in terms of how many samples were taken for each candidate item, was applied to the clustered forms (two-word clusters to six-word clusters) differently than the single forms. For the clustered forms, all the samples (occurrences/concordance lines) of the clustered forms were analysed due to the lack of existing literature on them which indicates the importance of looking at all of them. Also, the occurrences of the clustered forms are not many when compared to the single forms. In terms of single forms, if any single form occurred less than 1000 times, then all of their occurrences were analysed. However, only the first 1000 occurrences were analysed for any single form which occurred with more than 1000 occurrences. These samples were taken and analysed equally across both corpora.

4.6 Selection of Candidate Items

As established above, some of the candidates forms (for the investigation of spoken grammar) occur either alone or in a cluster of other words. The single word itself can also be a cluster due to reduplication, e.g. such as *yeah yeah yeah* which has the same function of *yeah* but with more emphasis (also *okay okay* which has the same function of *okay* but with more emphasis in ICE-Canada). We also note that some clusters appear in different word clusters but do not become complete until, but never beyond the six-word cluster. For example:

at the

at the end

at the end of

at the end of the

at the end of the day

This coincides with O'Keeffe *et al*'s (2007) observation that clusters typically run to six word strings at a maximum. These "subsumed clusters" will be noted within the analysis. These clusters are presented in candidate lists in this research. Candidate lists are the lists of the pragmatic items/forms which have been found as a part of spoken grammar. In other words, items that have pragmatic functions in the discourse. They are extracted based on the top 100 most frequent items from both varieties (ICE-Ireland and ICE-Canada) after undergoing quantitative and qualitative analyses. For the purposes of the candidate lists which are presented fully in Appendix N and Appendix O, clusters are logged at the point where they are complete (e.g. in Table 4.9, *at the end of the day* is listed under "six-word cluster"). Tables 4.9 and 4.10 below show only a sample of the candidate items extracted from both datasets (for the full amount see Appendix N and O):

Table 4.9: *The candidates list of Spoken ICE-Ireland*

One word	Two-word cluster	Three-word cluster	Four-word cluster	Five-word cluster	Six-word cluster
okay	You know	You know what	So I don't know	You know what I mean	at the end of the day
so	That's right	I don't know	That's what I'm saying	It seems to me that	you know what I mean yeah
oh	I mean	And so on	At the same time	to be honest with you	yeah yeah yeah yeah yeah yeah
well	Oh yeah	I'd like to	You know the way	I'm just going to	do you know what I mean
now	Kind of	At the moment	I don't think so	All that kind of stuff	and so on and so forth

Table 4.10: *The candidates list of Spoken ICE-Canada*

One word	Two-word cluster	Three-word cluster	Four-word cluster	Five-word cluster	Six-word cluster
really	I know	No I don't	Well I don't know	you know what I mean	and so on and so forth
yeah	It's like	You know what	I don't think so	No I don't think so	you know what I mean like
so	Oh ya	I'm not sure	uhm I don't know	But at the same time	hmm mm hmm mm hmm mm
like	I think	uh you know	ya I don't know	I don't know I think	
well	I mean	Is that so?	At the same time	you know I don't know	

One other note in the process is that a candidate identified in one of the sub-corpora which does not appear in the top 100 of the other dataset, will still be examined in the other dataset for comparison purposes. This adds to the richness of the results.

4.7 Applying an Iterative Approach

This iteration between top-down and bottom-up approaches to the data in this study aligns with the notions of corpus based versus corpus driven linguistics (as mentioned in section 4.2.4) which have both been combined in the new sub field of Corpus Pragmatics (CP) (as discussed in Chapter 2, in section 2.6). CP makes the best use of CL for pragmatics research so as to provide a more accurate analysis of data through the integration of both fields in terms of methodological approaches, as has been increasingly recognised (Jucker 2013; Amador-Moreno *et al* 2015; Rühlemann & Aijmer 2015; O'Keeffe *et al* 2020). To briefly summarize what has been elaborated upon in Chapter 3, section 3.4, CL moves from frequencies of forms to their functions whereas pragmatics moves in the opposite direction, from specific pragmatic functions to the forms that are typically used for that pragmatic function (both utilizing the inductive process) (Amador-Moreno *et al* 2015; Rühlemann &

Aijmer 2015; O'Keeffe 2018; O'Keeffe *et al* 2020). Using either the form-to-function

approach (CL) or the function-to-form approach (pragmatics) on its own opens the door for

criticism as each approach has its own strengths and weaknesses. CP, according to

O'Keeffe (2018), employs both approaches (see also O'Keeffe *et al* 2020; refer to Chapter

3, section 3.4 for a detailed explanation on CP). This mixture of methodologies allows

linguistic items to be examined both structurally (syntactic) and interpersonally

(pragmatic), enabling us to understand how words, utterances, and text work together to co-

construct meaning (Farenkia 2012; Walsh 2013; O'Keeffe 2018).

In this study, in summary, the data was approached first from a top down, research-

informed manner to arrive at a framework for forms of spoken grammar. Then the data was

scrutinised from a bottom up route (through qualitative analysis of frequency lists). This is

described as an *iterative approach* because the bottom up searches are guided by the top

down framework while the top down framework is also informed by data. For example,

some of the features of spoken grammar that appear as items in the top down framework

(such as ellipsis, tails, and headers) did not appear on the top 100 items and so the

framework was re-scaled and narrowed to the most frequent items. Within this narrowing,

of course we recognise that features such as ellipsis, tails, and headers are less easy to

identify because they are either omitted (in the case of ellipsis or are at turn level in the case

of tails and headers and not so much aligned with single or clustered forms). In this

analysis, the following spoken grammar features will be the main focus as they have been

found to be the most salient as a result of the iterative analysis: discourse markers, response

tokens, questions and tags, hedges, and stance markers.

4.8 Outline of Analysis Chapters

The analysis chapters have been organised based on the investigation of the similarities and the differences between Irish and Canadian Englishes in their spoken grammar in relation to pragmatic functions and forms in order to answer the research questions clearly. As detailed in Chapter 1, the core research questions of this study are as follows:

MAIN RESEARCH QUESTION:

> How different and similar are Irish English and Canadian English in their spoken grammar in terms of form and function?

SUB-QUESTIONS

1. How different and similar are Irish English and Canadian English in their Discourse Markers in terms of form and function?

2. How different and similar are Irish English and Canadian English in their Response Tokens in terms of form and function?

3. How different and similar are Irish English and Canadian English in their Questions and Tags in terms of form and function?

4. How different and similar are Irish English and Canadian English in their Hedges in terms of form and function?

5. How different and similar are Irish English and Canadian English in their Stance Markers in terms of form and function?

These questions were refined using the iterative approach detailed above and elsewhere and they will each be examined in the analysis chapters that follow:

Chapter Five: discourse markers found in both datasets are presented and analysed comparatively (with their quantitative and qualitative analyses) in order to answer the question as to how similar and different both varieties are in relation to discourse markers in terms of forms and functions.

Chapter Six: examines response tokens in both datasets comparatively. This will answer the question of how similar and different both varieties are in relation to response tokens in terms of forms and functions.

Chapter Seven: is about the Questions and Tags of both varieties so as to answer the question of how similar and different both varieties are in relation to Questions and Tags in terms of forms and functions.

Chapter Eight: focuses on hedges of both varieties and will answer the question of how similar and different both varieties are in relation to hedges in terms of forms and functions.

Chapter Nine: is about the stance markers of both varieties. It will answer the question of how similar and different both varieties are in relation to stance markers in terms of forms and functions.

4.9 Conclusion

In this chapter, we have covered the main concepts and aspects of the methodology and the approaches used in this current study. It has also detailed the iterative approach that is being used in the analysis phase of this study. We will now move to the first chapter of analysis which comparatively studies and investigates the discourse markers in Irish and Canadian discourse, see Chapter 5.

CHAPTER 5

DISCOURSE MARKERS IN IRISH AND CANADIAN ENGLISHES

5.1 Introduction

In this chapter, I will present the analysis of the discourse markers (DMs) of Irish and

Canadian Englishes captured in the ICE-Ireland and ICE-Canada corpora. This is done by

detailing the main quantitative findings which resulted from both datasets through the

theoretical and methodological framework taken in this study (as explained in Chapters 3

and 4) in order to unpack and analyse these findings qualitatively and comparatively. Thus,

the analysis of the pragmatic variation that occurs in the use of DMs between Irish English

and Canadian English will be investigated from two perspectives: form and function. DMs

are one of the fundamental linguistic elements in communication and interaction that have

various functions or pragmatic variations in the discourse (i.e. one single form like *well* or

uh can generate multiple pragmatic functions as will be shown in the analysis sections

below).

5.2 Previous Research

5.2.1 Different perspectives on discourse markers

Discourse markers have been studied for more than 30 years, beginning with the work of

Schiffrin (1987), from different theoretical frameworks and approaches in numerous fields

such as Discourse Analysis, Conversation Analysis, Pragmatics, Semantics, Syntax, and

Computational Linguistics, as well as many others. Additionally, the diversity of functions

that DMs can perform in the discourse has led to a major disagreement in relation to the

terms used by scholars and researchers to refer to them (Schiffrin 1987; Fraser 1999; Dér

2010; Beeching & Detges 2014; Das & Taboada 2018; Villegas 2019). As a result, we find

that the term "discourse markers" is used by Schiffrin (1987) (and this is the term adopted

in this study as mentioned in Chapter 4 and Appendix A); the term "pragmatic markers" is

used by Brinton (1996) and Fraser (1996); "discourse connectives" is used by Blakemore (1989); "discourse operators" is used by Redeker (1991); "cue phrases" is used by Knott (1993); "discourse particles" is used by Abraham (1991), Kroon (1995) and Schourup (1985); "pragmatic particles" is used by Ostman (1983); and the term "pragmatic expressions" is used by Erman (1987). In addition, there are many other terms that are less commonly used such as: discourse signaling devices, indicating devices, phatic connectives, pragmatic connectives, pragmatic operators, pragmatic formatives, semantic conjuncts and sentence connectives (Alami 2015). This distinction and dispute has occurred not only in the terms referring to DMs but also, in the definitions used and given by the aforementioned researchers and linguists in regards to their respective DMs. Therefore, Schourup (1999) and many other researchers such as Urgelles-Coll (2010) have remarked that the notion of discourse marker is problematic due to the fact that there is no agreed upon terminology or definition. This has caused the excess of terms and definitions in this area (all of which cannot be said to be created on whims or a desire to stake a name in the field) which has led to controversies within the field. Consequently, in general,

> the term and definition used in each framework are chosen to reflect theoretical preoccupations, to avoid unwanted associations, or to rule in or out particular linguistic items or functions. Such variation is to be expected in an area that has only recently become a focus of intensive study and which bears on many different areas of discourse research, cognitive, social, textual, and linguistic. On the other hand, so long as such uncertainties exist, DM must remain a term with theoretical aspirations, but whose precise reference remains at issue."
>
> (Villegas 2019, p 2)

 Therefore, it is extremely crucial to assign the definition of DMs by which this analysis has been approached and this is discussed in the section below.

5.2.2 Definition and function of discourse markers

Before talking about the definition embraced in this present study, it is important to see how DMs have been defined by other scholars. One of the prominent studies on DMs is the work of Schiffrin (1987) in which she defines discourse markers as "sequentially dependent elements which bracket units of talk" (p.31). Based on this definition, she claims that DMs have properties which form structure, convey meaning, and accomplish actions. These properties are interdependent which means they cannot be understood solely without the others. On the other hand, we have the definition proposed by Fraser (1999) who provides a relatively comprehensive definition of discourse markers:

> A class of lexical expressions that signal a relationship between the interpretation of the segment they introduce, S2, and the prior segment, S1. They have a core meaning which is procedural, not conceptual and their more specific interpretation is 'negotiated' by the context both linguistic and conceptual.
>
> (Fraser 1999, p.831)

Based on this definition, DMs must have a core meaning while their specific meaning is negotiated by the context, and they signal a relationship between the interpretation of the segment they introduce, S2, and the prior segment, S1 (Alami 2015). However, Fraser (1999) further argues that "DMs do not have to signal any relationship between S1 and S2. A DM can relate the segment it introduces with any other previous segment in discourse" which he calls "global coherence" as contrasted to Schiffrin's "local coherence" (p.938; Schiffrin 1987; Alami 2015). We also have the definition of Brinton (2008) who proposed that DMs are "phonologically short items that have no or little referential meaning but serve pragmatic or procedural purpose" (p.1). This means DMs act mainly at the pragmatic level of talk and have little or no propositional contribution to the meaning of the discourse.

Thus having presented some (but not all) of the definitions of DMs from different scholars coming from different perspectives, it has been shown that these definitions have

entailed either minor or major differences in the way DMs have been interpreted and classified, in terms of form and function especially; this will have an impact on how the analysis is conducted and examined (for more information see Schiffrin 1987; Fraser 1990a, 1996, 1999; Brinton 1996). In this study (as explained in Chapter 4 and shown in Appendix A), the analysis framework of DMs is based on the definition adapted and introduced by Carter and McCarthy (2006) where they defined DMs as "words and phrases which function to link segments of the discourse to one another in ways which reflect choices of monitoring, organisation, and management exercised by the speaker" (p.208). This definition involves the considerations of the following pragmatic functions:

1. Opening up and closing down the discourse. This is done by managing the discourse "in terms of launching and concluding topics, opening, concluding or temporarily closing a whole conversation, re-opening previously closed or interrupted conversations." (Carter & McCarthy 2006, p.214).

2. Sequencing which "indicates explicitly the order in which things occur or how different segments of a discourse are being organised" (Carter & McCarthy 2006, p.216).

3. Marking boundaries and linking segments of the topic. This is done by "indicating the beginning or end of a topic or a transition from one topic or bit of business to another." (Carter & McCarthy 2006, p.218).

4. "Focus[ing] the attention of the listener on what the speaker feels is "important" (Carter & McCarthy 2006, p.218). This is done by diverting, shifting, or resuming the topic of the discourse.

5. Monitoring and managing the ongoing discourse through reformulations and alternative expressions indicating that "the speaker has not selected the most appropriate way of expressing things and is adding to or refining what they say with a more apt word or phrase" (Carter & McCarthy 2006, p.220).

6. Monitoring and managing the ongoing discourse through "monitoring the state of shared knowledge in the conversation" (Carter & McCarthy 2006, p.221).

Note that DMs, overlap in terms of function and form due to the many classifications and terminology used (as mentioned above) to refer to their functions and forms. For example, Fraser (1996) in his earlier work categorised DMs as a subclass of "commentary pragmatic markers," but later in 1996 he considered DMs as a separate class of pragmatic markers (1996, 1999, p. 938; Alami 2015). Nevertheless, the definition used in this study to analyse DMs goes along with Brinton (1996) who classified the functions of DMs as interpersonal functions and textual functions. Carter and McCarthy's (2006) classification of functions of managing, monitoring, and organising the discourse in terms of the six functions categories outlined above, can be aligned with Brinton's textual function classification. At the textual level, DMs are used to mark various kinds of boundaries (by initiating and ending the discourse or by diverting, shifting, and resuming the topic of the discourse) and to assist in turn taking in oral discourse, or chunking in written discourse (Brinton 1996; Alami 2015; Villegas 2019). At the interpersonal level, proposed by Brinton, DMs are used subjectively to express attitude and interactively to achieve intimacy between speaker and addressee. Brinton's interpersonal function classification, therefore, is more useful in relation to items such as hedges, response tokens, and stance markers. For further illustration see Table 5.1 (adapted from Brinton 1996, pp.35-40) below.

Table 5.1: *Pragmatic functions of discourse markers*

	To initiate discourse, including claiming the attention of the hearer	Opening frame marker
Textual functions	To close discourse	Closing frame marker
	To aid the speaker in acquiring or relinquishing the floor.	Turn takers (Turn givers)
	To serve as a filler or delaying tactic used to sustain discourse or hold the floor	Fillers Turn keepers
	To indicate a new topic or a partial shift in topic	Topic switchers
	To denote either new or old information	Information indicators
	To mark sequential dependence	Sequence/relevance markers
	To repair one's own or others' discourse.	Repair markers
Interpersonal functions	Subjectively, to express a response or a reaction to the preceding discourse including also back-channel signals of understanding and continued attention while another speaker is having his/her turn.	Response/reaction markers Back-channel signals
	Interpersonally, to effect cooperation or sharing, including confirming shared assumptions, checking or expressing understanding, requesting confirmation, expressing difference or saving face (politeness).	Confirmation-seekers Face-savers

Now after looking at the functions of DMs, we will look at their general characteristics.

5.2.3 General characteristics of discourse markers

As mentioned above, DMs have been defined and viewed through many frameworks; on

the other hand, there are many basic characteristics and features which appear to be under

general agreement across most linguistic literature. These generally agreed upon

commonalities are as follows:

- DMs can cover a wide range of items from a variety of word (or grammatical)

  classes such as adverbs (*frankly*, *well*), lexical phrases (*you know*, *I mean*),

  conjunctions (*but*, *since*, *and*), filler words (*uh*, *oh*, *uhm*), and adjectives (*right*,

okay) (Schiffrin 1987; Fraser 1990a; Brinton 1996; Yilmaz 2004; Alami 2015; Villegas 2019).

- DMs can appear in almost all languages in the world (Lenk 1998; Yilmaz 2004).

- DMs are syntactically independent (Schiffrin 1987).

- DMs are syntactically flexible due to their appearance in the discourse (or the utterance) (i.e. they can appear at the beginning, in the middle, or at the end). This flexibility has resulted in a high frequency of DMs in the discourse and enormous usefulness in terms of function (Fujita 2001).

- DMs do not affect the propositional meaning of utterance; in other words, they do not change the truth value of utterance (Schiffrin 1987; Brinton 1996).

- DMs mainly function and deal with the pragmatic aspects of the discourse (Fraser 1990a; Andersen 2001; Yilmaz 2004).

- DMs are meaningful but non-truth conditional, and they make no contribution to the informational content of the discourse (Lam 2008)

- DMs are multifunctional to the point that one single form or item can have multiple pragmatic functions in the discourse (as will be seen in the analysis section below) (Schiffrin 1987; Fraser 1990a; Yilmaz 2004).

- DMs are optional to use and hard to translate (Brinton1996).

- DMs are used more in spoken language (oral discourse or conversation) than written language (Brinton 1996)

- According to Lenk (1997), DMs are short and consist of one to three syllables; however, they can be four, five, or even six-word clusters (as shown in the analysis section below).

- DMs often occupy the initial clause position (Schiffrin 1987, 2003)

- DMs can cover a range of prosodic contours (Schiffrin 1987, 2003)

- The meaning or function of DMs cannot be understood or obtained without referring to the context in which they occur. This can involve the integration of structural, semantic, pragmatic and social factors (Schiffrin 1987; Fraser 1990a; Brinton 1996; Yilmaz 2004; Alami 2015; Vellagasi 2019).

We will now detail the analytical approach taken in the study of DMs in the corpus data.

5.3 Methodology

As explained in detail in Chapter 4, an iterative approach was taken in this study where the bottom up searches are guided by the top down framework which in turn is also informed by data. For the purpose of examining DMs in this chapter, the data was approached first from a top down, research-informed direction to arrive at a framework for DMs. Following that, the data was investigated from a bottom-up route based on micro-analysis of the data (both quantitative and qualitative) where the top 100 most frequent words and word clusters lists for DMS were identified and examined. Then the concordance tool was employed (resulting in a "KWIC" (Key Word In Context format)), allowing me to see how DMs are commonly used in a corpus of texts.

In-depth concordance and source text analysis was crucial in determining whether or not items in the top 100 wordlists and cluster lists were actually functioning as discourse markers. Using this process, the pragmatic meanings or functions of only those items that are a part of spoken grammar were extracted. In other words, concordancing these top 100 most frequent single word and cluster list items manually helped to eliminate those that were not used as DMs in spoken grammar. This process resulted in the candidate list for

both corpora as shown below in Tables 5.2- 5.7. As a reminder before moving to the

analysis section, there are eight levels of pragmatic analysis that have been distinguished

based on an integrative model of spoken discourse (as previously stated in Chapter 3)

(Schneider 1988, 2001; Barron & Schneider 2005). Our focus in the present framework is

the formal level which is concerned with the analysis of linguistic forms and their

pragmatic (or communicative) functions in the discourse. Therefore, the pragmatic

variation analysis of the discourse markers use between Irish English and Canadian English

will be approached and investigated from two perspectives: form and function.

5.4 Results and Analysis

5.4.1 Forms

The Tables 5.2- 5.7 below are a summary of the comparative analysis of forms in ICE-

Ireland and ICE-Canada which demonstrate where Canadian English and Irish English

converge and diverge in terms of form. These tables have been constructed by the

qualitative analysis (from a bottom up route) of the top 100 most frequent word and cluster

lists in which each single word and cluster concordances has been manually examined. This

was done in order to affirm which items from the quantitative findings resulting from the

top 100 most frequent word and cluster lists (presented in Appendices B, C, D, E, F, G, H,

I, J, K, L, M) are actually acting as DMs within their pragmatic functions in both datasets.

Note that the light gray shaded cells in the tables below indicate the distinctive forms

(discourse markers) between Irish English and Canadian English, and the unshaded cells

indicate the forms (discourse markers) that both varieties of English share in common.

However, this commonality is only in terms of the forms of DMs. In other words, within

the data it was found that some DMs in the two varieties have common forms but, on closer

analysis, differ in the function, or the type of function within the same category of functions. For instance, the DM *well* in Irish English is also found in Canadian English as a DM; however, the DM *well* in Irish English has more pragmatic functions (or types of functions) within the function domain of DMs. Yet, this does not mean that the higher number of pragmatic functions presented by *well* in Irish English are not found in Canadian English, but they can be presented by a different form of DM (this will be further explained in section 5.4.2).

Table 5.2: *ICE-Ireland and ICE-Canada single-word discourse markers occurring within the top 100 most frequent word lists (frequency per 10,000 words)*

ICE-Ireland	Freq[1]	ICE-Canada	Freq
uh	86	uh	106
yeah	82	yeah	23
so	57	so	70
like	52	like	56
uhm	51	uhm	45
well	39	well	43
okay	16	okay	31
right	19	right	28
oh	35	oh	28
just	34	ya	53
now	39		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 5.3: *ICE-Ireland and ICE-Canada two-word discourse markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
You know	33	You know	47
That's it	1	That's it	2
I mean	12	I mean	14
And then	12	And then	13
It's just	2	Oh ya	4
Oh yeah	5	Sort of	9
Oh well	1	It's like	3

---

[1] The order of the forms presented in Tables 5.2- 5.7 is random and does not indicate rank.

I'm sure	2	Or so	1
Like what	1	And uh	13
Like that	7	And so	7
		Well let's	0.31

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 5.4: *ICE-Ireland and ICE-Canada three-word discourse markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I'd like to	1	I'd like to	1
And so on	2	You know like	2
At the end	2		
At the moment	3		
By the way	0.44		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 5.5: *ICE-Ireland and ICE-Canada four-word discourse markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
At the same time	1	At the same time	1
You know the way	1	And then after that	0.074
well I don't know	1	But oh ya anyways	0.014
That's what I'm saying	0.15		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 5.6: *ICE-Ireland and ICE-Canada five-word discourse markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
You know what I mean	1	You know what I mean	1

Table 5.7: *ICE-Ireland and ICE-Canada six-word discourse markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
and so on and so forth	1	and so on and so forth	1
at the end of the day	1	You know what I mean like	0.059
do you know what I mean	0.12		
you know what I mean yeah	0.031		
but you know what I mean	0.094		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

As demonstrated above in Tables 5.2-5.7, a comparative analysis of DMs has been conducted in which we can see the similarities and the differences between Irish and Canadian Englishes in relation to the DMs forms and their occurrences (which will all be analysed from the formal level below and the functional level in section 5.4.2). However, some of the DMs mentioned above appear in ICE-Ireland and ICE-Canada as other pragmatic markers presenting different categories of spoken grammar (see Table 5.8); in other words, with other pragmatic functions that are not included in the DM category. For example, the DM *yeah* in Irish English and Canadian English has been observed to act like a discourse marker, a question (tag), and also as a response token. On the other hand, we have *ya* which has been found only in Canadian English and has demonstrated more functions than *yeah* in Canadian English within these three different categories (mentioned above). In summary, an overall scan based on the quantitative and qualitative analyses, it has been found that:

- *ya* (jɑ) is more frequent than *yeah* (jeə) in Canadian English. However, we cannot inferably make the claim that *ya* in Canadian English is equivalent to *yeah* in Irish English (even though they differ phonologically as illustrated above). This can only be done if *yeah* in Irish English has shown more frequency than *yeah* in Canadian English not only in terms of its occurrences but also in terms of its numerous pragmatic functions; and, in this case it has (*yeah* in Irish English has shown more frequency (formally and functionally) than *yeah* in Canadian English).

- *Yeah* in Irish English has the equivalence of both of the discourse markers *yeah* and *ya* in Canadian English. This outcome has been determined by observing its frequency in the ICE-Ireland corpus (compared to both of the discourse markers

yeah and *ya* in ICE-Canada), and also by the various pragmatic functions performed by *yeah* in Irish English which overlap with the different pragmatic functions displayed by both *ya* and *yeah* in Canadian English (see Figure 5.1 and section 5.4.2)

Yet, a limitation is noted, the data underlying the analysis of such a claim is not comparable on all levels which leads to a possible influence (on the findings) resulted from some factors and measurements (see Chapter 3, section 3.2.1 on methodological and theoretical issues in variational pragmatics). For example, the principle of comparability requires the data used in investigating such a claim to be controlled data. The data is considered controlled when the social variables/contextual equivalence (e.g. males/ females, age, etc.) and text types are the same (because in the case of *yeah* vs *ya, all* different text types are used which may potentially cause a lack of equivalence if more of these forms are taken from one text type over another across varieties) (see Chapter 10, section 10.3 for a discussion of limitations) (Barron 2017a, 2021; Staley 2018).
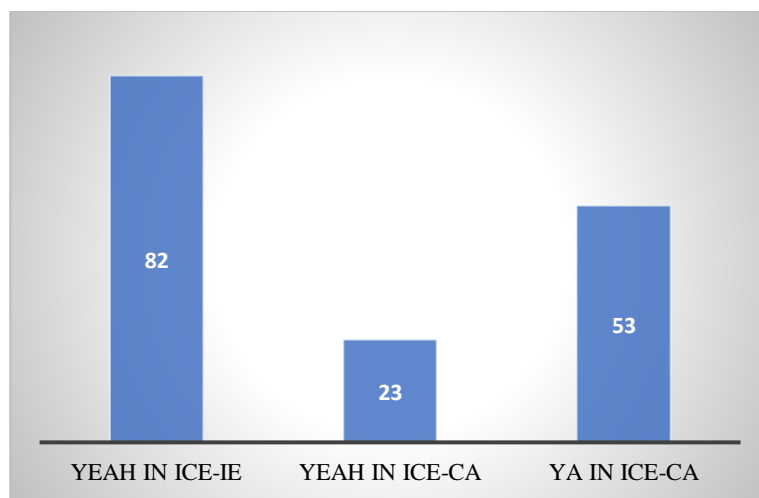

Figure 5.1 The frequency of *yeah* and *ya* (per 10,000)

From the perspective of overlapping, it is important to highlight and present the distribution of only those discourse markers of Irish English and Canadian English that have also appeared in the data as other pragmatic markers before going in depth with the analysis of forms and functions. Knowing the similarities and the differences between the DMs of Irish English and Canadian English in terms of what overlaps is the key to having a full understanding of the distinction between forms (as shown in Table 5.8). This knowledge will provide a better view (in terms of form) and analysis (in terms of pragmatic functions) than if we had only known their contribution (forms and functions) in one specific domain of spoken grammar (which have all been discussed in their specific analysis chapter in this study). Note that the DMs in Table 5.8 below that are only found in one variety but not in the other (as shown in the tables above) will be colour-coded according to the category in which the DM is found. For example, if the DM is found only in Irish English, it will take the light gray shading, while the dark gray shading indicates that it is only found in Canadian English.

Table 5.8: *The distribution (overlap) of the discourse markers of Irish English and Canadian English across other pragmatic markers*

Discourse Markers	ICE-Ireland		ICE-Canada	
uh	Response token		Response token	
yeah	Response token	Questions & tags	Response token	Questions & tags
okay	Response token		Response token	
right	Response token		Response token	
oh	Interjection		Interjection	
just	Hedge		Hedge	
oh well	Response token		Response token	
oh yeah	Response token		Response token	
well I don't know	Response token		Response token	Hedge

at the end of the day	Hedge		
do you know what I mean	Questions & tags		
you know what I mean yeah	Questions & tags		
so		Response token	Hedge
ya		Response token	Questions & tags
I mean		Stance marker	
and then		Response token	
sort of		Response token	Hedge
it's like		Hedge	
I'd like to		Stance marker	Response token

Having presented all the forms of DMs in both Irish English and Canadian English and their distribution in the tables above, it can be stated that:

- Irish English and Canadian English clearly vary in their DMs in relation to form.

- The data has shown that the most similarities and commonly shared items (forms) are at the single-word discourse markers level, with frequencies decreasing as the length of multi-word unit increases (in line with O'Keeffe *et al* 2007).

- At the single-word discourse markers level, we can see some notable distinction in frequency of some items which (as will be shown in section 5.4.2) have been found to have more pragmatic functions in the discourse, such as *okay* and *now* (see Figures 5.2 and 5.3 and section 5.4.2.1).
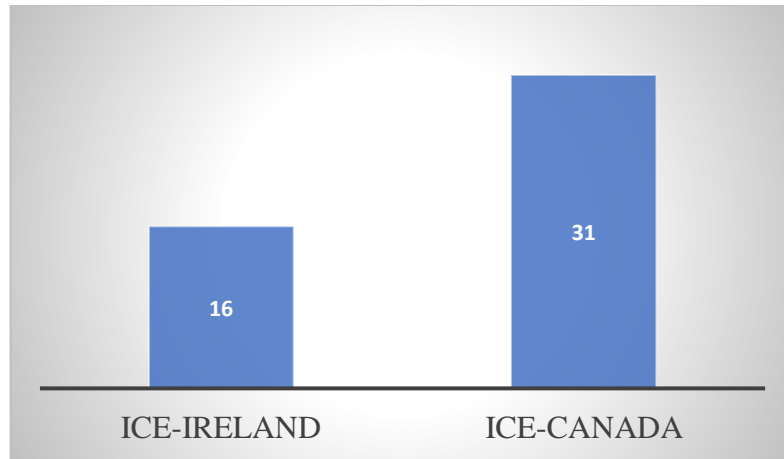
Figure 5.2 The frequency of *okay* (per 10,000)
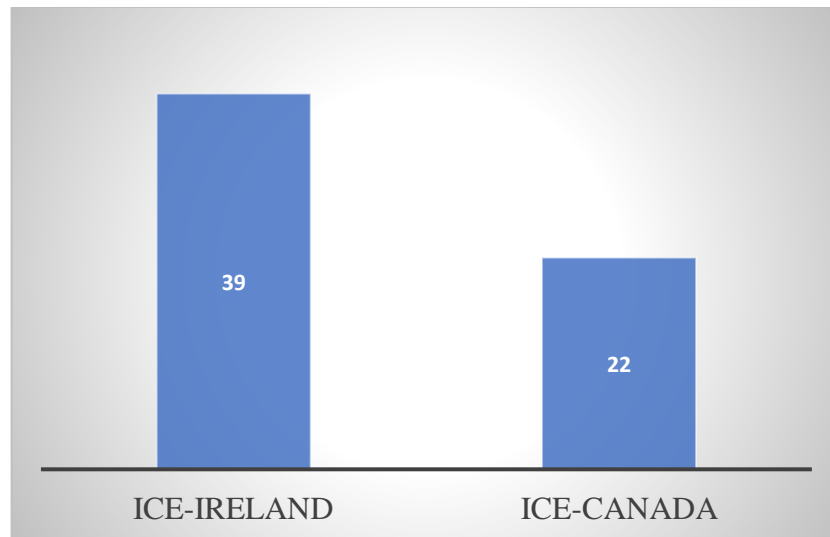
Figure 5.3 The frequency of *now* (per 10, 000)

- Additionally, the data has shown that Irish English and Canadian English in terms of form started diverging noticeably at the two-word discourse markers level (as seen in Table 5.3).

- At the two-word level, we can see the same forms of DMs occurring in both datasets (with different and similar frequency) where they were shown to reflect their pragmatic functions and use at the function level. Two examples of this are

you know and *I mean*. *You know* in Canadian English is more frequent than it is in Irish English and tends to implement more pragmatic functions than *you know* in Irish English, whereas, *I mean*, which has a very similar frequency in both datasets was shown to be serving the same pragmatic functions in both datasets (see section 5.4.2.2 for further discussion; see also Figures 5.4 and 5.5 below for frequency data).
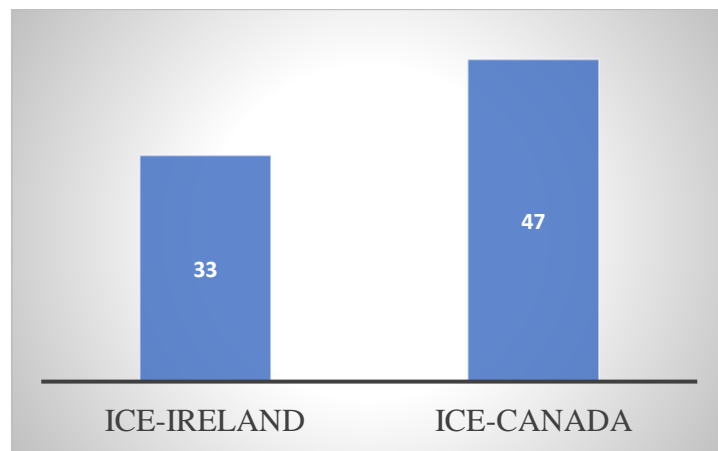
Figure 5.4 The frequency of *you know* (per 10,000)
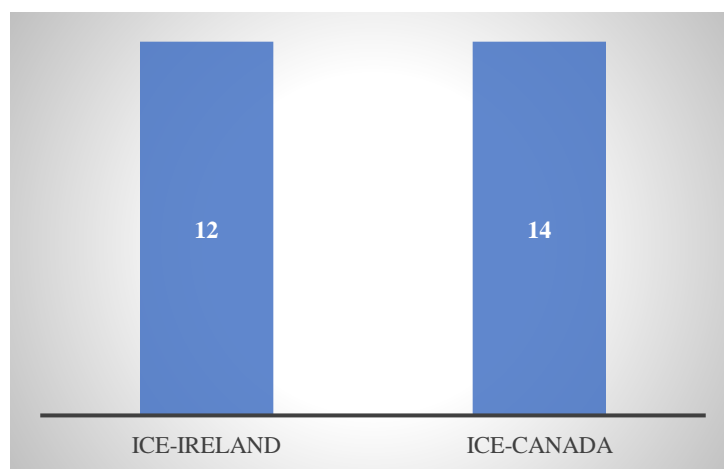
Figure 5.5 The frequency of *I mean* (per 10,000)

As a result, in terms of comparison at the level of forms based on single-word discourse markers up to six-word discourse markers with their frequency and distribution, the corpus search brings to light a number of points:

- Irish English and Canadian English have shown the most convergence at the single-word DMs level in comparison to the other multi-word levels; not only that, but they are also closer in terms of the DMs distribution at the single-word level compared to the other levels.

- At the single-word DMs level, there is only one DM item in Canadian English that is not found in Irish English, which is *ya*; in contrast, we have two discourse markers items which appear in Irish English that are not found in Canadian English, which are *just* and *now* (as illustrated in Table 5.2).

- The item *just* appears as a hedge in Canadian English but does not appear as a DM; on the other hand, the item *just* in Irish English appears as a DM and a hedge (for more on *just* as a hedge see Chapter 8, section 8.4.1; for *just* in Irish English see Kirk & Kallen 2009).

- It has been found that there is convergence in how these single-word DMs *uh*, *yeah*, *okay*, *right*, and *oh* are used in both varieties (Irish and Canadian) due to their distribution (overlap) and variation across other pragmatic markers (as shown in Table 5.8).

- It has been found that Canadian English has shown a different range of forms distribution at the two and three-word level (see Table 5.8). For example, these items *I mean*, *and then*, and *I'd like to*, which all occur in both varieties, have appeared in Canadian English as other pragmatic markers other than a DM whereas

these items appear in Irish English only as DMs (see Table 5.8 for other items such as *sort of* and *it's like* which are found only in Canadian English).

- At the two, three, four, and six-word DMs level, there is a clear distinction between Irish English and Canadian English in terms of their selection of DMs, whereas the five-word DMs level has the least amount of DMs occurring in both datasets.

- At the six and three-word DMs level, Irish English has demonstrated a wide range of DMs in comparison to Canadian English which only has a few.

Overall, we can say that Irish English and Canadian English have demonstrated a varying range of DMs in relation to forms at all levels except (inconclusively) at the single-word DMs level. This result was somewhat expected due to the geographical distance, historical factors and other factors that have helped to shape the two varieties (as explained in Chapter 1). However, so far we have only discovered the forms of DMs which do not offer enough conclusive evidence to provide an overall claim about how similar or different both varieties of English are in relation to DMs. Therefore, we must analyse the DM items that we have analysed in terms of forms above in relation to their pragmatic functions which will be covered in the section below.

5.4.2 Functions

The pragmatic functions, as we mentioned earlier, are based on the definition of DMs adapted in this present study from Carter and McCarthy (2006), and they are concerned with three main pragmatic functions: managing, monitoring, and organising the discourse. Other pragmatic functions can be categorised under these three main functions due to the fact that there are various and optional choices practised by the speaker in order to link the segments of the discourse to one another in ways that reflect their choices in relation to

managing, monitoring, and organising the discourse (Brinton 1996; Carter & McCarthy 2006). So, as mentioned earlier (in section 5.3 and in section 4.5.1.2 in greater detail), all the generated forms as single/clusters have been studied and looked at in terms of their pragmatic functions, through the concordance searches (the concordance lines) in which the (right and left) context have been analyzed based on the seven steps provided by Sinclair (as mentioned in Chapter 4) (2003). After that, the pragmatic functions were categorized based on the findings (which will be shown below). However, the overall summary of the pragmatic functions analysis has shown that both Irish English and Canadian English have performed or adopted almost the same pragmatic functions displayed by various and similar forms of DMs. Note that as we mentioned earlier, the pragmatic functions analysis is based on three main pragmatic functions: managing, monitoring, and organising the discourse in which they contain some sub-pragmatic functions (within themselves) as presented in Table 5.9 below. These sub-pragmatic functions will be unpacked and analysed in the following sections: 5.4.2.1, 5.4.2.2, 5.4.2.3, and 5.4.2.4.

Table 5.9: *An overall summary of the pragmatic functions found in ICE-Ireland and ICE-Canada in relation to managing, monitoring, and organising the discourse, and others*

Sub-pragmatic functions of managing	Sub-pragmatic functions of monitoring	Sub-pragmatic functions of organising	Other pragmatic functions
To open up the discourse. To close down the discourse. To focus the attention of the listener on the discourse by resuming the	To monitor the ongoing discourse through reformulations and alternative expressions. To monitor the ongoing discourse by elaborating and	To mark sequential dependence. To focus attention usually by giving or requesting an example.	To serve fillers or hesitation markers indicating various functions (i.e. thinking, hesitation, surprise, having a hard time communicating). To signal that something surprising

topic of the discourse.	expanding on the discourse.	To mark boundaries and link segments in the discourse.	is about to be announced.
To focus the attention of the listener in the discourse by diverting the topic of the discourse. To focus the attention of the listener in the discourse by shifting the topic of the discourse. (see section 5.4.2.1 for more information)	To monitor the ongoing discourse by softening a segment in the discourse. To monitor the ongoing discourse by denoting similar entities. To monitor the ongoing discourse by monitoring the state of shared knowledge in the conversation. To repair a segment in the discourse. (see section 5.4.2.2 for more information)	(see section 5.4.2.3 for more information)	To mark reported speech. To mark emphasis on a segment of discourse. To qualify a preceding statement. To mark sarcasm on a segment of discourse. To mark that everything has been taken into consideration. (see section 5.4.2.4 for more information)

These sub-pragmatic functions (presented above in Table 5.9) were derived through

qualitative analysis of the corpus data, based on the bottom-up approach. This involved

examining concordance lines and source files and categorising the main functions of DMs

in terms of managing, monitoring, and organising or "Other" the discourse based on the

existing literature. The sub-functions presented in Table 5.9 as regards managing,

monitoring, and organising the discourse, align with Carter and McCarthy's six function

categories (outlined in section 5.2.2). A number of the additional pragmatic functions

(classified in the "Other" column) can be said to be informed by Brinton's (1996)

interpersonal function (cf. Table 5.1).

5.4.2.1 The pragmatic functions: Managing the discourse

Managing the discourse can be done in various ways and can even be categorised differently from Table 5.10 (displayed below) due to the disputes which exist in relation to categorization, terminology, and definition (as discussed earlier in sections 5.2.1 and 5.2.2). According to Carter and McCarthy (2006), the discourse can be managed through "launching and concluding topics, opening, concluding or temporarily closing a whole conversation, re-opening previously closed or interrupted conversations" this even includes the focus of a segment of the topic within the conversation in which the speaker will not close down the whole topic completely, but will rather focus the attention of the listener on a segment within the topic which the speaker feels is important (p.214). This is usually done by diverting, shifting, or resuming the topic of the discourse (Carter & McCarthy 2006). In relation to the function of managing the discourse, Irish English and Canadian English have demonstrated all the pragmatic functions by which the discourse has been managed (as will be discussed below and seen in Table 5.10).

Here (in Table 5.10) are the forms used by Irish English and Canadian English for managing the discourse (categorized with their sub-functions). They are listed without frequencies because all frequencies of all the forms have been mentioned earlier (in section 5.4.). Also, the frequency indicates how frequently the forms have been used in the data in general across all other functions and as all other pragmatic markers but not how frequent they are in terms of their specific pragmatic function. The order of the forms (presented in Table 5.10) is random; in other words, it does not have any significance in the analysis. Some of the forms in Table 5.10 are repeated due to their frequent occurrences based on the sub-functions indicated in the table. These forms have been gathered through the quantitative analysis based on the wordlists and clusters lists and the qualitative analysis

based on concordance lines in which their pragmatic functions have been determined in the discourse, serving the pragmatic function of managing the discourse.

Table 5.10: *Comparative analysis of the forms used in Irish English and Canadian English in relation to managing the discourse*

Pragmatic functions	ICE-Ireland	ICE-Canada
To open up the discourse	well now right by the way okay	okay right well let's
To close down the discourse	so well that's it now	so okay that's it and so
To focus the attention of the listener in the discourse	oh yeah well so now	and so oh ya well but oh ya anyways so and so

Opening up and closing down the discourse

Now, based on Table 5.10 above, let us examine how the management of openings and closings is done in Irish and Canadian discourse through their DMs. In this respect, and based on the qualitative analysis we find that:

- The single form of *well* (with the frequency of 39 per 10,000 words in Irish English) is used more in Irish English than Canadian English, even though *well* (in Canadian English) has been found in the top 100 with the frequency of 43 per 10,000 words which is almost the same frequency/occurrences. On the other hand, in Canadian English the two-word cluster *well let's* has been adopted. This cluster appears in Canadian English but it does not appear in the top 100 of the cluster lists in Irish

English. This has been guided by the qualitative analysis and supported by the fact that *well let's* has been found among the top 100 of the cluster lists in Canadian English (see Appendices I and O)**.**

- The DM *now* (with the frequency of 39 per 10,000 words in Irish English) has been found to be in use much more in Irish English when compared with Canadian English generally (as *temporal*), and specifically in initiating (opening up) or closing down the discourse or changing, clarifying or refocusing/focusing (a topic in the interaction) the attention of the listener in the discourse. *Now*, in terms of its frequency in Irish English across other varieties, has the highest frequency of occurrences (Migge 2015; also see Clancy & Vaughan 2012 who compared *now* in (LCIE) to other spoken corpora from three other varieties of English: *The Scottish Corpus of Texts and Speech* (SCOTS), the *British National Corpus* (BNC), and the *Corpus of Contemporary American English* (COCA)). On the other hand, *now* does not appear as a DM in Canadian English (but as a lexical item with the frequency of 22 per 10,000 words, see Figure 5.3 for the distinction in frequency between their occurrences in both varieties). In addition to that, the pragmatic marker *now* has variety-specific uses/functions (the *hedging now* and the *deictic presentative now*) that are associated specifically with Irish English (see chapter 8, section 8.4.1 for more details in relation to the present study).

- Similarly, the same indication can be observed with the DM *okay* which occurs twice as much in Canadian English than in Irish English (see Figure 5.2).

To exemplify this here are some examples extracted from both datasets demonstrating how some of these DMs are used in relation to opening up the discourse.

Example 5.1[2]

A: Uh, uh coming on, you know, knowing that I was coming on this programme I did a little bit of background research on this whole embedding concept and we've discovered that uh it was a public relations woman uh, she works for Donald Rumsfeld uhm +uh, who coined this phrase embedded.

B: Who coined it?+ **Well** can can you explain to viewers what ++what's the difference between an embedded journalist and a non-embedded journalist? You're you're a non-embedded journalist.

A: Tory Clarke apparently is her name++.

                                            [ICE-Ireland, Broadcast discussions, S1B-037$A][3]

Example 5.2

A: Uhm **now** let's look at this novel as the, as the last one that that we've looked at. Uhm well he says that he's trying to discover something about the twentieth century but he sets it in the early part of the nineteenth century.

                                            [ICE-Ireland, Unscripted speeches, S2A-043$A]

Example 5.3

A: Gerry Collins you could, you could create a seat for him Gerry **by the way.** You're in Europe now. You could design [unclear speech] and he could stand in Limerick West.

                                            [ICE-Ireland, Broadcast discussions, S1B-033$A]

As seen in Example 5.3, the three-word cluster *by the way* has been used to open up a

segment of a topic, an idea or a suggestion, in the discourse that had not been brought up

before in the discourse.

Example 5.4

A: Okay just remember. Remind me because sometimes +I- I tend to forget.

B: Right, right.+

A: **Okay** how are things going?

B: Well, I just came back on Monday. Mosta, ^Most of^ last week I was out except for one day I tried to go in well I did go in, and uh. So I'm a bit out of it.

                                            [ICE-Canada, Classroom lessons, S1B-007 #132:1: A]

---

[2] The transcription of the mark-up symbols for all examples taken from ICE-Ireland and ICE-Canada used in this study are provided in Appendix P.

[3] This research is not concerned with social factors (such as gender and age) which can be identified by speaker IDs. However, all references for the examples provided throughout the thesis, along with the sub-corpora and the text-type, include the speaker IDs (as presented in the corpora) for the speaker who uses the pragmatic marker/item being displayed. This is done to help other researchers conduct further investigation within the domain of sociolinguistics based on the findings of this research. Note that there may be a discrepancy between the speaker ID in the example and the reference, this is because the examples are extracts from lengthier conversations in the corpora.

Example 5.5

A: And it'll feel great +at first cos it's a new ski but what happens is because it will be receiving so much abuse they're not designed that way they're not ++made to [unclear word] so

B: Right.+ **Right**.++ What about this section over here? Is there anything there that…

A: Uh generally those skis are probably a little bit more +expensive. If you wanna-want to ++take a look?

B: No, it's okay,+ it's okay.++

                                    [ICE-Canada, Business transactions, S1B-075 #157:1: A]

In Example 5.5, the single-word DM *right* has been used twice. The first *right* is a response

token; whereas, the second *right* has been used to redirect the focus of the topic onto

another focus in the discourse.

Example 5.6

A: Like the shower curtain you know. There wasn't enough air or something for it to uh breathe. I don't know about mould.

B: [yawns]

A: Mouldy, Mouldy.

B: Okay **well let's** change the clocks so that we don't wake up an hour late tomorrow.

A: Okay.

                                    [ICE-Canada, Face to face conversation, S1A-085 #96:1: B]

In Example 5.6, above, the two-word cluster *well let's* is used to close down the previous

focus of the discourse by opening up another focus in the discourse.

In terms of closing down the discourse, we found that *so* and *that's it* are both similar in

terms of their pragmatic functions in both varieties. However, in terms of the single form of

so occurring in clusters, it is apparent that Canadian English acquires more clusters with *so*

than in Irish English. Not only that, but it has been noticed that the single-word DM which

occurs more in clusters (in any variety of English) tends to be in a higher frequency than

the one that does not. For example, (in Canadian English) the DM *so* is at a higher

frequency than it is in Irish English as well as other DMs such as *uh* and *oh* (see Tables 5.2

and 5.3). Here are some examples illustrating *so* and *that's it* in both varieties of English.

Example 5.7

A: But you know you've ^you're^ near the college and you've got all the facilities you need, like a washing-machine and so on. **So** it's grand like. It does- does the job.

[ICE-Ireland, Face to face conversation, S1A-048$B]

Example 5.8

A: You said "shehbin" or something.
B: Oh I meant "shelves" I think.
A: Okay.
B: Anyway. So, **that's it.**
A: Cool.

[ICE-Canada, Face to face conversation, S1A-091 #195:1: A]

Again, the dominant usage of *now* and the single form of *well* appear more in Irish English

than Canadian English. Conversely, the dominant usage of *okay* appears in Canadian

English more than Irish English in relation to closing down the discourse (this can be done

through closing down a point in the discourse to opening another) (see examples 5.9- 5.11

below).

Example 5.9

A: So this type of software is specially written because the needs of companies vary from one company to another. Okay. That is special-purpose. Okay. **Now,** we've dealt with the uhm word processing, okay? So, in a nutshell uhm Anna, what are the basic features of a word processor?

[ICE-Ireland, Unscripted speeches, S2A-042$A]

Example 5.10

A: There really wasn't, there wasn't any management team meeting, and when there isn't a management team meeting, there isn't a whole pile it was mostly about the OT service which is down further on. **Well,** it's actually next on the document anyway.
B: Okay.

[ICE-Ireland, Business transactions, S1B-078$D]

Example 5.11

A: He didn't want that, he just wanted matching shirts.
B: Ya, well he can get that.
A: Oh, so in other words he can order team shirts for
B: Oh ya, no problem he can go there and get them right away.

A: **Okay,** thank you so much.
B: Okay, you want the phone number?
A: Sure.

                    [ICE-Canada, Broadcast discussions, S1B-037 #212:4: B]

[In example 5.11, the first *okay* is illustrative of a closing remark].

Focusing the attention of the listener in the discourse

Now we will look at the sub-pragmatic function of focusing the attention of the listener in the discourse which can be done by diverting, shifting, or resuming the topic (or a partial segment) of the discourse (Carter & McCarthy 2006). In relation to this pragmatic function "focusing the attention of the listener in the discourse" in general, Irish English and Canadian English have shown resemblance in terms of the usage of forms presenting this type of pragmatic function such as *so* and *well*. On the other hand, both varieties have also exhibited different DMs serving the pragmatic function of focusing the attention of the listener in the discourse with *oh ya* and *but oh ya anyways* clusters which are found only in Canadian English and with *oh yeah* which is found only in Irish English within the top 100 most frequent cluster list. See examples 5.12- 5.18 below for more illustration.

    Example 5.12

A: Yeah well I mean I'm hardly going to wear it now seeing +everyone thinking I've big hips! [laughs] Hip girl! I'll be called hippy. [laughs] Hippo! [laughs]. Ah, **so** how are you anyway?
B: Yeah+. Fine [laughs].

                    [ICE-Ireland, Telephone conversation, S1A-099$A]

In example 5.12, a telephone conversation, speaker A talks about how they will look wearing a certain type of clothing. The same speaker stops talking about their clothing and changes the topic on their own. They do this even though the hearer is engaging in the conversation very well (using *yeah*) and there is no indicator or signal from speaker B that

speaker A should stop. However speaker A seems to shift the topic focus to something else, using *so* at the beginning and supporting the focus shift by using *anyway* at the end. Interestingly, this is in the context of a telephone call where the use of *how are you* is well-documented as part of the canonical of call opening sequences (see Drew & Chilton 2000). In this context *how are you?* is seen as part of phatic communication within the opening sequence. It is interesting that Speaker A seeks to reset the call by (re)introducing *how are you?*. Normally, the business of the call follows the *how are you?* phrase so it is possible that *so + how are you?* is being used strategically by speaker A as a means of staging the main business of the call. This example stands in comparison to example 5.4, where we see that the speaker in example 5.4 proposed something that was confirmed and understood by the hearer. Thus, there is no point for the speaker to elaborate more since the hearer is fully acknowledging the situation and sending signals which indirectly trigger the speaker to stop or open another topic. This is because their message has been well received whereas this indication cannot be seen clearly in regard to focusing the attention. Yet, both sub-functions (focusing the attention and opening a topic) can overlap as seen in example 5.5.

Example 5.13

A: Ya ya a lot of the cold ya comes in from the bottom too. There's nothing heated underneath.
B: Well less now because we have the door to the shed but
A: Less
B: **So** let's mention that some time.
A: To Brian?
B: No just throw it out in, throw it out into the dinner conversation when he's had a few beers.

[ICE-Canada, Face to face conversation, S1A-080 #32:1: A]

In example 5.13, Speaker B is referring to how he is not the main person to talk to about the matter at hand so he is shifting the focus of the discourse to the landlord, Brian, who would be in charge of it.

Example 5.14

A: Why do you think that he has to have a BMW and keys and money?
B: **Well** a l- a lady needs to b, to be driven around. She she can't be expect- she can't
       be expected to sit on the back of a bicycle. [laughs]
                                    [ICE-Ireland, Face to face conversation, S1A-061$C]

The DM *well* in example 5.14 has been used to focus the attention of the listener in the

discourse by diverting to an unexpected segment within the topic for the sake of sarcasm.

Example 5.15

A: Why, have you made your dinner?
B: No. I have, I told you, I have to go get milk.
A: **Well** they're eating dinner right now.
B: Oh. **Well** it doesn't matter. I'll like, do you know when they'll be back?
A: From where?
B: You just finished saying that's ^that^ they're not there!
A: They are here.
B: Oh my god.
                                    [ICE-Canada, Face to face conversation, S1A-098 #194:1: B]
The DM *well* is used for diverting the focus of the discourse (shown mainly in the first use

of *well* as well as the second).

Example 5.16

A: I wanna ^want to^ put you on mute. [laughs]
B: You bastard.
A: Ya. That's the oh boy. I hate. **Oh ya** actually the uhm uhm stereo guys are coming
       today.
B: Are they?
                                    [ICE-Canada, Face to face conversation, S1A-073 #204:2: A]

Example 5.17 [Two friends are discussing what happened over a drunk weekend]

A: Who jumping on top of who?
B: Well she ate the face off me.
A: **Oh yeah**. Well she was always a bit like that though anyway, when I was there
       even.
B: Yeah she was, wasn't she?
                                    [ICE-Ireland, Face to face conversation, S1A-086$B]

The two-word cluster *oh yeah* occurs more frequently as a RT; however, here it seems that

it is a combination of both where it is acknowledging the focus of the topic and resuming

by expanding on what has been said earlier. However, what establishes this cluster as not

being a RT is the definition of RTs which states they should not take over the turn (as

discussed in Chapter 6 analysing RTs).

  Example 5.18

A: They love her. They eat her.
B: [laughs] eau de F one generation. [laughs] I don't know.
A: Ya. **But oh ya anyways** what I was saying, the computers that uh I'll be working
       with next year, uhm silicon graphics…
                              [ICE-Canada, Face to face conversation, S1A-020 #54:1: B]

The four-word cluster *but oh ya anyways* (in example 5.18) is used as a diverting marker,

diverting (or shifting) the focus of the discourse to the original focus of the conversation.

5.4.2.2 The pragmatic functions: Monitoring the discourse

According to Carter and McCarthy (2006), some DMs enable speakers to monitor the

ongoing discourse through different processes such as reformulations (alternative

expressions) indicating that the speaker has not selected the most appropriate way of

expressing themselves, or the idea (or the topic) being talked about in the discourse. Also,

monitoring the ongoing discourse can be through expanding on the topic of the discourse

for more clarification, softening a segment in the discourse, monitoring the state of shared

knowledge in the conversation (signalling that the speaker is sensitive to the needs of the

listeners being on the same platform in terms of what has been said), repairing a segment in

the discourse (and this is usually done by self-correction), and finally it can be done

through monitoring the ongoing discourse by denoting similar entities in terms of quantity

or quality (as shown in examples 5.34 and 5.35) which can help in reducing and shrinking

the length of the conversation or the discourse (see Table 5.11 below; also see Carter &

McCarthy 2006, pp.107, 108, 221).

Table 5.11 (below) shows all the forms used by Irish English and Canadian English for monitoring the discourse (categorized with their sub-functions); these are listed without frequencies because all frequencies of all the forms have been mentioned earlier (see section 5.4 for more details on frequencies in the data and how they are presented here).

Table 5.11: *Comparative analysis of the pragmatic functions and their forms in relation to monitoring the discourse*

Pragmatic functions	ICE-Ireland	ICE-Canada
To monitor the ongoing discourse through reformulations and alternative expressions.	I mean that's what I'm saying	I mean it's like
To monitor the ongoing discourse by expanding the discourse.	I mean	I mean
To monitor the ongoing discourse by softening a segment in the discourse.	I mean	I mean
To repair a segment in the discourse.	well I don't know I mean oh	I mean
To monitor the ongoing discourse by monitoring the state of shared knowledge in the conversation.	you know you know the way you know what I mean do you know what I mean you know what I mean yeah but you know what I mean	you know you know what I mean right
To monitor the ongoing discourse by denoting similar entities.	and so on and so on and so forth	or so and so on and so forth

Now, we will look at each sub-function (presented in Table 5.11) more closely which all serve to monitor the discourse. In terms of monitoring the ongoing discourse through reformulations and alternative expressions, we see from Table 5.11 that Irish English and Canadian English have both demonstrated the usage of the cluster *I mean* in almost the same way. This is not only in relation to the pragmatic function of monitoring the ongoing

discourse through reformulations and alternative expressions, but also, at most of the

different levels of pragmatic functions monitoring the discourse in general (see Figure 5.5)

(Carter & McCarthy 2006). See Examples 5.19 and 5.20 extracted from both datasets

showing the ongoing discourse being monitored through reformulations and alternative

expressions:

Example 5.19

A: Yeah, I thought it wouldn't be a bad holiday. Uh, uh we've had quite good weather
       though so I thought well, I mean I would be going to Galway for Christmas and it's
       quite **I mean,** it's pleasant enough. My sister is a great cyclist. We go out cycling
       around.
                                        [ICE-Ireland, Face to face conversation, S1A-047$A]

Example 5.20

A: I think more public access to really good routes in the city would be a nice thing. **I
       mean** it's a healthy thing.
                                        [ICE-Canada, Broadcast interviews, S1B-046 #81:1: A]

In terms of monitoring the ongoing discourse through elaborating and expanding on the

discourse (or a segment in the discourse), see the examples 5.21 and 5.22 below:

Example 5.21

A: I I don't think we quite got that. I I I think we we we got an opinion that it might
       have been the gentleman who's now deceased. But **I mean**, can I ask you the
       question like, **I mean** who was it? You're you're the former chairman.
                                        [ICE-Ireland, Legal cross-examinations, S1B-068$A]

Example 5.22

A: [laughs] Well well Rob Roy tended to show, it it it wasn't so gruesome, or so hard.
       I mean a a lot of the killing in there was totally without mercy, and I think they
       showed that to show that like the king's lack of feeling and desire for power but uh
       the uhm…
                                        [ICE-Canada, Face to face conversation, S1A-022 #124:1: A]

In terms of monitoring the ongoing discourse through softening a segment in the discourse,

see examples 5.23 and 5.24:

Example 5.23

A: Is Niagara Falls Canadian?
B: Well the, we went +to the Canadian side. It's meant to be more ++more nicer.
C: Oh right. Oh.+
A: What's it actually++ on the border anyway.
B: Yeah nicer than well **I mean** [unclear speech]
C: And is it impressive?

<div align="right">[ICE-Ireland, Face to face conversation, S1A-065$C]</div>

The DM *I mean* in example 5.23 is meant (and used) to soften the interrupted segment (as

she was about to insinuate criticism about the United States but refrained).

Example 5.24

A: It's kind of a cranberry colour.
B: Well it's a nice uh colour for Christmastime, isn't it?
A: Ya, it is. **I mean** it's not my favourite colour but Mel loves it.
B: Oh that's good!

<div align="right">[ICE-Canada, Face to face conversation, S1A-064 #321:2: A]</div>

In terms of monitoring the ongoing discourse through repairing a segment in the discourse,

see examples 5.25 and 5.26 below where the speaker corrects themselves by using the

cluster *I mean. I mean* has been known for this pragmatic function in many other varieties

of English as well, such as British English according to Carter and McCarthy (2006), as

well as Irish and Canadian English (as seen below):

Example 5.25

A: Do you reckon? Has there been anything like?
B: Yeah yeah. Now do Jeanette.
A: I swear on my mother's life, **I mean**
B: Swear on the holy Bible.
A: Listen.
B: Yeah.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-049$B]</div>

Example 5.26

A: Uh they're, the way that they judge their peers, the way they judge their
       accomplishments is completely different from mine. But I learned a lot about them.
       I mean, I learned from them- from them if they didn't if they stayed out of trouble it
       was something.

<div align="right">[ICE-Canada, Broadcast interviews, S1B-049 #59:1: B]</div>

Here (in example 5.26), the speaker repaired a segment in the discourse through self-correction where the speaker changes *about them* to *from them*. In addition to that, both varieties have adopted the two-word cluster *I mean* in the five-word cluster *you know what I mean* which has appeared within the top 100 frequent clusters in both datasets with the pragmatic function of primarily checking comprehension and indicating that the speaker and hearer share a similar point of view. However, despite the similarities found above, Irish English and Canadian English show some contrasts in relation to pragmatic functions monitoring the discourse. For example, in terms of self-correction in the discourse, we see the tendency in Irish English to use the four-word DM *well I don't know* not only as a RT but also as a DM. In Canadian English, however, this DM tends to be used more as a RT (as explained and discussed in Chapter 6). Similarly, the single-word DM *oh* has been found more in Irish English as a RT and a DM with the function of repairing a segment in the discourse as demonstrated in the examples 5.27 and 5.28 below:

Example 5.27

A: Did you ever call out? [laughs] He was lovely. He was also engaged, but anyway. Uhm [laughs] anyway, the parents never cooked for him. They never really ate at all. **Well I don't know** what they did, but I mean I was the only one that ever seemed to cook. Oh yeah, fast food kind of thing. But anyway, they just let him do whatever he wanted and they used to buy him paper plates and paper and you know plastic knives and forks because he wouldn't wash up. And this was all he had to do then afterwards.

[ICE-Ireland, Face to face conversation, S1A-056$B]

The DM *well I don't know* is a form of self-correction here as she had exaggerated in the previous statement and is now trying to soften and correct that exaggeration.

Example 5.28

A: You're trying to like cut up your pizza and suddenly you're cutting through the plate. [laughs]
B: Oh yeah.
C: Disgusting.

A: Orla's just so domesticated.
C: That's true.
B: Oh well
A: Orla, you cook a lot?
B: Only when I absolutely have to, **oh** no I do actually if I have the time I will.

[ICE-Ireland, Face to face conversation, S1A-056$D]

Furthermore, Irish English and Canadian English have shown an obvious contrast in relation to monitoring the state of shared knowledge in the discourse. This distinction is more at the form level in which Irish English uses more clusters to serve this pragmatic function in the discourse whereas Canadian English uses much fewer based on the quantitative findings (as shown in Table 5.11) and the qualitative analysis of the data. See examples 5.29- 5.33 below where the speaker monitors the ongoing discourse by monitoring the state of shared knowledge in the conversation and to keep the listener involved in the conversation to make sure that both share the same point of view or both are on the same track.

Example 5.29

A: We were in work today **you know,** Mark said to me uhm he was they, we, he lives out near Tallaght **you know.** And he was saying something like uhm, "oh me da when he drives around at night." And Mark is the one who's my assistant, **you know**. And the guy is really, he laughs all the time. He drives me crazy but he's okay…

[ICE-Ireland, Face to face conversation, S1A-070$A]

Example 5.30

A: Yeah, I gue- yeah I guess, it's going to depend on what specifically you're doing. **Do you know what I mean?** So that if for example, you are doing vocabulary and you decide to do geography words we'll say whatever and like you do five words per session.

[ICE-Ireland, Classroom lessons, S1B-017$A]

Example 5.31

A: And so uhm he was like he's been I've been meeting him like all week. I've me-met him all last week. I'd say we've been uh like meeting and he'd be we we'd be drinking coffee and whatever you know?
B: Yeah.

A: And uh what else then uhm +then
B: Well I uhm uhm I+ knew I had you sussed.
A: Yeah mm mm.
B: Knew it. But uhm I wasn't sure at the same time like. **You know the way?**
A: Yeah, oh I know. No I mean I've been always good friends with with uhm with Jim right.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-049$B]</div>

Example 5.32

A: They look good, de dah duh duh and they have the book and the teaching and duh de duh de duh. And after a month **you know** you see them start coming with their jeans and and their t-shirt **you know.**

<div align="right">[ICE-Canada, Unscripted speeches, S2A-040 #6:1: A]</div>

Example 5.33

A: Like you're in a family **right**+ an- and you're, you think you're the weird one++ in the family because you're the artist and everyone else does, you know, other stuff that involves you know nine to five jobs or something, and families ya
B: [laughs] This sounds strange.+
A: [laughs] Okay.++

<div align="right">[ICE-Canada, Face to face conversation, S1A-032 #124:1: B]</div>

The DM *right* is used more often in Canadian English to monitor the ongoing discourse by monitoring the state of shared knowledge in the conversation. It also marks boundaries in the discourse – because it goes from talking about the family in general to the specific place and feeling one holds while being in that family.

Moreover, based on the quantitative findings, the four-word DM *that's what I'm saying* has been found on the top 100 frequent clusters in Irish English but not in Canadian English with the pragmatic function of monitoring the ongoing discourse through reformulations as seen in example 5.34 below.

Example 5.34

A: Uh I think so yeah. Well we we don't know what to do because we're invited to Sarah's as well on the same night. And uhm, and we we're just a bit torn between the two of them.
B: Ah no, don't not go now. I knew you wouldn't come!
A: Why?
B: Cos I [tut-tut] I bet you now, you're not going to come, are you?

A: I don't know, **that's what I'm saying.** We're torn. Why?
B: Cos I just knew you wouldn't come.
A: She said I knew you wouldn't come.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-099$A]</div>

On the other hand, we have the two-word DM *or so* which appears in the top 100

frequent clusters in Canadian English. It can be said that it is similar to *and so on* (but not

the same) which has appeared in the top 100 frequent clusters in Irish English but not in

Canadian English. Both present similar pragmatic functions in which the speaker is

denoting similar entities either in terms of quantity (something that can be counted) or

quality (as ideas expressed usually by the phrase *or something like that* etc.) as seen in the

following examples (5.35 and 5.36).

Example 5.35
A: Uhm ya. I'll talk to you in about a week **or so** and let you how I'm doing.
B: Okay.

<div align="right">[ICE-Canada, Classroom lessons, S1B-008 #237:1: B]</div>

Example 5.36
A: But you know you've, ^you're^ near the college and you've got all the facilities you
        need like a washing-machine **and so on**, so it's grand like. It does does the job.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-048$B]</div>

5.4.2.3 The pragmatic functions: Organizing the discourse

In relation to the pragmatic functions of organizing the discourse as a whole, there are a

wide range of different approaches to serve this type of pragmatic function in the discourse.

One of the clearest methods to do that is sequencing which "indicates explicitly the order in

which things occur or how different segments of a discourse are being organised" (Carter &

McCarthy 2006, p.216). In other words, sequence markers function as markers marking

sequential dependence (Brinton 1996). In relation to this type of pragmatic function, we can

see clearly that Irish English and Canadian English have shown some similarities and

differences according to the quantitative findings extracted from the top 100 word clusters

and the qualitative analysis (Table 5.12) and the examples below, in which the speaker is

organising the discourse through sequencing.

Table 5.12: *The discourse markers marking sequential order in ICE-Ireland and ICE-Canada*

Pragmatic function	ICE-Ireland	ICE-Canada
Indicating explicitly the order in which things occur or how different segments of a discourse are being organised	And then At the same time At the moment At the end	And then At the same time And then after that

Example 5.37

A: And of course you have to dig it up and it's the root that you actually eat. So all you have to do is get the root like that. **And then** peel off the outer skin. Of course give it a good wash and make sure you get the clay off first. **And then** take a grater and just grate the horseradish finely like that, just like that. As much as you want. And I've already got some grated here so I'll use this.

[ICE-Ireland, Demonstrations, S2A-057$A]

Example 5.38

A: You might think of the word "self" for the breath. Slow. Even. Long. Full breaths. [pause about 5 seconds] Adjust the position. This time we're going to raise both legs **at the same time**. Completely exhale press the belly into the mat. Inhaling slowly raise both legs.

[ICE-Canada, Demonstrations, S2A-051 #236:1: A]

Example 5.39

A: No like there's always you know, you know there's the first time **and then after that** it sort of breaks the ice **and then after that** I feel a bit more comfortable.

[ICE-Canada, Face to face conversation, S1A-013 #267:2: A]

Moreover, one of the approaches used to organize the discourse is through focusing

the attention in the discourse, usually by giving, or requesting, an example. This has also

been expressed similarly, but not in exactly the same manner, in both varieties due to the

fact that more clusters which serve this pragmatic function have been found on the top 100

in Canadian English than Irish English. Some examples of these types of clusters used in

Canadian English are *sort of* (which can also act as a hedge with the pragmatic function of

downtoning the assertiveness of a segment in the discourse which has been found

equivalent to *kind of* in Irish English in terms of function), *you know like,* and *it's like*. In

Irish English, the pragmatic function of focusing the attention in the discourse is mainly

performed by *like* or some of the forms of *like* such as *like that* (*like what* can be classified

as a RT) (for more information about *like* in spoken English, see Carter & McCarthy 2006,

p.101; Kallen 2006; Schweinberger *et al* 2009; Lucek 2011; Amador-Moreno 2012, 2015;

Diskin 2017; and Schweinberger 2020). Also, one of the pragmatic functions (found as a

hedge) of *like* is to act as a mitigator in clause-final position, and this use is a variety-

specific use of *like* in Irish English (see Chapter 8 for more on this). Table 5.13 below

summarizes the forms appearing more frequently in both varieties for this pragmatic

function.

Table 5.13: *Discourse markers focusing attention by providing examples in ICE-Ireland and ICE-Canada*

Pragmatic function	ICE-Ireland	ICE-Canada
To focus attention by giving or requesting an example	Like Like that Like what	Like Sort of You know like

Additionally, here are some examples (5.40-5.43) extracted from the ICE-Canada and ICE-

Ireland corpora illustrating the performance of this pragmatic function shown by some of

the forms mentioned above.

Example 5.40

A: I mean if I could maybe do afternoons in the Blue Note. But then you see that would clash with my schedule in college, but I don't want any late nights **like,** working. You know the late nights, I'm going to spend on going out.

[ICE-Ireland, Face to face conversation, S1A-057$A]

Example 5.41

A: Then we'll get uhm a marker, and uh we draw a shape which rather resembles a big tongue. So uh we draw round **like that**, and uh then we get our scissors and we cut it out. Okay and uh cut around the end…

[ICE-Ireland, Demonstrations, S2A-058$A]

The cluster *like that* can be used to focus attention by giving an example verbally or like in this case, physically.

Example 5.42

A: And last week they had to write ah some **sort of** a a short essay about this creative person. And this week they had to do something artistic that represented the activities of the creative person.

[ICE-Canada, Face to face conversation, S1A-059 #158:1: A]

Example 5.43

A: Uh what do you call them like ah "aparro?" [nonsense word]? Uh, what do you call uh?
B: A cider? No.
A: Uhm **you know like** a Cinzano or port or something. What's that? **You know like** when you have a drink before dinner?

[ICE-Canada, Face to face conversation, S1A-063 #306:1: A]

Furthermore, one of the most crucial pragmatic functions by which the discourse can be organized is marking boundaries and linking segments of the topic in the discourse. This is done by "indicating the beginning or end of a topic or a transition from one topic or bit of business to another" (Carter & McCarthy 2006, p.218). In other words, there are some DMs used in the discourse to generate other segments of the topic in the discourse based on the boundaries and the linking of segments of the topic. This pragmatic function has been expressed and achieved in both varieties of English almost in the same way with a small degree of distinction with the interchangeable form of *ya* (which is only found in

Canadian English) and *yeah* (which is found in both Irish English and Canadian English) (note that *ya* and *yeah* are interchangeable only in Canadian English because both forms are found in Canadian English as illustrated in Table 5.14 and examples 5.44- 5.51).

Table 5.14: *The discourse markers marking boundaries and linking segments of the topic in the discourse*

Pragmatic function	ICE-Ireland	ICE-Canada
To mark boundaries and linking segments of the topic in the discourse.	yeah so right okay oh well	yeah and so right okay well ya

Here are some examples extracted from both datasets demonstrating the pragmatic function of marking boundaries and linking segments of topics in the discourse:

Example 5.44

A: I think we genuinely had a nice time.
B: Ya. Okay good.
A: I I usually feel it, you know. I don't even feel like Tommy wanted to rush out and leave, you know?
B: **Ya,** as I said, usually when we go to his house he doesn't even talk to…
                              [ICE-Canada, Face to face conversation, S1A-076 #315:2: B]

Example 5.45

A: And when Rick took that knife it was in the sheath?
B: When Rick took the knife, **yeah** it had a sheath.
A: And how does a sheath for a knife work? Does that go on the, on a belt somewhere?
B: Well it's just like a, like a, all sorts of knives are like that okay. It doesn't matter where you put the sheath.
                              [ICE-Canada, Legal cross-examinations, S1B-063 #144:1: B]

Example 5.46

A: Just like you, only wearing Escape for women instead.
B: **Yeah**. Is that, is that nice? Is that nice, all that stuff or would you know?
A: What? Escape for women?
B: Mm.
A: It's kind of old-womany. It's not that nice.

B: Yeah. That wouldn't be nice at all really.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-100$A]</div>

Example 5.47

A: No, but they used their names there.
B: Okay.
A: **And so, and so** Lisa finally proved it and then they they, so they brought him to court and uh so they had this scene from A Few Good Men where they hafta ^have to^ [takes a breath] where they ask him, the lawyer's like [laughs] "did you rig the election?" and he goes "no." [laughs]

<div align="right">[ICE-Canada, Face to face conversation, S1A-092 #76:1: A]</div>

Example 5.48

A: And were you in college today?
B: Yeah.
A: Oh **so** she got up?+ She said she was up in college about ten past ten.
C: She was up.+

<div align="right">[ICE-Ireland, Face to face conversation, S1A-068$A]</div>

The DM *so* or (*and so* which occurs more in Canadian English than in Irish English) is

often used with the pragmatic function of connecting what has been said earlier in the

conversation or the discourse.

Example 5.49

A: We'll just cut into it so we get an opening. Open it out and we keep a little piece there and then we'll get some glue and we put some glue on the end of our centre piece and then get the black and we start to roll around like that. Okay. And uh when you come to the very end then just get another dab of glue and stick it like that there. Alright. Now, at this stage we go back to our petals. And as I said we have a right bunch of them there. And we take each individual petal in our hand and we put our two thumbs in the centre and we just stretch the paper slightly. Do the same with this one. Stretch and stretch. And you do that altogether with about twelve petals. **So** all we do then is put some glue on the end of each petal. Glue it on to the centre. Let it set. Get some more glue on to the petal and glue it on there and let it set.

<div align="right">[ICE-Ireland, Demonstrations, S2A-058$A]</div>

In addition to the aforementioned function of *so*, this DM can be used with the pragmatic

function of summarising what has been said previously in the conversation or the discourse.

These sub-pragmatic functions can be under the general categorization of linking segments

of the topic in the discourse.

Now, look at how *well* as a single form and cluster is used to mark boundaries and link segments of the topic in the discourse as illustrated in examples 5.50 and 5.51.

Example 5.50

A: Ah there'll probably be a fair few people around or, I wonder will Eithne feel she has to go out?
B: Probably not. It's on at seven.
A: **Oh well** she's going to the gym at seven… or half seven+
B: Well+ that finishes at, well that doesn't finish until, uhm about half eight
A: Well she obviously isn't planning on going but somebody will probably tell her today about it.

[ICE-Ireland, Face to face conversation, S1A-060$B]

Example 5.51 [Two siblings are discussing what to get their mom as a gift].

A: So we can get her everything
B: Yeah. **Well**, why don't we get her then either the Le Poison or the L'Air du temps and dad'll get her the other one.

[ICE-Canada, Face to face conversation, S1A-096 #364:2: A]

Based on the qualitative analysis, it has been noticed that the single-word DM *well* has been used much more in Canadian English with the same meaning of the single-word DM *so* and we can thus say that they are interchangeable in relation to the pragmatic function of marking boundaries and linking segments of the topic in the discourse (refer to section 5.4.1 and Table 5.2, as well as section 5.4.2.1 and Table 5.10 for more on *well*).

Lastly, in relation to organising the discourse, sometimes speakers tend to mark emphasis on a segment of the topic in the discourse for the sake of showing the importance of what has been said or the importance of what will be said by preparing the listener to that by paying more attention. This goes back to the organization of ideas in the discourse and how they should be said in the discourse which surely has an impact on how they will be perceived and received by the listener. Therefore, it has been found that Irish English uses a wider range of forms and clusters in order to mark (or wave away) emphasis on a segment of the topic in the discourse, whereas in Canadian English, it occurs less often. In Irish

162

English, we see that the pragmatic marker *just* can act more as a hedge and a DM when occurring by itself or with the cluster *it's just* whereas in Canadian English, the cluster *I'd like to* serves that pragmatic function more, as well as in Irish English, by marking emphasis on a segment of a topic in the discourse for the listener to pay more attention to what is coming next (see Carter & McCarthy 2006, p.98 for more information about *just* in spoken English). See the examples 5.52- 5.55 below for more illustration on how this is used practically.

Example 5.52

A: Uhm say for example uhm a person where one of the spouses is working and they're **just** not eligible for unemployment assistance.
                                            [ICE-Ireland, Legal presentations, S2A-066$B]

Example 5.53

A: There's no variety and nobody really wants to talk and everybody's just drinking and dancing and posing and everything. **It's just** no good.
B: So, you don't really go out any more down in Cork?
                                            [ICE-Ireland, Face to face conversation, S1A-057$A]

Example 5.54

A: We had a couple of rules which helped us along but we did need some thinking as well. Now in practice if you have a large and complicated graph you very often will not be able to, or will not want to, find out the chromatic number. **It's just** too complicated a problem to colour the graph with with great effort using the minimum number of colours.
                                            [ICE-Ireland, Unscripted speeches, S2A-037$A]

Example 5.55

A: There comes a point where everything you touch becomes cute when Rachel seems, +ah no, cute describes you.
B: Ah no it's a lovely+ dress.
A: **It's just** short.
C: What colour is it?
                                            [ICE-Ireland, Face to face conversation, S1A-058$C]

The DM *just* in Example 5.55 has been used to take the emphasis and the attention away from the comment.

On the other hand, we have examples 5.56 and 5.57 in which the three-word cluster *I'd like to* is used to mark emphasis on a segment of a topic in the discourse for the listener to pay more attention to what is coming next. This type of usage with *I'd like to* can be found more apparently in politicians' speech in which politicians bring emphasis through their meta statements (as shown in the following examples 5.56 and 5.57).

Example 5.56

A: It's James Casey Caherciveen. **I'd like to say** that it it's quite obvious that we need the immediate implementation of a Freedom of Information Act and a referendum on cabinet confidentiality.

[ICE-Ireland, Broadcast discussions S1B-033$C]

Example 5.57

A: And I want to ask the minister to be clear with Canadians. His policy clearly follows the Tory Royal Commission on transport. And **I'd like to ask** the minister to explain how his plan to commercialize is different from the Tory's plan to privatize which he opposed so vehemently in the last Parliament.

[ICE-Canada, Parliamentary debates, S1B-053 #60:1: I]

5.4.2.4 Other pragmatic functions found in ICE-Ireland and ICE-Canada

Irish English and Canadian English have performed some specific pragmatic functions which have been hard to categorise (in terms of managing, organising, or monitoring the discourse) due to interpersonal/textual meanings some of which can reflect the cultural differences (Brinton 1996; Halliday 2002). They are also difficult to categorise due to the lack of previous work focusing on bigger clusters. The present study examines clusters as seen below, such as *I'm sure; well I don't know; at the end of the day;* and *you know what I mean like*). In addition, it has been found that the Irish English data contains more instances of sarcasm and humour than the Canadian English data, so there seems to be a tendency in Irish English towards using these devices in a humorous or joking manner. This is seen through a wide range of DMs which appear on the top 100 such as *well* (as shown in

Example 5.14) and *I'm sure* (as shown below in Example 5.58) as well as others such as *by*

the way.

Example 5.58

A: How do you mean he was revealing? What'd did he do?
B: On his thickness, you know?!
A: On his thickness! [laughs]
C: They say he's a very bright honest man really.
A: Oh **I'm sure** he's honest. Honest yeah yeah.
B: Well he certainly isn't bright but he's not a good politician though.

[ICE-Ireland, Face to face conversation, S1A-073$C]

Moreover, both varieties of English have presented a wide range of DMs acting as fillers or

hesitation markers indicating various pragmatic functions in the discourse. For example, the

DMs *uh* and *uhm* perform the pragmatic function of signaling to others that in some way

the speaker is having trouble communicating what they want to say (see the examples 5.59-

5.61 below demonstrating that type of pragmatic function).

Example 5.59

A: Chairman, just following on from what Deputy Bruton has said, **uh** I'd like to also
      draw attention to the whole **uh** evaluation of **uh** the various **uh** Fa/s schemes and
      training and education uh generally. A~ as has been stated we've had the National
      Economic and Social Forum **uh** report which examined this issue **uh** or should I say
      we've had **uh** newspaper reports **uh** of their evaluation of **uh** these various
      schemes. **uh** Unfortunately the paper itself **uh** which dealt with this issue is is is not
      generally available to the public or indeed to Oireachtas members **uh** because I s~ I
      sought it but I understand that once the final report of the NESF is out, **uh** that this
      paper then **uh** could be made available but according to newspaper reports anyway
      uh there was **uh** serious **uh** concern indeed about the **uh** benefit of **uh** some of the
      training **uh** which we are undertaking at the moment **uh** particularly **uh** through
      Fa/s. **uh** And **uh** I think that was probably **uhm** the start of of this **uh** general
      debate. **uh** I note also **uh** Minister that you accepted a report from IBEC yesterday
      uh dealing with this same issue and maybe you might use this occasion to **uh**
      respond to their concerns. **uh** I gather that **uh** IBEC believe that **uh** more training
      and education should be **uh** given to people already in employment or certainly that
      seemed to be a a a summary of what the the newspaper said today **uh** in relation to
      to that report.

[ICE-Ireland, Legal presentations, S2A-066$B]

Example 5.60

A: But **uh** she decided that Sylvie decided that she was **uh** going to take languages and
        uh she specialized in **uh** Mandarin. And **uh** speaks it fluently.
                                    [ICE-Canada, Face to face conversation, S1A-058 #13:1: B]

Example 5.61

A: We have to register, **uhm** on January second for our second semester.
                                    [ICE-Canada, Face to face conversation, S1A-099 #243:1: B]

In terms of the DMs serving as fillers marking hesitation and pauses in the speech, it

has been found that the two-word DM *you know* which has been known with the pragmatic

function of monitoring the ongoing discourse by monitoring the state of shared knowledge

in the conversation (according to Carter & McCarthy 2006), occurs often in Canadian

English as a filler marking hesitation and pauses in the speech, whereas in Irish English, the

discourse marker *well* and *well I don't know* occur more often than in Canadian English

where the speaker shows hesitation (see examples 5.62 and 5.63 below).

Example 5.62

A: And uhm, these occurrences of men dying in the household and and and no record
        of, **you know**, why except that they just had **you know** a a little s- smell and and
        went away.
                                    [ICE-Canada, Face to face conversation, S1A-031 #110:1: B]

Example 5.63

A: Nineteen? Oh! Not that much younger than you, is he? **Well I don't know** well
        how old you are so
B: I'm twenty-two. He's nineteen.
                                    [ICE-Ireland, Face to face conversation, S1A-062$B]

Also, we have the six-word DM, *at the end of the day,* which appears on the top 100 in

Irish English but not in Canadian English which performs the pragmatic function of

marking that everything has been taken into consideration (see example 5.64 below).

Example 5.64

A: Even though they are a bit foggied, they are confused by the by the bewildering statements coming out nevertheless there are better issues or bigger issues involved in this. **At the end of the day**, I do hope we'll have a more democratic society.

[ICE-Ireland, Broadcast discussions, S1B-033$E]

Lastly, we have the pragmatic function of organising the narrative in which the speaker marks out examples of behaviours in narratives and it has been presented almost equally by the single-word DM *like* in both varieties. However, Canadian English shows more associated clusters with the DM *like* which do have some pragmatic functions in the discourse (i.e. *it's like*, *you know like*, and *you know what I mean like* are all found in the top 100), see the examples 5.65-5.67 below.

Example 5.65

A: He stares at you for a minute. And then **like** he eventually decides oh well I better give her her change like. And he's an hour then looking for change.

[ICE-Ireland, Face to face conversation, S1A-087$C]

Example 5.66

A: And I'm going down there, cos she needs me. Like basically since last week she's been **like** crying on the phone.
B: Right.

[ICE-Canada, Face to face conversation, S1A-087 #314:1: B]

Example 5.67

A: But **you know what I mean like** when, when there's like two chromosomes on the board and the guy goes "and how many chromosomes do we have up here?" [laughs]
B: Yeah exactly.

[ICE-Canada, Face to face conversation, S1A-092 #177:1: A]

In relation to qualifying a preceding statement, *like* can be placed in an end position in order to serve this kind of pragmatic function (see examples 5.68 and 5.69 below for more illustration) (Carter & McCarthy 2006, p.101; see also section 8.4.1).

Example 5.68

A: I don't balance out my diet at all no. But uhm there's nothing wrong with me. Like I'm, I'm not anaemic. Yet but I'm not anaemic or anything like that. I'm totally healthy **like.**

[ICE-Ireland, Face to face conversation, S1A-080$A]

Example 5.69

A: I mean you will have learned more in four years than you would have in two.
B: That's true.
A: **You know what I mean like?**
B: Mm.

[ICE-Canada, Face to face conversation, S1A-033 #229:1: A]

5.5 Conclusion

In this chapter, we have presented the analysis of DM use in Irish and Canadian Englishes in terms of forms and functions that have been captured in ICE-Ireland and ICE-Canada corpora. Methodologically, word and cluster lists were generated for both corpora to identify and compare the forms used in both datasets, and from these lists, DMs with pragmatic functions were identified manually by cross-checking qualitatively using concordancing. This results in interesting quantitative findings at the form level in which Irish English and Canadian English match the most at the single-word discourse markers level with some prominent DMs appearing more in one variety than the other such as *okay*, *now*, and *ya*. Both varieties start to contrast noticeably at the two-word discourse markers level and onwards. In terms of the qualitative findings, both varieties have displayed common pragmatic functions in relation to monitoring, organising, and managing the discourse. However, both varieties have shown clear contrasts in terms of functions and the forms that are typically deployed for these functions. For example, the DM *now* has been found to be in use much more in Irish English generally, and specifically in opening up and closing down the discourse, than in Canadian English. Similarly, the DM *okay* occurs twice

as much in Canadian English as in Irish English with wider distribution, especially in terms of the pragmatic function of opening up and closing down the discourse. Also, Canadian English acquires more clusters with *so* than in Irish English which perform distinctive pragmatic functions as in the pragmatic marker *or so*. On the other hand, Irish English and Canadian English have shown many similarities in regards to pragmatic functions. For example, the use of the DM *I mean* has been displayed in both varieties almost in the same way, not only in relation to function but also its occurrences in both datasets. Lastly, this chapter has revealed the pragmatics of both varieties of English which demonstrate difference and similarity in Irish English and Canadian English in relation to pragmatic functions and forms in their use of DMs. Now, in the next chapter, we will explore how response tokens are used in Irish and Canadian Englishes.

CHAPTER 6

RESPONSE TOKENS IN IRISH AND CANADIAN ENGLISHES

6.1 Introduction

In this analysis chapter, I will present the analysis of response tokens (RTs) used in Irish

and Canadian Englishes which have been captured in the ICE-Ireland and ICE-Canada

corpora. This has been done in the same way as the analysis of discourse markers (DMs) in

Chapter 5, by detailing the main quantitative findings resulting from both datasets through

the theoretical and methodological framework taken in this study (as explained in Chapters

3 and 4) in order to unpack and analyse these findings qualitatively and comparatively (as

done in section 6.4). Thus, the analysis of the pragmatic variation that occurs in RT use

between Irish English and Canadian English is approached and investigated from two

perspectives: form and function (as demonstrated in section 6.3). RTs are one of the

fundamental linguistic items used intentionally, or unintentionally, in communication and

interaction and they perform various pragmatic functions in spoken discourse.

RTs can be defined generally in discourse such as vocal, verbal, or non-verbal items

such as *mm*, *yeah*, a head nod, and so forth, which demonstrate degrees of engaged

listenership without changing the speaker turn (O'Keeffe & Adolphs 2008; also see section

2.3.2 for more information). Based on this definition (and others introduced in section

6.2.1), the separation has been made between "discourse markers" and "response tokens" in

terms of categorization. This is due to two main distinctions found between DMs and RTs.

The first is that DMs focus on and are more concerned with, monitoring, organising, and

managing the discourse as a whole whereas RTs focus on and are more concerned with the

listenership in the discourse so that the listener can be more engaging and effective. The

second distinction is that DMs can occur in the middle of the turn as the speaker is speaking

or as a response in which the speaker takes over or changes the speaker turn (Carter &

McCarthy 2006; O'Keeffe & Adolphs 2008). As a result, in this present study RTs have

been classified as a separate category from DMs which are both classified as pragmatic

markers.

As previously stated in Chapter 2, there is no general agreement on most of the

categorization of pragmatic markers. However, scholars seem to "live with" the lack of

consistency in nomenclate by being precise in their definitions, terms, and limitations

within their individual studies. Consequently, the term response tokens was given its own

categorization in the work of Carter and McCarthy (2006) following Gardner (2001); the

term response tokens was treated within and as a type of DM under the pragmatic markers

categorization as a whole (Carter & McCarthy 2006; Gardner 2001). Additionally, in

Longman Grammar of Spoken and Written English, the term response tokens has been

subdivided and organised with distinctive terms based on the pragmatic functions that they

play such as "response forms" which are more general, and "backchannels" which are more

specific, and "response elicitors" which include fixed tags (see Biber *et al* 1999, p.1089 for

more information). Thus, knowing that there is a difference in terminology regarding

pragmatic markers in terms of categorization and inclusion helps to choose and determine

the terminology which will be used consistently throughout the whole research (as

mentioned briefly in Chapter 1 and in details in Chapter 4 and Appendix A).

6.2 Previous Research

6.2.1 Response tokens

RTs reflect the notion that conversations actually contain listener responses, or signals

produced by the listener, in order to keep the conversation going efficiently. In other

words, according to Tottie (1991) these signals "grease the wheels of the conversation but

constitute no claim to take over the turn" (p.254). Thus, they can be short utterances and non-verbal surrogates (e.g. head nods) (see Fries 1952; Kendon 1967; Yngve 1970; Maynard, D. 1989, 1997; Maynard S. 1986, 1997; Tottie 1991; Drummond & Hopper 1993a, 1993b; McCarthy 2002; Gardner 2001; O'Keeffe & Adolphs 2008). Additionally, as some evidence shows and as Kendon (1967) suggests, RTs are seen as an accompaniment to a speaker and are signs of guidance upon which the speaker can rely in terms of how the message has been received, perceived, and understood by the listener (Kendon 1967; O'Keeffe & Adolphs 2008). Furthermore, RTs mark how the listeners indicate their involvement with what has been said and how they manage their own responses. As we discuss in section 6.2.3 below, RTs operate with different pragmatic functions; some of them are general and some are more specific such as acknowledgments (*mm, yeah*), negation (*definitely not*), premodification (*most definitely*), continuers (*huh, I see*), change-of-activity tokens (*alright*), assessments (*great, how interesting*), collaborative completions, and newsmaker-like objects (*really*) and so forth (Biber *et al* 1999; Gardner 2001; Carter & McCarthy 2006).

RTs have been introduced in research literature with a plethora of terms, often depending on discipline and definition, which set the limitations as previously mentioned. For example, the term backchannel has been introduced by Yngve (1970) to refer to the "short messages" received by the speaker while holding the floor and this term is commonly used by many researchers. Also, we have the term "minimal response" used by Fellegy (1995) as well as the broader term "listener response" used by Roger *et al* (1988). In this present research, we will use the term "response token" to refer to the items or markers used for the listenership activity and the term "listener response" as a general

term, if needed, to refer to the actual activity of listenership. This goes along with the definition of RTs adopted in this present study as items that fill response slots without taking over the speaker's turn (as mentioned in section 6.1 and Appendix A). Hence, in our analysis, the RTs that take over the speaker turn are not included or viewed as RTs, as RTs are the ones seen as turn yielding (as discussed in section 6.1).

At the discourse level, Mott and Petrie (1995) believe that RTs are the opposite of interruptions. On the other hand, Duncan and Niederehe (1974) find it difficult to draw the line between brief utterances and proper turns where the "listener" becomes the "speaker." However, this problem exists more for the analyst than for the participants in the conversation in which the latter is able to draw on clues (i.e. prosodic features, facial expressions, and gestures) indicating whether an interlocutor is attempting to take the floor or show listenership in a given context (O'Keeffe & Adolphs 2008). This confusion between brief utterances and proper turns can be measured by corpus linguistic tools which provide the analyst with the RTs occurring as or in clusters taken from the data and viewed linguistically as a fixed expression that cannot be taken apart. Also, making the distinction between the similar pragmatic markers that commonly overlap in terms of function and form (i.e. DMs vs RTs) would help the researcher to identify which is which.

6.2.2 Forms of response tokens

In this chapter we will compare and contrast RTs in terms of forms and functions in two varieties of spoken English: Irish and Canadian, using data from two sub-corpora of ICE: ICE-Ireland and ICE-Canada, which are well designed for a comparative study of spoken discourse because they both use the same design matrix (as detailed in Chapter 4). The

existing literature on RTs show that the forms of RTs can be divided into two types, or categories: minimal and non-minimal RTs (Zimmerman & West 1975; Fishman 1978; Schegloff 1982; Maynard, D. 1989, 1997; Maynard S. 1986, 1997; Fellegy 1995; Gardner 1997, 1998, 2001; Tottie 1991; McCarthy & Carter 2000; McCarthy 2002). The distinction between minimal RTs and non-minimal RTs is that the minimal responses have been described as short utterances such as *yeah* or non-word vocalisations such as *mm* and *umhum* while non-minimal RTs are mostly from the word classes of adverbs or adjectives functioning as pragmatic markers (such as *good*, *really great*, and *absolutely*). Also, non-minimal RTs can be short phrases or minimal clauses such as: *you're not serious*, *is that so*, *by all means*, *fair enough*, *that's true*, and *not at all*. Yet, that distinction is not necessarily clear cut, especially given that both minimal and non-minimal RTs can occur in pairs or clusters (McCarthy 2002). Also, non-verbal RTs such as head nods and shoulder shrugs cannot usually be captured while using a corpus of transcribed recordings (O'Keeffe & Adolphs 2008). Nevertheless, in this present research we have looked at any item occurring within the top 100 most frequent word lists and cluster lists which fill the response slots with the condition of not taking over the speaker's turn. Therefore, non-verbal responses are not considered in this analysis because both corpora were collected as audio recordings and unfortunately this precludes any focus on non-verbal items.

6.2.3 Functions of response tokens

As discussed in Chapter 2, section 2.3.2 and shown in Appendix A illustrating the top-down framework for the analysis of the forms and pragmatic functions of spoken grammar,

RTs perform numerous pragmatic functions in the discourse such as (and some of these overlap):

- Making arrangements and reaching decisions which can be displayed, for example, with the form *fine* (Carter & McCarthy 2006).

- Replying to a request for a service or favour which can be performed with the forms *certainly* and *definitely* (Carter & McCarthy 2006).

- Offering positive feedback to the speaker and often marking the boundaries of topics where speakers show their satisfaction which can be expressed with a wide range of RTs such as *excellent*, *fine*, *great*, *good*, *lovely*, *right*, and *perfect* (Carter & McCarthy 2006).

- Acting as continuers which hand the floor back to the immediately prior speaker which can be demonstrated with *mm*, *hm*, *uh*, *huh*, and *I see* (Gardner 2001).

- Acknowledgments, which claim agreement or understanding of the prior turn (i.e. *mm* and *yeah*) (Gardner 2001).

- Change-of-activity tokens, which mark a transition to a new activity or a new topic in the talk which can be voiced with the forms *okay* and *alright* (Gardner 2001).

- Assessments, which evaluate the speech of the prior speakers and can be expressed with a wide range of RTs such as *great*, *how interesting*, and *what a load of rubbish* (Gardner 2001).

- For clarification or other types of repair which seek to clarify mishearing or misunderstandings. This is usually done by brief questions which can be expressed with one single-word response tokens (*who*) or clusters (*which book do you mean*) or the very generalised (*huh*) (Gardner 2001).

- Newsmarkers and newsmarker-like objects marking the prior speaker's turn as newsworthy in some way which can be expressed with *really*, the change-of-state token *oh*, and the idea-connector *right* (Gardner 2001).

- Indicating further emphasis in response. This is done through premodification by intensifying adverbs (i.e. *jolly good*, *most definitely*) and negation by adding not (i.e. *absolutely not*, *definitely not*) (Carter & McCarthy 2006).

There are also many other pragmatic functions played in the discourse by RTs which are strongly associated with their particular contexts which "refer to a whole preceding utterance rather than their word-class identity as adjectives or adverbs" (Carter & McCarthy 2006, p.189). They also often occur in pairs for more emphasis. This can be performed by a large number of RTs which vary in terms of forms as they vary in terms of functions such as *indeed*, *yeah exactly*, *possibly*, *precisely*, *yeah definitely*, *absolutely*, *brilliant*, *fabulous*, *cheers*, *bye*, *that's excellent*, *thank you very much*, *thank you so much*, *thanks*, *wonderful*, *is that so*?, *by all means*, *fair enough*, *not at all*, *true enough*, *of course*, and *what a pity*! (Carter & McCarthy 2006). O'Keeffe and Adolphs (2008) analysed the forms and functions of RTs across British and Irish English and concluded that there are four main, or general, pragmatic functions performed by RTs in the discourse under which all the other sub-functions can be covered. These are continuer RTs, convergence RTs, engagement tokens, and information receipt tokens (see Chapter 2, section 2.3.2.1 for more details as well as section 6.4.2). Additionally, there are some RTs formed or viewed as other pragmatic markers with the pragmatic function of a RT in terms of the way they fill the response slots, occur initially, and do not take the floor from the speaker. This can be seen with interjections (which have been classified as a separate category from RTs, see

Carter & McCarthy 2006, pp.208-224) which express positive or negative emotional reactions to any segment in the discourse; social routines markers such as *thanks* or *thank you*; questions and tags, see the examples 6.1 and 6.2 below with *oh* which occurs often in both datasets as a RT (based on the definition adopted in this research) and an interjection simultaneously which reflects on the idea mentioned above on how the pragmatic functions of RTs are strongly associated with their particular contexts which "refer to a whole preceding utterance rather than their word-class identity as adjectives or adverbs" (Carter & McCarthy 2006, p.189).

Example 6.1

A: They got this crane right, and it was kind of like Santa arriving in Brown Thomas. Because now you're never gonna guess Dad, Switzer's window is no more.
B: **Oh?**
A: Ah it's, it's horrible.
B: Switzer's window isn't any more?
A: No, because you know they were renovating it, or they were redoing it, and they can't do the windows.

[ICE-Ireland, Face to face conversation, S1A-074$B]

Example 6.2

A: Why, have you made your dinner?
B: No. I have- I told you, I have to go get milk.
A: Well they're eating dinner right now.
B: **Oh**? Well it doesn't matter. I'll like, do you know when they'll be back?
A: From where?
B: You just finished saying that's ^that^ they're not there!
A: They are here.
B: Oh my god.

[ICE-Canada, Telephone conversation, S1A-098 #195:1: A]

6.3 Methodology

This section outlines the main framework of analysis for this chapter. As explained in prior chapters (Chapter 3 and Chapter 5, section 5.3), the current study makes use of the formal level of pragmatic analysis; so, I will be looking at the form and function of RTs in Irish

and Canadian Englishes. Methodologically, corpus linguistics is used as a tool in order to conduct this research using the iterative approach in which top-down and bottom-up processes are used to compile RTs and their pragmatic functions in the discourse. The top-down analysis was based on a framework for spoken grammar generally, and RTs specifically, based on the existing literature while the bottom-up process was based on micro-analysis of the data.

In terms of form, the quantitative findings of RTs across both varieties (Irish English and Canadian English) is presented comparatively based on their word clusters (from one single-word form to six-word clusters form). As a result, as elaborated upon in Chapter 4, wordlists and cluster analyses were generated to identify and compare the RTs used in both datasets (ICE-Ireland and ICE-Canada). Within the form analysis, we will also present the RTs across both varieties which appear as questions and tags. However, these will not be discussed in this chapter but will be elaborated on in Chapter 7 which is dedicated to the analysis of the forms and functions of questions and tags. This is because questions and tags display a variety of pragmatic functions in spoken grammar found across different categories (i.e. they can be DMs or RTs). However, in the selection process, RTs are classified as items that fill a response slot but do not take over the speaker's turn (as stated earlier in section 6.2.1). Therefore, in our analysis, pragmatic markers or "response tokens" that form part of a turn are not included as RTs. This marks the distinction between DMs and RTs in this present research (see section 5.4.2 and Table 5.9 in Chapter 5, which provides a brief overview on the pragmatic functions of DMs in terms of opening up, closing down, resuming, shifting, and diverting the topic of the discourse which can be similar to the pragmatic functions of RTs). For further illustration of this point, see

examples 6.3 and 6.4 from ICE Canada. In these two examples, *okay* was counted as a RT.

However, examples 6.5 and 6.6 were not counted as RTs but instead were considered as

DMs because they take over the speaker turn. We note that this decision is in opposition to

some studies where such instances (as examples 6.5 and 6.6) are considered RTs (see the

definition of RTs adopted by McCarthy 2002 and this study as mentioned in section 6.2.1).

Example 6.3

A: Uhm ya, I'll talk to you in about a week or so and let you how I'm doing.
B: **Okay.**
                                        [ICE-Canada, Classroom lessons, S1B-008 #238:1: A]

Example 6.4

A: No, but they used their names there.
B: **Okay.**
A: And so, And so, Lisa finally proved it and then they, they, so they brought him to
        court, and uh, so they had this scene from A Few Good Men where they hafta ^have
        to^ [breath] where they ask him, the lawyer's like, [laughs] "Did you rig the
        election?" and he goes, "No." [laughs]
                                        [ICE-Canada, Telephone conversation, S1A-092 #75:1: B]

Example 6.5

A: So Granny said well **okay** let me help you out a little bit and that way we can
        invite them.
B: **Okay** so you're happy for that helth ^help^? You don't find it an intrusion?
                                        [ICE-Canada, XXX (Unknown), #212:2: A]

Example 6.6

A: Remind me because sometimes I, I tend to forget.
B: Right, right.
A: **Okay,** how are things going?
B: Well I just came back on Monday
                                        [ICE-Canada, Classroom lessons, S1B-007 #132:1: A]

In terms of function, the quantitative findings of RT forms will be unpacked and

analysed qualitatively by concordancing them manually (as discussed in Chapter 4) in order

to identify and determine their pragmatic functions in the discourse. This has been mainly

approached by the form-to-function approach (Aijmer 2018) with the help of function-to-

form approach (O'Keeffe 2018) which both present the iterative approach (which has been

discussed here briefly and in details in Chapter 4, also see Appendix A) (Jucker *et al* 2018).

6.4 Results and Analysis

6.4.1 Forms

Tables 6.1 to 6.4 below provide a summary of the comparative analysis of RT forms in

ICE-Ireland and ICE-Canada which demonstrates where Canadian English and Irish

English converge and diverge in terms of forms. These tables below have been derived by

the qualitative analysis (from a bottom up route) of the top 100 most frequent word and

cluster lists in which each single word and cluster has been 181ay a181dance manually in

order to affirm which items from the quantitative findings resulted from the top 100 most

frequent word and cluster lists (presented in Appendices B, C, D, E, F, G, H, I, J, K, L, M)

are actually acting like RTs with their pragmatic functions in both datasets. Note that the

light gray shaded cells in the tables below indicate the distinctive forms (response tokens)

between Irish English and Canadian English, and the unshaded cells indicate the forms

(response tokens) that both varieties of English share in common.

Table 6.1. *ICE-Ireland and ICE-Canada single-word response tokens occurring within the top 100 most frequent word lists (frequency per 10,000 words)*

ICE-Ireland	Freq[4]	ICE-Canada	Freq
yeah	82	yeah	23
no	43	no	40
right	19	right	28
really	19	really	25
okay	16	okay	31
uh	86	uh	106
what	53	what	55
uhm	51	uhm	45

---

[4] The order of the forms presented in Tables 6.1- 6.4 is random and does not indicate rank.

well	39	well	43
oh	35	oh	28
when	21	when	24
who	23	who	19
which	23	so	70
		ya	54
		hmm	30
		mm	32
		Because	22

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 6.2. *ICE-Ireland and ICE-Canada two-word response tokens occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
yeah yeah	14	ya ya	7
that's right	2	that's right	3
I know	7	I know	7
do you?	13	do you?	8
did you?	6	you can	10
are you	6	sort of	9
is it?	7	and then	13
		so I	8
		I didn't	4
		You know	47
		mm mm	0.44
		mm hmm	23

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 6.3. *ICE-Ireland and ICE-Canada three-word response tokens occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I don't know	8	I don't know	8
no I don't	1	no I don't	1
yeah that's right	0.47	you don't know	1
that's right yeah	0.47	I'd like to	1
yeah yeah yeah	4	But I don't	1
		mm hmm mm	1
		hmm mm hmm	1

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 6.4. *ICE-Ireland and ICE-Canada four-word response tokens occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I don't think so	0.22	I don't think so	0.31
well I don't know	0.22	well I don't know	0.34
yeah I don't know	0.18	ya I don't know	0.13
so I don't know	0.37	that's right that's right	0.26
it's not too bad	0.18	mm hmm you know	1
I know I know	0.44	thank you very much	1
it's not it's not	0.12	mm hmm mm hmm	4
yeah yeah yeah yeah	2	hmm mm hmm mm	1

Note: There are no lexicalised response tokens occurring as five and six-word response tokens. They occur only by the duplication of *yeah* and *I know* which appeared as five and six-word clusters in Irish English and a duplication of *mm* and *hmm* has appeared as five and six-word clusters in Canadian English (see Table 6.5). Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

As seen above in Tables 6.1 to 6.4, the comparative analysis of RTs has been conducted at the form level in which we can see the similarities and the differences between Irish and Canadian Englishes in relation to their RT forms and their rate of occurrences (which will be analysed below). As seen in the previous chapter regarding DMs, Irish English and Canadian English demonstrate the most resemblance at the single-word level and after that the contrasts in relation to forms begin to appear more clearly. Based on the RTs which have occurred in both varieties on the top 100, we see that Canadian English tends to use non-word vocalisations such as *hmm* and *mm* with their duplications. These minimal non-word vocalisation items can perform various pragmatic functions such as: continuers, which hand the floor back to the immediately prior speaker; or hesitation markers, which are more likely to be associated with head nods (O'Keeffe & Adolphs 2008). This form (reduplicated minimal non-word vocalisation items) is not found in Irish English. Instead, Irish English uses the short utterance of *yeah* much more (even in clusters) than is found in Canadian English in general (as RTs or DMs). See Table 6.5

below illustrating the duplications of *yeah*, *hmm*, and *mm*, and other reduplication forms found within the top 100 most frequent cluster lists.

Table 6.5. *ICE-Ireland and ICE-Canada response tokens with reduplication forms found within the top 100 most frequent cluster lists*

ICE-Ireland	ICE-Canada
yeah yeah	mm hmm
yeah yeah yeah	mm hmm mm
yeah yeah yeah yeah	hmm mm hmm
I know I know	mm hmm mm hmm
yeah yeah yeah yeah yeah	hmm mm hmm mm
yeah yeah yeah yeah yeah yeah	That's right that's right
I know I know I know	mm hmm you know
yeah yeah yeah yeah yeah mm	mm hmm mm hmm mm
	hmm mm hmm mm hmm
	mm hmm mm hmm so
	mm hmm mm hmm mm hmm
	hmm mm hmm mm hmm mm

Table 6.5 clearly shows the contrasts in the reduplication of forms in both datasets presenting Irish English and Canadian English in relation to RTs. Not only that, but also the reduplication of forms (extracted from both varieties) indicates the pragmatic functions being used more in one variety than the other. Here we see the pragmatic function of acknowledgments (found more in Irish English) versus continuers (found more in Canadian English) (see the examples 6.7 and 6.8 below). However, both of these forms with their duplication are interchangeable with their pragmatic functions (for example, *yeah* has been proven to be a continuer RT as well, as elaborated and shown in sections 6.4.2.1 and 6.4.2.2).

Example 6.7
A: They're, do you know where uhm…
B: Oh 'tis you're related to Eithne O'Byrne's is it?
C: **Yeah yeah yeah yeah yeah.**

B: Oh yeah, I was going to say while ago how did you know Eithne O'Byrne? Do you
        know?
C: Yeah, I met her last weekend.
                                [ICE-Ireland, Face to face conversation, S1A-090$D]

    Example 6.8

A: They're not, nuh- virtually none of them are helping me with what I'm, with what
        my field is, except because I make them.
B: **Mm hmm. Mm hmm.**
A: You know?
B: Mhmm ya, well…
A: And some of them I can't.
                                [ICE-Canada, Face to face conversation, S1A-037 #77:1: B]

        Also, in terms of comparison at the level of forms, we can see again (as noticed in

the previous chapter) that *so* and *ya* are prominent markers, not only as DMs, but also as

RTs at the single and two-word level in which *so* has been found quantitively to be in use

more in Canadian English (even in clusters such as *so I*) and qualitatively to have more

distribution as a RT, see Figure 6.1 and the examples 6.9 and 6.10 below.
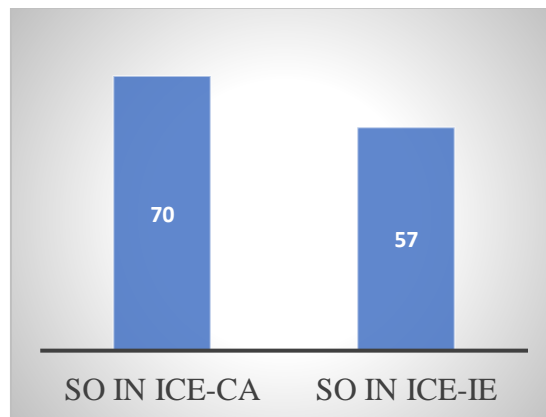
Figure 6.1 The frequency of *so* (per 10,000)

    Example 6.9

A: So uhm that's probably one of the reasons that he's interested in, in Innes.
B: Hmm.
A: **So**?
B: So uh let's talk the bucks! [laughs]
                                [ICE-Canada, Face to face conversation, S1A-024 #37:1: A]

Example 6.10

A:　　That's the problem with my letters?
B:　　That's the problem.
A:　　**So**?
B:　　Different [unclear word] for different stuff.

<div align="right">[ICE-Canada, Telephone conversation, S1A-096 #189:1: A]</div>

At the two-word RTs level, it is very recognizable that Irish English adopts much more RTs in the form of questions and tags. This will be elaborated upon further in Chapter 7 which highlights the differences and the similarities of Irish English and Canadian English in terms of their questions and tags and their pragmatic functions in spoken grammar. However, we can see that Irish English displays a larger number of questions and tags such as *do you? Did you? Are you?* And *is it?,* whereas in Canadian English only *do you?* Is found within the top 100 most frequent clusters which both varieties share. See examples 6.11- 6.15 below for illustration.

　　　Example 6.11

A:　　No they wouldn't get away with that kind of thing. They wouldn't be getting away with it I mean. They… they wouldn't be getting away with all the appearances and all the… you know all the uh kind of show and everything. They're definitely married. I'd say so.
B:　　Maybe, maybe. I wrote a letter to Simon today for Christmas.
A:　　**Did you?**
B:　　Because I don't think he'll be over before. Would you consider writing? He's been very good to you.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-059$B]</div>

　　　Example 6.12

A:　　But you're not going to?
B:　　I don't know? I'm half thinking. But…
A:　　**Are you?**
B:　　I'm half thinking and yet I'm not. I mean when I go out with Ciaran I, I say no way definitely not. And then when I'm out with Jim, I say definitely yes. You know that way?

<div align="right">[ICE-Ireland, Face to face conversation, S1A-049$B]</div>

Example 6.13

A: I thought of it today when I was looking for am I got one of the counselling books.
       It's really good
B: **Is it?**
A: Nelson-Jones Richard.
B: Nelson-Jones.
A: None of the psychiatry books are in James'. That's great load of use to us.
                              [ICE-Ireland, Face to face conversation, S1A-050$B]

Example 6.14

A: No I wouldn't make, I've never [unclear speech] make it in five. On a Saturday of
       course it's, it's really quick because the traffic isn't you know. I'd be going in at
       maybe like, like half twelve and the traffic isn't +bad.
B: Yeah+ I go in then. Although sometimes
A: **Do you?**
C: When there's football it's bad enough sometimes.
B: Yeah
                              [ICE-Ireland, Face to face conversation, S1A-065$C]

Example 6.15

A: I have her first album.
B: **Do you?**
A: The pop one.
B: **Do you?**
A: Used to play it all the time.
B: Did you?
A: That was in my pop days, before I became cultured and listened to Gys- Gypsy
       music instead.
                              [ICE-Canada, Face to face conversation, S1A-074 #206:1: A]

Furthermore, at the two-word RTs level, we see that in Canadian English the two-word

cluster *you know* is not only appearing more as a DM (as stated in Chapter 5) but also, as a

RT as we can see in example 6.8 (above) and in example 6.16 below. Additionally, see

Figure 5.4 which shows the distinctive frequency of *you know* (per 10,000) across both

varieties with *you know* appearing much more in Canadian English with various pragmatic

functions.

Example 6.16

A: And if they do, well then that's uh competition for the type of school that we have. I still would rather send my daughter to an [indigenous speech] because I think it's a different experience and I think that she'll have more the type of experience that, that he was mentioning

B: Mm hmm.

A: **You know?**

B: That's right.

                              [ICE-Canada, Business transactions, S1B-071 #107:1: B]

Irish English and Canadian English have exhibited a wide range of RT forms, some of which they both share and more in which they contrast and differ (as illustrated above). What has been observed while conducting the qualitative and quantitative analysis for this chapter and the previous chapter is that Irish English tends to use more distinctive forms serving as either DMs or RTs at the two-word level, whereas in Canadian English, we often see identical forms serving as DMs in one context and as RTs in another context. This can be seen with pragmatic markers of *so*, *and then*, and *sort of*. At the single-word level, items are usually distributed across different pragmatic markers in which one single form can be a DM, RT (i.e. *right* and *okay*), or an interjection. An example of this is *oh* which was the only common interjection (that did not easily fit into the major word classes: noun, verb, adjective, adverb) shared by Irish and Canadian Englishes based on the datasets with the pragmatic function of expressing positive or negative emotional reactions to any segment of discourse (Carter & McCarthy 2006; also see Chapter 2 section 2.3.5). This utterance occurs more in Irish English than in Canadian English as can be observed in Table 6.1 for frequency, along with the examples (6.17- 6.22) provided below, and also examples 6.1 and 6.2 (above).

Example 6.17

A: Nah, they taste a little bit different I find.

B: **Oh.**

A: They're not as fine tasting.

B: Hmm.

A: Uhm, do you know what, uh, what Sandy's having at ^for^ dinner on uhm on Friday?

<div align="right">[ICE-Canada, Face to face conversation, S1A-013 #238:2: B]</div>

Example 6.18

A: I might be put in jail for doing illegal things [laughter] and I never do anything illegal or bold.

B: Do you not ever?

A: Never.

C: Ever?

B: +Ever, ever?

A: Ever, ever+… ever.

C: Ever ever ever? Aisling would you take that book out of the library? [laughter]

A: Yeah, Aisling was actually thinking of doing something very bad.

B: But I didn't do it. Actually I thought of it alright again today.

C: **Oh!** [laughter]

<div align="right">[ICE-Ireland, Face to face conversation, S1A-050$A]</div>

In terms of the occurrences of the vocalisation *oh* in clusters, we see that both varieties adopt *oh well* and *oh yeah* almost equally. We see more tendency to use the two-word response token *oh really* versus the single-word RT *really* in Canadian English. Also, there are many forms which have appeared in the data that do not appear on the top 100 most frequent word and cluster lists. Here are some occurrences of *oh*-clusters found in both datasets which have not appeared in the top 100. See Table 6.6 and the examples 6.19-6.22 extracted from both datasets for more illustration.

Table 6.6: *Some examples of oh-clusters found in ICE-Ireland and ICE-Canada*

ICE-Ireland	ICE-Canada
oh yeah	oh well
oh yes	oh ya
oh brilliant	oh really
oh right	oh shoot
oh my god	oh yeah
oh well	

Example 6.19

A: Oh do you know what's over there belonging to +you? Uhm, Oscar and Lucinda?
B: Mm.+
C: **Oh brilliant**.
B: It's her life story, you know her family…
A: What Gran said reminded me to give it to you. Uhm beside, the computer there.
                                    [ICE-Ireland, Face to face conversation, S1A-077$D]

Example 6.20

A: +Centra!
B: Centra!+
C: Where?
A: You know just by, +oh uhm…
B: Do you know+ there on O'Connell Bridge.
C: **Oh my God!**
A: Yeah. Very embarrassing.
C: What, but like ye hadn't. Ye were just drinking when ye went to the pub like wasn't it?
B: Okay, we went.
                                    [ICE-Ireland, Face to face conversation, S1A-068$B]

Example 6.21

A: So, it's not too bad.
B: I found a shirt for your dad for Christmas.
A: **Oh ya?**
A: Bright pink.
B: No!
                                    [ICE-Canada, Face to face conversation, S1A-086 #52:1: B]

Example 6.22

A: She never bought a place up at Sardona.
B: **Oh really?**
A: She said she and Heath bought a log cabin uhh, I don't know how many years ago she said.
                                    [ICE-Canada, Face to face conversation, S1A-002 #60:1: B]

Also, Canadian English is distinctive in its tendency to use the single-form of

because as a RT marking hesitation in the discourse, which can also take the pragmatic

function of *you know* as a RT which is used quite often in Canadian English. Refer to

example 6.23 below and note its frequency in Figure 6.2 in comparison to Irish English.

More information on *because* has been covered in relation to its pragmatic function in the discourse in section 6.4.2.4.

    Example 6.23

A: Ya, oh ya, ya.
B: I +mean you will have learned more in four years than you would have in two.
A: That's true.
B: You know++ what I mean like?
A: Mm++
A: Mm.
B: **Because.**
A: Ya, ya.
A: Sort of the long haul.
                  [ICE-Canada, Face to face conversation, S1A-033 #234:1: A]
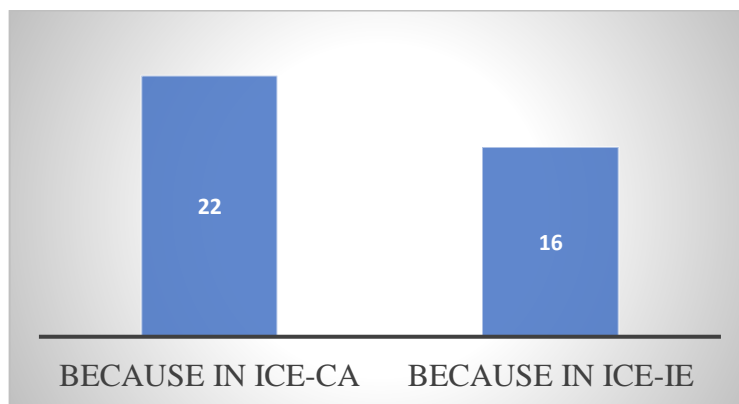
Figure 6.2 The frequency of *because* (per 10,000)

6.4.2 Functions

RTs have been classified and categorised in relation to their pragmatic functions broadly and specifically. In other words, some scholars use, or provide the classifications which suit the aim of their research focus. For example, the different terms and classifications found in the work of Carter and McCarthy (2006) and Gardner (2001), (presented in Appendix A) show the top-down framework for the analysis of the forms and pragmatic functions of spoken grammar used in this present research. On the other hand, we have the broad pragmatic functions classifications or types of RTs presented by O'Keeffe and Adolphs

(2008) in which they cover most, if not all, the sub-pragmatic functions found based on the existing literature review. Here are the broad types of RTs customised by O'Keeffe and Adolphs (2008).

Continuer response tokens' pragmatic function is to maintain the flow of the discourse by encouraging the current speaker to continue because speakers perceive continuer RTs as floor-yielding signals that mark the addressee's desire for the talk to continue. This type of pragmatic function is associated with the minimal forms such as *yeah* and *mm* (O'Keeffe & Adolphs 2008).

1. **Convergence response tokens** are markers that mark agreement, topic boundary, closure, or convergence on an understanding of what is common ground or shared knowledge between participants in the discourse. This type of RT is the most frequently occurring type of RT found in conversations where

> Participants agree, or simply converge on opinions or mundane topics and this leads them to negotiate topic boundary points collaboratively, where a topic can be shifted or changed. Convergence can also be followed by a conversational closure point. In this way, response tokens have a pragmatic function in that they help bring about agreement and convergence leading sometimes to topic shifts.
>
> (O'Keeffe & Adolphs 2008, p.18).

And there are many discourse forms that perform this pragmatic function such as single-word items: *yeah*, follow-up questions: *did you*? *Is she*? And short statements: *yeah it's pretty sad* (agreeing statement) (O'Keeffe & Adolphs 2008, pp.16, 18).

2. **Engagement response tokens** are markers of high engagement in which addressee(s) respond on an affective level to the content of the message. Engagement tokens "thus signal the addressee's enthusiasm, empathy, sympathy, surprise, shock and disgust at what the speaker is saying, without taking over the turn" (O'Keeffe & Adolphs 2008, p.20). These engagement tokens are usually non-

minimal responses and common items include single-word forms such as *excellent*,

absolutely, and *brilliant*; short statement repetitions such as *that's nice*, *oh wow*, *oh*

really, *wow*, *cool*, *gosh*, *really*; and short phrases such as *that's tough*, *you're not*

serious; and follow-up questions such as *did you*? *Is that so*? (O'Keeffe & Adolphs

2008).

3. **Information receipt tokens** are markers of points indicating when information in

   the discourse flows and stops. In other words, they draw a boundary in the discourse

   and signal a point of topic transition or closure, and they can be indicative of

   asymmetrical discourse. Thus, they are commonly used with the discourse forms

   right and *okay* as "self-imposed pragmatic markers at which the storyteller marks a

   boundary where the narrative can begin now that the contingent details are clear for

   the participant" (O'Keeffe & Adolphs 2008, p.21).

Therefore, the analysis of the pragmatic functions of the RTs found across Irish English

and Canadian English (as shown above) will be mainly approached from these four types as

well as the types discussed briefly in section 6.2.3 where relevant.

6.4.2.1 Continuer response tokens

Irish English and Canadian English show a clear contrast in relation to the RTs that help to

maintain the flow of the discourse, where the listener uses signals seen by the speaker as

floor-yielding signals that mark the addressee's desire for the talk to continue. In Canadian

English, we see that there is a higher tendency to adopt the minimal response tokens of *mm*

and *hmm* as continuer RTs, whereas this is not the case in Irish English to the point that

these minimal RTs of *mm* and *hmm* did not appear within the top 100 most frequent items

in the word and cluster lists of Irish English. Instead, we have the RT *yeah* which appears

much more in Irish English as a continuer RT as well as other minimal RTs such as *uh*, *huh*, *hum*, and *uh-huh* (for more information about the comparative frequency see section 5.4.1, Figure 5.1). However, there are some occurrences of *mm* and *hmm* used in Irish data, but still they are infrequent compared to Canadian English (see Figure 6.3 and examples 6.24- 6.27 and 6.30 below for more illustration).
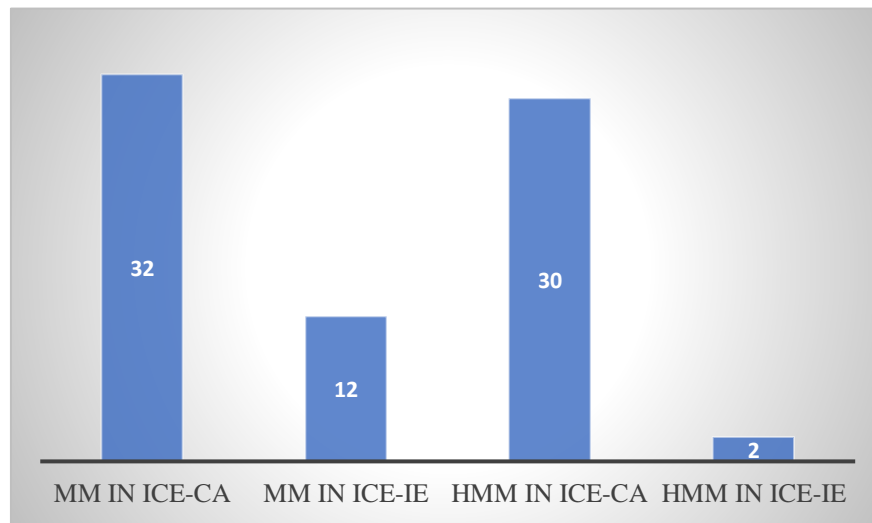
Figure 6.3 The frequency of *mm* and *hmm* (per 10,000)

Example 6.24

A: We took him to the Okanagan one time and uh, we were sittin, sitting down the beach and he said, "ya I think I could stay here. I like it here."
B: +**Mm**
A: And I+ said, "oh Chris, how come you like it here so much?"
                          [ICE-Canada, Face to face conversation, S1A-054 #250:1: A]

Example 6.25

A: Nah, they taste a little bit different I find.
B: Oh.
A: They're not as fine tasting.
B: **Hmm**.
A: Uhm, do you know what, uh, what Sandy's having at ^for^ dinner on uhm on Friday?
                          [ICE-Canada, Face to face conversation, S1A-013 #240:2: B]

Example 6.26

A: What are they going to charge her? Flat rate?
B: I can't remember the quote. I don't remember.
A: **Hmm**.
B: She's got a lot of stuff, and heavy stuff. [laughs]
                              [ICE-Canada, Face to face conversation, S1A-026 #150:1: B]

Example 6.27

A: Ya, oh ya, ya.
B: I mean you will have learned more in four years than you would have in two.
A: That's true.
B: You know+ what I mean like?
A: **Mm+**
A: **Mm**.
B: Because.
A: Ya, ya.
A: Sort of the long haul.
                              [ICE-Canada, Face to face conversation, S1A-033 #232:1: B]

Example 6.28

A: Well he was messing with the aerial.
B: **Yeah**.
A: Is it, is it broken now?
C: It isn't broken at all. Have you any brain in your head at all have you?
A: Who?
C: You.
                              [ICE-Ireland, Face to face conversation, S1A-087$C]

Example 6.29

A: Beginning of the summer sometime. And he came over there uhm in September.
B: **Yeah**.
A: And he's getting on well. He was home for a week.
B: Really?
                              [ICE-Ireland, Face to face conversation, S1A-086$A]

Example 6.30

A: There'll be war.
B: There's going to be war?
A: **Mm**.
B: As long as I'm not there I don't mind.
A: Yeah, I know.
                              [ICE-Ireland, Face to face conversation, S1A-070$A]

In Examples 6.19 and 6.30 above, we can see that *mm* is also used for acknowledgment purposes or functions which can be perhaps associated with a head nod; this reflects on the usage of *yeah* in Irish English which can serve as a continuer RT and an acknowledgment RT as well. This can be inferred from Table 6.5 in which the reduplication forms of *yeah*, *mm*, *hmm*, and their comparative analysis in both varieties have been presented. Yet, this type of pragmatic function in the discourse performed by continuer RTs has been demonstrated in both varieties similarly by other RT items such as *uh* and *hum*.

Example 6.31

A: Has she thanked you ah again for her present?
B: **Uh uh**.
A: She didn't, You don't think she liked it?
                    [ICE-Canada, Face to face conversation, S1A-011 #39:1: A]

Example 6.32

A: …Yeah and went back again. It's uh yeah you have to have two to three years'
        experience minimum and then…
B: **Uh-huh**.
A: You stay there for a year and then you come home.
B: Very good.
                    [ICE-Ireland, Face to face conversation, S1A-086$A]

6.4.2.2 Convergence response tokens

As previously noted, convergence RTs are the most frequent RT found in conversations and they display multiple pragmatic functions. Both varieties of English have shown a varying wide range of RTs serving as convergence RTs. For example, in terms of marking agreement we see that, based on the quantitative findings and the qualitative analysis, there are clear distinctions between Irish English and Canadian English in relation to forms, see the examples below.

Example 6.33

A: But I mean like, I don't know I mean some- sometimes for myself, the issue just becomes… I'm just worried that when I do have the time that, it's not gonna ^going to^ +happen or it's not gonna ^going to^.
B: **Ya+**
A: But I mean it +will, so, you know I mean?
B: **Ya+**
A: I mean maybe that is a sort of an interesting year to, I don't know, lay some groundwork for something +after at the same time you're…
B: **Ya+**

                         [ICE-Canada, Face to face conversation, S1A-033 #251:1: B]

Example 6.34

A: But they're all quite old and feeble.
B: **197ay a**.
A: I wonder why she likes being away.
B: Ya I don't know.
A: She's probably g-got her life centred+ there now
B: **Ya**.+ But anyway…

                         [ICE-Canada, Face to face conversation, S1A-040 #230:1: B]

Example 6.35

A: You see that, that would maybe be a bridge where you could say well for one thing the +reviews might be selective, ++you know, ah the person may particularly ah focus on… or there might be a hypothesis confirmation of, of process.
B: **Right**.+
B: **Right**.++

                         [ICE-Canada, Classroom lessons, S1B-001 #329:4:A]

Example 6.36

A: Eventually
B: I understood your evidence is you didn't wanna ^want to^ spend fifteen or twenty years in jail you wanted to get out you had a life you had a kid.
A: **Right**.
B: **Right**.

                         [ICE-Canada, Legal cross-examinations, S1B-066 #127:1: B]

Example 6.37

A: I don't want her to come with me.
B: **Right**.
A: Let her go with her friends+ and I go with my friends.
B: **Right**.+

                         [ICE-Canada, Face to face conversation, S1A-062 #272:1: A]

It has been found that using *ya* is more commonly found in Canadian English as demonstrated in the examples above; and using the equivalent form *yeah* is more common in Irish English and these have both been found in clusters in both varieties. *198ay a* exists only in Canadian English (as seen above in Examples 6.33 and 6.34) and *yeah yeah* is only in Irish English (as shown below in Examples 6.38- 6.41); however, both of these clusters display a significant difference in their frequency (as displayed in Table 6.2). *Yeah* is a more dominant marker marking agreement (acknowledgement) or understanding of the prior turn whereas this is not the case in Canadian English; Canadian English uses more RTs for this pragmatic function, including *yeah*, which is not used as commonly as it is in Irish English.

This can be explained as follows: firstly, *yeah* in Irish English appears as ranging from a single word to a six-word cluster while *yeah* in Canadian English appears only as a single word within the top 100 most frequent clusters; even *ya* appears only once in a cluster at the two-word level. Secondly, there is significant contrast in their frequency in both corpora. We see a variety of RTs in Canadian English used to mark acknowledgement or agreement such as *you can* which appears within the top 100. Also, it has been found that the RT *right* in Canadian English is used more to mark agreement or understanding of the prior turn and also in higher frequency than *right* in Irish English which is used more as an information receipt RT than Canadian English (see section 6.4.2.4 and see Table 6.1 for frequency). Here are some examples (6.38-6.41) of the RT *yeah* extracted from the Irish data:

Example 6.38

A: Are you sure?+ Would you like some orange juice Michelle?
C: No, I'm fine thanks.
A: Are you sure?

C: **Yeah**.
A: Okay.

Example 6.39

A: Is this Maureen Murphy?
B: Morgan Stanley.
C: In England?
B: **Yeah**.
A: She has a boyfriend?
B: **Yeah**.
A: Since when?

Example 6.40

A: They're, do you know where uhm…
B: Oh 'tis you're related to Eithne O'Byrne's is it?
C: **Yeah yeah yeah yeah yeah**.
B: Oh yeah, I was going to say while ago how did you know Eithne O'Byrne? Do you know?
C: Yeah, I met her last weekend

Example 6.41

A: Did ye see the coloured fellas in uhm Cormac's?
B: Yeah. Didn't I point them out to you? Y~ you'd swear it as something new to see a +black person like.
C: **Yeah yeah yeah**.+
A: No, but like [unclear speech] was saying they're bouncers…

On the other hand, both varieties of English have revealed similar RTs marking

agreement, topic boundary, or understanding of the prior turn in the discourse with almost

the same frequency such as *that's right*, *I know*, *I don't know*, *no I don't*, *I don't think so*,

well I don't know, *yeah I don't know*, and *ya I don't know* (see also Table 6.1 to 6.4).

Example 6.42

A: But I can't wait til after Halloween. Hopefully I'm going to make my fortune
B: Exactly Angela. +Yeah.
A: And then+ I'm off to Brussels on the ++eighth or the eleventh of December.
B: **That's right**.++ Excellent!
A: So I'm looking forward to that now.

Example 6.43

A: And if they do, well then that's uh competition for the type of school that we have. I still would rather send my daughter to an [indigenous speech] because I think it's a different experience and I think that she'll have more the type of experience that, that he was mentioning

B: Mm hmm.

A: You know?

B: **That's right**.

                                          [ICE-Canada, Business transactions, S1B-071 #108:1: A]

Example 6.44

A: Well, I would prefer if it was fucking freezing than that wet that's out there +tonight and that. Cos it's not even well it was a wee bit chilly this evening there getting a bit chilly ++but, do you know the way it was pissing down, what day was it? Wednesday down at home I think it was or Tuesday? +++and it would've fucking boiled you alive ++++outside you know? You know warm. It's shit like.

B: **I know**.+ Yeah.++ Yeah.++ **I know I know**.++++ Well, come here. I texted you during the week there.

A: Yes, aye, what was that about?

B: Uhm just to see whether you were busy over Halloween but I'm sure you probably are.

                                  [ICE-Ireland, Telephone conversation, S1A-098$A]

Example 6.45

A: Do I have to go to your meeting on your year…?

B: Not unless you want to.

A: Well you didn't invite me.

B: Okay you're invited.

A: **Well I don't know**.

B: I'll tell you…

                                    [ICE-Canada, Face to face conversation, S1A-085 #248:1: A]

Example 6.46

A: But they're all quite old and feeble.

B: Ya, ya.

A: I wonder why she likes being away.

B: **Ya I don't know**.

A: She's probably g-got her life centred+ there now

B: Ya.+ But anyway…

                                    [ICE-Canada, Face to face conversation, S1A-040 #232:1: B]

Example 6.47

A: Ah sure you don't. [laughter]
B: Do we have to take her biccies?
C: Yeah. Don't you think? No? Oh.
B: Okay.
A: **I don't think so**.
B: Would Dawn bring us some if she came here?
A: Uh-huh uh-huh. I don't know. Probably not. [laughter] I don't know.
                              [ICE-Ireland, Face to face conversation, S1A-066$C]

In addition to the convergence RTs mentioned above, we can see some RTs found

within the top 100 of one variety but not the other, acting as convergence RTs. For

example, *you don't know* and *you know* which mark what is common ground or shared

knowledge between participants in the discourse. This appears more in Canadian English

than it does in Irish English, see the examples 6.48-6.50 below for a demonstration of this.

Example 6.48

A: They're not, nuh- virtually none of them are helping me with what I'm, with what
      my field is, except because I make them.
B: Mm hmm. Mm hmm.
A: **You know**
B: Mhmm ya, well…
A: And some of them I can't.
                              [ICE-Canada, Face to face conversation, S1A-037 #78:1: A]

Example 6.49

A: And if they do, well then that's uh competition for the type of school that we have. I
      still would rather send my daughter to an [indigenous speech] because I think it's a
      different experience and I think that she'll have more the type of experience that,
      that he was mentioning
B: Mm hmm.
A: **You know**
B: That's right.
                              [ICE-Canada, Business transactions, S1B-071 #107:1: B]

Example 6.50

A: She didn't get out of her cart at that point?
B: No.
A: And you say the other, +the other people that you were golfing with were also
      there?

B: Yes. Frank and Mark+ and Steve were nearby as well. Yeah. Mm hmm.
A: Was, ^Is^ it possible that someone was purchasing something from her, something to drink at that time? Or, Or you just didn't see anything like that happen?
B: No, No.
A: **You don't know**!
B: No, I don't know.
A: Okay.

[ICE-Canada, Legal cross-examinations S1B-061 #171:1: A]

Note that example 6.50 can be considered debatable and problematic in terms of whether it is an actual RT or a full turn held by Speaker A making a formulation. This is a legitimate argument which relates to how *you don't know* has been expressed (whether it is a surprise statement or not) and to what extent that turn is taking over (in Example 6.50, there is no object after *know* even though it is a transitive verb); also, it is arguable whether or not the statement *you don't know* can be considered to be one single ready-made chunk Therefore, we see different views on what is considered to be full turns in relation to RTs. For example, Gardner (2001) gives a broader definition for what he considers to be RTs as opposed to what he considers to be full turns. This can be seen with questions seeking clarification like *which one* as in example 6.51 below which is taken as a RT rather than a turn. Nevertheless, in this study (as illustrated in Appendix A), Gardner's position has been adopted; according to this position, any statement or item used for clarification or other types of repair which seeks to clarify mishearing or misunderstandings is considered to be a RT with the condition of it not taking over the floor from the prior speaker. This is usually done by brief questions and therefore short questions seeking clarification or as part of repair sequences have been included (Gardner 2001). Another issue is the missing audio or visual component which leaves the researcher guessing in some cases as to whether a short question is a clarification-check, repair, or a follow up engagement question (as in example 6.52). Thus, broadening out the definition of questions as RTs to

include Gardner can be legitimate. However, it does not take away the possibility that the second *what?* in example 6.52 could be a genuine question. Speakers A and B are on the phone, A's roommate walks in and bothers or teases Speaker A and Speaker B is trying to makes sense of the change in conversation, and asks *what*?

Regardless of the issue of categorization, Irish English has shown more questions and tags than Canadian English. There are a number of convergence RTs in the form of follow-up questions found only in Irish English in the top 100 but not in Canadian English (i.e. *did you*? *are you*? *is it*? etc. see examples 6.11-6.14 presented above). There are also some that both Irish English and Canadian English share such as *what?* and *who?* (see Table 6.1 and examples 6.28 and 6.51- 6.53). Brief questions appear as RTs for several pragmatic functions. According to Gardner (2001), this can be for clarification or a type of repair which seeks to clarify mishearing or misunderstanding. Also, according to Carter and McCarthy (2006), the pragmatic functions of what they refer to as "follow-up questions" is to serve as a signal of engagement and attention by the listener. They are very similar to RTs of *yeah* and *really*. In other words, follow-up questions in informal spoken language often simply function to keep the conversation going by inviting more information (as discussed earlier in this chapter and also in Chapter 7 which will focus on questions and tags in more detail). See the examples 6.51- 6.53 below for more illustration.

Example 6.51

A: Being a hazard to every~ to everyone. [laughter]
B: But they have people walking across and everything don't they?
C: Yeah, they do yeah.
D: Sharon's from Cork, isn't she?
E: Mm. Fine woman.
C: But the guards have those as well. In Templemore.
F: **Which one**?
C: Yeah, that's right.
A: **Is it**?

D: I want to learn how to put the old handbrake on…

<div align="right">[ICE-Ireland, Face to face conversation, S1A-064$H]</div>

Example 6.52

A: Interesting how people…
B: **What**?
A: Nothing. How you doin ^doing^?
B: Good, how're you?
A: [laughs]
B: What were you sayin ^saying^?
A: My roommate just walked in.
B: Ahh… I see. That was Dana, she's calling me back.
A: [laughs]
B: I'm just a popular person today.
A: Arrrrgh, get off! [laughs]
B: **What**?
A: I'm on the phone long distance, get off, okay? [laughs]
B: … Anyways like I was saying, Ryan left this morning.
A: Mm hmm.
B: Mm hmm.
A: How you doin ^Doing^?
B: I'm doin, doing pretty good actually.
A: Good.

<div align="right">[ICE-Canada, Telephone conversation, S1A-097 #220:2: B]</div>

Example 6.53

A: Is it coffee for everyone?
B: +Please.
C: I'm fine, yeah.+
A: Are you sure?
C: Yeah. If I've coffee now I'll never get to sleep.
A: Deborah?
B: Uhm actually do you know what I'd love?
A: **What**?
B: A mug of hot water please.
A: Okay.
C: That's what Edina does there.
B: +Mm.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-056$B]</div>

6.4.2.3 Engagement response tokens

Based on the top 100 most frequent words and clusters, there are not many RTs marking

the addressee's enthusiasm, empathy, sympathy, surprise, shock, and disgust at what the

speaker is saying compared to the convergence RTs. However, we can see that Irish

English and Canadian English have shown agreement on fulfilling the engagement process

in the discourse with some single-word RTs such as *really*, *right*, *no*, and *oh.* These RTs

mark the prior speaker's turn as newsworthy in some way or surprise as in the RTs *really*,

right, and *no* which often occur with *oh* for more emphasis as discussed earlier with the *oh*-

clusters which are not found at the top 100 most frequent words and clusters (see Table

6.6, as well as some examples extracted from the datasets expressing engagement in the

discourse).

Example 6.54

A: That's like one of the girls in my class. She, you know Siobhan, the one I'm doing
       my thesis with?
B: Yeah.
A: She is the best singer. You would +oh my God it's unbelievable. She'll play the, she
       plays the guitar like and she sings and writes her own songs.
C: **Really?+ Oh**.
A: She is just, oh her voice is just like, oh my God. It's brilliant. Cos there's a Mass, a,
       there's Mass on tomorrow night. Do you remember I told you about that girl who
       died?
C: Oh yeah.

[ICE-Ireland, Face to face conversation, S1A-068$B]

Example 6.55

B: Is that just Vancouver though?
A: Uhh, I don't know? It's Western Canadian I guess.
B: Sounds a bit Ottawa Valley.
A: **Really?**
B: Ya. A little bit.
A: Oh well, I've mh got a Welsh and Scottish background so maybe that's got
       something to do with it?

[ICE-Canada, Classroom lessons, S1B-008 #200:1: B]

Example 6.56

A: Nine kids or something.
B: Will you be heard over there now? You will.
A: Not sure. Okay bit more. Uh no I've only one brother and one sister.
B: The sister's older.
A: No.

B: **No?**
A: Uh uh wrong one. Try again [laughter].
B: Brother's older.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-062$B]</div>

Example 6.57

A: So, it's not too bad.
B: I found a shirt for your dad for Christmas.
A: Oh ya?
A: Bright pink.
B: **No!**

<div align="right">[ICE-Canada, Face to face conversation, S1A-086 #54:1: A]</div>

However, above the single-word level, it has been noticed that Irish English has exhibited more clusters used as engagement RTs, especially with the follow-up questions (as shown in examples 6.11, 6.12, 6.13, and 6.14), see example 6.58 for the cluster of *it's not too bad* which appears in the top 100.

Example 6.58

A: Do you like it?
B: Yeah, but ah you get tired of it after six hours like +you know?
A: Yeah.+ And do they pay you well or what?
B: Fiver an hour.
A: **It's not too bad**.
B: Yeah.
A: Sure 'tisn't.
B: No it's grand. One of those hours I was lifeguarding so I got three pounds an hour. That's grand like. But teaching is better you know. Passes the time much more quickly.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-046$B]</div>

In relation to engagement RTs, there are some engagement RTs appearing in the datasets that are not in the top 100. Regardless, Irish English still shows more engagement RTs such as *that's disgusting* and *that's excellent* and many others as presented in Table 6.6. Conversely, Canadian English also uses some engagement RTs that do not appear in the top 100 such as *that's a good idea* and *is that so*? but still remain less frequent compared to Irish English.

6.4.2.4 Information receipt tokens

In terms of RTs serving as markers of points where the information in the discourse either flows or stops, we find the forms *right* and *okay* commonly functioning as "self-imposed pragmatic markers at which the storyteller marks a boundary where the narrative can begin now that the contingent details are clear for the participant" (O'Keeffe & Adolphs 2008, p.21). We see that Canadian English has distinctive RTs which can be used for this type of pragmatic function. For example, *because* as displayed in example 6.59 shows that speaker B finished his/her point and was waiting for a type of expected response which did not occur; this led to the use of *because* by Speaker B in case Speaker A is expecting more information even though there is no information to be added by Speaker B. This makes the RT *because* mark hesitation as well as to what type of information should be added for Speaker A to be satisfied in terms of receiving enough information.

Example 6.59
A: Ya, oh ya, ya.
B: I +mean you will have learned more in four years than you would have in two.
A: That's true.
B: You know++ what I mean like?
A: Mm++ Mm.
B: **Because**.
A: Ya, ya. Sort of the long haul.

[ICE-Canada, Face to face conversation, S1A-033 #234:1: A]

Also, we have the RT *so* in Canadian English (as shown in examples 6.9 and 6.10); this expresses the pragmatic function of handing the floor back to the immediately prior speaker due to the fact that the speaker has not satisfied the recipient's expectation in terms of providing enough information or completing the discourse which causes vagueness (and this pragmatic function also has been expressed with the response token *and then* in Canadian English). As a result, *so* has been used a lot in Canadian English as the RT *okay* would be used which means *what's next*? which makes *so* and *okay* interchangeable with

their meanings. However, in Irish English, *right* has been used interchangeably with the meaning of *okay* as *so* does in Canadian English. However, this is not to say that *right* is used exclusively in Irish English as it is found in Canadian English as well.

Example 6.60

A: You see you know your own one inside there, Sally right?
B: Yeah.
A: That's her name. She went out with Jim, okay?
B: **Right**.
<div align="right">[ICE-Ireland, Face to face conversation, S1A-076$B]</div>

Example 6.61

B: And okay I talked to my mom on Sunday and I told her that I wanted to go down to Windsor when he came home in November +for a couple weeks.
A: Yeah, mm hmm.+
B: Well two weeks actually. Cos his parents o~ ^his whole family^ was having like a Christmas for him.
A: **Okay**.
B: Cos he won't be home for Christmas.
A: **Right**.
B: And so I wanna ^want to^ go ho~ ^go down^ there on a weekend so I wouldn't miss any school or anything but on a weekend.
<div align="right">[ICE-Canada, Telephone conversation, S1A-097 #236:2: A]</div>

6.5 Conclusion

In this chapter, we presented the quantitative and qualitative analysis of RT use in Irish and Canadian Englishes in relation to forms and functions that have been captured in the ICE-Ireland and ICE-Canada corpora. Methodologically, word and cluster lists were generated for both corpora to identify and compare the forms used in both datasets, and from these lists, RTs with pragmatic functions were identified manually by cross-checking qualitatively using concordancing. This results in interesting quantitative findings at the form level in which Irish English and Canadian English show the most similarities at the single-word RTs level (as they did with their DMs) with some prominent RTs appearing

more, or only, in one variety over the other. For example, *ya*, *so*, *because*, *mm*, and *hmm* have been found more in Canadian English as RTs than Irish English. On the other hand, both varieties begin diverging noticeably at the two-word and beyond level (of RT). In terms of the qualitative findings, we saw that both varieties have displayed common pragmatic functions in relation to acknowledgment, engagement, and drawing a boundary in the discourse presented by the same RTs such as *that's right*, *I know*, *do you*? *no I don't*, *I don't think so* and so forth. Nevertheless, both varieties have shown clear contrasts in terms of forms. For example, questions and tags were found to be in use much more in Irish English generally and specifically as convergence RTs and engagement RTs than in Canadian English. In contrast, Canadian English uses more RTs marking what is common ground or shared knowledge between participants in the discourse such as *you don't know* and *you know*. Also, Canadian English and Irish English reveal a special tendency towards using information receipt tokens indicating when the information in the discourse would flow and/or stop. These are commonly used with discourse forms of *right* and *okay; so* in Canadian English is used more than in Irish English, and *right* and *okay* are interchangeably in use more in Irish English than in Canadian English. Lastly, this chapter has revealed the pragmatics of both varieties of English which demonstrate how different and similar Irish English and Canadian English are in relation to pragmatic functions and forms in their RTs. We will now explore the use of questions and tags in Irish and Canadian Englishes in the following chapter (Chapter 7).

CHAPTER 7

QUESTIONS AND TAGS IN IRISH AND CANADIAN ENGLISHES

7.1 Introduction

In this analysis chapter, we will present the quantitative findings and the qualitative analysis of the questions and tags found across Irish and Canadian Englishes which have been captured in the ICE-Ireland and ICE-Canada corpora. Although we have discussed some of their forms and pragmatic functions in the previous chapters (Chapters 5 and 6), due to their occurrences as discourse markers (DMs) or response tokens (RTs), they merit a separate and more in-depth treatment in both datasets. This will be done in the same manner DMs and RTs were covered in Chapters 5 and 6 by detailing the main quantitative findings resulting from both datasets through the theoretical and methodological framework taken in this study (as explained in Chapters 3 and 4) in order to unpack and analyse these findings qualitatively and comparatively (as will be demonstrated in section 7.4). Hence, the analysis of the pragmatic variation resulting from the way questions and tags are used in the spoken grammar of Irish English and Canadian English will be approached and investigated from two perspectives: forms and functions (as will be demonstrated in section 7.3).

Questions, generally, have been defined as "utterances which require a verbal response from the addressee" (Carter & McCarthy 2006, p.715). In terms of grammar, questions can be viewed and approached either from the grammar of written English or the grammar of spoken English due to the fact that each mode (spoken or written) has its own characteristics and distinctive forms (as discussed in Chapter 2, section 2.3.2) (McCarthy & O'Keeffe 2014). In relation to spoken grammar, there is a range of question types or formations which perform different pragmatic functions (as discussed in section 2.4.4) in terms of formality, politeness, directness, dependence on immediate context, projections of

211

degrees of shared knowledge and so forth (Andersen 1998; Carter & McCarthy 2006; Columbus 2010; Axelsson 2011; Kimps *et al* 2014, 2018; Barron *et al* 2015). Thus, we will explore the formation of questions that appear at the top 100 most frequent word and cluster lists extracted from the ICE-Ireland and ICE-Canada corpora (as discussed in Chapter 4 and shown in Appendices B-M).

7.2 Previous Research

7.2.1 The definition of questions

There has not been a consensus on the definition of questions that has been agreed upon by all the scholars of linguistics. Questions have been viewed and used as a semantic category, an illocutionary act, or to refer to requests or verbal directives, or simply as something that expects an answer (Tsui 1992). As a result, questions have been classified with three semantic classes based on the type of answer required, as outlined below (Quirk *et al* 1985):

- Questions requiring yes/no answers.

- Wh-questions which require an answer from a range of possible answers.

- Alternative questions which expect a reply from two or more options

  presented in the question.

  However, in relation to the spoken grammar or language, questions can be more complex in terms of forms and pragmatic functions. Consequently, the problem of question forms and pragmatic functions is addressed by many researchers. Bolinger (1957) points out that:

    the Q[uestion] is an entity that is often assumed but seldom defined… the difficulty in definition betokens
    a complex which is not only made up of a number of ingredients, but whose ingredients may vary as to

212

presence or absence or proportionate weight. If there were no such variation there would not be a complex in the linguistic sense (Bolinger 1957, cited in Weber 1993, pp.4-5).

Weber (1993) believes that the term *question* applies "ambiguously to interactive function and morphosyntactic form" with the correlation between morphosyntactic form and the communicative function of what she terms "doing questions" (p.4). Tsui (1992) also proposes a functional description of questions in which she characterises any utterance which proposes an obligatory verbal response as an "elicitation" irrespective of its syntactic form to be considered "a question." Taking this definition as a base would avoid the inconsistency of using syntactic criteria for some utterances and discourse criteria for others. This goes against the "expected answer" model provided by Quirk *et al* (1985) on the basis that it confuses issues of syntactic form and pragmatic function (for example, a "declarative question" versus a request or an exclamatory question). However, Uwajeh (1996), in a detailed discussion on the difficulty of distinguishing questions clearly from certain other pragmalinguistic phenomena, concludes that "communicative *context*, not sentence structure, is the ultimate basis for determining a sentence's communicated illocutive intent, and therefore for its possible classification as a question" (p.108; O'Keeffe 2003; also see section 7.4.2 for more illustration).

7.2.2 Forms and functions of questions in spoken grammar

Question tags (QTs, also known as tag questions TQs) are one of the types of questions that repeatedly occur in spoken language (Biber *et al* 1999; Carter & McCarthy 2006). From the formal level (or perspective), TQs have been defined as a combination of two clauses, an anchor clause and a tag clause, uttered by the same speaker (as discussed in detail and demonstrated in Chapter 2, Table 2.1). The anchor clause can contain different types of

clauses such as: a declarative, imperative, exclamative, or interrogative clause (Axelsson 2011, p.30). Whereas the tag clause hosted by the anchor clause, on the other hand, is made of a clause with interrogative syntax, consisting of an auxiliary verb operator and a pronominal subject (as shown in Chapter 2, Table 2.2.). This type of TQ is the prototypical TQ, known as the **canonical TQ** in which the tag clause takes the form of a "concordant mini clause" (Allerton 2009, p.310). Canonical TQs can be formed with reversed polarity as in *they've been* affected by it, *haven't* they? or constant polarity as in *he's gone back, has he?* In summary, canonical TQs can generate the following anchor-tag combinations: (positive anchor – negative tag), (positive anchor – positive tag), (negative anchor – positive tag), and lastly (negative anchor – negative tag) which are known be uncommon and unusual to occur (Allerton 2009; Axelsson 2011, pp.33–35, Barron *et al* 2015; for more information on different types of polarity, see Table 2.1 in Chapter 2).

The second group of TQs that fit in the definition of TQs provided above is the **invariant TQs**. Invariant TQs consist of an interrogative tag consisting of an operator and a pronominal subject as in *is it?* The difference between invariant TQs and canonical TQs here is the non-concordant tag which is not dependent on the syntactic properties of the anchor in the invariant TQs (Allerton 2009, p.310; cf. Andersen 2001, p.104). In other words, the concordant tags are more flexible and changeable with their forms which can be positive as in *do you?*, negative as in *don't you?*, or even *is it?* within the same clause. There are other types of invariant TQs known as **fixed tags** which are a very prominent type or formation of questions in spoken language. This type of question, fixed tags, are served in the discourse and spoken grammar in general with a wide range of forms (including single-word tags, phonological sequences, fixed phrases containing lexical

material) such as: *(all) right, okay, yeah, eh, huh, (do) you know, see, I think, and don't you think?* (Biber *et al* 1999; Carter & McCarthy 2006; Barron *et al*. 2015; for more examples of fixed tags, see Chapter 2, section 2.4.4, examples 2.8- 2.11).

Lastly, studies have shown that there is a clear correlation between turn position, function, and the role of TQs in turn-taking organisation. TQs have been classified and viewed as turn-final devices which signal that the speaker is yielding the conversational floor and projecting a response by the addressee (Barron et al. 2015; also see examples 7.12 -7.35; cf. Sacks *et al*. 1974). However, some recent empirical studies have shown that not all TQs are turn-final projecting the next turn (Andersen 1998; Axelsson 2011; Kimps *et al*. 2014, 2018, Barron et al. 2015). However, in this empirical study, TQs are viewed as turn-final devices, signaling that the speaker is yielding the conversational floor and projecting a response by the addressee. Yet, there are some forms of pragmatic markers (or questions) which adopt more than one type of question such as: *what do you mean?* which can be a TQ (an invariant TQ, as in example 7.43) and a discourse marker (or a follow-up question, as in example 7.42), performing the same pragmatic function. Therefore, as mentioned earlier in section 7.2.1, questions in general, including TQs, overlap with the other types of questions.

From the functional perspective, TQs have been studied semantically and pragmatically from different approaches in relation to analysis and categorisation. Some of these works are Algeo (1988, 1990, 2006); Axelsson (2011); Holmes (1982, 1995); Tottie and Hoffmann (2006, 2009); Kimps *et al*. (2014, 2018); and Barron *et al*. (2015). Some of these works have shown an overlap in terms of how questions (generally) and TQs (specifically) have been categorized and analysed. They also simultaneously highlight some

differences in regards to the type of analytical unit selected and investigated, whether tag, or anchor and tag (for an overview, see Axelsson 2011, pp.41-58). For example, some (Holmes 1982, 1995; Andersen 2001; Allerton 2009) consider only the tag; whereas (Axelsson 2011; Kimps *et al*. 2014, 2018; Barron *et al*. 2015) consider the whole TQ. Others, however, alternate between the pragmatic functions of TQs and tags such as Algeo (1988, 1990) and Tottie and Hoffmann (2006).

The pragmatic functions of TQs have been extended by Barron *et al*. (2015), using the functional model of Kimps *et al.* (2014) and Axelsson (2011) to enable researchers to investigate the speech and interactional functions of TQ utterances as well as the relationship between function and surface characteristics. The main advantage of this combined functional model is the distinction between **information-oriented TQs** (which are about the exchange of information between speaker and addressee) and **action-oriented TQs** (which are used to "give or demand goods and services and include requests, offers and suggestions") (Barron *et al*. 2015, p.6; Axelsson 2011; Kimps *et al*. 2014, p.81-82) See Figure 7.1 below on the functional classification of the information-oriented TQs and the action-oriented TQs which have been adopted in this study to analyse TQs of Irish and Canadian Englishes based on ICE-Ireland and ICE-Canada (Barron et al. 2015, p.6):

Questions

Statement-question blends
— Establishing common ground
— Topic-initiating
— Surprised reactions

Information

Statements
— Stating a fact or opinion (incl. self-monitors)
— Acknowledging responses
— Challenging
— Conversational joking

TQs

Desired action
— Give ———— Offers, advice, suggestions
— Demand ———— Requests, commands

Figure 7.1 The functional classification of TQs

Furthermore, TQs, whether they are information-oriented TQs or action-oriented TQs, have been further sub-differentiated based on the relative knowledge status of the interlocutors and adjacency (Kimps *et al*. 2014, 2018; Labov & Fanshel 1977). Based on the two primary criteria of knowledge status and adjacency (mentioned above), there are two question types that are relevant to the analysis in the present study. The first of these are questions by which information is sought from the hearer after the next turn is projected by the speaker. The second are statement-question (S-Q) blends in which the speaker is seeking a response from the addressee on a proposition that the speaker is certain about (Kimps *et al.* 2014, 2018; Barron *et al*. 2015). Therefore, Kimps *et al*. (2014) define S-Q blends as follows (p.74):

> [S-Q blends] have some recognizable features of both statement and question but cannot be reduced to either. Labelling them as a question would be stretching the limits of the category since the speaker is not seeking information, but is a primary knower making an assertion or evaluation. Categorizing them as a statement, on the other hand, is also problematic since these TQs intrinsically expect a response from the co-participant(s).

Having discussed TQs, we must also briefly remind the reader of **echo questions** which are another common type of questions used in spoken language to request further clarification about a segment in the discourse (for a more expansive definition and examples see Chapter 2, section 2.4.4, examples 2.12- 2.15 and for the analysis in relation to ICE-Ireland and ICE-Canada see section 7.4.2.2 below).

In addition to echo questions, we have **follow-up questions** which are used in the discourse with the typical pragmatic function of serving as signals of engagement and attention by the listener to encourage continuation of the conversation (Carter & McCarthy 2006). Follow-up questions occur in speech with different types (for a more extensive definition and examples see Chapter 2, section 2.4.4, and examples 2.16- 2.20, and section 7.4.2.3 in relation to ICE-Ireland and ICE-Canada).

Lastly, we have the **two-step questions** and responses which are one of the question constructions formed and used in speech for politeness purposes. They involve a two-step process (questions) by which the first question acts like "a preface" for the upcoming question. Speakers usually use this type of question in order to avoid being rude, too direct, or too general (Biber *et al* 1999; Carter & McCarthy 2006; for more discussion and examples see Chapter 2, section 2.4.4, and examples 2.22- 2.26).

7.3 Methodology

This section outlines the main framework of analysis for this chapter. The pragmatic variation analysis (using the formal level of analysis) of the questions used between Irish English and Canadian English will be approached and investigated from two perspectives: form and function (for further details on this refer to Chapters 3 and 4).

In terms of form, we will comparatively present the quantitative findings of the questions and tags across both varieties (Irish English and Canadian English) based on their word cluster (from single-word form to six-word clusters form). As elaborated upon in Chapter 4, wordlists and cluster analyses were generated to identify and compare the forms of questions used in both datasets (ICE-Ireland and ICE-Canada). Within the form analysis, we have already come across some forms of questions which have appeared as either DMs or RTs in Chapters 5 and 6 (i.e. the form *yeah* which can act as a DM, RT, fixed tag, and a follow-up question). Questions are very broad in relation to spoken versus written grammar in terms of their forms, concepts, and definitions (as mentioned earlier in section 7.2.1). Our concern in this chapter is the spoken grammar forms of questions presented by the top-down framework for the analysis of the forms and pragmatic functions of spoken grammar (demonstrated in Appendix A) and resulting from the bottom-up framework for the analysis of the forms and pragmatic functions of spoken grammar.

In terms of function, the quantitative findings presented with their forms will be unpacked and analysed qualitatively by concordancing these items manually (as discussed in Chapter 4). This results in determining the pragmatic function(s) of these forms of questions in the discourse. As a result, at the level of the questions and tags, the *iterative approach* has been taken. The bottom-up searches are guided by the top-down framework while the top-down framework is also informed by data in order to comparatively analyse and present the questions used by Irish English and Canadian English in the most accurate and clear manner.

7.4 Results and Analysis

7.4.1 Forms

Tables 7.1 to 7.6 below are a summary of the comparative analysis of question forms found in ICE-Ireland and ICE-Canada which demonstrate where Canadian English and Irish English both converge and diverge in terms of form. These tables below were derived from the qualitative analysis (from a bottom up route) of the top 100 most frequent word and cluster lists in which each single word and cluster has been concordanced manually in order to affirm which items from the quantitative findings resulting from the top 100 most frequent word and cluster lists (presented in Appendices B, C, D, E, F, G, H, I, J, K, L, M) are actually manifesting as questions within their pragmatic functions in the discourse. This usually involves going back to source files to check the extended turns that surround a given form. Note that the light gray shaded cells (in Tables 7.1-7.3 below) indicate the distinctive forms (questions) between Irish English and Canadian English, and the unshaded cells indicate the forms (questions) that both varieties of English share in common. However, this is only in terms of the forms of questions; in other words, it has been found in the data that some forms are in both ICE-Ireland and ICE-Canada but differ in their function or the type of function within the same category of functions (e.g. *did you?* or *do you?* which can serve as a follow-up question or a question tag) (Biber *et al* 1999; Carter & McCarthy 2006; Kimps *et al.* 2014, 2018, Barron *et al*. 2015).

Table 7.1. *ICE-Ireland and ICE-Canada single-word questions occurring within the top 100 most frequent word lists (frequency per 10,000 words)*

ICE-Ireland	Freq[5]	ICE-Canada	Freq
yeah	82	yeah	23
no	43	no	40

---

[5] The order of the forms presented in Tables 7.1- 7.6 is random and does not indicate rank.

really	19	really	25
right	19	right	28
what	53	what	55
when	21	when	24
who	23	who	19
okay	16	okay	31
which	23	how	17
		ya	54

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 7.2. *ICE-Ireland and ICE-Canada two-word questions occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
do you?	13	do you	8
you know	33	you know	47
did you?	6		
is it?	6		
you don't	3		
are you?	6		
isn't it?	2		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 7.3. *ICE-Ireland and ICE-Canada three-word questions occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
you know what	2	you know what	2
do you know	3	you don't know	1
do you think	2		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 7.4. *ICE-Ireland four-word questions occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq
do you know what	1
what do you mean	1

Table 7.5. *ICE-Ireland five-word questions occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq
you know what I mean	1
what do you call it	0.22

Table 7.6. *ICE-Ireland six-word questions occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq
do you know what I mean	0.12
you know what I mean yeah	0.12

Note: There are no forms of questions found in ICE-Canada as four, five, and six-word clusters; these are only found in ICE-Ireland as shown in Tables 7.4- 7.6 above and as will be explained below.

As seen above in Tables 7.1 to 7.6, the comparative analysis of questions and tags has been conducted at the form level in which we can see the similarities and the differences between Irish and Canadian Englishes in relation to their question forms and their occurrences (which will all be analysed below). As we have seen so far in Chapters 5 and 6 with DMs and RTs, Irish English and Canadian English demonstrate the most resemblance at the single-word level and after that the contrasts in relation to forms begin to appear more clearly. Based on the quantitative findings of questions occurring in both varieties at the Top 100, we see that both varieties of English (Irish and Canadian) match not only in forms but also in types at the single-word level. For example, fixed tags (elaborated upon below with examples extracted from the data in section 7.4.2.1) are displayed in both datasets with almost the same forms, with the exception of the distinctive form *ya* which appears only in Canadian English. Even this can be debated due to the phonological concept of phonemes and allophones. This can be the case with *yeah* and *ya* which are both interchangeable in their pragmatic functions as DMs, RTs, and QTs (yet, they are both

different in terms of form and interchangeable in terms of function with the underlying

form of *ya* in Canadian English). Furthermore, at the single-word level we can see that Irish

English and Canadian English demonstrate very similar echo questions and follow-up

questions performed with wh- words such as *what? who?* and *when?*. However, at the two-

word level, follow-up questions are not visible in Canadian English but remain in Irish

English with different forms of questions such as*: is it? are you?* and *did you?* (which are

all explained in details in section 7.4.2.3). At the single-word level, we see that *which?* has

been found in the top 100 serving as an echo question and a follow-up question in Irish

English but not in Canadian English. On the other hand, we have *how?* which has been

found in the top 100 (in ICE-Canada but not in ICE-Ireland) with the total of 1134

occurrences. Yet *how* as a single form is not found in the data as a question. It occurs only

in two and three-word clusters such as *how come?* and *how about you?* serving as an echo

question and a follow-up question in Canadian English, see sections 7.4.2.2 and 7.4.2.3 and

examples 7.1 and 7.2 below:

Example 7.1

A: You should, you should come out with us.
B: I'm not going to out with you guys Garfield. Well you guys+ are a bit too serious about it
C: **How come?+**
A: We're not too serious
B: I, well I mean, to get up at five thirty in the morning to go jogging, three days a week *coughs* I agree within principle, but I have a hard enough time getting up at seven and walking the dog.
C: Mm
B: You know…
                                   [ICE-Canada, Face to face conversation, S1A-027 #81:1:C]

Example 7.2

A: Uh, I don't, I can't see anything that's very important. Uhm, I wouldn't mind a little lie down
B: Mm hmm
A: **How about you**?

B: I'm very tired.
A: Ya well why don't we do that, and then go out fairly early for dinner.
B: I was up quite early today too. Not as early as you were, but uh…
                        [ICE-Canada, Face to face conversation, S1A-067 #244:1: A]

Note that the two and three-word clusters: *how come?* and *how about you?* are not in the top 100 of ICE-Canada based on the (quantitative) bottom-up analysis; however, they were driven by top-down analysis and analysed qualitatively. In comparison, *how* as a single form is not in the top 100 for ICE-Ireland. However, in relation to the two and three-word clusters *how come?* and *how about you?*, *how come?* occurs only twice in the ICE-Ireland data as a two-word cluster. While the three-word cluster *how about you?* does not occur at all in ICE-Ireland. It is noteworthy also that *where?* has not been found in the top 100 most frequent words in both datasets. At the single-word level, we have the question form *no?* Semantically, it cannot be classified or described as an indicator of questioning or looking for more information (as it can be pragmatically, as seen in examples 7.3- 7.5). Rather, it is a determiner as in: *there is **no** class* or an exclamation used to give a negative response as in: *is anything wrong? **No***. However, the pragmatic analysis and meaning can give more information based on the context, by which we know how *no* has been used. Examples 7.3 - 7.5 below, which have been extracted from both datasets, reveal that *no* as a form can be a question with the pragmatic function of checking and making sure what has just been said in the discourse (a segment of the topic in the discourse) is actually the case. *No,* combined with a rising tone can serve a similar pragmatic function to the pragmatic marker *really*, in marking surprise along with requesting more clarification on a segment of the topic in the discourse. Also, it has been noted that *no* in Irish English has a higher distribution based on the way it appears in the data (see example 7.18 in which *no* is used as a fixed tag).

Example 7.3

A: The quality of the tape uh disimproves when you uhm record over it.
B: It doesn't.
A: Oh does it not?
B: No.
C: **No?**
A: Well that's good.

[ICE-Ireland, Face to face conversation, S1A-052$B]

Example 7.4

A: About those chairs, you don't~ do you remember what they look like?
B: No.
A: **No?**
B: Which chairs are we talking about?
A: Uhhm, well I guess you never really saw those chairs did you?

[ICE-Canada, Face to face conversation, S1A-004 #192:1: B]

Example 7.5

A: Nine kids or something.
B: Will you be heard over there now? You will.
A: Not sure. Okay bit more. Uh no I've only one brother and one sister.
B: The sister's older.
A: No.
B: **No?**
A: Uh uh wrong one. Try again [laughter].
B: Brother's older.

[ICE-Ireland, Face to face conversation, S1A-062$B]

At the two-word level, Irish English displays a lot more questions, especially as RTs (as mentioned in Chapter 6), whereas Canadian English shows a lower frequency of questions at the two-word level, and in general, as opposed to DMs and RTs use in Canadian English which appears slightly more at the two-word level (as shown in Chapters 5 and 6). There is a wide range of question forms which have been displayed much more in Irish English than in Canadian English. This finding can be supported by Barron (2015) who found that there are more extensive use of questions (and interrogative anchors) in the ICE-Ireland corpus which relates to the confirmation-eliciting function. It has also been found that this use is more prominent in Irish English than it is in British English which

represents a variety-preferential form of Irish English relative to British English and

American English (Tottie & Hoffmann 2006; Borlongan 2008; Barron *et al.* 2015).

Nevertheless, both Irish English and Canadian English serve the same range of pragmatic

functions in the discourse but with different pragmatic markers or different types of

questions. For example, Canadian English has a variety-preferential use of declaratives

sentences functioning as questions (based on the top 100 and also based on general

concordance items taken and tested in order to see whether this is the case or not); whereas

in Irish English, this does not occur as much as it has been noted in Canadian English. See

the examples 7.6 - 7.8 below for illustration.

Example 7.6

A: It, you know, it seems to me that this would be a big enough space to hold it cos it-
      it's wide here.
B: Oh ya, it is.
A: **You know?**
B: Yup.
A: So it it really could take all of the books and I think that would be a better place for
      them cos they're close to the washing machine.
                              [ICE-Canada, Face to face conversation, S1A-004 #254:1: B]

Example 7.7

A: It's $665, heated.
B: Wow.
A: It's really cheap compared to some of the oth~. We looked at one that was lovely
      but it was $790, **you know?**
B: Yeah.
                              [ICE-Canada, Face to face conversation, S1A-007 #135:1: B]

Example 7.8

A: Or, or you just didn't see anything like that happen?
B: No. No.
A: **You don't know?**
B: No, I don't know.
A: Okay.
                              [ICE-Canada, Legal cross-examinations, S1B-061 #171:1: A]

As can be seen, (based on Tables 7.2- 7.6, the quantitative findings,) Irish English has a variety-preferential use of interrogative sentences. This is the case even with some clear direct questions which are inserted in the discourse as a hedge, a filler, or a hesitation marker as shown with the five-word cluster *what do you call it?* in the examples 7.9- 7.11 below:

Example 7.9

A: Daddy used to ah, **what do you call it?** He used to wear this wig, cos he lost his hair through ringworm when he was really young. So uh [laughs] he used to uhm so he couldn't really cope with being bald at such a young age and that so he got a wig right? And he used to sellotape this wig to his head you know [laughs], double-sided sticky tape.

                                        [ICE-Ireland, Face to face conversation, S1A-051$A]

Example 7.10

A: I was reading my book.
B: What book?
A: My… **what do you call it?**... Coulthard.
C: That is a stupid book. It really is.
A: It makes nice bedtime reading.
C: Oh right yeah. [laughs]

                                        [ICE-Ireland, Face to face conversation, S1A-066$B]

Example 7.11

A: Mm. I won't take it I don't think.
B: And those, the waitressing things are, the waitressing jobs are in- are, **what do you call it?** advertised in college.

                                        [ICE-Ireland, Face to face conversation, S1A-086$B]

Finally, the comparative analysis of questions and tags at the form level (as seen in Tables 7.1 to 7.6) illustrates that Irish English shows a wide range of questions taking place in the Irish discourse based on the analysis conducted through ICE-Ireland and ICE-Canada; in addition, this analysis has shown how Irish English and Canadian English contrast and match in displaying various types of questions as illustrated above.

7.4.2 Functions

As mentioned earlier (in section 7.2.2), there are a number of types in relation to questions serving different pragmatic functions. According to Carter and McCarthy (2006) and Barron *et al.* (2015), engaging and exchanging information between the listener and the speaker is the main focus of many forms of questions, QTs in particular (as discussed in section 7.2). According to Barron *et al.* (2015), information-oriented TQs occur much more (in general) than TQs as desired action-oriented. Information-oriented TQs trigger the exchange of information between speaker and addressee to flow by marking agreement, topic boundary, convergence on an understanding of what is common ground or shared knowledge between participants in the discourse, surprise, and shock. On the other hand, desired action-oriented TQs mark requests, commands, offers, advice, and suggestions (Columbus 2010; Axelsson 2011; Kimps *et al.* 2014, 2018, Barron *et al.* 2015). These pragmatic functions can all be presented in the discourse in forms of questions with either interrogative or declarative statements (as explained above and seen below in this section).

Several identification criteria and clues were used for the functional analysis and interpretation of all questions (and their types) alongside the bottom-up analysis of corpus analysis resulting in an inventory of questions forms (a candidates list) (as mentioned above in section 7.4.1) as follows:

- First is the cooperation of the top-down analysis of the existing literature on questions conducted for this study (see Appendix A).

- Second is the propositional content and the wider linguistic context preceding and following the questions forms.

- Third is the relative knowledge status of speaker and addressee and how the next turn is projected and triggered by the power of the question form.

- Fourth is intuitive values (intuitions) which were used to analyse the intonations of some of the questions forms (such as TQs) (based on the first two identification criteria). This is due to the fact that the ICE corpora do not include phonetic/prosodic annotation which would provide information about prosodic features like stress, intonation, and pauses. Intonation is a very useful and valuable clue in the assignment of function; however there was no access to that crucial information and it had to be inferred. Accordingly, this has been recognised as a limitation of the present study (see Chapter 10, section 10.3).

- Fifth is hearer uptake. Staley (2008, p.5) points out that context and the way in which the hearers react to the discourse is key in interpreting and categorising the pragmatic force of an utterance. Hearer uptaker is used as evidence that the speaker's communicative intention is recognised by the hearer. Barron points out that verbal or non-verbal hearer uptake may be evidenced in corpus data (Barron 2017b, p.97).

Now, the sections below will discuss the types of questions displayed by both varieties of English marking multiple pragmatic functions captured from ICE-Ireland and ICE-Canada.

7.4.2.1 Question tags in ICE-Ireland and ICE-Canada

As discussed earlier, QTs are very common forms in spoken language with a wide range of pragmatic functions (as discussed in section 7.4.2 and shown in Figure 7.1). QTs can be expressed in the discourse in different manners or structures (as discussed in sections 7.2.2). One of the popular types of QTs is fixed tags (or invariant TQs) which can be performed in the discourse, and spoken grammar in general, with a wide range of forms and functions. Thus, it has been found that Irish English and Canadian English use almost the

same forms and functions of information-oriented TQs and desired action-oriented TQs

(with some exceptions as seen in examples 7.22, 7.23, and 7.25). This match between Irish

English and Canadian English is established at the single-word level as seen in examples

7.12- 7.27 below. Yet, with regard to QTs, Irish English has been found to serve the

pragmatic function of QTs with more clusters (i.e. *do you think? do you know? what do you*

mean? you know what I mean?) which do not appear in the top 100 list collected for

Canadian English. Here are some examples (examples 7.12 - 7.21) extracted from ICE-

Ireland with their pragmatic functions identified based on the scheme of the functional

model as presented in Figure 7.1 and Appendix A.

Example 7.12

A: Yeah but the thing is go in and ask
B: Yeah, maybe. If things get much worse maybe.
C: Do you just all want tea **yeah?**
B: And they're fairly bad.
A: Yeah can I've the, one of the herbal ones?
C: Oh.

[ICE-Ireland, Face to face conversation, S1A-063$C]

[*yeah*, in example 7.12, is used by Speaker C to check and make sure that everyone wants tea]

Example 7.13

A: My brother. +But, that's just probably just she's she's around more and she's
       younger than he is so you know. She caused a bit of hassle ++if you know what I
       mean. So it's kind of a, he was a really perfect child. +++You know what I mean
       yeah?
B: Yeah.+ Yeah yeah yeah.++ Yeah yeah yeah yeah.+++ I see. And what school did
       you go to in Dublin?
A: What school?
B: Yeah.

[ICE-Ireland, Face to face conversation, S1A-048$B]

[*yeah*, in example 7.13, is used by speaker A to establish common ground with speaker B who kept saying yeah yeah…]

Example 7.14

A: Yeah I'd say you'd need to be working on part of a team.
B: Yeah yeah I suppose they work in places like Saint John of God's.
C: It's a very academic uhm academic course isn't it **really?**
D: Very.
A: Yeah.
D: It is indeed, yeah yeah.
C: It's a very difficult course.
D: Definitely.

[ICE-Ireland, Face to face conversation, S1A-051$B]

[*really*, in example 7.14, is used by speaker C to mark surprise/shock]

Example 7.15

A: So it's up to your man to let them or not?
B: Yeah.
A: God that's shit **right?**
B: Yeah it's terrible.
A: Uhm.
B: And what do you think of Castlewhite anyway? What's it like living here?

[ICE-Ireland, Face to face conversation, S1A-048$B]

[*right*, in example 7.15, is used by speaker A to establish common ground by stating an opinion about the situation]

Example 7.16

A: But it was the pre~ it was the predominant cause in terms of time of the delay, isn't that **right?**
B: I think probably that is the case.

[ICE-Ireland, Legal cross-examinations, S1B-066$C]

[*right*, in example 7.16, is used by speaker A to state a fact or an opinion]

Example 7.17

A: You owe me about one fifty **okay?**
B: Yeah

[ICE-Ireland, Face to face conversation, S1A-088$C]

[*okay*, in example 7.17, is used by speaker to make a request/command for speaker B to pay]

Example 7.18

A: Well Eleanor what do you think of that?
B: I'm flattered Leah [laughs].

A: So does Michael tell you such nice things, **no?**
B: No [laughs].
A: And uhm were you doing anything nice today? Were you studying or, were you in
        the library?

<div align="right">[ICE-Ireland, Face to face conversation, S1A-061$C]</div>

[*no*, in example 7.18, is used by speaker A to state a fact or an opinion about Michael]

Example 7.19

A: Sure we'll be only there for uh…
B: Oh yeah yeah yeah yeah just a few days yeah mm hmm. Will you be able to count
        that on your TAB account **do you think?**
C: Mm?
A: I suppose so.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-074$C]</div>

[*Do you think*, in example 7.19, is used by speaker B to make a challenge with speaker A
who didn't respond right away]

Example 7.20

A: We wouldn't do a thing like that! [laughs] Oh.
B: Right, yeah.
A: We must get some biscuits to bring to Dawn's house **do you think?**
C: +How long's this meeting going to be on for?
A: I know [unclear speech] have a meeting [unclear speech]+. I don't know. We have
        to have a meeting though.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-066$A]</div>

[*Do you think?* in example 7.20, is used by speaker A to make a suggestion]

Example 7.21

A: I'd say that. There's great characters there like [unclear speech].
B: Oh yeah.And cash is so important to them **do you know?**
A: Right. I suppose.
B: And still.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-055$B]</div>

[*Do you know?* in example 7.21, is used by speaker B to state a fact/opinion to establish
common ground]

On the other hand, we have Canadian English which also exhibits fixed tags in

almost the same way as in Irish English but only at the single-word level (except with *ya*

which appears in the top 100 in ICE-Canada and *eh* which occurs frequently in the ICE-

Canada corpus, but it is not found in the top 100 in ICE-Canada). Now, see examples 7.22-

7.27 below extracted from ICE-Canada.

Example 7.22

A: See if he stays into music. You can also turn it into a a little music room, **ya?**
B: A conservatory.
A: And if you had the computer set up in there you can eventually put in synthesizers.
B: No I just don't think, see^ why everything has to always s-stay the same.
                              [ICE-Canada, Face to face conversation, S1A-079 #178:1: B]

[*ya*, in example 7.22, is used to make a request, acting like *okay*]

Example 7.23

A: Most people would say well it's a little ridiculous to wait half a year.
B: Ya.
A: I'll go find someone else. Like this guy he said "okay I'll wait till spring." Should I
        book you now **ya?**
B: Oh ya.
                              [ICE-Canada, Face to face conversation, S1A-084 #186:1: B]

[*ya*, in example 7.23, is used to make a suggestion]

Example 7.24

A: Eating with people I don't know.
B: Well you'll get used to it after a few meals [laughs].
A: Will I?
B: Shovel it all in like the rest of us do [laughs]. Real country eating **yeah?**
A: Yup [drawls 'yeah-urp'].
B: [laughs]
A: Pass us the mashed potatoes [pronounced 'potaters'].
                              [ICE-Canada, Face to face conversation, S1A-013 #284:2: B]

[*yeah*, in example 7.24, is used to make a conversational joke]

Example 7.25

A: That's what's going on in self-reflection so what you come in with should influence
        what's going to come out.
B: Ya?
A: Ya.
B: Okay. "The extent to which…" All right. It's nice **eh?**
A: Ya it does feel better.
B: Ya it's like, like nice.
A: Mm it has a flow. It has that, more^ of a thread.
                              [ICE-Canada, Classroom lessons, S1B-001 #121:4:B]

[*eh*, in example 7.25, is used to establish common ground, but note that although *eh* occurs a lot in Canadian discourse, it did not appear in the top 100 in ICE-Canada based on the wordlist].

Example 7.26

A: They were right by their cart I know.
B: So everybody was in a fairly small area **right?**
A: Yeah yeah.
B: Within hearing distance of one another?
A: Yeah, yeah should, yeah.

[ICE-Canada, Legal cross-examinations, S1B-062 #108:1: A]

[*right*, in example 7.26, is used to state an opinion to establish common ground]

Example 7.27

A: Yes, but I know we've got it.
B: I haven't seen it in a couple of weeks because I've looked for it. You know, I wonder if we used it camping?
C: I think I remember using it camping.
B: No **really?**
C: Ya.
A: Did we forget it there?

[ICE-Canada, Face to face conversation, S1A-043 #276:2:C]

[*really*, in example 7.27, is used to make or show surprise]

Based on the quantitative findings and qualitative analysis, we can say that Irish English and Canadian English demonstrate very similar forms (or single-word tags) at the single-word level shown with *yeah*, *right*, *really*, and *okay* and also some distinctive forms as well (as seen in Table 7.1). Also, in relation to QT type, we found that both varieties demonstrate only the invariant TQs. As mentioned earlier in section 7.2.2, the difference between invariant TQs and canonical TQs is the non-concordant tag which is not dependent on the syntactic properties of the anchor in the invariant TQs (Allerton 2009, p.310; cf. Andersen 2001, p.104). In other words, the concordant tags have more flexibility in which they can be changeable with their forms (as discussed in section 7.2.2) (Columbus 2010;

Axelsson 2011; Kimps *et al.* 2014, 2018, Barron *et al.* 2015). In relation to invariant TQs,

Canadian English shows only one invariant TQ (*do you?*) existing in the top 100. Whereas

Irish English, on the other hand, shows a wide range of forms of invariant TQs such as, *do*

you? did you? is it? are you? isn't it?, existing in the top 100 (as shown in Table 7.2). For

more illustration see examples 7.28 - 7.35 below with their intonation being inferred based

on the context.

Example 7.28

A: In this letter he's stating he is sexually abused as a child.
B: Correct. Correct.
A: You have no better knowledge than me as to which one of those is true **do you?**
B: Exactly.
A: And what you did point out that there were other allegations of sexual abuse in these letters referring to what happened to him at the institution.
                         [ICE-Canada, Legal cross-examinations, S1B-068 #149:1: A]

[Checking whether what has been said or proposed is true or not by stating a fact or opinion about the situation]

Example 7.29

A: Ya but then I uh, taking down walls means that you gotta, got to^ move shelf space doesn't it?
B: But you don't have any shelves on those walls, **do you?**
A: I do.
B: Oh, you do?
A: On the inside ya.
B: Okay.
A: I've got uhh those plastic milk crates you know, all lined up you know here, and I've got stuff in them.
B: Oh ya? Well those could still be lined up against the back wall.
                         [ICE-Canada, Face to face conversation, S1A-004 #230:1: A]

[Stating a fact or opinion for the sake of checking whether what has been said or proposed is true or not]

Example 7.30

A: Uh I think so yeah. Well we we don't know what to do because we're invited to Sarah's as well on the same night. And uhm, and we we're just a bit torn between the two of them.
B: Ah no, don't not go now. I knew you wouldn't come!

A: Why?

B: Cos I [tut-tut] I bet you now, you're not going to come, **are you?**

A: I don't know, that's what I'm saying. We're torn. Why?

B: Cos I just knew you wouldn't come.

A: She said I knew you wouldn't come.

                    [ICE-Ireland, Telephone conversation, S1A-099$B]

[*Are you* used here, in example 7.30, to mark a challenge based on their feeling about the situation]

    Example 7.31

A: Oh you have to keep them the whole year **do you?**

B: No I think they bought them about four or five weeks ago.

A: God I should have get my father to do that.

                    [ICE-Ireland, Face to face conversation, S1A-069$C]

[Stating a fact or opinion for the sake of checking whether what has been said true or not]

    Example 7.32

A: We were just all over the town and everything dresses

B: Oh you went shopping **did you?**

A: Yeah depression session. +Nothing

B: [laughs] Why? Was there not?++

                    [ICE-Ireland, Telephone conversation, S1A-099$B]

[It is clear for speaker B that speaker A went shopping but *did you* is used, in example 7.32, to establish common ground and more engagement]

    Example 7.33

A: Not a big bald patch **is it?**

B: Well you know that kind of hair that goes from here to here though. Does he have that?

                    [ICE-Ireland, Face to face conversation, S1A-051$E]

[Stating a fact or opinion/guess for the sake of checking whether what has been said true or not]

    Example 7.34

A: Yeah, someone in the department **is it?**

B: Yeah. We've got two s~ examiners. Sweat.

C: Mm.

                    [ICE-Ireland, Face to face conversation, S1A-063$C]

[Stating a fact or opinion for the sake of checking whether what has been said true or not]

Example 7.35

A: So, it say if you weren't repeating the exams what would you, would you do? Say at the weekends, or, it's a Thursday night to go out in Cork usually **isn't it**?

B: No, no no no. Well, during college yeah, Thursday night is the, night because all the other kids go home uhm, uhm at the weekends. But uhm during the summer it's more Saturday night. Everybody goes out on a Saturday night.

[ICE-Ireland, Face to face conversation, S1A-057$B]

[Stating a fact or opinion for the sake of checking whether what has been said true or not]

Now, we can say that at the single-word level Irish English and Canadian English (from the formal level), are very similar in relation to single-word tags (fixed tags or invariant TQs) (as discussed above in section 7.4.2.1); however, this is not the case at the higher formal level. Irish English has shown some variety-specific/preferential clusters used as TQs (i.e. *do you think? do you know? did you? is it? are you? isn't it? what do you mean?* and *you know what I mean?*) which are not found in the top 100 most frequent clusters list collected for Canadian English (see Appendices D and I). This suggests that Irish English speakers are very listener-oriented in that they seem to check in with the listener a lot to ensure that the listener understands or converges with their message. From the functional level, both varieties have revealed much more information-oriented TQs than desired action-oriented TQs which are equally low in both functional samples. Furthermore, as mentioned earlier, information-oriented TQs have been expressed in Irish English with a wider range of forms, while desired action-oriented TQs were expressed in Irish English with *okay* (as in example 7.17) and *do you think?* (as in example 7.20) and Canadian English with *ya* (as in examples 7.22 and 7.23). Lastly, it was found that almost all the TQs in the corpora (ICE-Ireland and ICE-Canada) occur in turn-final position as seen in the provided examples (7.12- 7.35). Now, we will turn to the second type of questions found in ICE-Ireland and ICE-Canada.

7.4.2.2 Echo questions in ICE-Ireland and ICE-Canada

Echo questions are also known to be one of the common type of questions occurring frequently in conversations (and spoken language in general). Echo questions (as mentioned earlier in Chapter 2, section 2.4.4) occur in the form of declarative word order and with a wh- word. Their typical pragmatic function in the discourse is to request more clarification about a segment in the discourse (Biber *et al* 1999; Carter & McCarthy 2006; for a more expansive definition see Chapter 2, section 2.4.4). These types of questions have been used more commonly as RTs in the discourse of Irish English and Canadian English based on ICE-Ireland and ICE-Canada. With this type of question, the prior speaker still has the floor, and it is not taken over by the addressee. As mentioned in Chapter 6, both varieties show very similar forms such as: *what? who?* and *when*? and some distinctive forms such as *which*? which appears in the top 100 in Irish English but not in Canadian English, and *how*? which appears in the top 100 in Canadian English but not in Irish English, see the examples 7.36- 7.40 below.

Example 7.36

A: Yeah, okay I'm going to Mass tonight.
B: I, I've to go to Mass tomorrow.
C: **When?**
B: **What?**
C: You'd nearly want want to check first I suppose with Julie to see on account of the weekend.

[ICE-Ireland, Face to face conversation, S1A-081$B]

Example 7.37

A: Daniel Delaney's working there.
B: Is she? Oh that's right, yeah.
C: **Who?**
A: Daniel Delaney.

[ICE-Ireland, Face to face conversation, S1A-089$B]

Example 7.38

A: No that's fine. I haven't gotten around to listening to that tape yet. My little brother
saw it   and nicked it on me.
B: **Which?**
A: The tape you gave me on Friday to to have a listen to. My little brother~ I just took
it in, I   just took my stuff out of my jacket.
B: Oh right.
                              [ICE-Ireland, Face to face conversation, S1A-052$C]

Example 7.39

A: You don't wanna~ want to take the dog do you?
B: **When?**
A: Can you take the dog for like, ehm a month [laughs].
                              [ICE-Canada, Face to face conversation, S1A-026 #122:1:C]

Example 7.40

A: [laughs] So when are you going to go down there?
B: **What?**
A: When are you gonna~ going to go down there, to Windsor?
                              [ICE-Canada, Telephone conversation, S1A-099 #71:1: B]

Now we can say that Irish English and Canadian English display similar forms in

relation to echo questions with the two distinctive forms (*which?* in Irish English and *how?*

in Canadian English which both display the same pragmatic functions as discussed earlier).

The third type of questions found in ICE-Ireland and ICE-Canada are the follow-up

questions which will be discussed below in section 7.4.2.3.

7.4.2.3 Follow-up questions in ICE-Ireland and ICE-Canada

Follow-up questions are commonly used in the discourse with the typical pragmatic

function of signalling engagement and attention. They are mainly performed by the listener

(but can sometimes be used by the speaker as well) to encourage a continuation of the

conversation or to expand the discourse by requesting further specification (see example

7.41) (Carter & McCarthy 2006). Follow-up questions contain a variety of different types

(as discussed in detail in Chapter 2, section 2.2.4 and briefly in section 7.2.2). In addition to

that, according to Biber *et al* (1999) and Carter and McCarthy (2006, 2017), follow-up

questions can be viewed and classified as DMs, RTs, QTs in terms of their position, forms,

and functions as opposed to echo questions which mainly adopt the response token position

as seen in examples 7.36- 7.40. Thus, follow-up questions can adopt the same pragmatic

function of QTs, marking surprise as shown in example 7.43 and establishing common

ground as illustrated in example 7.48 (with the first *you know*). This is due to the fact that

they occur as turn-final devices which signal that the speaker is yielding the conversational

floor and projecting a response by the addressee (which is the adopted classification of QTs

in this present study). However, follow-up questions are usually more like the RTs

indicating supportive responses (see *what do you mean?* in examples 7.43 versus 7.42 and

the second occurrence of *you know* in example 7.48). Thus, based on the quantitative

findings and qualitative analysis, Irish English demonstrates more follow-up questions

compared to Canadian English which use less follow-up questions. For more information

see Chapter 6, sections 6.4.2.2 and 6.4.2.3; also see the examples 7.41 and 7.43 below.

    Example 7.41

A: Ah we went up to Canada alright, to Niagara Falls. Canadian side, lovely. Just for a few hours, on one of our weekends off.
B: **Did you?**
C: Is Niagara Falls Canadian?
A: Well the we went to the Canadian side. It's meant to be more- more nicer.
        [ICE-Ireland, Face to face conversation, S1A-065$A]

    Example 7.42

A: I've to stay in and do this project and then I'm going to watch Peter's Friends.
B: Yeah me too. Take off my boots and dry my feet.
A: **What do you mean** dry your feet? Are they wet from walking home still?
B: Yeah [laughs].
A: My God, and Caitriona and and Angela are coming up tomorrow.
B: Are they?
A: Yeah.
        [ICE-Ireland, Face to face conversation, S1A-070$B]

Example 7.42 is an interesting example and appears to be made in jest –Speaker B is telling Speaker A that he needs to take off his boots and dry his feet – Speaker A clearly understands the propositional content of what Speaker B has stated so the first question is a signal of engagement (bordering on fake disbelief) all in the context of a humorous sequence. In comparison to example 7.42, we have example 7.43 below in which *what do you mean?* acts as a QT (a fixed tag/invariant TQ) which both act to manifest an attitude of surprise to an unexpected segment in the discourse.

Example 7.43

A: Did you get the yoke fixed on the phone?
B: Yoke? **What do you mean?**
A: The code switch.
B: Well we've had one false call today.
A: What?
B: We've had one false call today.

[ICE-Ireland, Face to face conversation, S1A-060$B]

Furthermore, we have some interrogative sentences/statements necessitating the addressee to respond. Although they appear as direct questions structurally, pragmatically they are not really questions; based on their pragmatic functions in the discourse they are more like DMs (see Chapter 5). However, according to Carter and McCarthy (2006, 2017), they are preface questions which occur with the pragmatic function of stating what the speaker considers newsworthy or important information for the listener which sometimes involves a shifting of the topic or the focus of the discourse (see the examples 7.44 and 7.45 below).

Example 7.44

A: Yeah. They must have thought you'd met a man inside and proposed and everything.
B: You wouldn't in a polling booth dear.
A: You mi~ well **do you know what?** You're not going to meet him in college. You might meet him in a polling booth.

C: Well it's a small enough place

Example 7.45

A: That was well put in there now.
B: Did you say anything? Does anybody want some jam?
A: Oh I'd love some. **Do you know what?** It's a pity there's no fireworks. [laughs] Do you want me to mark down our soundplay?
C: No.

Also, in relation to the pragmatic function of showing what the speaker considers newsworthy or important information to the listener and primarily checking comprehension and indicating that the speaker and hearer share a point of view (or common ground), it has been found that Irish English adopts more clusters as questions with both types of sentences: interrogative and declarative. Whereas Canadian English has only the clusters *you know?* and *you know what?* appearing in the top 100 most frequent clusters; see the examples 7.46 and 7.47 below.

Example 7.46

A: In oh yeah yeah. Better being on a bike than in a bus anyway.
B: Yeah.
A: Cos you can dodge them, like you can go up on the footpath. **You know what I mean?**
C: You can hang onto trucks.

Example 7.47

A: Because they may not be claiming, paying their taxes or, see most the majority^ of people on Rent Allowance are okay but there are some people and they just give the others a bad name. It's like in all walks of life.
B: That's al~ that's always the way like.
A: **Do you know what I mean?**
B: Yeah.

These examples also underscore a strong orientation to the listener on the part of the speaker.

As discussed before (in section 7.4.1), according to Barron *et al* (2015) Irish English employs more interrogative sentences (as questions) than declarative sentences compared to British English. This is also the case when compared with Canadian English, as found in the present study. Canadian English leans towards using declarative sentences as questions (in general) (as shown in examples 7.6- 7.8) rather than using actual questions or interrogative clauses (as mentioned briefly in section 7.4.1). Comparatively, Irish English generally uses both structures (declarative sentences and interrogative sentences) in relation to questions (i.e. [declarative + rising intonation] *you know?* vs [interrogative clause] *do you know?* and [declarative + rising intonation] *you know what I mean?* vs [interrogative clause] *do you know what I mean?*). Keeping this in mind, *you know?* and *you know what I mean?* sometimes occur as questions in Irish English with the expectation of receiving an answer from the addressee. This is inferred based on the context in which these utterances have occurred due to the rising or falling intonation used in that context. However, not every declarative statement used for the same pragmatic functions with the same clusters can be viewed as a question. For example, in example 7.48 *you know* acts as a QT in the first occurrence and a RT in the second occurrence and in example 7.49 *you know* acts as a DM (and not as a question) below for more illustration:

Example 7.48
A: That could've been the drink as well now Michelle.
B: What?
A: The whole oh he's a friend like. You know it's fine. It's fine. **You know?**
B: No.
A: **You know**.
B: No.
A: No. Not anymore.
B: Oh I do. I do but I think it's actually helped me to get over him.
                                        [ICE-Ireland, Face to face conversation, S1A-068$C]

Example 7.49

A: What time do you usually go down for re- rehearsals?
B: Uh in the evening. A-after [unclear speech].
C: Sure you don't go to the rehearsals
B: No, I usually +go down
A: After lunch+ After lunch til four, or half three.
B: See they're taking more classes now like **you know**. They're under pressure now to
        get finished.
A: Oh right.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-072$B]</div>

[For more examples refer back to example 7.6 in which *you know* acts as a RT and example

7.7 in which *you know* acts as a QT with the pragmatic function of monitoring the state of

shared knowledge in the conversation].

7.5 Conclusion

In this chapter, we have presented the quantitative and qualitative analysis of questions and

tags used in Irish and Canadian Englishes in relation to forms and functions that have been

captured in the ICE-Ireland and ICE-Canada corpora. Methodologically, word and cluster

lists were generated for both corpora to identify and compare the forms used in both

datasets. From these lists, questions (and their types) with pragmatic functions were

identified manually by cross-checking qualitatively using concordancing. This resulted in

interesting quantitative findings at the formal level. Irish English and Canadian English

match the most at the single-word questions level as has been seen frequently at the single-

word level of DMs and RTs. The single-word questions have been mainly used as tags and

echo questions. It is clear that Irish English displays a wide range of questions whereas

Canadian English shows far fewer; this has resulted in significant contrasts seen in

question tags and follow-up questions (supporting the findings of Barron *et al* 2015 as

mentioned above in sections 7.4.2.1 and 7.4.2.3). This reflects on the notion of engagement

and convergence in the discourse in which questions are not the common forms to perform these typical pragmatic functions, especially at the high level of clusters, due to the fact that Canadian English does not show four, five, or six word question clusters. Additionally, Canadian English has the tendency to use declarative statements as questions rather than interrogative clauses, whereas in Irish English both types of question formations are used commonly (supporting the findings of Barron *et al* 2015 as mentioned above in sections 7.4.1 and 7.4.2). However, in relation to echo questions with wh-words, both varieties have demonstrated very similar usage in their discourse. Finally, Irish English tends to perform the pragmatic functions of all the types of questions (under investigation) with much more variety-specific/preferential clusters (as displayed in Tables 7.2 -7.6) which are not found in the top 100 most frequent clusters list collected for Canadian English (see Appendices D and I).

CHAPTER 8

HEDGES IN IRISH AND CANADIAN ENGLISHES

8.1 Introduction

This chapter presents the comparative analysis of the usage of hedges in Irish and Canadian

discourse which have been captured in the ICE-Ireland and ICE-Canada corpora. This is

done by detailing the main quantitative findings which resulted from both datasets through

the theoretical and methodological framework taken in this study (as explained in Chapters

3 and 4) in order to unpack and analyse these findings qualitatively and comparatively. The

analysis of the pragmatic variation that occurs in the use of hedges between Irish English

and Canadian English will be investigated mainly from the perspective of forms, since the

core function of hedges is the same, regardless of variety, as opposed to the pragmatic

markers (PMs) discussed in the preceding chapters, which vary in terms of forms as well as

pragmatic functions. Hedges appear in many different forms (as noted by Clancy 2010)

such as:

- Closed class grammatical sets:

  1. Modal verbs: *could*, *might*

  2. Nouns: *possibility*

  3. Adjectives: *possible*

  4. Adverbs: *possibly*, *maybe*

- Syntactic markers: Question tags, passives.

- Pragmatic markers: *I think*, *just*, *sort of*

- Rhetorical devices: Understatement, vagueness/approximation.

- Paralinguistic features: Stutter, hesitation, false start.

Also, in addition to the grammatical, syntactic, and pragmatic markers which have

been the focus in most analyses of hedging across many different linguistic disciplines, it is

important to include other forms such as negation, reporting devices, and prefaces to speech acts, further broadening the range of forms with the potential to function as a hedge (Carter & McCarthy 2006). Hedges are one of the fundamental linguistic elements in communication and interaction in the discourse. According to Skelton (1988):

> Without hedging, the world is purely propositional, a rigid (and rather dull) place where things are either the case or are not. With a hedging system, language is rendered more flexible and the world more subtle... Language without hedging is language without life. (p.38)

We will now look at the existing literature on hedges.

8.2 Previous Research

A lot of work has been done on hedges and hedging from theoretical, empirical, and applied perspectives. This work was a part of the larger focus on the interactional language features (existing in over a thirty-year period) (Hyland 1996a, 1996b; Clemen 1997; Schröder & Zimmer 1997; Crompton 1997, 1998; Lindemann & Mauranen 2001; Mauranen 2004). For the present purposes of this study, hedging/hedges will be reviewed in terms of definition, forms, possible sub-pragmatic functions, different approaches and perspectives on hedges, and lastly psycho-affective aspects.

8.2.1 Hedging: Definition and origins

First of all, throughout the research literature, hedges have not adopted any widely accepted definition due to the divergence in approach to the nature and realisation of hedging. However, traditionally, hedges were considered to be semantic modifiers or approximators in the spirit of the original definition by George Lakoff (1972) who introduced hedges to linguistic research through his influential work (p.195; Farr & O'Keeffe 2002). Lakoff introduced the term "hedge" in order to describe the lexical expression or phrase "whose

job it is to make things fuzzier or less fuzzy" (1972, p.195). Lakoff was primarily concerned with the semantic contribution that hedges make to the phrases or statements in which they occur especially given that hedges can weaken or strengthen category membership (Loewenberg 1982, p.196). In other words, they affect assertions of category membership. Thus, when it is claimed that *a penguin is a **sort of** bird*, the penguin's semantic connection to the category of birds is more true than the unhedged statement *a penguin is a bird* (Farr & O'Keeffe 2002; Clancy 2010). This goes along with Rosch's (1973, 1978) view of hedges as linguistic devices that modify prototypical category membership. Nevertheless, Lakoff was also interested in other hedges such as *regular* which can be illustrated in the examples below:

Example 8.1:   Esther Williams is a fish.

Example 8.2:   Esther Williams is a regular fish.

Lakoff, claims that example 8.1 is false since clearly Esther Williams is a human being and not a fish. However, in example 8.2 it can be viewed differently due to the use of the word *regular* which invokes characteristics attached to the word fish while simultaneously negating the literal meaning. In doing so, Lakoff drew attention to the relationship between meaning and connotation, thus beginning the process of establishing that any adequate treatment of hedges must consider the context within which they occur (Clancy 2010).

Consequently, based on the work of Lakoff, Prince *et al* (1982) used data taken from Physician-Physician interaction and suggested that hedges should be classified into approximators and shields. Approximators are the hedges that affect the truth conditions of propositions as in the following example: *His feet were sort of blue*. This phrase indicates that the speaker is fully committed to the truth of the proposition s/he is conveying. Shields,

on the other hand, do not affect the truth conditions; rather, they reflect the speaker's commitment to the truth value of the whole proposition as seen in the following proposition: *I think his feet were blue* in which the phrase *I think* marks a level of uncertainty on the part of the speaker in that s/he does not fully believe what they are saying. Similarly, Hübler (1983) approached hedges by drawing a similar distinction between understatements and hedges in which understatements are corresponding to approximators, and hedges are corresponding to shields (Markkanen & Schröder 1997). For example, these following sentences have been viewed differently:

Example 8.3:   It's a bit cold in here.

Example 8.4:   It's cold in Alaska I suppose.

Example 8.3 contains an understatement, whereas example 8.4 contains a hedge. However, in examining and reviewing these categories, a concern was raised by Markkanen and Schröder (1997) who questioned the usefulness of these divisions. As a result, Skelton (1988) claims that the distinction between approximators and shields is only maintainable in the abstract since shields appear to have an indefinitely large potential domain (which may include approximators) and can comfortably extend over more than one sentence (p. 39). Similarly, Hyland (1994) completely removes shields from his categorisation of hedges.

8.2.2 Hedging within the scope of pragmatics and discourse

The emergence of research focusing on the pragmatic aspect of hedges in discourse has been reflected implicitly as we saw with the approach adopted by Rosch (1978) mentioned earlier. Such an approach is rooted in cognitive science in which "semantic grasp" has preceded analysis at the level of discourse, and therefore discounts language function

(Clemen 1997; Farr & O'Keeffe 2002). Therefore, we see research questions focus more on *why* hedges are used and offering reasons for such a phenomenon such as politeness, indirectness, vagueness and understatement. This is when the concept of hedging explicitly moved from its origins to the area of discourse analysis and pragmatics. According to Markkanen and Schröder (1997), the most frequent motivating factor or trigger for hedges is politeness (p.10). As a result, we see the work of Brown and Levinson (1978) on politeness strategies in which they provided a framework for investigating the role of hedging in domains, such as mitigation and indirectness. This results in the understanding that hedges are context-dependent and are integral to face saving strategies based on this approach. We also have the work of Channell (1990), Clemen (1997), and Markannen and Schröder (1997) who examine pragmatic strategies and their linguistic components in terms of hedges from various perspectives.

In addition to that, we see many researchers attempting to reclassify and subcategorise what have traditionally been collectively known as "hedges" in order to reflect on the pragmatic component of hedges. As mentioned earlier in section 8.2.1, Prince *et al* (1982) tend to the classification under which hedges should be divided into *shields* (those performing a pragmatic function) and *approximators* (those performing a semantic function) and Rounds (1982) adds *diffusers* to this (Rounds views them as dispersers or "diffusers" of disagreement) (Farr & O'Keeffe 2002). As we said in section 8.2.1, hedges have been categorised using various terms and definitions based on the perspective they are being viewed from. Thus, Markannen and Schröder (1997) state that "through extension the concept has lost some of its clarity and sometimes seems to have reached a state of definitional chaos, as it overlaps with several other concepts" (p.15).

Nevertheless, the use of hedges has been investigated and studied in relation to gender in which gender has become, and is considered to be, essential to the nature and use of hedges. The study of Preisler (1986), following in the Lakoff (1975) tradition, argues that women hedge more than men because their speech is more tentative and less assertive. However, this claim or viewpoint has become more contested over time and is now considered controversial due to the amount of research based on real speech data which has failed to support such a claim (Holmes 1986, 1990, 1993; Bradac *et al* 1995; Dixon & Foster 1997). Lakoff's original proposals were based primarily on hypothesis and personal observation which have not only been challenged, but many findings now suggest that the contrary may in fact be true (Farr & O'Keeffe 2002). Therefore, evidence on the effect of gender on the use of hedges remains inconclusive and inadequate (Farr & O'Keeffe 2002). However, many recent works have been done on cultural constraints on the use of hedges. An example is found with Crismore *et al*'s (1993) study across socio-cultural borders comparing the American and Finnish contexts. There is also the fascinating and innovative study of Hinkel (1995) examining the use of modals on a comparative and contrastive basis between native and non-native users of English in a written context which revealed that there are considerable socio-cultural constraints on the pragmatics associated with modality.

Also, the use of hedges and intensifiers has been examined and studied by several researchers in relation to their impact and effects on the listener in terms of features such as: attractiveness, authority, credibility and so forth. However, the results are contradictory and inconsistent which made it hard to compare and contrast the findings due to the different approaches and dissimilarities in empirical procedures adopted by the researchers

such as Bradac *et al* (1995), Holmes (1990) and Hosman (1989). Nevertheless, these studies are highly systematic in approach and comprehensive in nature (Farr & O'Keeffe 2002). Additionally, there are some specific language domains that have formed test-beds for looking into how and why hedges are employed by native speakers. This results in considerable research into the use of hedges in academic texts (Rounds 1982; Fahnestock 1986; Myers 1989, 1992; Salager-Meyer 1994; Hyland 1994, 1996a, 1996b; Hinkel 2005; Algi 2012; Takimoto 2015; Wang & Tatiana 2016). For example, we have the significant corpus-based study into the use of hedging in a professional spoken context conducted by Prince *et al* (1982). Based on their corpus of 12 hours of physician to physician talk, the study shows that the most salient linguistic feature, in terms of frequency, is hedges whose definition is a word or phrase "whose job it is to make things fuzzier" and that definition of hedges was taken from Lakoff (1972, p.195). Based on that definition, they found between 150 to 450 hedges per hour and more than one every fifteen seconds.

In addition to that, we have the approach of Brown and Levinson (1987) on hedges, in which they view a hedge as "a particle, word, or phrase that modifies the degree of membership of a predicate or noun phrase in a set; it says of that membership that it is partial, or true only in certain respects, or that it is more true and complete than perhaps might be expected" (p.145). For instance, the two-word cluster hedge *I think* (labelled a quality hedge by Brown & Levinson 1987, p.164) permits the speaker to avoid full responsibility for the truth of their utterance by distancing both themselves and the listener from the act in order to satisfy or redress the hearer's negative face. Therefore, hedges downtone the illocutionary force of an utterance in which they allow the speaker to weaken their commitment to its propositional content. Yet, hedges have a lesser role in terms of

positive politeness in which linguistic actions are targeted at building on indices of solidarity such as in-group membership or changing extremes on the value scale such as *beautiful* or *revolting*. Thus, Brown and Levinson (1987) claim that in the utterance *it's really beautiful in a way*, the hedge *in a way* allows the speaker to avoid the precise communication of their attitude by leaving it up to the addressee to figure out how to interpret it (pp.116-117). This type of pragmatic application in the discourse is applied and maintained by using one of these hedges in which the speaker calls upon the hearer to use the common knowledge between them to interpret speaker attitude, while maintaining positive face (of speaker, i.e. being liked and approved) while at the same time appealing to both the hearer's positive and negative face (Clancy 2010) by, for example, allowing them to agree or disagree, or acknowledge shared knowledge or not.

Moreover, the focus of research on hedging has shifted from casual, everyday spoken language to both spoken and written language in the academic context. For example, we have the work of Myers (1989) who found that hedges (in academic writing) function in a similar manner to spoken discourse. This is due to the findings that hedging in academic writing is dependent on the same variables that govern everyday spoken interaction such as social distance, power difference, and rank of imposition. Thus, hedging is seen as an essential linguistic tool not only in the casual, everyday spoken language but also in the art of academic and scientific discourse. In terms of function, we see that the main function of hedging in academic writing is one of negative politeness in which the presentation of new knowledge and ideas is downtoned or mitigated. So, according to Myers, academic knowledge establishes a face threatening act to other researchers in the field because it encroaches on their freedom to act (1989, p.13). Consequently, hedges can

function as signals, signalling that new knowledge is being presented as being provisional, pending acceptance in the literature, or acceptance by the community (Myers 1989). As Hyland (2000) states, writers in general seek modification and adjustments for the assertions that they make through toning down the uncertain or potentially risky claims, emphasising what they believe to be correct, and conveying appropriate collegial attitudes to the readers. Therefore, Myers (1989) claims that a sentence bearing a claim without hedging/hedges is unlikely to be a statement of new knowledge. In other words, writers simultaneously save their own faces through hedging and the use of hedges which help avoid imposing on the reader's faces. Additionally, hedges can also increase the credibility and reliability of a statement (i.e. in academic texts) due to their mitigating effect (Clemen 1997). This has triggered the raised question regarding the reason and the motivation behind the use of hedges among writers. Hence, Markkanen and Schröder (1997) argue that "hedges can be manipulated by using them to disguise writer attitude in the sense that the reader is left in the dark as to who is responsible for the truth value for what is being expressed" (1997, p.9; Clancy 2010, p.214). Research done in this domain has also contributed to an increased understanding of the multiple functions that hedges can perform. So, Hyland supports the contention that "hedges are polypragmatic, conveying a range of different meanings often at the same time. As a result, they do not fit into a neat scheme of discrete categories which allows one meaning to be clearly distinguished from others" (1996b, p.437). This can be seen in other research contexts as well. According to Mauranen (2004) who conducted a micro-level examination of the functions of epistemic and strategic hedges in spoken academic discourse, some hedges fulfil the criteria for one category while the context induces the other interpretations or functions.

8.2.3 Hedging within the model of positive and negative politeness

Politeness theories (which have been known as a set of linguistic theories that relate linguistic action or behaviour captured in social settings) can play a big role in the way hedges are viewed or analysed (Al-Hindawi & Alkhazaali 2016). For example, speakers of Irish English tend towards indirectness in answering questions, and this can be seen in the way Irish people rarely answer a polar question with a single word answer (*yes* or *no*) because it is considered to be too direct and impolite (Asián & McCullough 1998; Farr & O'Keeffe 2002). This goes back to the socio-cultural norm of avoiding over assertiveness. Thus, very frequently in our data, we find that both Irish and Canadian speakers (Irish speakers to a greater extent) tend to mitigate their speech. However, traditional politeness theories (the notion of face-saving) have been criticized by many scholars on different grounds. Thus, there are many critiques, and the most prominent critiques are as follow (O'Keeffe *et al* 2020):

One: Some researchers (Brown & Levinson 1978, 1987; Ide 1989; Werkhofer 1992; Marriott 1993; Spencer-Oatey 2000; Fukada & Asato 2004; Chang 2008; Cutrone 2011; LoCastro 2012; Wijayanto *et al* 2013) criticize traditional politeness theories on the basis that they have a Western-centric bias and therefore do not account for the nuances that are present in other cultures, societies, and groups. These critics claim that this "face-saving model" is based on the philosophy of individualism whereby each partaker in a conversation is concerned with protecting their own "individualistic" rights (Brown & Levinson 1978). However, the same is not true for Eastern communities which place a heavier emphasis on collectivism (Werkhofer 1992; LoCastro 2012). An example of this would be Japanese societies where politeness strategies are positive ones used in everyday utterances (Marriott 1993; Cutrone 2011). This is because there is a greater importance

placed on people's interdependence and reciprocal relationships and the judgement of appropriate behaviour and speech is directly related to this notion (Ide 1989). So, while an individualistic society celebrates the self, a collective society celebrates social relationships. The maintenance of those relationships is directly related to the negative politeness culture where speakers use hedges to mitigate any sense of hassle or burden on the listener.

Two: Another criticism of more current theories on politeness is their neglect of "impoliteness" research (Locher & Bousfield 2008; Culpeper 2008; LoCastro 2012; Chapman 2013). Proponents of this criticism argue that politeness studies must include the absence of acts of politeness language and expressions (Culpeper 2008). Sociolinguists and pragmatists are especially interested in this branch in order to understand hostile, conflictual and argumentative use of language (Archer *et al* 2012). Researchers have coined this branch "Impoliteness Theory" and hope to use it in a multidisciplinary and transdisciplinary approach (Chapman 2013).

Three: Another criticism of this theory involves its lack of exploring and distinctively identifying related notions and concepts (tact, civility, respect etc.) and the need for it to differentiate between linguistic and non-linguistic politeness (Fraser 1990b).

Four: Similar to the first criticism, there are also critics who take issue with the correlation made between politeness and indirectness. A common belief is that indirect speech is thought to be polite and this is certainly the case in most English communities (Leech 1983). However, other communities use directness and indirectness in their speech to convey different things. Mexican, German, and Polish speakers adopt directness in their speech to show affiliation and solidarity with whomever they are speaking with (Felix-

Brasdefer 2005; Ogiermann 2009). In Arabic, indirect speech is often considered to be impolite and directness is preferred (Archer *et al* 2012).

Five: Yet another criticism of politeness or "face-saving" theories is that they assume that certain linguistic impressions are always representative of polite speech while failing to account for irony, sarcasm, or tone. The word *please* is an example of this where the intonation can drastically change the meaning of the word (Fraser 1990b; Watts 2003; LoCastro 2012).

This literature review has shown how complicated any investigation of hedging can be due to the various perspectives and approaches on hedging, especially an empirical, corpus-based one. Therefore, any investigation of hedging should be guided by a caveat, especially considering that hedging can be achieved in an indefinite number of surface forms (Brown & Levinson 1987). However, determining the function of the PMs of hedges (in other words, PMs that function discretely as hedges) can be assisted through a number of features such as: syntactic (see Holmes 1985, 1990; Lenk 1998; Oh 2000), prosodic (see Holmes 1985, 1990), lexical (see Lindemann & Mauranen 2001; Aijmer 2002) and stylistic (see Miller & Weinert 1995; Cheng & Warren 2000). We also have a number of socio-pragmatic indicators such as ethnicity (see Cheng & Warren 2000; Youmans 2001; O'Sullivan 2004; Fung & Carter 2007), sex (see Maltz & Borker 1982), socio- economic background (see Huspek 1989), age (see Erman 2001; Macaulay 2002) or speaker relationship (see Östman 1981; Lee 1987; Markkanen & Schröder 1997; Ruzaitė 2007) that may impact on the hedging function. Nonetheless, all the above-mentioned studies agree on supporting the belief of Fraser (1999) that hedges (and PMs in general) have a specific core meaning that is procedural and not conceptual, and any other (or more)

specific interpretations (or meanings) are negotiated by the context, both linguistic and conceptual. (Also, see section 8.3 below for the research and the analysis limitations on hedges adapted in this current study).

8.3 Methodology

This section outlines the main framework of analysis and methodology for this chapter. Our focus in the present framework (as explained in Chapter 3) is the formal level; thus, the pragmatic variation analysis of the use of hedges in Irish English and Canadian English will be approached and investigated from two perspectives: forms and functions. In terms of functions in this present study, we adopted the core pragmatic function of hedges introduced by Carter and McCarthy (2006), presented in the top-down framework for the analysis of the forms and pragmatic functions of spoken grammar (as shown in Appendix A). Thus, hedges are the expressions and markers used by speakers in order to downtone the assertiveness of a segment of discourse (Carter & McCarthy 2006). In other words, these expressions (or PMs) enable speakers to be less assertive and blunt in formulating their message. Methodologically, corpus linguistics is used as a tool in order to conduct this analysis quantitatively and qualitatively (as explained in Chapter 3), by using the iterative approach. Top-down analysis was based on a framework for spoken grammar generally (hedges specifically) based on the existing literature while the bottom-up process was based on micro-analysis of the data.

In terms of form, the quantitative findings of the hedges across both varieties, Irish English and Canadian English, will be presented comparatively based on their word-cluster lists (from one single-word hedges to six-word cluster hedges). As a result (as elaborated in

Chapter 4), wordlists and cluster analyses were generated to identify and compare the hedges used in both datasets (ICE-Ireland and ICE-Canada). Within the form analysis, we will also present the hedges across both varieties which appear as other PMs at the same time.

In terms of function, the quantitative findings of the forms of hedges will be unpacked and analysed qualitatively by concordancing them manually (as discussed in Chapter 4) in order to identify and determine their core pragmatic function in the discourse adopted in this present analysis. (To review the iterative approach in detail, see Chapter 4, also see Appendix A).

8.4 Results and Analysis

The Tables 8.1-8.6 (below) are a summary of the comparative analysis of the forms of hedges in ICE-Ireland and ICE-Canada which demonstrate where Canadian English and Irish English can meet and depart in terms of hedges. As detailed in Chapter 4 and shown in Appendices N and O, these tables below have been derived by the qualitative analysis (from a bottom up route) of the top 100 most frequent word and cluster lists in which each single word and cluster has been concordanced manually in order to affirm which items from the quantitative findings (from the top 100 most frequent word and cluster lists presented in Appendices B, C, D, E, F, G, H, I, J, K, L, M) are actually acting like hedges in their pragmatic function in both datasets. Note that the light gray shaded cells in the tables below indicate the distinctive forms (hedges) between Irish English and Canadian English and the unshaded cells indicate the forms (hedges) that both varieties of English share in common.

Table 8.1. *ICE-Ireland and ICE-Canada single-word hedges occurring within the top 100 most frequent word lists (frequency per 10,000 words)*

ICE-Ireland	Freq[6]	ICE-Canada	Freq
just	34	just	42
like	52	like	56
now	39		

Table 8.2. *ICE-Ireland and ICE-Canada two-word cluster hedges occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
You know	33	You know	47
I think	21	I think	19
Kind of	14	Kind of	11
I'd say	3	Sort of	9

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 8.3. *ICE-Ireland and ICE-Canada three-word cluster hedges occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq	ICE-Canada	Freq
I don't know	7	I don't think	3
		I'm not sure	1

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 8.4. *ICE-Ireland and ICE-Canada four-word cluster hedges occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
Or something like that	1	Or something like that	1
That kind of thing	1	Well I don't know	0.34
I'm just going to	0.34	uhm I don't know	0.22
		But I don't know	0.22

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

---

[6] The order of the forms presented in Tables 8.1- 8.6 is random and does not indicate rank.

Table 8.5. *ICE-Ireland and ICE-Canada five-word cluster hedges occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq	ICE-Canada	Freq
to be honest with you	0.22	I'm just gonna (going to)[7]	0.28
what do you call it	0.18		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common.

Table 8.6. *ICE-Ireland and ICE-Canada six-word cluster hedges occurring within the top 100 most frequent cluster lists*

ICE-Ireland	Freq
And all that kind of stuff	0.22
at the end of the day	1

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

As seen above in Tables 8.1 to 8.6, the comparative analysis of hedges has been conducted at the form level in which we can see the similarities and the differences between Irish and Canadian Englishes in relation to their forms of hedges and their occurrences (when needed for comparison purposes) which will all be analysed below based on their word clusters.

8.4.1 Single-word hedges

As we have seen in the previous chapters (Chapters 5-7) with discourse markers (DMs), response tokens (RTs), and questions and tags, Irish English and Canadian English demonstrate the most resemblance at the single-word level and after that the contrasts in relation to forms begin to appear more clearly. Thus, at the single-word level and based on the quantitative analysis, we see the common forms *just* and *like* (which have appeared

---

[7] *I'm just gonna* is referred to as a five-word cluster because it embeds five words within its phonological contraction: *I am just going to.* Also note that *I'm just gonna* is not found in the top 100 list of clusters found for Canadian English. The similar phrase appears as *I'm just going to* in Irish English while in Canadian English it only shows as *I'm just gonna going to* in the data. Refer to section 8.4.5 below for more information on this cluster.

previously as DMs) being shared by both varieties of English (Irish and Canadian).

However, based on the qualitative analysis, we see that *just* and *like* agree in forms across

both varieties but differ in their distribution. For example, the single-word hedge *just* has

been found to be more frequent in distribution as a hedge in Canadian English than it is in

Irish English. In Irish English, the hedge *just* is associated and found more within the four-

word cluster *I'm just going to*, while *just* in Canadian English demonstrates more variety in

its distributions as illustrated below with the following examples from the data.

Example 8.5

A: And he passed away like it was, awwww.
B: [laughs]
A: But when you think about it [laughs].
B: Yeah I was **just** a little devastated when the little one went cos I I kept, because I had kept nursing her.

[ICE-Canada, Face to face conversation, S1A-059 #226:2: B]

Example 8.6

A: So the end result was there used to be three courses, there are still three courses— they're **just** a little different. So the students in Office Management Technology at our place now do this, this plus the three oh eight.

[ICE-Canada, Unscripted speeches, S2A-024 #44:1: A]

Example 8.7

A: When it comes time to…
B: You gonna ^going to^ roll it up first?
A: No I'm **just** gonna ^going to^ throw it downstairs. You know how thick that thing is?

[ICE-Canada, Face to face conversation, S1A-013 #157:2: A]

Example 8.8

A: My name is Mike Murphy and I'm a student in UCC College and I'm here staying in Castlewhite Apartments which are apartments on campus. And uhm I'm talking to a friend of mine, Aisling. So **I'm just going to** ask her a few questions as to what she's doing down in Cork, although she's leaving now however. Uh, so what's your name? Your full name.
B: Uhm, Aisling Ruane.

[ICE-Ireland, Face to face conversation, S1A-057$B]

Nevertheless, *just* in Irish English is displayed as a DM much more than it is in

Canadian English performing the pragmatic function of marking emphasis (by boosting or

downtoning emphasis or attention) on a segment of the topic in the discourse (as explained

in Chapter 5). For more illustration, see the examples below. Examples 8.9, 8.10 and 8.11

are boosting while example 8.12 shows *just* functioning to downtone.

Example 8.9

A: Uhm say for example uhm a person where one of the spouses is working and they're **just** not eligible for unemployment assistance.

[ICE-Ireland, Legal presentations, S2A-066$A]

Example 8.10

A: There's no variety and nobody really wants to talk and everybody's just drinking and dancing and posing and everything. It's **just** no good.
B: So, you don't really go out any more down in Cork?

[ICE-Ireland, Face to face conversation, S1A-057$A]

Example 8.11

A: We had a couple of rules which helped us along but we did need some thinking as well. Now in practice if you have a large and complicated graph you very often will not be able to, or will not want to, find out the chromatic number. It's **just** too complicated a problem to colour the graph with with great effort using the minimum number of colours.

[ICE-Ireland, Unscripted speeches, S2A-037$A]

Example 8.12

A: There comes a point where everything you touch becomes cute when Rachel seems, +ah no, cute describes you.
B: Ah no it's a lovely+ dress.
A: It's **just** short.
C: What colour is it?

[ICE-Ireland, Face to face conversation, S1A-058$C]

Similarly, we have the pragmatic marker *like* (included in Table 8.1) which has

been considered as a DM in this study with the pragmatic functions of focusing attention,

organizing the narrative, and qualifying a preceding statement (as elaborated upon in depth

in Chapter 5) as well as a hedge. *Like* as a pragmatic marker generally, and as a DM/a

hedge specifically, has been the subject of several studies in which its uniqueness has been highlighted. For example, Schweinberger (2012) who found (in an analysis of ICE-Ireland) that gender differences in the use of clause-final *like* appear noticeably in which older male speakers employ this form more than older female speakers. The peak age for using such a DM was 26-33 years. Schweinberger (2015) also observes how clause final *like* occurs "substantially more often" in Irish English than it does in other varieties (p.117) and is the most frequent form of *like* in this variety; he concludes the clause-final *like* can serve as a marker of "Irishness" (p.132).

One of the pragmatic functions found for *like* in clause-final position is to act as a mitigator, and this function is a variety-specific use of *like* in Irish English (see Kallen 2006; Schweinberger *et al* 2009; Lucek 2011; Diskin 2017; and Schweinberger 2020). Thus, based on the analysis of ICE-Ireland and ICE-Canada, this pragmatic function has been expressed and performed only in Irish English (but not in Canadian English which proves the uniqueness of the variety-specific use of *like* in Irish English) as demonstrated in the following examples 8.13- 8.15.

Example 8.13

A: But you know you've, ^you're^ near the college and you've got all the facilities you need like a washing-machine and so on, so it's grand **like**. It does does the job.
[ICE-Ireland, Face to face conversation, S1A-048$B]

Example 8.14

A: I I don't think we quite got that. I I I think we we we got an opinion that it might have been the gentleman who's now deceased. But I mean, can I ask you the question **like**, I mean who was it? You're you're the former chairman.
[ICE-Ireland, Legal cross-examinations, S1B-068$A]

Example 8.15

A: Do you reckon? Has there been anything **like**?
B: Yeah yeah. Now do Jeanette.
A: I swear on my mother's life, I mean
B: Swear on the holy Bible.

A: Listen.
B: Yeah.
[ICE-Ireland, Face to face conversation, S1A-049$B]

It is notable that in examples 8.14 and 8.15, *like* is used after interrogative structures, where it softens the directness of the questions. This is in line with Kallen (2006) who observes how clause final *like,* as a mitigator, can be used in interrogatives as well as affirmatives.

In Canadian English, on the other hand, *like* (as a single form) is used much more as a filler marking hesitation and pauses in the speech and a DM with the pragmatic functions of focusing attention by giving or requesting an example. This is in comparison to its use as a hedge in Irish English in which *like* performs more pragmatic functions (as illustrated above). Nevertheless, hedges (and PMs in general) have a specific core meaning that is procedural and not conceptual, and any other (or more) specific interpretations (or meanings) are negotiated by the context, both linguistic and conceptual (Fraser 1999), see the following example 8.16 from ICE-Canada:

Example 8.16
A: And it got. It finally got to the midterm and he literally. I think he passed out in
class or he…
B: Really? Wow.
A: **Like** he was an old sort of…
B: [laughs]
[ICE-Canada, Telephone conversation, S1A-095 #48:1: A]

In example 8.16, *like* can be seen as a hedge where Speaker B does not want to directly say that the person is old, so hedges at the beginning of the sentence with *like* and at the end of the sentence with the two-word cluster *sort of.* On the other hand, *like* can also be viewed as a hesitation marker in which Speaker B was hesitant about whether or not he should say that the person is old. Therefore, *like* can be difficult to interpret in terms of its

266

core pragmatic function from its secondary functions/meanings. (See also section 5.4.2 for the pragmatic functions and analysis for the DM *like*). There are many studies that have been conducted on individual hedges pointing towards the fact that hedges can perform different pragmatic functions, often at the same time. For example, Holmes (1985, 1986, 1990, 1993) has shown how commonly identified hedges such as *I think*, *you know*, and *sort of* can perform a number of different, though closely related, functions in casual conversation between men and women. In relation to the current study, we can see many PMs across both varieties displaying various pragmatic functions simultaneously, as will be illustrated in the section below (8.4.2) with *you know*, *I think*, *sort of*, and so forth. Therefore, we can conclude that PMs such as *like* and many others can act as hedges and also as other PMs once the distinction between their core pragmatic functions and their minor or side pragmatic functions are made. This type of distinction will be discussed more with the two-word PM *you know* and *I think* in section 8.4.2.

However, before moving on to the two-word cluster hedges, we will look at the pragmatic marker *now* which is known to be a salient feature in Irish English with variety-specific uses and functions (which does not appear as a hedge in Canadian English). The pragmatic marker *now* has been investigated by many researchers (i.e. Brown & Levinson 1987; Schiffrin 1987; Bolinger 1989; Hirschberg & Litman 1993; Biber *et al* 1999; Aijmer 2002; Carter & McCarthy 2006) and still further by Clancy and Vaughan (2012) whose analysis of LCIE reveals that *now* in Irish English shares many pragmatic functions existing in Irish and British Englishes (for a list of these functions and to see extracts of *now* as a DM from ICE-Ireland, see section 5.4.2.1). However, *now* in Irish English has two specific pragmatic functions viewed as variety-specific uses/functions that are denoted only to Irish

English. The two variety-specific forms are the *hedging now* and the *presentative now* (Clancy & Vaughan, 2012). The *hedging now* is used by Irish English speakers (in clause-final position, usually in informal contexts) to downtone the illocutionary force of face-threatening acts. These can include: challenges as in *that's not fair now*; disagreements as in *you can't say that now*; evaluations as in *I'd say I'm crap now;* direct questions as in *How many Euros would that be now?;* and orders as in *Hold on a minute now* (Clancy & Vaughan, 2012, pp.236-238). Also, Clancy and Vaughan (2012) argue that such uses of the pragmatic functions of *now* are performed to minimise power and add to "the emphasis on solidarity and corollary downtoning of power, both actual and conversational" in Irish society (p.240). In relation to the present study based on the analysis of ICE-Ireland, the *hedging now* is used with a downtoning function (downtoning the illocutionary force of face-threatening acts), as in the following examples:

Example 8.17

A: Uh I think so yeah. Well we we don't know what to do because we're invited to Sarah's as well on the same night. And uhm, and we we're just a bit torn between the two of them.
B: Ah no, don't not go now. I knew you wouldn't come!
A: Why?
B: Cos I [tut-tut] I bet you **now**, you're not going to come, are you?
A: I don't know, that's what I'm saying. We're torn. Why?
B: Cos I just knew you wouldn't come.
A: She said I knew you wouldn't come.

[ICE-Ireland, Telephone conversation, S1A-099$B]

Example 8.18

A: Six what?
B: Six+
C: Sixteen+
B: Six zero+
C: Ugh yuck! Jimmy you're never getting married boy.
B: Six, six zero [laughs]
D: Wait a minute **now** uhm

[ICE-Ireland, Face to face conversation, S1A-089$D]

Example 8.19

A: Next I'm going to show you how to make a completely different kind of starter. I'm going to do a sorbet. **Now** a sorbet is actually a water-ice like an ice-cream but made completely without cream. So you can serve it either as a starter. You can serve it as a dessert for example if you had a very creamy main course or something spicy like a curry.

[ICE-Ireland, Demonstrations, S2A-056$A]

In example 8.17, *now*, is used in clause final position to downtone what would otherwise be a very direct challenge to the interlocutor (i.e. *I bet you now; you're not going to come*); similarly in example 8.19, it is used in the same position to mitigate an order suggesting disagreement with the interlocutor in the form of imperative clause + *now* (Clancy & Vaughan 2012, p.238); *Wait a minute now* (see Clancy & Vaughan 2012, pp.236-238). By contrast, in example 8.19, *now* is used in clause final position (see Aijmer 2002); it is not used to mitigate an FTA but rather it is used to soften the directness of the instructional language and thereby to minimise power (Clancy & Vaughan 2012, p.240) (***Now** a sorbet is actually a water-ice*). As proposed by Aijmer (2002), *now* is used in this way to promote a friendly overton in conversation (Clancy & Vaughan 2012, p.239).

On the other hand, we have what is called the *deictic presentative now* (Grenoble & Riley 1996, p.820), which is associated with Irish English (Clancy & Vaughan 2012, p.241); this function of *now* involves the use of *now* as a lexicalised pointing device as in the following example 8.20 (in which *now* makes *money* salient):

Example 8.20

Speaker 1:   How much is that?
Speaker 2:   One fifty. **Now**. Thanks.
(Clancy & Vaughan 2012)

This function of *now* as shown in example 8.20 was found by Clancy and Vaughan (2012, p.241) to account for only a small number of occurrences in LCIE, nevertheless they highlight it as a function that merits further investigation.

Now, we will turn to the two-word cluster hedges.

8.4.2 Two-word cluster hedges

At the two-word hedges level, we can see clearly where Irish English and Canadian English meet and part in terms of forms as presented in Table 8.2. However, Irish English and Canadian English both show distinctive characteristics in relation to hedges not only with the forms that they do not share but also with the forms that both varieties share in common. This can be seen with the two-word-hedge *you know* which can serve a variety of different, though closely related, functions in the discourse such as: expressing speaker confidence or certainty, expressing uncertainty of various kinds, or monitoring and managing the ongoing discourse through "monitoring the state of shared knowledge in the conversation" (Holmes 1986; Carter & McCarthy 2006, p.221). So, based on the qualitative analysis and the adopted definition of hedges (as presented in Appendix A), Canadian English uses *you know* as a hedge more frequently than Irish English. In Canadian English, *you know* can act as a hedge in slots of RTs and DMs with the pragmatic functions of being less assertive and blunt in formulating their message or a segment of discourse (see examples 8.22- 8.24 below). This can be done through the speaker and addressee sharing their mutual background and experience which may eventually result in confidence between the speaker and addressee. On the other hand, Irish English uses *you know* as a hedge in slots of DMs but not in slots of RTs. As mentioned in the previous chapters and shown in Table 8.2, the frequency of the two-word PM *you know* varies across both

varieties. See the examples below illustrating the distribution of *you know* acting as a hedge across both varieties (for more information on *you know*, see section 5.4.2.2 and Figure 5.4 in Chapter 5 and section 6.4.2.2 in Chapter 6):

Example 8.21

A: They're not, nuh- virtually none of them are helping me with what I'm, with what my field is, except because I make them.
B: Mm hmm. Mm hmm.
A: **You know**
B: Mhmm ya, well…
A: And some of them I can't.

[ICE-Canada, Face to face conversation, S1A-037 #78:1: A]

Example 8.22

A: Like you're in a family right+ an- and you're, you think you're the weird one++ in the family because you're the artist and everyone else does, you know, other stuff that involves **you know** nine to five jobs or something, and families ya
B: [laughs] This sounds strange.+
A: [laughs] Okay.++

[ICE-Canada, Face to face conversation, S1A-032 #124:1: B]

Example 8.23

A: Do you like it?
B: Yeah, but ah you get tired of it after six hours like +**you know**
A: Yeah.+ And do they pay you well or what?
B: Fiver an hour.
A: It's not too bad.
B: Yeah.
A: Sure 'tisn't.
B: No it's grand. One of those hours I was lifeguarding so I got three pounds an hour. That's grand like. But teaching is better **you know**. Passes the time much more quickly.

[ICE-Ireland, Face to face conversation, S1A-046$A]

Example 8.24

A: And so uhm he was like he's been I've been meeting him like all week. I've me- met him all last week. I'd say we've been uh like meeting and he'd be we we'd be drinking coffee and whatever **you know**
B: Yeah.

[ICE-Ireland, Face to face conversation, S1A-049$A]

Based on the quantitative and qualitative analyses of *you know*, we see that Irish English uses it less often with the frequency of 33 per 10 000 words, whereas Canadian

English uses it more often with the frequency of 47 per 10 000 words (see Table 8.2). In addition, Canadian English adopts more (functional) distributions of *you know* (as mentioned above), whereas in Irish English, it usually appears either in the middle or at the end of the clause. Also, it can be said that (as discussed in Chapter 5), Canadian English uses *you know* more as a hedge that draws on mutual shared knowledge as well as downtoning the assertiveness of a segment of discourse while Irish English uses it more to downtone the assertiveness of a segment of discourse as well as to draw on mutual shared knowledge.

Also (as can be seen from Table 8.2), the hedge *I think* is one of the shared forms existing in both varieties of English. The PM *I think* is one of the PMs that perform multiple pragmatic functions by which it can be classified as either a hedge or a stance marker (as will be explained in details in Chapter 9 on the use of stance markers across both varieties). However, the PM *I think* in general (in terms of propositional knowledge) can express both certainty and uncertainty which can be classified under two types; the first is the **deliberative** *I think* which may convey a speaker's careful deliberation in regards to what they are saying with some type of authority associated with it (cf. Preisler 1986; Aijmer 1997; Vaughan 2009) and the second is the **tentative** *I think* which may express tentativeness and uncertainty and can also weaken a segment or an assertion in the discourse that might be too direct (Holmes 1985, 1990; Aijmer 1997, p.21). Holmes (1985, 1990) mentions that *I think* varies demonstrably in terms of intonation as well as syntactic position in an utterance. For example, *I think* can add weight to an utterance when it is at the initial position of the clause with level stress on *think*. It can also be used tentatively with a fall-rise intonation (Vaughan 2009). *I think* in medial and final positions (as often

found in ICE-Ireland and ICE-Canada) were classed as tentative, even if they had prosodic prominence (Aijmer 1997, p.21). Aijmer found that the tentative *I think* was more frequent than deliberative *I think* in the Lund Corpus (LLC) (as in our finding in ICE-Ireland and ICE-Canada). Therefore, both Irish English and Canadian English have shown the same usage of the PM *I think* as hedges and *I think* as stance markers in relation to function. But, in relation to *I think* as a hedge, both varieties have shown the semantic category of tentative more as opposed to the semantic category of deliberative which has been associated more with stance markers based on the datasets. See examples 8.25 and 8.26 below in which *I think* is used as a tentative hedge in both varieties.

Example 8.25
A: Another honeymoon.
B: Great.
A: Huh?
B: Where to?
A: Egypt. [laughs] Cairo, **I think**+
C: Morocco.+
B: Seriously, he's gone away again.

[ICE-Ireland, Face to face conversation, S1A-087$B]

Example 8.26
A: Out of that, I didn't get a call until August, fifteenth **I think**? Around August
       fifteenth. So the whole summer passed before I got a call. And then I got two job
       interviews.

[ICE-Canada, Unscripted speeches, S2A-040 #44:2: A]

Furthermore, as can be seen in Table 8.2, the hedge *kind of* is one of the prominent forms of spoken grammar used as a hedge in everyday spoken language in order for speakers to downtone the assertiveness of a segment of discourse (Carter & McCarthy 2006, p. 223). Irish English and Canadian English commonly use the hedge *kind of;* however, based on the qualitative and quantitative analysis, it appears more in Irish English than Canadian English. This can be inferred from the frequency variation of the hedge *kind*

of (as shown in Table 8.2) and the use of the hedge *sort of* which exists only in Canadian English but not at the top 100 most frequent clusters in Irish English in which both *kind of* and *sort of* have been found to be interchangeable with their pragmatic functions as hedges in Canadian English. So, based on the frequency of *sort of* (9 per 10k) and *kind of* (11 per 10k) in Canadian English in comparison to *kind of* (14 per 10k) in Irish English, we can conclude that Canadian English uses *kind of* twice more than it is used in Irish English. See the examples 8.27- 8.33 below:

Example 8.27

A: Her neck muscles have all gone locked… No it's nothing like a cramp. It's like uhm…
B: Sit very close.
A: Uhm what's it like? If you've not had it, it's really **kind of** hard to explain it. It's like something **kind of** just going so tight that you can't actually move. You know that it, it just locks. You+ know if you ever get something locked?
C: Right+. Yeah.

<div align="right">[ICE-Ireland, Face to face conversation, S1A-053$A]</div>

Example 8.28

A: Yeah and I think it is, it's a tricky issues alright+ because. Yeah…
B: No cos you.+ Cos you don't, are you the therapist or are you **kind of** a dictator?
A: Exactly and that you kind of assume maybe that at that age they will be able to participate in a group without resorting [laughs] to kicking you and kicking each other.
C: Yeah.

<div align="right">[ICE-Ireland, Classroom lessons, S1B-017$C]</div>

Example 8.29

A: You know with the creaky wood floors and they've had you know they could buy housecoat there that **kind of** a thing.
B: Oh, ya.

<div align="right">[ICE-Canada, Broadcast discussions, S1B-037 #150:4: B]</div>

Example 8.30

A: The paper say this happened at around eleven P M on Saturday night.
B: Kinda ^**Kind of**^ reckless isn't it?
A: Well it s- sounds like it.

<div align="right">[ICE-Canada, Broadcast discussions, S1B-037 #5:1: A]</div>

Example 8.31

A: So the fact that they were prohibiting this stuff suggests that they thought it was occurring.
B: Uh huh.
A: It's kind of difficult to say, I guess how uhm, sorta ^**sort of**^ bad priests were by the, by the, the standards of the church hierarchy so…
B: Uh huh.

<div align="right">[ICE-Canada, Face to face conversation, S1A-025 #173:1: A]</div>

Example 8.32

A: She's, she just **sort of** decided to do the uhm Florida thing only over there. Like+ six months of winter over there and six months here.
B: Yeah yeah.+ Yeah.

<div align="right">[ICE-Canada, Face to face conversation, S1A-011 #114:1: A]</div>

Example 8.33

A: And it got. It finally got to the midterm and he literally. I think he passed out in class or he…
B: Really? Wow.
A: Like he was an old **sort of**…
B: [laughs]

<div align="right">[ICE-Canada, Telephone conversation, S1A-095 #48:1: A]</div>

Lastly, at the two-word hedges level, we have the hedge *I'd say* which exists in Irish English but does not exist at the top 100 most frequent clusters in Canadian English. This aligns with Farr and O'Keeffe's (2002) finding, in relation to LCIE, that *I would say* and its contracted form *I'd say* are used as hedges more frequently by Irish speakers than by British or American speakers.

Example 8.34

A: Is she old?
B: She, she'd be in uh uh uh, **I'd say** her late forties, early fifties.
A: And she's riddled with arthritis already?

<div align="right">[ICE-Ireland, Face to face conversation, S1A-056$D]</div>

8.4.3 Three-word cluster hedges

Irish English and Canadian English, at the single and two-word hedge levels, demonstrate common forms and similarities in the way those hedges are used in their discourse. However, at the three-word level, we can see that each variety displays distinctive forms of hedges (see Table 8.3). At this level, we can see that the pragmatic function of hedges has been performed by interesting PMs which have been previously explored as other PMs such as RTs and DMs. For example, in Irish English (as discussed in Chapter 6), the three-word cluster *I don't know* has been displayed as a RT in which it fills the response slot, occurs initially, and does not take the floor from the speaker. Yet, the same three-word cluster appears in Irish English to act as a hedge with the pragmatic function of downtoning the assertiveness of a segment of discourse. For illustration, see the example 8.35 below:

Example 8.35

A: Depends lighter, lighter. He wanted me to. He said, "ah your hair'd look really good if you if you dyed it."
B: What, black?
A: Uhm no, uhm kind of.
C: Highlighted?
B: Green?
A: More, more red or something, **I don't know**.
                                        [ICE-Ireland, Face to face conversation, S1A-071$B]

On the other hand, we have Canadian English showing other distinctive forms of hedges such as *I don't think* and *I'm not sure.* These perform the pragmatic function of downtoning at a more epistemic level where the assertiveness of a clause or longer segment in the discourse is downtowned in terms of certainty of the speaker rather than protecting the face of the listener. In other words, it marks uncertainty more than mitigation of pragmatic force. See the examples 8.36 and 8.37 below:

Example 8.36

A: I mean the weekends were gonna ^going to^ be when we do something else.

B: I could take maybe a couple of extra days off. Okay? **I don't think** I'm gonna
      ^going to^ take a whole week off. Okay? I, **I don't think** I have [unclear speech]. I
      got four weeks of vacation okay this year.
A: Mm hmm.
B: I don't feel like giving up a whole week of my vacation to uhm do up our house.
                    [ICE-Canada, Face to face conversation, S1A-005 #113:1: A]

      Example 8.37
A: Well then how do you know if she's going to Burlington with Bill?
B: Well I don't know. I, she… I think she mentioned it or something. **I'm not sure.**
      I'm not sure.
A: Well what are you gonna ^going to^ do with that place, with that bed?
                    [ICE-Canada, Face to face conversation, S1A-011 #126:2: B]

      In summary, based on the quantitative and the qualitative analysis of ICE-Ireland

and ICE-Canada, we see that at the three-word level, Irish English and Canadian English do

not share any hedges. This is also the case at the five and six-word levels which will be

elaborated upon in the sections below.

8.4.4 Four-word cluster hedges

In relation to the four-word hedge level, it was found that both Irish English and Canadian

English have only one PM in common, namely *or something like that* (as displayed in

Table 8.4). The four-word cluster *or something like that* can also be classified as a DM

functioning like *and so on* or *and so on and so forth* (which have both been discussed in

Chapter 5), as a vagueness marker. However, based on the context on which pragmatic

function heavily relies, the distinction has been made between the four-word cluster *or*

something like that and *and so on* or *and so on and so forth*. The four-word cluster *or*

something like that goes further than closing down the ongoing discourse (or the idea) by

denoting and including similar entities in terms of quantity or quality to a hedge, thereby

downtoning the assertiveness of a segment of discourse; whereas the phrases *and so on* or

and so on and so forth do not indicate this. Both varieties have shown the four-word cluster

or something like that in the discourse acting as a DM as well as a hedge (as explained above). Yet, on the surface, while *or something like that* is a DM, it appears also to be a straightforward vagueness marker, which could be substituted by *and so on*. Therefore, it has the same function as *apples, oranges and so on* but the speaker, in choosing to use it in these examples, is doing so also because they want to sound more purposefully vague. Essentially, these examples would be adequate without *or something like that* but by using it the speaker is mitigating the force of the message through vagueness (O'Keeffe 2003; Cheng & O'Keeffe 2014). For more illustration, see the following examples 8.38- 8.40 extracted from both datasets:

Example 8.38

A: No it's like you know we've clinics every week. Well as far as I know they come in and observe you for a cli- full clinic, and then I think maybe they ask you questions **or something like that.**

B: Yeah.

[ICE-Ireland, Face to face conversation, S1A-063$B]

Example 8.39

A: Black area only, and all of us whites, a crowd of whites going in to clean it up.

B: Are you serious?

A: We'd, we'd to go up like, it was up on the tenth floor or whatever, and you you'd go in on the lift in the morning, past all the blacks and stuff. You'd get a lift up and there'd be a special metal door protecting us from the rest of them, below us you know. We'd clean in packs, in the morning. We went in at s- about what is it seven in the morning **or something like that** and left by about three in the afternoon because all, before all the gun-toting crazies got up out of bed.

[ICE-Ireland, Face to face conversation, S1A-065$A]

Example 8.40

A: Like it's easier for us to pick up. I find it's easier for TESL people to pick up a book on, on uh ecology and and, or biology **or something like that** and read it and figure out what it is we have to do and know how to teach a lesson on it than it is for a biology person to think that oh my god how is someone who doesn't speak my language gonna ^going to^ learn this.

[ICE-Canada, Unscripted speeches, S2A-038 #107:1: A]

Furthermore, this type of pragmatic function has also been performed more in Irish

English with the four-word cluster hedge *that kind of thing* which is not found at the top

100 most frequent clusters in Canadian English as seen in the example 8.41 below:

Example 8.41

A: What do you mean you don't like watching the box? I love watching the box. I love
        watching soaps, and uh [coughs] what else do I like? I like, I don't actually watch
        that much. But what I do watch I really enjoy. I watch soaps and I watch things like
        what Beverley Hills Cop nine-o-two-one-o and **that kind of thing,** but I mean,
        apart from that I know, I'm certainly not glued to it.
B: No.
                                                    [ICE-Ireland, Face to face conversation, S1A-059$B]

On the other hand, as Table 8.4 shows, Canadian English presents distinctive forms

of hedges that are not occurring as hedges in Irish English based on the qualitative analysis

of ICE-Ireland and ICE-Canada as in the examples below. These all include either a

vocalisation (e.g. *uhm, uhh, uh)*, conjunction (e.g. *but)*or discourse marker (e.g. well) used

with *I don't know* and suggest a pattern where the force of a disagreement introduced by *I*

don't know is being hedged. Examples 8.42 to 8.44 show vocalisations that prime *I don't*

know:

Example 8.42

A: So, so can you tell quickly just what, what happens next?
B: Well **uhm I don't know**. The people have been getting letters, The CLSC [Centres
        Locaux de Services Communautaires] there's over fif- fifty communities, over fifty,
        well over.
                                                    [ICE-Canada, Parliamentary debates, S1B-050 #24:2: B]

Example 8.43

A: Look at how the Lord has blessed us. They're so beautiful, aren't they?
B: They are.
A: Ya.
A: **Uhm I don't know** there it was, it just grabbed me by the throat. I said to Morris
        write this, I- I'm going to follow on Monday and I'm going to volunteer on Friday.
                                                    [ICE-Canada, Face to face conversation, S1A-003 #38:1: A]

Example 8.44

A: Uh, it's uh. **Uhm I don't know**. I just think Zellers or something like that. Stupid thing.

[ICE-Canada, Face to face conversation, S1A-074 #85:1: A]

In example 8.45, we see the disagreement turn begins with a vocalisation (*uhh*), follow by an assertion that is then mitigated with *but I don't know* to downtone the force of the assertion.

Example 8.45

A: I, like, drums in Let It Be?
B: Uhh I think they're light ones **but I don't know**. But something else we could see.

[ICE-Canada, Face to face conversation, S1A-009 #146:2: A]

Example 8.46 places *Well I don't know* at the beginning of the turn and this is followed by the assertion. This is further hedged with a repetition of *I'm not sure.*

Example 8.46

A: Well then how do you know if she's going to Burlington with Bill?
B: **Well I don't know**. I, she… I think she mentioned it or something. I'm not sure. I'm not sure.
A: Well what are you gonna ^going to^ do with that place, with that bed?

[ICE-Canada, Face to face conversation, S1A-011 #125:2: B]

The four-word cluster *well I don't know* (in example 8.46) is a shared PM in both varieties of English. It appears in Irish English as a DM with the pragmatic function of repairing a segment in the discourse and a RT with the pragmatic function of marking agreement, topic boundary, or understanding of the prior turn in the discourse. On the other hand, *well I don't know* in Canadian English has been used only as a RT with the pragmatic function of marking agreement, topic boundary, or understanding of the prior turn in the discourse and a hedge. As a hedge, *well I don't know* marks less assertiveness in relation to a statement in the discourse or downtones the assertiveness of the emphasised segment of the topic in the discourse, making it less direct, as in example 8.46. Comparatively, the

four-word cluster *uhm I don't know* appears only in Canadian English in the top 100 most

frequent clusters but not in the top 100 of word clusters in Irish English based on the

quantitative analysis.

8.4.5 Five-word cluster hedges

At the five-word hedges level, Irish English has shown hedges that are not found in the top

100 most frequent clusters in Canadian English (see Table 8.5 and Appendices G and L).

See the examples below extracted from ICE-Ireland:

Example 8.47

A: Six inches [laughs]. Uhm yeah I don't know. Well sure you could try it on. When is
       it? Next Wednesday?
B: Yeah. Does it fall from the hip?
A: Ye~ uhm I can't remember **to be honest with you.**
B: You know my obsession with that.

[ICE-Ireland, Telephone conversation, S1A-099$B]

Example 8.48

A: Daddy used to ah **what do you call it**[8] he used to wear this wig cos he lost his hair
       through ringworm when he was really young. So uh [laughs] he used to uhm, so he
       couldn't really cope with being bald at such a young age and that so, he got a wig
       right? And he used to sellotape this wig to his head you know [laughs] double-sided
       sticky tape!

[ICE-Ireland, Face to face conversation, S1A-051$A]

Based on the qualitative analysis, these two five-word hedges have been found in

Irish English acting as other PMs. For example, the five-word cluster *to be honest with you*

performs as a stance marker with the pragmatic function of indicating the speaker's stance

or attitude towards the messages (or a segment in the discourse) (as will be discussed in the

---

[8] Note that the five-word cluster, we acknowledge that *what do you call it* could function as a discourse
marker marking hesitation and or as a hedge (or a vagueness marker) in which the speaker knows very well
what they want to say but they are trying to use the most appropriate phrase for what is coming next as a
politeness strategy. As an analysis of a transcript, it can sometimes be a matter of speculation based on the
context

upcoming Chapter 9). Also, we have the five-word cluster *what do you call it* which acts as

a DM marking hesitation (indicating that the speaker is thinking), a filler, (as illustrated in

examples 8.49 and 8.50 and discussed in Chapter 7), or a hedge (as shown in example 8.45)

in which the speaker is trying to use the most appropriate phrase for what is coming next

for politeness purposes (as discussed earlier on politeness and vague category markers, see

section 8.4.4).

    Example 8.49

A: I was reading my book.
B: What book?
A: My, **what do you call it?** Coulthard.
C: That is a stupid book. It really is.
A: It makes nice bedtime reading.
C: Oh right yeah. [laughs]

                [ICE-Ireland, Face to face conversation, S1A-066$B]

    Example 8.50

A: Mm. I won't take it I don't think.
B: And those, the waitressing things are, the waitressing jobs are in- are, **what do you call it?** advertised in college.

                [ICE-Ireland, Face to face conversation, S1A-086$B]

In relation to Canadian English at the five-word cluster level, *I'm just gonna going*

to acts exactly like *I'm just going to* in Irish English which has raised the question as to

why is it presented in this form based on the generated cluster list. The answer based on the

qualitative analysis is that *I'm just gonna going to* is always repaired to *I'm just going to* in

the course of the utterance in the Canadian discourse (see example 8.7 as an example)[9].

Also, the qualitative analysis has brought to light the fact that Canadians frequently self-

correct their speech as they speak where this has not been noticeable with Irish speakers

---

[9] The phrase *I'm just gonna going to* is marked with self-correcting normalization mark-ups in the data, showing that it is a phrase being corrected by the speaker (Nelson 2002, p. 9).

based on both the corpora used in this study (ICE-Ireland and ICE-Canada). (For more illustration see examples 8.7, 8.36, 8.37, 8.40, and 8.46).

8.4.6 Six-word cluster hedges

 At the six-word clusters level and even at the five-word clusters level, we can see clearly throughout the previous chapters that Irish English has repeatedly shown more clusters as pragmatic markers such as DMs, questions, and hedges. In relation to hedges, there were only two six-word clusters functioning as hedges (see Table 8.6) and they were both found in ICE-Ireland, namely *and all that kind of stuff* and *at the end of the day*. *And all that kind of stuff* occurs in Irish English not only as a DM but also as a hedge with the pragmatic function of marking less bluntness in closing down the idea, the topic, or the conversation as a whole. On the other hand, the DMs *and so on* and *and so on and so forth* are acting more as a DM with the core pragmatic function of closing down the idea by denoting similar entities. They do so without considering the core pragmatic function of hedges, which has more concerns attached to it such as politeness, indirectness, vagueness, understatement and assertiveness. For more illustration see example 8.51 below in which Speaker B uses the phrase *and all that kind of stuff* in order to give the impression of having more knowledge than he or she actually possesses. This stands in juxtaposition to the phrase *and so on and so forth* where the connotation is that the speaker does have more knowledge on the subject but does not have the time in that moment to elaborate. This pragmatic marker *and all that kind of stuff* can be used for the sake of politeness in order to keep the conversation going amongst all the speakers. This is illustrated in example 8.51, with speakers C and A responding to speaker B's statement, with their knowledge about

283

occupational therapists. However, at the end, Speaker A, who initiated the question, simply

concluded that "they must have some other job." Yet, this item *and all that kind of stuff* is

a vague category marker that is also deployed sometimes to not just form a vague category

but to use vagueness as a way of downtoning the force or directness of an assertion (as

discussed in section 8.4.4):

Example 8.51

A: I don't really know what a- an occupational therapist does. I know that they're
       meant to get you into…
B: They help people with feeding **and all that kind of stuff.**
C: Yeah rehabilitation.
A: Yeah get you back to everyday life. But like what, they must have some other job.
                                        [ICE-Ireland, Face to face conversation, S1A-051$D]

Similarly, we have the six-word cluster *at the end of the day* which can be used as a

hedge (as one of the politeness strategies) or triggered by other factors which may shape

some elements in the conversation such as: confidence between speaker and addressee or

politeness, indirectness, vagueness, understatement and assertiveness (as mentioned before,

see section 8.2.2). Thus, this expression (or pragmatic marker) can enable speakers to be

less assertive and blunt in formulating their message by marking that everything has been

taken into consideration. In other words, this item *at the end of the day* is an idiom that

often marks an evaluation on the part of the speaker (O'Keeffe *et al* 2007). As we can see

from examples 8.52 and 8.53, assertions can be made less blunt through the formulation of

a summary, by marking (with *at the end of the day*) that everything has been taken into

consideration:

Example 8.52

A: I think the speaker in in the audience has said that we need more accountability
       more openness regard- regarding politicians in our country. Even though they are a
       bit foggied, they are confused by the, by the bewildering statements coming out,
       nevertheless there are better issues or bigger issues involved in this. **At the end of**

the day, I do hope we'll have a more democratic society, well accountable society. These are the issues that are important. Yeah I do.

[ICE-Ireland, Broadcast discussions, S1B-033$E]

Example 8.53

A: But but but that equals a pacif~ That actually equals a pacifist state **at the end of the day.** I mean.
B: Wouldn't it be wonderful?
A: It but have you, can you give an example of it in history?

[ICE-Ireland, Broadcast discussions, S1B-039$B]

8.5 Conclusion

In this chapter, we have presented the quantitative and qualitative analysis of the use of hedges in Irish and Canadian Englishes in relation to forms and core function (embraced for this current study) that have been captured in the ICE-Ireland and ICE-Canada corpora. Methodologically, word and cluster lists were generated for both corpora to identify and compare the forms used in both datasets, and from these lists hedges with their pragmatic function adopted in this study were identified manually by cross-checking qualitatively using concordancing. This results in interesting quantitative findings:

- At the single-word level, Irish English and Canadian English continue to show the most frequent pragmatic markers acting as hedges (as we found across the other pragmatic markers explored thus far: discourse markers, response tokens, and questions and tags).

- Yet, In Irish English, the hedge *just* is found to occur more within the four-word cluster *I'm just going to*, while *just* in Canadian English demonstrates more distributions as a single form hedge.

- At the two-word level, we see both varieties of English still hold some similar hedges as well as some distinctive hedges.

- At the two-word level, it has been noticed that the distinction has not appeared only in relation to forms but also in the frequency of the same form. For example, Irish English and Canadian English display the hedge *kind of* which is based on the qualitative and quantitative analysis, yet it appears more in Irish English than it does in Canadian English. The higher frequency of *kind of* in Irish English could be explained by the fact that Canadian English also uses the hedge *sort of* which appears in its top 100 list of most frequent clusters (but does not appear in Irish English). Thus, because *kind of* and *sort of* are interchangeable with their pragmatic functions as hedges in Canadian English, they divide the usage frequency between each other.

- At the three-word hedges level and onward, Irish English and Canadian English have both demonstrated distinctive hedges, with the exception of *or something like that* which is common to both varieties.

- Lastly, Irish English still shows more PMs (such as: DMs, questions, and hedges) at the five-word clusters level and at the six-word clusters level not only as hedges but also as other PMs (as seen in the previous Chapters 5-7). In relation to hedges, this finding (Irish English speakers use more hedges than Canadian English speakers) can be supported by the findings of O'Keeffe and Farr (2002) where they suggest that Irish people seem to hedge more than American and British speakers.

CHAPTER 9

STANCE MARKERS IN IRISH AND CANADIAN ENGLISHES

9.1 Introduction

In this chapter, we will present the analysis of stance markers (SMs) in the ICE-Ireland and ICE-Canada corpora. As with previous chapters, the theoretical and methodological framework described in Chapters 3 and 4 will be used to generate the main quantitative findings from both datasets and these will then be unpacked and analysed qualitatively and comparatively. Thus, the analysis of the pragmatic variation that occurs in the use of SMs between Irish English and Canadian English has been approached mainly from the perspective of forms since the core function of SMs is the same: "expressions that express the speaker's stance or attitude towards the message" (note that this is the adopted definition in this study as shown in Appendix A) (Carter & McCarthy 2006, p.222; also see section 2.3.3). This is slightly different from the way pragmatic markers (PMs) (discourse markers, response tokens etc.) were discussed in the preceding chapters in which they varied both in terms of forms and pragmatic functions as well. SMs can be categorised differently because they are complex forms that can serve as a response token, a hedge, and a stance marker all at once (see section 9.4.3 and examples 9.19 and 9.21 discussing the trickiness of categorising multifunctional items and how overlaps and multifunctionality have been treated in this chapter). Nevertheless, SMs (in this study) have been categorised as the pragmatic markers defined with the core pragmatic function of signaling stances, attitudes, or points of view towards some of the segments in the discourse (regardless of their clause positions) (Carter & McCarthy 2006). Corpus linguistics has helped to provide a clearer picture of how spoken grammar and written grammar differ in relation to the frequency and the use of SMs, along with other features (Timmis 2012; Jones 2016). This has been proven from key findings from the main corpus-based grammars (McCarthy &

Carter 1995; Biber *et al* 1999; Carter & McCarthy 2006, 2015; Leech 2000). According to Carter and McCarthy (2006) and Biber *et al* (2002), SMs are much more frequent and common with their occurrences in speech and conversational discourse than they are in written language. SMs can be individual words, often adverbs, or they can be phrasal items. Carter and McCarthy (2006) offer the following range of examples: *actually, admittedly, amazingly, basically, certainly, clearly, confidentially, doubtless, essentially, frankly, to be frank, fortunately, honestly, to be honest, hopefully, ideally, if you ask me, I'm afraid, I must admit, I must say, I think, in fact, indeed, literally, naturally, no doubt, obviously, of course, predictably, putting (or, to put) it mildly/bluntly, (quite) rightly, really, sadly, seriously, (I'm) sorry, strictly speaking, surprisingly, thankfully, to tell you the truth, understandably, undoubtedly, and unfortunately* (Carter & McCarthy 2006, p.222). In this chapter, we will see how both varieties of English (Irish and Canadian) converge and diverge with regard to their forms of SMs (which have both already shown some similar and unique forms of SMs, see section 9.4).

9.2 Previous Research

Stance markers are one of the pragmatic markers that researchers have long been interested in. They have been used as a linguistic tool by which both speakers and writers express their personal attitudes, emotions, and stance; their evaluations and assessments; and their level of commitment towards the message (the propositions) (Aijmer & Rühlemann 2014). Much work has been done on stance in writing, in particular in English for academic purposes contexts and most notably by Hyland and his associates (Hyland & Guinda 2012; Hyland 2016; Hyland & Jiang 2018). SMs, similar to any of the other linguistic devices

such as response tokens (RTs), discourse markers (DMs), hedges and so forth, have been investigated under a variety of terms and approaches: intensity (Labov 1984), posture (Grabe 1984), disjuncts (Quirk *et al* 1985), hedges (Brown & Levinson 1987), modality (Palmer 1986 ; Bybee & Fleischman 1995), and (inter) subjectivity (White 2003; Fitzmaurice 2004; Lyons 1993). Today, the frameworks of evaluation (Hunston & Thompson 2000; Hunston & Sinclair 2000; Hunston 1994), appraisal (Martin 2000, 2003; Martin & White 2005) and stance (Biber *et al* 1999; Biber & Finegan 1988, 1989) have been particularly productive in helping researchers understand this pragmatic function in natural discourse (Aijmer & Rühlemann 2014). Therefore, based on the two earlier works of Ochs and Schieffelin (1989) and Besnier (1990) and evidentiality (Chafe 1986; Chafe & Nichols 1986), there are two primary types of meaning attached to SMs which have been differentiated:

- a speaker/writer's personal attitudes, emotions, and assessments.
- evaluations of the epistemic status of an entity or a proposition.

These types of meanings have been explored through a variety of methods, including detailed analyses of individual texts and descriptions of quantitative patterns across texts in large collections of authentic texts "corpora" (see Hunston 2011 for a book-length discussion on this topic, especially Chapters 2 and 4). Corpus-based approaches investigating stance have traditionally focused on lexical and grammatical patterns that mark stance. Such investigations range from analyses of a single lexical item (e.g. Diani 2008 on *really* as an emphasizer, Aijmer 2009 on *seem*), a lexical class (e.g. Swales & Burke 2003 on evaluative adjectives), or a particular grammatical structure (e.g. Biber & Finegan 1988 on stance adverbials; Baratta 2009 on passives, Charles 2006, 2007 and

Hyland & Tse 2005 on *that*-clauses, Hewings & Hewings 2002 on extraposed clauses).

Furthermore, there are some analyses of stance that have been conducted, taking the

comparative register approach (e.g. Biber *et al* 1999: Chapter 12; Biber 2006a, 2006b;

Biber & Finegan 1988, 1989) which have indicated that stance is marked more frequently

in spoken registers than it is in written registers which are interpersonal or persuasive in

nature, and is marked less frequently in informational writing (as in academic prose)

(Aijmer & Rühlemann 2014).

9.2.1 What are stance markers?

Stance markers have been classified as a form of pragmatic marker which do not indicate a

propositional meaning but rather perform a pragmatic function in which they "indicate the

speaker's stance or attitude vis-a-vis the message" (Carter & McCarthy 2006, p.208; Jones

2016). In other words, they signal how the speaker feels about the message they are

attempting to convey rather than the content of the message itself. Thus, stance can be

performed and marked through evaluative or value-laden lexis, in which individual words

indicate attitudes or emotions (e.g. *happy*, *angry*, *surprised*) and evaluations (e.g., *good*,

wonderful, *lovely*):

Example 9.1:   You climb the mountain because it's a **nice** hike

This type of stance (illustrated in Example 9.1 above) is grammatically embedded,

which requires reference to the context and shared background in order for the

reader/hearer to recognize that this is an attitude/evaluation of the speaker/writer (Aijmer &

Rühlemann 2014; see Biber *et al* 1999, pp.968-969 and Hunston 2011 for further

discussion). SMs as a sub-category of PM categorization sometimes interchange with their

pragmatic functions (in other words, an item that acts as a SM in one context can have a different pragmatic function in another). For example, Jones and Carter (2014) have suggested that the DMs *I mean* and *well* are optional in speech and can have multiple functions. This has been recognised in the present study where the same pragmatic marker (in terms of form) can act as a hedge in one context and a SM in another context (such as *I think*) or a DM in one context and a question (as RT) in another context (such as *you know*). Also, SMs are known to be optional items in speech as in the examples 9.2 and 9.3 below with the word *clearly*:

Example 9.2:   "**Clearly**, you should tell her the truth" (Cambridge Advanced Learner's Dictionary 2015).
Example 9.3:   I can see **clearly**.

In example 9.2, *clearly* is acting as a SM, and it is an optional item to be used. While *clearly* in example 9.3 acts as an adverb of manner and has a propositional meaning (Jones 2016).

Nevertheless, SMs have been approached and analysed in different ways. For example, some researchers (such as Biber & Finnegan 1988) have attempted to classify different forms of SMs based on their functions. Biber and Finnegan (1988), who provided the term "stance adverbials" for SMs, suggest that SMs can be categorised under one of six functions (1) *honestly* adverbials (*frankly*), (2) *generally* adverbials (*roughly*), (3) *surely* adverbials (*clearly*), (4) *actually* adverbials (*in fact*), (5) *maybe* adverbials (*possibly*) and (6) *amazingly* adverbials (*amazingly*). The core goal of such classification is to distinguish between items with their literal meaning and (pragmatic meaning) those marking stance, such as the example *clearly* provided previously (Jones 2016). Biber and Finegan have studied a large number of SMs based on spoken and written corpora which resulted in the

previous exemplars for each functional category listed above (Biber & Finegan 1988).

Therefore, this initial analysis can be a useful framework for those who are interested in

comparing and analysing SMs in relation to each other, whereby SMs can be identified

with clear categories based on how they function in context (taking the function-form

approach). However, in this present study, our focus is more on the core pragmatic function

adopted and provided by Carter and McCarthy (2006) (through the iterative approach

applying both function-form approach and form-function approach), as demonstrated in

Appendix A, to which all functional categories listed above belong (see section 9.3 for

further details).

Moreover, Biber *et al* (1999) and Biber *et al* (2002), went further than this by

suggesting that according to large spoken and written corpora, epistemic stance (the largest

category) gives "the speaker's judgments about the information in a proposition" (p.382),

allowing speakers to express aspects such as the certainty with which they view the

proposition, while an attitude adverbial "tells the speaker's attitude towards the

proposition" (p.384) and style adverbials "comment on the manner of conveying the

message" (p.385) (Biber *et al* 2002, p.383; Jones 2016, p.84). The examples they provided

for each category were *probably* (epistemic), *hopefully* (attitude) and *to tell you the truth*

(style) (Biber *et al* 2002, pp.383-385). Furthermore, their research suggested that SMs are

most frequent in conversational discourse, with the most common functional category

expressing epistemic stance. Yet, SMs may function in more than one way according to

Biber *et al* (2002). Therefore, we see the example of *honestly* which can be "perfectly

possible to suggest that it can comment on the manner in which a speaker is creating a

message (style adverbial, emphasising that *I am being honest*) whilst also conveying an

attitude towards the message (attitude adverbial, meaning that is a surprising or even ridiculous idea)" (Jones 2016, p.85). It is also possible, as mentioned above, that SMs can belong to other pragmatic markers categories such as RTs (Biber *et al* 2002; Carter *et al* 2011; Jones 2016). In relation to this present study, the limitation and the definition of SMs have been set through the methodological framework: the top-down framework for the analysis of the forms and pragmatic functions of spoken grammar presented in Appendix A.

9.2.2 Investigations of stance markers

Most studies investigating SMs have a similar focus and approach, which is usually the corpus-based approach. SMs have often been examined in terms of how they are used by non-native speakers in comparison to native speakers, and this kind of investigation has tended to focus on usage within the realm of written language (Jones 2016). For example, Hyland and Milton (1997) built a corpus containing almost a million words from students' writings. These were compiled from essays of non-native speakers (speakers of Cantonese as an L1) and native speakers (speakers of English as an L1) of equivalent age and educational level. The findings indicate that non-native speakers show a limited range of stance markers, and they are less able to express differing degrees of certainty with precision. We also have the work of Aijmer (2002) who used a corpus to examine the writing of high-level Swedish learners in comparison with native speakers; the study found that the learners overused forms of modality in their writing when compared to native speakers when expressing epistemic stance. Additionally, there is the study of Precht (2003) which is relevant to the present study in relation to investigating two different geographical zones. She examined two varieties of English, British and American, in regard to their usage of SMs, using the British and American conversation elements of the

Longman Corpus of Spoken and Written English. Her findings propose that American

speakers tend to use more SMs as what she calls "affect markers" such as *cool* and *wow*,

while British speakers tended to the use of what is called "evidential markers" such as *a bit*

in order to hedge propositions (Jones 2016). Moreover, Gabrielatos and McEnery (2005)

compared epistemic stance in native and non-native speakers by using a corpus of MA

dissertations. Their findings show that considerably more modal auxiliaries, adjectives, and

adverbs are used by native speaker writers than non-native speaker writers when expressing

epistemic stance.

As mentioned earlier (section 9.2), substantial work has been done on stance

markers in relation to English for academic purposes. Such studies also offer insights in

relation to second language acquisition and pedagogy. The work of Fordyce (2009) who

used a corpus of Japanese EFL students' language to examine and compare the means by

which epistemic stance was expressed in spoken and written modes is one such example.

This investigation led to interesting findings in which L2 learners were found to favour

lexical ways to express stance, to avoid modal verbs, and to have a heavy reliance on the

verb *think*. As such, studies investigating native and non-native speakers use of SMs expect

(and find to be true) that native speakers generally use a wider range of SMs (Jones 2016).

Considering this in our present study, we can expect that both Irish and Canadian speakers

will show a wider range of SMs (since they are both native English speakers) although the

corpora used are not large.

Within another spoken language study, Gablasova *et al* (2015) investigated spoken

epistemic stance in a corpus of advanced English taken from second languages speakers of

mixed nationalities. The data was collected from standardised speaking tests containing

different tasks in which interaction patterns varied. For instance, students would be given

the choice of a topic to present on, and later they would be required to undertake an

interactive discussion with the examiner on the same topic. The results revealed that the

distribution of SMs varied depending on whether the task was a monologue or dialogue. It

showed that there are much fewer SMs used in the monologic tasks. This led to the

suggestion that the interactional demand of the task has a significant impact upon the usage

of SMs. Not only that, but the use of pragmatic markers can also be influenced by how

learners wish to present themselves in their L2. In relation to the teaching and learning of

SMs, Fordyce's comparative analysis study (2014) looks at epistemic stance in the written

work of Japanese EFL learners, comparing explicit and implicit intervention when teaching

items such as *probably*, *seems*, and *believe*. The results suggested that explicit instruction is

likely to have a stronger effect upon acquisition of targeted forms as well as alongside

meta-analyses of instructed SLA in general (Norris & Ortega, 2000; Spada & Tomita,

2010; Jones 2016).

In summary, it is clear that SMs are one of the main pragmatic forms studied from

different perspectives (as discussed earlier in section 9.2), holding two primary meanings:

the speaker/writer's personal attitudes and emotions, and assessments and evaluations of

the epistemic status of an entity or a proposition (Chafe 1986; Chafe & Nichols 1986; Ochs

& Schieffelin 1989; Besnier 1990). The definition of SMs then, is defined based on stance

which is performed and marked through evaluative or value-laden lexis, in which

individual words indicate attitudes or emotions (e.g. *happy*, *angry*, *surprised*) and

evaluations (e.g. *good*, *wonderful*, *lovely*). Consequently, some DMs can act as a hedge in

one context and a SM in another context (such as *I think*) or a DM in one context and a

question (as RT) in another context (such as *you know*). Also, SMs are known to be optional items in speech (as discussed earlier in section 9.2.1) Jones and Carter (2014). Now, in the next section, we will talk about framework of analysis and methodology for SMs conducted in this chapter.

9.3 Methodology

This section outlines the main framework of analysis and methodology for this chapter. Similar to the previous analysis chapters (Chapters 5-8) our focus in the present framework (as explained in Chapter 3) is the formal level and thus, the use of SMs in Irish English and Canadian English is approached and investigated from two perspectives: forms and functions. In terms of functions in this present study, we adopted the core pragmatic function of SMs introduced by Carter and McCarthy (2006) in which SMs serve the pragmatic function of signaling stances, attitudes, or points of view towards some of the segments in the discourse. Methodologically, corpus linguistics is used as a tool in order to conduct the analysis of SMs quantitatively and qualitatively (as explained in Chapter 3), by using the iterative approach (for more details on this approach see Chapter 4 and Appendix A).

    In terms of form, we will present comparatively the quantitative findings of the SMs across both varieties (Irish English and Canadian English) based on their word-cluster lists (from one single-word SMs to six-word cluster SMs). As a result, as elaborated upon in Chapter 4, wordlists and cluster analyses were generated to identify and compare the SMs used in both datasets (ICE-Ireland and ICE-Canada). Within the form analysis, we will also

present the SMs across both varieties which appear as other pragmatic markers at the same time.

In terms of function, the quantitative findings of the forms of SMs will be unpacked and analysed qualitatively by concordancing them manually, as discussed in Chapter 4, in order to identify and determine their core pragmatic function in the discourse adopted in this present analysis.

9.4 Results and Analysis

The Tables 9.1-9.5 below are a summary of the comparative analysis of SMs in ICE-Ireland and ICE-Canada which demonstrate where Canadian English and Irish English converge and diverge in terms of SMs. Tables 9.1- 9.5 below have been derived from the qualitative analysis (from a bottom up route) of the top 100 most frequent word and cluster lists. Each single word and cluster has been concordanced manually in order to affirm which items from the quantitative findings resulted from the top 100 most frequent word and cluster lists (presented in Appendices B, C, D, E, F, G, H, I, J, K, L, M) are actually acting like SMs with regard to their pragmatic function in both datasets because some of these SMs were found to act as hedges in one context and SMs in another context (such as *I think*) or DMs in one context and SMs in another context (such as *I mean*). Thus, these items were concordanced manually while carefully setting the possible categorizations in which each item could possibly occur. This was done in order to make the distinction and the right category selection for those items which appeared as similar forms or were counted twice under different categorizations (in this chapter, and in the other relevant analysis chapters). Note that the light gray shaded cells in the tables below indicate the distinctive forms (SMs)

between Irish English and Canadian English, and the unshaded cells indicate the forms

(SMs) that both varieties of English share in common (we note overall that the instances of

SMs is low across both corpora within the top 100 items):

Table 9.1 *ICE-Ireland and ICE-Canada single-word stance markers occurring within the top 100 most frequent word lists (frequency per 10,000 words)*

ICE-Ireland	Freq[10]	ICE-Canada	Freq
really	19	really	25

Table 9.2 *ICE-Ireland and ICE-Canada two-word stance markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I think	21	I think	19
I mean	12	I mean	14
I'd say	3		

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 9.3 *ICE-Ireland and ICE-Canada three-word stance markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I think that	3	I think that	3
		I don't think	3

Note: Shaded cells indicate the distinctive forms; unshaded cells indicate items in common

Table 9.4 *ICE-Ireland and ICE-Canada four-word stance markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I don't think so	0.22	I don't think so	0.31

Table 9.5 *ICE-Ireland and ICE-Canada five-word stance markers occurring within the top 100 most frequent cluster lists (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
I don't know I think	0.28	I don't know I think	0.16
No I don't think so	0.062	No I don't think so	0.16
I think it's important that	0.094		
It seems to me that	0.25		

Note: No six-word stance markers were found in both datasets.

---

[10] The order of the forms presented in Tables 9.1- 9.5 is random and does not indicate rank.

As seen above in Tables 9.1 to 9.5, the comparative analysis of SMs has been conducted at the form level. We can see the similarities and differences between Irish and Canadian Englishes in relation to their forms of SMs and their occurrences (when needed for comparison purposes) which are analysed below based on their word clusters.

9.4.1 Single-word stance markers

As seen in the previous chapters (Chapters 5-8) with DMs, RTs, questions and tags, and hedges, Irish English and Canadian English demonstrate the most resemblance at the single-word level and after that the contrasts in relation to forms begin to appear more clearly. However, at the single-word level and based on bottom-up analysis (the quantitative and qualitative analysis) and top-down analysis, we see that SMs do not adopt many single forms, compared with DMs and RTs. Therefore, at the single-word level of SMs in Irish and Canadian Englishes based on ICE-Ireland and ICE-Canada, we see only *really* which has been considered to be amongst the SMs (Carter & McCarthy 2006). However, *really* in relation to SMs is more about intensifying the stance being made but not about making the stance itself. Both varieties of English, Irish and Canadian, have displayed the pragmatic marker *really* across other domains of spoken grammar such as DMs, RTs, and also questions, as has been discussed previously in their dedicated chapters (Carter & McCarthy 2006). In regards to SMs, it has been found that both varieties (Irish and Canadian) use the SM *really* in the same manner with the pragmatic function of intensifying the degree of stances, attitudes, or points of view towards the topic being talked about or focused on in the discourse. Yet, while *really* is intensifying, it is also a part of evaluative statements which can act as intensifying evaluations marker (Swales & Burke 2003) as seen in examples (9.4- 9.7) below extracted from both datasets.

Example 9.4

A: Yeah and I don't know what she had visioned but she reckoned this wasn't what a
       spasm would feel like
B: I mean I never had…
C: That's right
A: I don't know what sh~ I think the girl thought she was dying at one stage. Like it
       was **really** bad.
C: She yeah, she was in an awful state.
A: Yeah.
D: God love her.
A: Hmm.

                                        [ICE-Ireland, Face to face conversation, S1A-053$A]

Example 9.5

A: Can I just tell her the story? This is where I came home right, last, what day was it
       last Thursday? I was lying in my bedroom, right. And in my bedroom right, the
       desk was there but in the corner there's a hole in the floor [laughs]. So this sounds
       really bad.
B: It's not a real hole okay.
A: It is a hole. You can see down.
C: The pipes go down.

                                        [ICE-Ireland, Face to face conversation, S1A-084$B]

Example 9.6

A: I hate clowns.
B: Me too.
A: No, I **really** hate clowns.
B: So, the apartment is now taken.

                                        [ICE-Canada, Face to face conversation, S1A-082 #146:2: B]

Example 9.6 is a good example illustrating the pragmatic function of *really* mentioned

earlier. Speaker A has already given his stance about clowns, and speaker B has followed;

however, Speaker A used the SM *really* to intensify the degree of hatred he has towards

clowns.

Example 9.7

A: Anyways+ I show you to th~ I should show this to you later.
B: Hmm+.
A: Can I borrow this again? I feel **really** bad.
B: Wow, I don't know if I can read this the print is really small.
A: It's **really, really** poor print but it's good.
B: Ya. Ya.

A: It's **really** good stuff.
B: Good stuff, ya.

<div align="right">[ICE-Canada, Face to face conversation, S1A-035 #10:1: A]</div>

Example 9.7 combines both types of *really* in which the first example of *really* is not intensifying evaluation; it is only intensifying the degree to which Speaker A feels bad whereas in Speaker A's subsequent turns, *really* is used as part of evaluative statements marking some evaluation in the discourse (see sections 9.2 and 9.4.1 (above) for more information on the two primary types of functions expressed by SMs).

Also as discussed before, *really* has been used in both varieties of English (Irish and Canadian) as engagement RTs which mark the prior speaker's turn as newsworthy in some way or to denote surprise. In addition, *really* can also act as a question in various ways (as discussed in Chapter 7); it can be combined with a rising tone, performing the pragmatic function of marking surprise, and requesting more clarification on a segment of the topic in the discourse. Also, the pragmatic marker *really* can be used as a question tag performing the pragmatic function of checking and making sure that something has been understood or to confirm that an action has been agreed upon (see section 7.4.2.1 and example 7.36 in Chapter 7).

9.4.2 Two-word stance markers

At the two-word SMs level, we see the cluster *I think* which has been found in both varieties as a hedge and a SM. The cluster *I think* is known to be one of the main prominent SMs according to the existing literature (as mentioned in Chapters 2 and 8). In general, *I think* (in terms of propositional knowledge) can express both certainty and uncertainty which can be classified under two types; **the first** is the deliberative *I think* which may convey a speaker's careful deliberation in regards to what they are saying with some type

of authority associated with it (cf. Preisler 1986; Aijmer 1997; Vaughan 2009) and **the**

second is the tentative *I think* which may express tentativeness and uncertainty and also can

weaken a segment or an assertion in the discourse that might be too direct (Holmes 1985,

1990; Aijmer 1997; Zhang & Sabat 2016; O'Keeffe *et al* 2020). Holmes (1985, 1990)

mentions that *I think* varies demonstrably in terms of intonation as well as syntactic

position in an utterance. For example, *I think* can add weight to an utterance when it is at

the initial position of the clause with level stress on *think*. Also, it can be used tentatively

with a fall-rise intonation (Vaughan 2009). *I think* in medial and final positions (as often

found in ICE-Ireland and ICE-Canada) were classed as tentative, even if they had prosodic

prominence (Aijmer 1997). Aijmer found that the tentative *I think* was more frequent than

the deliberative *I think* in the Lund Corpus (LLC) (going along with our finding in ICE-

Ireland and ICE-Canada). Therefore, both Irish English and Canadian English have shown

the same use of the pragmatic marker *I think* as hedges and SMs. But, in relation to SMs,

both varieties have shown the semantic category of tentative more than the semantic

category of deliberative based on the datasets as illustrated in the following examples.

Example 9.8

A: Do you think a melody should have tunes or
B: No I don't think so. No, I I really I I I don't necessarily think so. I **I think** it's
        possible to have, yeah **I think** it's possible to have music without, without tunes.
A: All your mu~ music has
B: Oh very yes it's very melodic. Yeah there's no question about that yeah.
                                        [ICE-Ireland, Broadcast interviews, S1B-049$B]

Example 9.9

A: I need to get away from that type of work ent~ entirely. But **I think** bo~ both types
        of work that you've been doing with terminally ill patients is ah, is very demanding.
B: It's very demanding. Ya.
                                        [ICE-Canada, Face to face conversation, S1A-003 #149:1: A]

The cluster *I mean* has been found in both varieties as a DM (monitoring the discourse with multiple pragmatic functions as discussed in section 5.4.2.2) and a SM in which the speaker demonstrates their stance with more details used as a justification for such a stance. As examples 9.10- 9.12 illustrate, it marks a degree of uncertainty of stance in doing so. This is particularly notable in comparison with *I think,* which provides the stance with no justification or detailed information preceding or following which can be viewed as a type of hedging in the discourse. See the following examples 9.10- 9.12.

Example 9.10
A: Yeah, I thought it wouldn't be a bad holiday. Uh, uh we've had quite good weather though so I thought well, **I mean** I would be going to Galway for Christmas and it's quite **I mean,** it's pleasant enough. My sister is a great cyclist. We go out cycling around.
                                        [ICE-Ireland, Face to face conversation, S1A-047$A]

Example 9.11
A: I think more public access to really good routes in the city would be a nice thing. **I mean** it's a healthy thing.
                                        [ICE-Canada, Broadcast interviews, S1B-046 #82:1: A]

Example 9.12
A: W'I, Well I don;t know if they've given up. I think they, they should give up. **I mean** it's quite clear the people of Alberta and B.C. [British Columbia] aren't interested.
                                        [ICE-Canada, Broadcast interviews, S1B-040 #25:1: A]

Both varieties tend towards presenting the cluster *I mean* in the same way not only as a SM but also as a DM (as demonstrated in Chapter 5 in section 5.4.2.2).

In contrast, at the two-word SMs level, we see that Irish English displays the pragmatic marker *I'd say* which is found as a hedge and also a SM (acting like *I think*) in Irish English only but not in Canadian English due to the fact that *I'd say* does not appear in the top 100 most frequent clusters in Canadian English. As example 9.13 illustrates, Irish English speakers use *I'd say* to introduce a proposition.

Example 9.13

A: All she plays them up full blast and she has no problem if anyone's in the house she will [laughs] sing to her heart's content. And oh my God it's like
B: Oh no
C: The confidence of her!
B: And why, is she in singing? or does she just like it?
A: She's a real drama queen like. +I mean she loves acting and stuff but she can't sing. She really can't sing now [laughs]. **I'd say** she'd be good at acting alright but she can't sing!
B: Oh yeah+. Yeah.
<div align="right">[ICE-Ireland, Face to face conversation, S1A-068$A]</div>

Also, with Irish English, we see the use of *I'd say so* which functions in the same way

as the form of *I think so* as in the following example 9.14. We can see that the speaker adds

it on to reinforce their stance:

Example 9.14

A: Well last week the, I think they were going to have Michael Jackson and
B: Lisa Presley.
A: Yes.
B: What a combo. Wonder will it last.
A: I wonder if they're married at all.
B: Oh yeah. Oh yeah, **I'd say so**.
A: So we're told at any rate.
B: No they wouldn't get away with that kind of thing. They wouldn't be getting away with it I mean. They, they wouldn't be getting away with all the appearances and all the, you know all the uh kind of show and everything. They're definitely married. **I'd say so**.
<div align="right">[ICE-Ireland, Face to face conversation, S1A-059$B]</div>

9.4.3 Three and four-word stance markers

Irish English and Canadian English demonstrate common forms and similarities in the way

three- and four-word SMs are used in their discourse such as *I think that* which is shared by

both varieties as illustrated in examples 9.15- 9.17.

Example 9.15

A: Why do you think that he has to have a BMW and keys and money?
B: Well, a l~ a lady needs to b~, to be driven around. She she can't be expect~, she can't be expected to sit on the back of a bicycle [laughs].
A: I don't even get to sit on the back of a bicycle.

B: Well **I think that** you know, a lady needs to be taken out and looked after uhm. Despi~ despite what the modern, I know there is a mo~ a modern trend towards equality of the sexes and while I am in favour of equality of the sexes **I think that**, I still believe in the old-fashioned uhm uhm, in the old-fashioned uhm, uhm standards standards. Like uh think a man should be able to drive h~ his his girlfriend home.
                                      [ICE-Ireland, Face to face conversation, S1A-061$A]

     Example 9.16

A: It was simply to recognize the fact that society in Canada as a nation we ought to deal with the young offender in a way different from a more mature offender that the penalties ought to be different, that the, where incarceration is required that there ought to be different facilities. And **I think that** Canadians generally recognize that and accepted that rationale. Unfortunately, I think as we know on all sides of the House and as Canadians from coast to coast to coast know the Young Offenders Act has in at least a minority of cases been somewhat badly abused.
                                      [ICE-Canada, Parliamentary debates, S1B-058 #6:1: A]

     Example 9.17

A: I don't want to spend my life carving ducks.
B: [laughs] I know.
A: So, he ya, he's going through a big adjustment now.
B: Well, I think, you know, **I think that** hobbies aren't enough to keep you, you you dream about having time to do that or things that you'd like to do but uh I think being off work this term is uhm, has really, has really shown me that first hand, is that despite the fact that I was so busy, stressed and, and uhm, very stressed, probably more stressed then I needed to be but uhm without it I didn't have the same sense of fulfillment at all.
                                      [ICE-Canada, Face to face conversation, S1A-002 #123:1: A]

However, at the three-word level, we see only one distinctive form of SM, *I don't think,*

found in the top 100 most frequent clusters in Canadian English but not in Irish English

which seems a very assertive stance as seen in example 9.18 (below):

     Example 9.18

A: I spose, ^suppose^ that would be really difficult to do, would it.
B: To put in a door?
A: Yeah.
B: **I don't think** that's a good idea there, cos then we'd have to extend that wall to the front.
A: But I would like to eventually wall that off you know. Is that, Is that a big…
B: That can be next year's project.
                                      [ICE-Canada, Face to face conversation, S1A-012 #204:1: A]

At the four-word SMs level, we see the cluster *I don't think so* is used in both

varieties not only as a SM but also as a RT. The distinction between *I don't think so* as a

RT and *I don't think so* as a SM is that *I don't think so* as a RT is more associated with

marking agreement and does not take the floor from the speaker (as in example 9.19) while

I don't think so as a SM is more associated with marking agreement and providing stance

and takes over the floor from the speaker (as in example 9.21; see the examples 9.19- 9.21

illustrated below.

Example 9.19

A: Vocabulary maps yeah. There is a programme as well, now I don't know if we have
       it any more, or maybe they have it in the school, called Inspiration. And it's just, it's
       doing that kind of vocabulary mapping on a computer programme. And it's very
       good cos you can organise.
B: Okay.
C: It's on computer here is it?
A: **I don't think so.**
B: I think, Felice had the programme or something.
A: Yeah and you could get it for a trial period or whatever, but then it r~ goes out. Uh
       like it self-destructs after however many days. [laughs].
                                           [ICE-Ireland, Classroom lessons, S1B-018$A]

Example 9.20

A: Okay. Uhm now. We can forget about doing the Old English question. Uhm, having
       said that, in Beowulf, the possible questions that could come up are uhm digressions
       in Beowulf, the digressions in Beowulf. Discuss the role of women in Beowulf, uh,
       uh… What part or what role do the monsters play in Beowulf? uh, I don't think
       you'll get a question on landscape.
B: Oh
A: [laughs] You all of a sudden, oh no! *Zut* [French]. Uhm that's one thing I **I don't
       think so**. Uhm that would be nice, 'twould be very nice, it'd be very nice for me,
       it'd be very nice for you. But uhm unfortunately that that may not come up.
C: And we will definitely be asked on Beowulf will we?
A: Oh yeah! That's, that's a cert.
                                           [ICE-Ireland, Classroom lessons, S1B-015$A]

Example 9.21
A: [laughs] Hmm.
B: And I'm sure there, there must be some jazz theory you know.
A: **I don't think so**. Hmm we'll have to see.
                                    [ICE-Canada, Face to face conversation, S1A-010 #144:1: A]

Example 9.22

A: Now is that equivalent to our ah Bachelor's degree here do you think?
B: I'm not so sure.
A: I don't think so either.
B: **I don't think so** but, don't quote me on that.
A: No but, I I just I, I can't see it.

[ICE-Canada, Face to face conversation, S1A-057 #192:1: A]

This distinction between *I don't think so* as a RT or SM can cease to exist or become unclear given that some researchers define RTs differently from the definition chosen for this research (see Chapter 6 for other common definitions for RTs, and section 6.4.2.2 for more information on *I don't think so*). Additionally, as mentioned previously, pragmatic markers can perform multiple pragmatic functions simultaneously within the same context.

9.4.4 Five-word stance markers

At the five-word SMs level, Irish English has shown more clusters of SMs, and some of these are distinctive SMs that are not found in the top 100 most frequent clusters in Canadian English such as: *It seems to me that* (as in example 9.23) and *I think it's important that* (as in example 9.24) (see Table 9.5 and Appendices G and L). Example 9.23 shows *it seems to me that* being used in Irish English in a discussion of the sensitive topic of the Northern Irish conflict. Speaker A (a well know political commentator on a radio interview) uses the cluster to mark stance in an assertive but non-face-threatening way.

Example 9.23

A: …There are arms on a lot of different sides and that process clearly has to be handled with some delicacy but if you simply put up for each side of that this kind of direct hurdle that you must give up the arms before you begin uh substantive political talks then **it seems to me that** that the process could be in very serious trouble.
B: Nell McCafferty what is your view? You were in Derry over the weekend. What's the view there?
A: Uh I take it that question refers to the arms held by the IRA?

B: Uh well it it refers, yes it refers to the arms held by the IRA as a prerequisite for involving them in inclusive talks.

[ICE-Ireland, Broadcast discussions, S1B-039$A]

In example 9.24, we also see an example from a political radio interview and again a public persona uses a SM to augment the force of their proposition in a non-face-threatening way. The alternative more assertive SM might have been *I think the government must…*

Example 9.24

A: …in their report to the Minister yesterday. I would hope, Minister, that that report will not be ignored and that some action will be taken on the recommendations of that report, and that they will be taken very very quickly. Uh, uh **I think it's important that** workers who began their working life uhm perhaps a long time ago uh wi~, uh will have to be assisted to keep pace with technology. Will have to have the kind of in-house in-service training that will enable them to have the to to to increase their level of skills and will give them the kind of versatility of skill and outlook…

[ICE-Ireland, Legal presentations, S2A-069$A]

On the other hand, we have some clusters of SMs at the five-word level that both varieties share as can be seen in the following examples. As examples 9.25 and 9.26 show, the *I don't know* component of the cluster attends to the previous turn as a RT but then adds *I think* + a proposition:

Example 9.25

A: Do you remember that?
B: Yeah, yeah. In, yeah in Wembley.
A: Yeah.
B: Yeah oh right. Oh right. Well sure that'd be grand for him.
C: When is Maire/ad coming over here?
A: **I don't know. I think** she's coming home today from [unclear speech].
C: Oh she's not home yet.

[ICE-Ireland, Face to face conversation, S1A-067$C]

Example 9.26

A: It was really good. It was fun.
B: Is this something they only did for the ten-year reunion or do they do this every year?
A: **I don't know. I think** it's only the ten-year anniversary.
B: Mm. I think it was a new, it was a new thing.

A: It was pretty interesting.

<div align="right">[ICE-Canada, Face to face conversation, S1A-026 #266:1:C]</div>

This overlap between RTs and SMs is also apparent in the case of *No I don't think so* (example 9.27). Based on the qualitative analysis, it has been found that it can be classified as a SM or a RT due to the way it occurs in the discourse of both Englishes. However, it has been noticed that *no I don't think so* occurs as a RT, following the definition adopted in this current study, much more in Canadian English than in Irish English. The five-word cluster *no I don't think so* can perform various pragmatic functions in the same way as in *I don't think so* which was mentioned earlier in section 9.4.3. For more illustration see the examples 9.27 and 9.28 below where we see that on one hand the cluster acts as a RT while also marking stance; in so doing it could be argued that it also mitigates the force of a bald *no* response. The speaker is saying *no* but by marking the stance with a degree of uncertainty, they are downtoning the pragmatic force of the negative response:

Example 9.27

A: Do you think a melody should have tunes or
B: **No I don't think so**. No, I I really I I I don't necessarily think so. I I think it's
        possible to have, yeah I think it's possible to have music without, without tunes.
A: All your mu~ music has
B: Oh very yes it's very melodic. Yeah there's no question about that yeah.

<div align="right">[ICE-Ireland, Broadcast interviews, S1B-049$B]</div>

Example 9.28

A: And a lightbulb.
B: Is that dangerous?
A: **No, I don't think so**.
B: Well then, that's simple, that's what I'll do.

<div align="right">[ICE-Canada, Face to face conversation, S1A-004 #121:1: B]</div>

As an overall observation throughout this chapter and the previous ones, speakers of Irish English seems to use or adopt more fixed clusters in general. This observation aligns

with the finding of O'Keeffe and Adolphs (2008) on RTs in which they propose that there is greater fixedness in Irish English. However, this is something that needs further exploration and careful investigation.

9.5 Conclusion

In this chapter, we have presented the quantitative and qualitative analysis of the use of SMs in Irish and Canadian Englishes in relation to forms and core function (embraced for this current study) that have been captured in the ICE-Ireland and the ICE-Canada corpora. Methodologically, word and cluster lists were generated for both corpora to identify and compare the forms of SMs used in both datasets. From these lists, SMs with their pragmatic function adopted in this study were identified manually by cross-checking qualitatively using concordancing. This results in interesting quantitative findings. Irish English and Canadian English are found to have the most SMs in common at the single-word level, as we have seen in other analysis chapters of DMs, RTs, questions and tags, and hedges. However, at the single-word level and based on both bottom-up analysis (the quantitative and qualitative analysis) and top-down analysis, we see that SMs do not deploy as wide a range of single-word forms compared with DMs and RTs for instance. Thus, *really* is the only SM found which is shared by both varieties. At the two-word SMs level, we can see that both varieties of English use *I think* and *I mean* which are both known to be amongst the prominent SMs according to the existing literature (see section 9.2). At the two-word level, Irish English and Canadian English differ only with the SM *I'd say* which appears in the top 100 most frequent clusters in Irish English but not in Canadian English. At the three- and four-word SMs level, Irish English and Canadian English show common forms

and similarities in the way those SMs are used in their discourse. However, at the three-word level, we see only one distinctive form of SM that is not found in the top 100 most frequent clusters in Irish English which is *I don't think*. At the five-word SMs level, Irish English has shown more clusters of SMs, and some of these are distinctive SMs that are not found in the top 100 most frequent clusters in Canadian English. Lastly, both Irish English and Canadian English do not display any SMs at the six-word level.

Overall, we can see that stance has received far more attention in writing, particularly in academic writing, and it has been much neglected in the context of spoken language. Looking at stance here across varieties suggests that it is a fruitful area rife for further comparative research. We have now reached the end of the analysis chapters in which the spoken grammar (as a whole) of Irish English and Canadian English in terms of forms and functions was comparatively investigated. This concludes the analysis chapters of this study; in the next chapter we will revisit the research questions (findings), discuss the limitations of the study, and provide directions for further research.

CHAPTER 10

CONCLUSION

10.1 Introduction

This research is a corpus pragmatics study examining the pragmatic variation of Irish English and Canadian English. It seeks to comparatively investigate the intra-varietal differences in the nature of spoken grammar usage in these two varieties of English in terms of their forms and pragmatic functions within the framework of variational pragmatics. Hence, the data (ICE-Ireland and ICE-Canada corpora) have been examined mainly from two different approaches. The first is through Corpus Linguistics (CL), moving from forms to functions (known as the form-to-function approach or alternatively as the bottom-up approach). The second is through Pragmatics which takes the opposite direction, starting with the specific pragmatic function to the forms (known as function-to-form approach or alternatively as the top-down approach). As a result, this mixture of methodologies and approaches allows the data to be examined, the research questions to be answered, and the research objectives to be achieved (i.e. a comprehensive description of the spoken grammar of Irish and Canadian Englishes in comparison to each other). Therefore, this chapter will revisit the research questions by summarising and discussing the key findings, review the research limitations, and provide some points and directions for future research.

10.2 Revisiting the Research Questions

The research questions of this present study are proposed (as discussed in Chapter 1) to serve the purpose and focus of the study which looks into the differences and similarities between Irish English and Canadian English in relation to pragmatic functions and forms of their spoken grammar. Therefore, the first main question **"How different and similar are**

Irish English and Canadian English in their spoken grammar in terms of form and function?" has been answered through the analysis chapters (from Chapters 5 to 9) which resulted in the following overall differences and similarities in the spoken grammar of Irish and Canadian Englishes:

Differences

- At the single-word level, there are some prominent pragmatic markers with various pragmatic functions appearing more in one variety than the other such as: *okay, ya, so, uh, hmm, mm, because, how, right and really* in Canadian English and *now, yeah, oh, and which* in Irish English.

- Canadian English uses more clusters with *so* than Irish English, such as: *or so, and so,* and *so I.* These perform more distinctive pragmatic functions, including vagueness and discourse marking in narratives.

- *So,* as a single-word form, is a prominent pragmatic marker (acting as a DM and a RT) appearing more in Canadian English (70 per 10,000 words) than Irish English (57 per 10,000 words*),* especially as a RT.

- While pragmatic markers are used widely in both varieties, it was found that certain single word items were more prevalent in Canadian English as a specific pragmatic marker than they are in Irish English. For example, the pragmatic markers *ya*, *so*, *because*, *mm*, and *hmm* are attributed more to Canadian English as RTs than to Irish English.

- Questions and tags have been found to be in use much more in Irish English generally and specifically as convergence RTs and engagement RTs than in Canadian English.

- Canadian English uses more RTs marking what is common ground or shared knowledge between participants in the discourse than Irish English, such as *you don't know* and *you know*.

- Canadian English and Irish English reveal a tendency towards using information receipt RTs indicating when the information in the discourse would flow and stop. Common exponents of this function are *right* and *okay* in both varieties. In Canadian English, *so* is used more than in Irish English. The RTs *right* and *okay* are found to be more interchangeable in Irish English than in Canadian English.

- Irish English displays a wide range of questions tags (especially at the two-word level) whereas Canadian English shows a much narrower range; therefore, significant contrasts between the two varieties can be seen in question tags and follow-up questions.

- Irish English shows a consistent pattern of greater use of clusters. In other words, there appears to be more fixedness in the patterning of spoken grammar items in Irish English than in Canadian English across all the pragmatic markers (such as discourse markers (DMs), hedges, and questions), as evidenced by these clusters which were investigated. For example, Canadian English does not show four, five, or six-word question clusters whereas Irish English does, as follows: *do you know what?, what do you mean?, you know what I mean?, what do you call it?, do you know what I mean?,* and *you know what I mean yeah?.* These clusters mark a high degree of convergence on the part of the speaker who seems to monitor understanding on the part of the listener through these markers (Carter & McCarthy 2006). It seems to suggest a strong listener-orientation in Irish English. This is in

316

line with the finding noted above that question tags were used more in this variety as convergence RTs and engagement response.

- Canadian English has a tendency to use more declarative statements (as questions) rather than interrogative sentences whereas in Irish English both types of question formations are used commonly (see Chapter 7, and sections within and including section 7.4.2).

- Overall, there appears to be more evidence of hedging in Irish English:

  o In Irish English, the hedge *just* is more associated and found within the four-word cluster *I'm just going to*, while *just* in Canadian English demonstrates more distributions as a single form hedge.

  o Irish English and Canadian English both display the (pragmatic marker) hedge *kind of*, yet it appears more in Irish English than it does in Canadian English.

  o The hedge *sort of* exists within the top 100 most frequent clusters list in Canadian English but it does not feature with a similar frequency in Irish English. In Canadian English *kind of* and *sort of* have been found to be used interchangeably in regard to their pragmatic functions as hedges.

  o The pragmatic marker *I'd say* appears in the top 100 most frequent clusters in Irish English but not in Canadian English.

Similarities

- Irish English and Canadian English have the most in common in terms of spoken discourse features at the single-word level across all the pragmatic markers (see Table 10.1 below):

317

Table 10.1 *Single-word pragmatic markers Irish English and Canadian English have in common (frequency per 10,000 words)*

ICE-Ireland	Freq	ICE-Canada	Freq
uh	86	uh	106
yeah	82	yeah	23
so	57	so	70
like	52	like	56
uhm	51	uhm	45
well	39	well	43
okay	16	okay	31
right	19	right	28
oh	35	oh	28
no	43	no	40
really	19	really	25
what	53	what	55
when	21	when	24
who	23	who	19
just	34	just	42

(See Tables: 5.2, 6.1, 7.3, 8.1, and 9.1 for their specific classification in the spoken discourse).

Irish English and Canadian English start to diverge at the two-word level across all the pragmatic markers.

- Both varieties display common pragmatic functions across all the pragmatic markers (yet, sometimes they differ in the way they deploy them in terms of form, frequency, precise pragmatic function, and distribution).

- The use of the pragmatic marker *I mean* is (almost) similar in both varieties, not only in relation to function but also in its occurrences in both datasets. It is speculated that this is because it serves an important core function common to all speakers: monitoring the discourse through reformulations and alternative expressions, expanding the discourse, softening a segment in the discourse, or repairing a segment in the discourse, and there are few alternative forms to perform

this function (for more information, see Figure 5.5, Table 5.11, and examples 5.19 and 5.20).

- The RTs *that's right*, *I know*, *do you? no I don't*, *I don't think so* are used similarly in both varieties, displaying common pragmatic functions in relation to acknowledgment, engagement, and drawing a boundary in the discourse.

- Both varieties demonstrate a use of echo questions with wh-words with a very similar usage in their discourse.

- At the single-word level, we see that stance markers (SMs) in both varieties adopt an equally lower number of forms compared to other pragmatic markers such as DMs and RTs. This suggests the importance of politeness markers in the form of hedges (as downtoners) in spoken grammar over SMs in general.

The sub-questions are as follows:

Sub-Question One: How different and similar are Irish English and Canadian English in their Discourse Markers in terms of form and function?

This question has been responded to in Chapter 5 which results in interesting findings. As discussed above in relation to overall findings, at the form level, Irish English and Canadian English have most in common at the single-word DMs level with some prominent DMs appearing more in one variety than the other such as *okay*, *now*, and *ya*. Both varieties, as noted, start contrasting noticeably from the two-word DMs level and onwards. In terms of the qualitative findings, both varieties display common pragmatic functions in relation to monitoring, organising, and managing the discourse. However, both varieties show clear contrasts in terms of functions and the forms that are typically deployed for these functions.

For example, the DM *now* has been found to be in use much more in Irish English generally, and specifically in opening up and closing down the discourse, than in Canadian English. Similarly, the DM *okay* occurs twice as much in Canadian English than in Irish English, especially in terms of the pragmatic function of managing talk in opening up and closing down the discourse. Also, Canadian English displays more clusters with *so* than in Irish English and these perform more distinctive pragmatic functions (such as: monitoring the ongoing discourse by denoting similar entities with the form *or so*, marking boundaries and linking segments of the topic in the discourse with *and so*, and *so I* acting as a RT). On the other hand, Irish English and Canadian English show many similarities in regards to pragmatic functions. For example, the use of the DM *I mean* is displayed in both varieties almost in the same way, not only in relation to function but also its occurrences in both datasets (as mentioned above). Lastly, as an overall picture in relation to DMs in Irish and Canadian discourse, Irish English (Irish speakers) seems to use more fixed clusters in general. This observation aligns with the finding suggested by O'Keeffe and Adolphs (2008) on RTs in which they propose that there is greater fixedness in Irish English. It could be speculated that this is perhaps related to Ireland being smaller and more concentrated as a society (The Republic of Ireland has a population of almost 5 million while Canada has over 37.5 million).

Sub-Question Two: How different and similar are Irish English and Canadian English in their Response Tokens in terms of form and function?

This question has been addressed in Chapter 6 resulting in noteworthy findings. Again here, we see that at the level of form, Irish English and Canadian English have most in common at the single-word RTs level, with some frequently-occurring RTs appearing

either more, or solely, in one variety over the other. For example, *ya*, *so*, *because*, *mm*, and *hmm* appear more often and, in some cases, solely in Canadian English as RTs. Both varieties begin contrasting noticeably at the two-word RTs level and onwards. In terms of the qualitative findings, we saw that both varieties display common pragmatic functions in relation to acknowledgment, engagement, and drawing a boundary in the discourse, using similar RTs such as *that's right*, *I know*, *do you? no I don't*, *I don't think so,* and so forth. Nevertheless, both varieties show clear contrasts in terms of forms. For example, questions and tags have been found to be in use much more in Irish English generally, and specifically as convergence RTs and engagement RTs, than in Canadian English.

In contrast, Canadian English uses more RTs marking what is common ground or shared knowledge between participants in the discourse such as *you don't know* and *you know*. Also, Canadian English and Irish English reveal a particular tendency towards using information receipt tokens indicating when the information in the discourse would flow and stop, especially using *right* and *okay.* In Canadian English, *so* is used more than in Irish English for this function. The forms *right* and *okay* were found to be used interchangeably more in Irish English than in Canadian English. Overall, we can conclude that there is a stronger focus in Canadian English discourse on the monitoring of shared knowledge and an attention to the management and flow of information whereas in Irish English the primary concern is on managing the force of the discourse so as to maintain relations and converge more with the listener. This aligns with other studies and can also be linked to findings found in this study such as the following: Irish English speakers use more hedges than Canadian English speakers; this aligns with the finding of O'Keeffe and Farr (2002) that Irish people seem to hedge more than American and British speakers.

Also, Irish speakers tend to indirectness in answering back and forth with the speaker/hearer (Asián & McCullough 1998; Farr & O'Keeffe 2002). This is how some of the findings mentioned above (such as managing the force of the discourse and the tendency to hedging) can be associated with the socio-cultural norm of avoiding over assertiveness (O'Keeffe *et al* 2020).

Sub-Question Three: How different and similar are Irish English and Canadian English in their Questions and Tags in terms of form and function?

This question has been answered in Chapter 7 which shows notable results. As has been the general pattern, we find that at the form level, Irish English and Canadian English have most in common at the single-word level. The single-word questions are mainly used as tags and echo questions. In relation to echo questions with wh-words, both varieties demonstrate very similar usage in their discourse. Nevertheless, it is clear that Irish English displays a wide range of questions whereas Canadian English shows a much narrower range which has resulted in significant contrasts seen in question tags and follow-up questions. This finding can be supported by Barron (2015) who found that the frequent use of positive constant polarity TQs (tag questions) is more prominent in Irish English than it is in British English; this represents a variety-preferential feature of Irish English relative to British English, American English and Filipino English (Tottie & Hoffmann 2006, p. 290; Borlongan 2008). This can reflect engagement and convergence in the discourse which is usually not expressed by questions (i.e. questions are not the common features used to perform these pragmatic functions in other varieties, especially at the high level of clusters), and this may explain why Canadian English does not show four, five, or six word

question clusters. However, it is important in Irish English to maintain convergence with the listener, and this is expressed even with questions. Additionally, in Canadian English, there seems to be a tendency to use declarative statements rather than interrogative sentences whereas in Irish English both types of question formations are used commonly. According to O'Keeffe (2006), declarative questions are more face-threatening and so this finding further underscores the point that hedging and negative politeness is a priority in Irish English.

Sub-Question Four: How different and similar are Irish English and Canadian English in their Hedges in terms of form and function?

This question corresponds to the findings explored in Chapter 8. At the single-word level, as noted in relation to DMs, RTs, and questions and tags above, Irish English and Canadian English show the most similarity in terms of pragmatic markers acting as hedges. In Irish English, the hedge *just* occurs more within the four-word cluster *I'm just going to*, while *just* in Canadian English demonstrates more distributions as a single form hedge. At the two-word level, we see that both varieties of English still display some similar hedges as well as some distinctive hedges. At the two-word level, it has been noticed that the distinction appears not only in relation to forms but also in the frequency of the same form. For example, while both varieties use the hedge *kind of*, it appears more in Irish English than it does in Canadian English. As mentioned earlier in this section, this relates to the fact that Canadian English uses the hedge *sort of* interchangeably with *kind of* (but is not in the top 100 most frequent clusters in Irish English). At the three-word hedges level and upwards, Irish English and Canadian English both demonstrate distinctive hedges except in

the case of *or something like that* which they both share. Lastly, as discussed elsewhere, Irish English still shows more pragmatic markers (such as: DMs, questions, and hedges) at the five-word clusters level and at the six-word clusters level not only as hedges but also as other pragmatic markers.

Sub-Question Five: How different and similar are Irish English and Canadian English in their Stance Markers in terms of form and function?

This question is addressed in Chapter 9. The main findings are summarised as follows. As with analysis of DMs, RTs, questions and tags, and hedges, at the single-word level, Irish English and Canadian English show the most similarity. However, at the single-word level (based on bottom-up analysis (the quantitative and qualitative analysis) and top-down analysis), we note that overall, there are fewer instances of SMs as single forms when compared to DMs and RTs. Thus, *really* is the only SM shared by both varieties. At the two-word SMs level, we can see that both varieties of English display *I think* and *I mean,* both known to be among the prominent SMs according to the existing literature (Brown & Levinson 1987; also see section 9.2); they differ only with the SM *I'd say* which appears at the top 100 most frequent clusters in Irish English but not in Canadian English. At the three- and four-word SMs level, Irish English and Canadian English show common forms and similarities in the way those SMs are used in their discourse. However, at the three-word level, we see only one distinctive form of SM, *I don't think,* which is not found in the top 100 most frequent clusters in Irish English. At the five-word SMs level, Irish English shows more clusters of SMs, and some of these are distinctive SMs that are not found at the

top 100 most frequent clusters in Canadian English. Lastly, both Irish English and Canadian English do not display any SMs at the six-word level.

We note also here by way of summary that the research questions were answered using a methodology which was two-way and novel. This study could have been conducted solely as a bottom-up corpus study where all single-word and cluster lists were analysed for salient items. They could then have been investigated through concordance lines to lead to functional findings. This "form-to-function" approach alone would have generated some of the findings of this study. However, it was complemented by a function-to-form approach which meant that the salient items that are already identified in the literature on spoken grammar also provided an inventory of candidates for corpus searches. The iteration between form-to-function and function-to-form approaches has made the comparison more robust and the findings richer. It is argued that this approach offers a methodology that can be replicated in other studies of this kind so that more Variational Pragmatic studies can explore spoken grammar as a whole in a systematic way rather than focusing on just one form or one feature in isolation.

10.3 Limitations of the Study

Having a common methodology for a world-wide study of varieties of English is a major advantage in terms of facilitating comparison across varieties. Therefore, having a common set of text types which have been collected according to the same corpus sample frame with the intention of gathering naturally-occurring data from similar contexts would surely facilitate international comparison. Further, it would allow for empirical tests of the degree to which it would be possible to speak of an international standard of English. This

would by definition, display uniformity within the English-speaking world, and have the equal possibility of presenting national and local variation within the empirically-defined standard (Kallen 2006). Yet, although the ICE methodology has much value, there are some limitations to it. The first of these is the issue of representativeness and generalisability which often arises in corpus linguistics (in general) and especially in the studies of pragmatic and sociolinguistic variation using corpora. So, the most obvious disadvantage regarding ICE is that it is not conceived of as a sociolinguistic database due to the fact its sample frame does not require geographical spread, gender balance, and other variables (in other words the data underlying the analysis are not comparable in relation to social factors, see Table 4.5). This impedes the demographic base for an ICE corpus to be representative of the standard English speaking population at large (Kallen 2006).

However, in relation to this study, this limitation was not an issue because the goal was not to conduct research from a perspective of sociolinguistic variables such as age, gender or the like. The most important goal for this study is comparability of varieties and this aligns with the main goal of the ICE project. Therefore, the basis for choosing ICE-Ireland and ICE-Canada was the robustness of the comparability that the ICE sample frame offers. When we compare the two datasets, we are comparing like with like. Other corpora, for example the *Limerick Corpus of Irish English* (LCIE) (Farr *et al* 2002) and the *Strathy Corpus of Canadian English* (Strathy Language Unit 2011) could also have been used but did not have comparability in terms of their corpus design. McCarthy (2002) states that corpora being built with a common design provide safer ground for generalisations. This is especially important for the study of variational pragmatics. Findings from the present study can be used as the basis for further exploration in LCIE and Strathy corpora but

because no baseline comparison of spoken grammar is available for Irish English and Canadian English specifically, this study will offer a starting point for further research.

Another limitation that is of relevance to this study is the lack of an interpretive mechanism in ICE transcripts. That is, orthographic transcripts and annotation schemes for ICE data do not tag discourse functions (Kallen 2006, 2012). However, this limitation is also not a problem because of how the data has been treated. Top-down and bottom-up processes are used to compile pragmatic markers and their pragmatic functions in spoken grammar in order to avoid inaccuracy in the results. Top-down analysis is based on a framework for spoken grammar founded on existing literature while the bottom-up process is based on micro-analysis of the data which both reflect the notion of the new sub-field of Corpus Pragmatics (Clancy & O'Keeffe 2015; O'Keeffe 2018) (as discussed in Chapter 3 and in Chapter 4, section 4.5). Therefore, this present study is derived from both CL and pragmatics (in which context is not ignored) and it is not restricted by the limitations of corpus-based or corpus-driven studies in terms of methodology and results. Also, the cooperation of both approaches to compile forms and functions for investigation can make the interpretation easier especially in the way spoken grammar exploits the deictic system to create interpersonal meanings.

Furthermore, the ICE corpora do not include phonetic/prosodic annotation which is known to be a very useful tool to be used in the assignment of functions of some of the pragmatic markers (such as questions and TQs). This requires that some intuitive values be proposed by the researcher in this study.

Another limitation for this study is the pragmatic annotation in which the information about the kinds of speech act that occur in a spoken dialogue is not provided. Thus, the

corpora (used in this study) do not represent comparable speech act data which enable like-for-like comparison.

In discussing comparability, it is important to refer briefly to the concept of the pragmatic variable. The broadening of Labov's definition of the linguistic variable to include application to other levels of linguistic analysis, e.g. discourse pragmatic analysis, has been proposed by linguists including Pichler (2013); Barron and Schneider (2009); Barron (2017b). Staley (2018, p.12) points out that in defining variables in pragmatics, we need to broaden the concept of sameness. When form is the starting point, the definition of a pragmatic variable must be able to account for multiple potential propositional, referential, procedural, and interpersonal meanings of each variant, in addition to their potential social meaning. Forms of pragmatic variables convey these different meanings, depending on the social context. In addition, the forms that can be exploited to achieve a particular illocutionary force will likewise vary according to the context. Therefore, the variants of a pragmatic variable cannot be defined without or before establishing the context. The complete interaction needs to be analysed to discover how individual utterances relate to the speech event (the entire interaction between the interlocutors) as a whole; this may be made up of multiple speech acts, or a speech act which is developed over several utterances. Barron (2017b) points out that a lack of situational information can present difficulties in interpreting the discourse. As a result, this can affect the analysis of response tokens (offer/ request/ narrative, etc.) which may depend on the speech act they represent. For example *no problem* can be a response to *thanks* or to a request. In addition, in this research **all** different text types are used for the pragmatic markers analysis which

may cause a potential lack of equivalence if more of these pragmatic markers are taken from one text type than another across varieties.

In addition, Pichler (2013) proposes that functional equivalence (variants that have the same function but different syntax e.g. the response tokens *no problem* and *you're welcome* can both function as a response to *thanks* but are syntactically different) is not enough in defining the variable context. In addition to functional equivalence, in analysing discourse pragmatic markers, Pichler takes derivational equivalence as a defining factor. She gives the example of "*I don't know* and *I don't think* as constituting variables due to the fact that their respective variants are derived from the same linear string of components" (2013, p.13). This allows for the narrowing of the variable context (if based on function only) of a very large set of variants that in only some contexts share the same meaning and function. Therefore it is important in variational pragmatics to explore variation on multiple levels of variables.

Another limitation of this study is that it cannot allow generalisation of the results as regards both Northern Ireland (NI) and the Republic of Ireland (ROI). ICE-Ireland is made up of data from NI and the ROI; both data-sets have the same corpus design and scheme for grammatical annotation, comprise the same type and number of texts and present Irish English. However, these datasets represent two different regions and each has different political climates and different language origins (as discussed in Chapter 4). Therefore, the focus here (as proposed in Chapter 4) is on the spoken component of ICE (ROI) because this study is a regional variation study investigating a language in one single zone, sharing the same political and linguistic climate. This is important because the pragmatic analysis depends heavily on context, which can be easily influenced by different factors triggered by

some of the facts mentioned below. The English of Northern Ireland and the English of Southern Ireland are in two different language zones representing a distinctive kind of cross-linguistic influence (Kallen & Kirk 2007; Kallen 2012). In other words, Ireland can be divided, in linguistic terms, into two sections; one section is the north (or the province of Ulster). This comprises the six counties of NI (which are part of the UK) and the ROI county of Donegal (Hickey 2004, p.30); the second section is that of the south which comprises the provinces of Munster, Connacht and Leinster, encompassing the remaining counties of Ireland. As a result, for the purpose of this study, the focus was only on the data collected from the Republic of Ireland. The rationale for this is based on the following facts: 1) the dialect of English spoken in NI (presented in ICE (NI)) derives from Lowland Scots and different English forms spoken in northern England which were taken to Ulster during the plantations of the seventeenth century (Hickey 2012); 2) according to Kallen and Kirk (2006), the different governmental, administrative, and economic environments which both sub-corpora present (linguistic areas), have reflected on the notion or the hypothesis that governments affect the development of standard language. For example, the presentations of the official terminology in both sub-corpora varies in terms of its linguistic origin which somehow has been triggered by political agendas. Also, the presentation of what is called "Irishness" or Irish words in terms of usage in both sub-corpora is quite different (Kallen & Kirk 2007).

This study on spoken grammar has been conducted based on the five types of pragmatic markers including DMs, RTs, questions and tags, hedges, and SMs (the core components of spoken grammar) (Carter & McCarthy 2006). It excludes some types of spoken grammar due to the fact that they are either not found at the top 100 in either variety

or they have a very limited number of occurrences in the data, (for example, interjections such as *oh* or deictic markers such as *that* which have been found in the top 100 most frequent wordlists in both the ICE-Ireland and ICE-Canada corpora, performing various pragmatic functions in both varieties). Ellipsis, for example, was not included as it was too challenging to examine the omission of items within the methodology that was used. Also, the deixis (deictic markers/system) in this study was ignored unless they occurred in clusters such as *you know* and *I mean* which are analysed all together as a cluster. This is because the nature of deixis often makes it difficult to interpret conversational transcripts due to the various interpersonal meanings generated (see Leech 2000; McCarthy & O'Keeffe 2014). For example, the usage of *we* can be used for solidarity and a collective sense. As McCarthy and Handford (2004) show, authoritative individuals in corporate settings can construe *we* to take on different meanings to reflect their more covert intentions. *We* can be used to take on a more collective, inclusive meaning when used to represent the whole corporation; or, it can be inserted into the highly occurring chunk *we need to* in order to lessen the authoritative command that is truly meant—*you must*, thereby creating a manufactured solidarity between the chains of command in the workplace (McCarthy & O'Keeffe 2014; also see Drew & Heritage 1992).

Analysing the entire spoken grammar of two distinctive varieties in terms of forms and functions comparatively is not an easy task and cannot be done without shortcomings/limitations especially if such an investigation was not done before. Thus, there is an underrepresentation of functions realised via a variety of different forms due to their multi-functionalities. This is the case especially after knowing that the top-down analysis can be expanded upon based on the existing literature which is still growing. Also,

using different corpora to investigate the multi-functional spoken grammar items can reveal more insights and results. As a result, there are multi-functional spoken grammar items in this study that have a better chance of being analysed further such as:

- *So* which can be used as an intensifier as in, *so good*; as a vague extended (vagueness marker), *and so on* (as discussed in Chapter 5, section 5.4.2); and as discourse marker, *so let us see* (which did not appear in the top 100 in ICE-Ireland or ICE-Canada).

- *Really* which can be an intensifier and follow-up question, *really?* And used in tags as in, *isn't it really?* (as shown in Chapter 7, section 7.4.2.1).

- *You know* which can be an invariant tag and question as in, *do you know* (as demonstrated in Chapter 7, section 7.4.2.1).

- *Like* which can be a downgrader (hedge) versus a focuser *like* (as discussed in Chapter 5, section 5.4.2 and Chapter 8, section 8.4.1).

This underrepresentation of functions for some of the forms can be seen as a bias towards spoken grammar items which are only multi-functional. As a result, this is potentially why a salient index of Irish identity, such as *shur*, is not analysed.

Lastly, this study is the first to look comparatively at Irish English and Canadian English in terms of spoken grammar forms and functions as a whole. Thus, it is necessarily broad, looking at the top 100 forms, across the core features of spoken grammar. Ideally, a study would look at all forms that manifest in spoken grammar across these varieties but as a starting point, this study had to limit itself to a manageable cut off point (of the top 100). The result is that figures presenting forms are general, and they present only the frequency of the form itself in the data as a whole and not the particular functions performed by the

forms in the data. Also, quantitative data at the functional level is not provided (due to the

huge number of forms and functions needed to be analysed across all the pragmatic

markers). Therefore, the analysis and findings (generalizations) arrived in this study are

limited to the samples of the most frequent forms/ clusters and do not provide an overview

of the overall use of a category (e.g. analysis of all responses to *thanks*). However, while

this is a limitation in terms of the depth of findings across any one feature, it does allow for

breadth of findings across spoken grammar as a whole in a comparative context.

10.4 Directions for Further Research

There are several directions that can be proposed for further research. This present study

claims that no research has examined or provided the full comparative description of

spoken grammar of Irish English and Canadian English in comparison to each other. This

highlights the originality of the present research. Given the lack of previous studies on this

topic, more work is essential to confirm the results of the present research, and this should

be done with other available corpora for Irish and Canadian Englishes such as LCIE and the

Strathy corpus. Another direction is that since the focus of this present research was on

spoken grammar as a whole across both varieties, there is still much scope to explore and

compare variables such as age, gender, setting and so forth. What this research offers is an

in-depth comparative description in relation to spoken grammar that will be a useful

starting point for traditional sociolinguistic studies. Future work can take the results of any

one chapter in this thesis and look at them in terms of form and function across specific

variables. Thus, prominent forms of spoken grammar across these two varieties can be

further explored from the perspectives of a macro-social perspective and a micro-social

perspectives such as age, gender, or social class and how these interplay with pragmatic choices.

This current research can be easily expanded through the primary framework of Variational Pragmatics taken in this study to look at other levels of pragmatic analysis such as: actional, interactional, topical, and organisational. There is also scope to expand this to the English of Northern Ireland; this may yield in interesting results due to the different climates in terms of politics, economy, and so forth. Moreover, the model of analysis in this study can pave the way for the comparison of other neglected English varieties.

10.5 Final Conclusion

This study involved much detailed and painstaking analysis but what this study points to in general is a stronger tendency in Irish English to soften the message, show stronger convergence as a listener, and generally use less face-threatening language. While these functions are also associated with pragmatic markers in Canadian English, there is evidence of more use of information and shared knowledge marking and monitoring that suggests more attention on the part of Canadian English speakers to the importance of clarity and the conveyance of the message, even if, at times, it means using more direct language forms that Irish English speakers would not tend to use (see the quantitative and qualitative findings presented in Chapter 5, section 5.4; Chapter 6, section 6.4; Chapter 7, section 7.4 and Chapter 8, section 8.4). In relation to variational pragmatics, it was found that while forms may seem to be in common in varieties of the same language, when we look at their pragmatic function, we can reveal that they often function in different ways. Therefore, the results offer a baseline description of the commonalities and differences in terms of spoken

grammar and pragmatics (in terms of forms and functions) across the two varieties of English which may have application to the study of other varieties of English. The study offers a methodology which will serve the new area of Corpus Pragmatics well because it incorporates both the corpus tradition of moving from forms to functions and the Pragmatics tradition of beginning with the function of language being explored. Most of all, it offers a comprehensive exploration of spoken language in two varieties of English that are important to me.

REFERENCES

Abraham, W. (1991) "Discourse particles in German: How does their illocutive force come about" in Abraham W. ed., *Discourse Particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German*, Amsterdam: John Benjamins, 203-252.

Ädel, A. and Reppen, R., eds. (2008) *Corpora and discourse: The challenges of different settings*, Amsterdam: John Benjamins.

Adolphs, S., Brown, B., Carter, R., Crawford, P. and Sahota, O. (2004) "Applying clinical linguistics: Corpus linguistics in health care settings" *Journal of Applied Linguistics,* 1, 9-28.

Aijmer, K. (1996) *Conversational routines in English: Convention and creativity,* New York: Addison Wesley Longman Limited.

Aijmer, K. (1997) *Conversational routines in English*, London: Longman.

Aijmer, K. (2002) "English discourse particles. Evidence from a corpus" *Studies in Corpus Linguistics,10,* Amsterdam: John Benjamins.

Aijmer, K. (2009) *Corpora and language teaching*, Amsterdam: John Benjamins.

Aijmer, K. (2018) "Corpus pragmatics: From form to function" in Jucker, A., Schneider, K. and Bublitz, W. eds., *Methods in pragmatics,* Berlin: De Gruyter Mouton, 555-585.

Aijmer, K. and Rühlemann, C. (2014) *Corpus pragmatics: A handbook,* Cambridge: Cambridge University Press.

Alami, M. (2015) "Pragmatic functions of discourse markers: A review of related literature" *International Journal on Studies in English Language and Literature,* 3(3), 1-10.

Algeo, J. (1988) "The tag question in British English: It's different, i'n'it?" *English World-Wide,* 9(2), 171-191.

Algeo, J. (1990) "'It's a myth, innit? Politeness and the English tag question" in Ricks, C. and Michaels, L. eds., *The State of Language*, Berkeley: University of California Press, 443-450.

Algeo, J. (2006) *British or American English? A handbook of word and grammar patterns*, Cambridge: Cambridge University Press.

Algi S. (2012) "Hedges and boosters in L1 and L2 argumentative paragraphs: Implications for teaching L2 academic writing," unpublished thesis (MA), Middle East Technical University.

Al-Hindawi, F. and Alkhazaali, M. (2016) "A critique of politeness theories" *Theory and Practice in Language Studies*, 6(8), 1537-1545.

Allerton, D. (2009) "Tag questions" in Rohdenburg, G. and Schlüter, J. eds., *One language, two grammars? Differences between British and American English*, Cambridge: Cambridge University Press, 306-323.

Andersen, G. (2001) *Pragmatic markers and sociolinguistic variation: A relevance theoretic approach to the language of adolescents*, Amsterdam: John Benjamins.

Amador-Moreno, C. (2005) "Discourse markers in Irish English: An example from literature" in Barron, A. & Schneider, K. eds., *The pragmatics of Irish English*, Berlin: Mouton de Gruyter, 73-100.

Amador-Moreno, C. (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. (2015) "There's, *like,* total silence again, roysh, and no one says anything: Fictional representations of 'new' DMs and quotatives in Irish English" in Amador-Moreno, C., McCafferty, K., and Vaughan, E. eds., *Pragmatic markers in Irish English*, Amsterdam: John Benjamins, 370-389.

Amador-Moreno, C., McCafferty, K. and Elaine Vaughan, E. eds. (2015) *Pragmatic markers in Irish English*, Amsterdam: John Benjamins.

Andersen, G. (1998) "Are tag questions questions? Evidence from spoken data" presented at *19th ICAME Conference*, Belfast, 20-24 May.

Anthony, L. (2020) Antconc (Version 3.5.9) [Computer Software], Tokyo: Waseda University.

Archer, D., Aijmer, K. and Wichmann, A. (2012) *Pragmatics: An Advanced Introduction for Students*, London: Routledge.

Asián, A. and McCullough, J. (1998) "Hiberno-English and the teaching of modern contemporary Irish Literature in an EFL context" *Letters and Links*, 5, 37-60.

Atkinson, D. and Biber, D. (1994) "Register: A review of empirical research" in Biber, D. and Finegan E. eds., *Sociolinguistic perspectives on register variation*, New York: Oxford University Press.

Atkinson, J. and Heritage, J. eds. (1984) *Structures of social action: Studies in conversation analysis*, Cambridge: Cambridge University Press.

Avis, W. (1972) "So *eh?* is Canadian, *eh?*" in Avis, W. ed., *Essays and articles*, Kingston: Royal Military College of Canada, 172-190.

Axelsson, K. (2011) "Tag questions in fiction dialogue" unpublished thesis (Ph.D.), University of Gothenburg.

Baker, P. (2010) "Corpus methods in linguistics" in Litosseliti, L. ed., *Research methods in linguistics*, London: Continuum, 93-113.

Baker, P., Gabrielatos, C. and McEnery T. (2013) *Discourse Analysis and Media Attitudes: The representation of Islam in the British Press*, Cambridge: Cambridge University Press.

Baker, P. and McEnery, T. eds. (2015) *Corpora and discourse studies: Integrating discourse and corpora,* Basingstoke: Palgrave Macmillan.

Banim, J. (1838) "The stolen sheep" in *The bit o' writin': And other tales*, Philadelphia: E.L. Carey and A. Hart, 171-184.

Baratta, A. (2009) "Revealing stance through passive voice" *Journal of Pragmatics*, 41(7), 1406-1421.

Barber, K. ed. (2004) *The Canadian Oxford dictionary,* Toronto: Oxford University Press.

Barbuto, R. (2013) *The Canadian theater, 1813: U.S. army campaigns of the War of 1812*, Washington: Center of Military History United States Army.

Barron, A. (2014) "Variational pragmatics" in Chapelle, C. ed., *The Encyclopedia of Applied Linguistics (EAL)*, New Jersey: Wiley Blackwell.

Barron, A. (2015) "And your wedding is the twenty-second <.> of June is it?: Tag questions in Irish English" in Amador- Moreno, C., McCafferty, K. and Vaughan, E. eds., *Pragmatic Markers in Irish English,* Amsterdam: John Benjamins, 203-228.

Barron, A. (2017a) "The speech act of offers in Irish English" *World Englishes*, 36(2), 224-238.

Barron, A. (2017b) "Variational Pragmatics" in Barron, A., Gu, Y., and Steen, G. eds., *The Routledge Handbook of Pragmatics,* Oxon/New York: Routledge, 91-104.

Barron, A. (2019) "Pragmatic development and stay abroad" *Journal of Pragmatics*, 146, 43-53.

Barron, A., Pandarova, I. and Karoline, M. (2015) "Tag questions across Irish English and British English: A corpus analysis of form and function" *Multilingua,* 34(4), 495-524.

Barron, A. and Pandarova, I. (2016) "The sociolinguistics of language use in Ireland" in Hickey, R. ed., *Sociolinguistics in Ireland*, Basingstoke: Palgrave Macmillan, 107-130.

Barron, A and Schneider, K. eds. (2005) *The pragmatics of Irish English*, Berlin & New York: Mouton de Gruyter.

Barron, A. and Schneider, K. (2009) "Variational pragmatics: Studying the impact of social factors on language use in interaction" *Intercultural pragmatics,* 6(4), 425-442.

Beeching, K. and Detges, U. (2014) *Discourse functions at the left and right periphery crosslinguistic investigations of language use and language change*, Leiden: Brill.

Bennett, G. (2010) "Using corpora in the language learning classroom: Corpus linguistics for teachers" *Michigan ELT*, available: https://www.press.umich.edu/pdf/9780472033850-part1.pdf.

Besnier, N. (1990) "Language and affect" *Annual Review of Anthropology,* 19, 419-451.

Bianchi, F. (2012) *Culture, corpora and semantics: Methodological issues in using elicited and corpus data for cultural comparison,* Lecce: Universita del Salento.

Biber, D. (1988) *Variation across speech and writin*g, Cambridge: Cambridge University Press.

Biber, D. (1990) "Methodological issues regarding corpus-based analysis of linguistic variation" *Literary and Linguistic Computing,* 5(4), 257-269.

Biber, D. (1993) "Representativeness in corpus design" *Literary and Linguistic Computing,* 8(4), 243-257.

Biber, D. (2006a) "Stance in spoken and written university registers" *Journal of English for Academic Purposes*, 5, 97-116.

Biber, D. (2006b) *University language: A corpus-based study of spoken and written registers,* Amsterdam: John Benjamins.

Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus linguistics: Investigating language structure and use,* Cambridge: Cambridge University Press.

Biber, D. and Finegan, E. (1988) "Adverbial stance types in English" *Discourse Processes*, 11(1), 1-34.

Biber, D. and Finegan, E. (1989) "Styles of stance in English: Lexical and grammatical marking of evidentiality and affect" *Text*, 9(1), 93-124.

Biber, D., Fitzpatrick, S. and Reppen, R. eds. (2002) *Using corpora to explore linguistic variation,* Amsterdam: John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman grammar of spoken and written English*, Harlow: Longman.

Bijeikienė, V. and Tamošiūnaitė, A. (2013) *Quantitative and qualitative research methods in sociolinguistics: Study guide: A resource book for students,* Kaunas: Vytautas Magnus University.

Blakemore, D. (1989) "Meaning and force: The pragmatics of performative utterances" *Mind and Language*, 4(3), 235-245.

Boberg, C. (2004) "Canadian English" in Kortmann, B. and Schneider, E. eds., *A Handbook of varieties of English: The Americas and the Caribbean*, Berlin: Mouton de Gruyter, 351-365.

Bolinger, D. (1957) *Interrogative structures of American English: The direct question. Publications of the American dialect society*, 28, Alabama: University of Alabama Press.

Bolinger, D. (1989) *Intonation and its uses: Melody in grammar and discourse,* Stanford: Stanford University Press.

Borlongan, A. (2008) "Tag questions in Philippine English" *Philippine Journal of Linguistics,* 39(1), 1-34.

Bradac, J., Mulac, A. and Thompson, S. (1995) "Men's and women's use of intensifiers and hedges in problem-solving interaction: Molar and molecular analysis" *Research on Language and Social Interaction,* 28(2), 93-116.

Brazil, D. (1995) *A grammar of speech,* Oxford: Oxford University Press.

Brinton, L. (1996) *Pragmatic markers in English. Grammaticalization and discourse functions*, Berlin: Mouton de Gruyter.

Brinton, L. (2008) *The comment clause in English: Syntactic origins and pragmatic development*, Cambridge: Cambridge University.

Brown, P. and Levinson, S. (1978) "Universals in language usage: Politeness phenomena" in Goody, E. ed., *Questions and politeness: Strategies in social interaction,* Cambridge: Cambridge University Press, 56-289.

Brown, P. and Levinson, S. (1987) *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bublitz, W. (1998) "Cohesion and coherene" *Handbook of Pragmatics Online,* 4, 1877-9611, available: https://doi.org/10.1075/hop.4.coh1.

Bybee, J. and Fleischman, S. (1995), *Modality in grammar and discourse,* Amsterdam: John Benjamnins.

Carleton, W. (1843) *Traits and stories of the Irish peasantry*, Dublin: William Curry, Jun, and Co.

Carter, R. and McCarthy, M. (1995) "Grammar and the spoken language" *Applied Linguistics,* 16(2), 141-158.

Carter, R. and McCarthy, M. (2002) "From conversation to corpus: A dual analysis of a broadcast political interview" in Sanchez-Macarro, A. ed., *Windows on the world: Media discourse in English*, Valencia: University of Valencia Press, 15-39.

Carter, R. and McCarthy, M. (2006) *Cambridge grammar of English: A comprehensive guide to spoken and written English grammar and usage*, Cambridge: Cambridge University Press.

Carter, R. and McCarthy, M. (2015) "Spoken grammar: Where are we and where are we going?" *Applied Linguistics*, *Online Advanced Access,* available: http://doi.org/10.1093/applin/amu080.

Carter, R. and McCarthy, M. (2017) "Spoken grammar: Where are we and where are we going?" *Applied Linguistics* 38(1), 1-20.

Carter, R., McCarthy, M., Mark, G. and O'Keeffe, A. (2011) *English grammar today: An A-Z of spoken and written grammar,* Cambridge: Cambridge University Press.

Chapman, S. (2013) *Pragmatics*, New York: Palgrave Macmillan.

Chafe, W. (1986) "Evidentiality in English conversation and academic writing" in Chafe, W. and Nichols, J. eds., *Evidentiality: The linguistic coding of epistemology,* Norwood: Ablex.

Chafe, W. and Nichols, J., eds., (1986) *Evidentiality: The linguistic coding of epistemology,* Norwood: Ablex.

Chambers, J. and Trudgill, P. (1998) *Dialectology,* Cambridge: Cambridge University Press.

Chang, W. (2008) "Australian and Chinese perceptions of (im)politeness in an intercultural apology" *Griffith Working Papers in Pragmatics and Intercultural Communication*, 1(2), 59-74.

Channell, J. (1990) "Precise and vague quantities in writing on economics" in Nash, W. ed., *The Writing Scholar*, Newbury Park: Sage, 95-117.

Charles, M. (2006) "Phraseological patterns in reporting clauses used in citation: A corpus based study of theses in two disciplines" *English for Specific Purposes*, 25(3), 310-331.

Charles, M. (2007) "Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions" *Journal of English for Academic Purposes*, 6(4), 289-302.

Charteris-Black, J. (2004) *Corpus approaches to critical metaphor analysis*, Basingstoke: Palgrave Macmillan.

Cheng, W. and Warren, M. (2000) "The Hong Kong corpus of spoken English: Language learning through language description" in Burnard, L. and McEnery, T. *Rethinking language pedagogy from a corpus perspective*, Frankfurt am Main: Peter Lang, 133-144.

Cheng, W. and O'Keeffe, A. (2014) "Vagueness" in Aijmer, K. and Rühlemann, C. eds., *Corpus pragmatics: A handbook*, Cambridge: Cambridge University Press, 360-378.

Cheshire, J. (2016) "Epilogue: The future of discourse-pragmatic variation and change research" in Pichler, H. ed., *Discourse-pragmatic variation and change in English*, Cambridge: Cambridge University Press, 252-266.

Clancy, B. (2005) "You're fat. You'll eat them all: Politeness strategies in family discourse" in Barron, A. and Schneider, K. eds., *The pragmatics of Irish English*, Berlin: Mouton de Gruyter, 177-199.

Clancy, B. (2010) "'Hurry up baby son all the boys is finished with their breakfast'": A sociopragmatic analysis of Irish settled and traveller family discourse," unpublished thesis (Ph.D.), Mary Immaculate College, University of Limerick.

Clancy, B. (2011a) "Complementary perspectives on hedging behaviour in family discourse: The analytical synergy of variational pragmatics and corpus linguistics" *International Journal of Corpus Linguistics*, 16(3), 371-390.

Clancy, B. (2011b) "'Do you want to do it yourself like?' Hedging in Irish traveller and settled family discourse" in Davies, B., Haugh, M. and Merrison, A. eds., *Situated Politeness*, London: Continuum, 129-146.

Clancy, B. and O'Keeffe, A. (2015) "Pragmatics" in Biber, D. and Reppen, R. eds., *The Cambridge handbook of corpus linguistics*, Cambridge: Cambridge University Press, 235- 251.

Clancy, B., O'Keeffe, A. and Adolphs, S. (2019) *Introducing Pragmatics in Use*, 2nd ed., London: Routledge.

Clancy, B. and Vaughan, E. (2012) "'It's lunacy now': A corpus-based pragmatic analysis of the use of *now* in contemporary Irish English" in Migge, B. and Ní Chiosáin, M. eds., *New Perspectives on Irish English*, 225-246.

Clarke, S. (2012) "From Ireland to Newfoundland: What's the perfect after doing?" in Migge, B., Chiosáin, M. eds., *New perspectives on Irish English,* Amsterdam: John Benjamins, 101-130.

Clemen, G. (1997) "The concept of hedging: Origin, approaches and definitions" in Markkanen, R. and Schröder, H. eds., *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts*, Berlin: Walter de Gruyter, 235-248.

Columbus, G. (2010) "A comparative analysis of invariant tags in three varieties of English" *English World-Wide*, 31(3), 288-310.

Cook, G. (1990). "Transcribing infinity: Problems of context presentation" *Journal of Pragmatics,* 14(1), 1-24.

Corrigan, K. (2015) "'I always think of people here saying 'like' after every sentence': The dynamics of discourse-pragmatic markers in Northern Irish English" in Amador-Moreno, A., McCafferty, K. and Vaughan, E. eds., *Pragmatic markers in Irish English,* Amsterdam: John Benjamins, 37-64.

Cotterill, J. (2004) "Collocation, connotation and courtroom semantics: Lawyers control of witness testimony through lexical negotiation" *Applied Linguistics,* 25, 513-37.

Crismore, A., Markkanen, R. and Steffensen, M. (1993) "Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students" *Written Communication*, 10(1), 39-71.

Crompton, P. (1997) "Hedging in academic writing: Some theoretical problems" *English for Specific Purposes,* 16(4), 271-287.

Crompton, P. (1998) "Identifying hedges: Definitions or divination" *English for Specific Purposes,* 17, 303-313.

Cruse, D. (2002) "Hyponymy and its varieties" in Green, R., Bean, C. and Myaeng, S. eds., *The semantics of relationships,* Information Science and Knowledge Management, vol 3, 3-21.

Culpeper, J. (2008) "Reflections on impoliteness, relational work and power" in Bousfield, D. and Locher, M. eds., *Impoliteness in language: Studies on its interplay with power in theory and practice*, Berlin: Mouton de Gruyter, 17-43.

Cutrone, P. (2011) "Politeness and face theory: Implications for the backchannel style of Japanese L1/L2 speakers" in Giannoni, D. and Ciarlo, C. eds., *University of Reading Language Studies Working Papers*, 3, 51-57.

Da Mota, C. and Herment, S. (2016) "The pragmatic functions of the final particle 'eh' and of high rising terminals in Canadian English: Quite similar, eh!" at *Speech Prosody,* Boston, 31 May- 3 June.

D'Arcy, A. (2008) "Canadian English as a window to the rise of *like* in discourse" *International Journal of English Studies,* 19(2), 125-140.

D'Arcy, A. (2017) *Discourse-pragmatic variation in context: Eight-hundred years of 'like'*, Amsterdam: John Benjamins.

Das, D. and Taboada, M. (2018) "Signalling of coherence relations in discourse" *Discourse Processes,* 55(8), 743-770.

Denis, D. and Tagliamonte, S. (2016) "Innovation, 'right'? Change, 'you know'? Utterance-final tags in Canadian English" in Pichler, H. ed., *Discourse-pragmatic variation and change in English: New methods and insights*, Cambridge: Cambridge University Press.

Dér, C. (2010) "On the status of discourse markers" *Acta Linguistica Hungarica*, 57(1), 3-28.

Diani, M. (2008) "Social movement theory and grassroots coalitions in the Middle East" Paper for the 2008 ASA Annual Meeting, Boston, August 1-4.

Dickinson, M. (2015) "Detection of annotation errors in corpora" *Language and Linguistics Compass*, 9(3), 119-138.

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.

Dixon, J. and Foster, D. (1997) "Gender and hedging: From sex differences to situated practice" *Journal of Psycholinguistic Research,* 26(1), 89-107.

Dolan, T. P. (2004) *A dictionary of Hiberno-English*, 2nd ed. Dublin: Gill & Macmillan.

Drew, P. and Heritage, J. eds. (1992) *Talk at work: Interaction in institutional settings,* Cambridge: Cambridge University Press.

Drew, P. and Chilton, K. (2000) "Calling just to keep in touch: Regular and habitualised telephone calls as an environment for small talk" in Coupland, J. ed. *Small talk*, Harlow: Longman, 137-162.

Drummond, K. and Hopper, R. (1993a) "Back channels revisited: Acknowledgment tokens and speakership incipiency" *Research on Language and Social Interaction,* 26(2), 157-177.

Drummond, K. and Hopper, R. (1993b) "Some uses of *yeah*" *Research on Language and Social Interaction,* 26(2), 203-212.

Duncan, S. and Niederehe, G. (1974) "On signalling that it's your turn to speak" *Journal of Experimental Social Psychology, 10,* 234-247.

Duranti, A. and Goodwin, C. (1992) *Rethinking context: Language as an interactive phenomenon*, Cambridge: Cambridge University Press.

Edwards, J. and Lampert, M. eds. (1992) *Talking data: Transcription and coding methods for language research*, Hillsdale: Lawrence Erlbaum Associates.

Eggins, S. and Slade, D. (2001) *Analyzing casual conversation*, London: Continuum.

Erman, B. (1987) *Pragmatic expressions in English: A study of "you know," "you see," and "I mean" in face-to-face conversation,* Stockholm: Almqvist & Wiksell.

Erman, B. (2001) "Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk" *Journal of Pragmatics*, 32(9), 1337-1359.

Fahnestock, J. (1986) "Accommodating science: The rhetorical life of scientific facts" *Written Communication,* 3(3), 275-296.

Farenkia, B. (2012) "Face-saving strategies in responding to gratitude expressions: Evidence from Canadian English" *International Journal of English Linguistics,* 2(4), 1-11.

Farr, F. (2002) "Classroom interrogations – How productive?" *Teacher Trainer,* 16(1), 19-23.

Farr, F. (2003) "Engaged listenership in spoken academic discourse: The case of student-tutor meetings" *Journal of English for Academic Purposes,* 2(1), 67-85.

Farr, F. and O'Keeffe, A. (2002) "*Would* as a hedging device in an Irish context: An intra-varietal comparison of institutionalised spoken interaction" in Reppen, R., Fitzmaurice, S. and Biber, D. eds., *Using corpora to explore linguistic variation*, Amsterdam: John Benjamins, 25-48.

Farr, F., Murphy, B. and O'Keeffe, A. (2002) "The Limerick corpus of Irish English: Design, description, and application," *Teanga*, 21, 5-29.

Fasold, R. (1990) *The sociolinguistics of language*, Oxford: Blackwell.

Félix-Brasdefer, J. (2005) "Indirectness and politeness in Mexican requests" *Select Proceedings of the 7th Hispanic Linguistics Symposium,* 6678, Cascadilla Proceedings Project.

Félix-Brasdefer, J. (2015) *The language of service encounters: A pragmatic-discursive approach*, Cambridge: Cambridge University Press.

Fellegy, A. (1995) "Patterns and functions of minimal response" *American Speech*, 70(2), 186-199.

Fillmore, C. (1992) "'Corpus linguistics' or 'Computer-aided armchair Linguistics'" in Svartvik, J. ed., *Directions in corpus linguistics. Proceedings of Nobel Symposium 82*, Stockholm, 4-8 August 1991, *Trends in Linguistics. Studies and Monographs* 65, Berlin: Mouton de Gruyter, 35-60.

Filppula, M. (2002) *The grammar of Irish English. Language in Hibernian style*, London: Routledge.

Fishman, P. (1978) "Interaction: The work women do" *Social Problems*, 25(4), 397-406.

Fitzmaurice, S. (2004) "Subjectivity, intersubjectivity and the historical construction of interlocutor stance: From stance markers to discourse markers" *Discourse Studies*, 6(4), 427-448.

Flöck, I. (2016) *Requests in American and British English: A contrastive multi-method analysis,* Amsterdam: John Benjamins.

Fordyce, K. (2009) "A comparative study of learner corpora of spoken and written discursive language: Focusing on the use of epistemic forms by Japanese EFL learners" *Hiroshima Studies in Language and Language Education*, 12, 135-50.

Fordyce, K. (2014) "The differential effects of explicit and implicit instruction on EFL learners' use of epistemic stance" *Applied Linguistics*, 35(1), 6-28.

Fraser, B. (1990a) "An approach to discourse markers" *Journal of Pragmatics*, 14, 383-395.

Fraser, B. (1990b) "Perspectives on politeness" *Journal of Pragmatics*, 14(2), 219-236.

Fraser, B. (1996) "Pragmatic markers" *Pragmatics*, 6(2), 167-190.

Fraser, B. (1999) "What are discourse markers?" *Journal of Pragmatics*, 31, 931-952.

Fries, C. (1952) *The structure of English: An introduction to the construction of sentences*, New York: Harcourt-Brace.

Fukada, A. and Asato, N. (2004) "Universal politeness theory: Application to the use of Japanese honorifics" *Journal of Pragmatics,* 36(11), 1991-2002.

Fujita, Y. (2001) "Functions of discourse markers in Japanese" *Texas Papers in Foreign Language Education,* 6(1), 147-162.

Fung, L. and Carter, R. (2007) "Discourse markers and spoken English: Native and learner use in pedagogic settings" *Applied Linguistics,* 28(3), 410-439.

Gablasova, D., Brezina, V., Mcenery, T. and Boyd, E. (2015) "Epistemic stance in spoken L2 English: The effect of task and speaker style" *Applied Linguistics*, 1-26.

Gabrielatos, C. and McEnery, T. (2005) "Epistemic modality in MA dissertations" in Fuetres, P. ed., Lengua y sociedad: Investigacionesrecientes en Lingüística Aplicada. *Lingüística y Filología* [Linguistics and society: Recent studies in Applied Linguistics], Valladolid: Universidad de Valladolid, 311-331.

Gabryś-Barker, D. (2016) "New approaches to multilingualism research: Focus on metaphors and similes" presented at *25th PASE Conference*, University of Silesia, April 2016.

Gardner, R. (1997) "The conversation object *mm*: A weak and variable acknowledging token" *Research on Language and Social Interaction*, 30(2), 131-156.

Gardner, R. (1998) "Between speaking and listening: The vocalisation of understandings" *Applied Linguistics*, 19(2), 204-224.

Gardner, R. (2001) *When listeners talk: Response tokens and listener stance,* Amsterdam: John Benjamins.

Gatto, M. (2015) "The web as corpus: Theory and practice" *Digital Scholarship in the Humanities,* 30(4), 611-614.

Goldschmidt, M. (1998) "*Do me a favor*: A descriptive analysis of favor asking sequences in American English" *Journal of Pragmatics,* 29(2), 129-153.

Grabe, W. (1984) "Towards defining expository prose within a theory of text construction" unpublished dissertation (Ph.D.), University of Southern California.

Grenoble, L. and Riley, M. (1996) "The role of deictics in discourse coherence: French *voici/voilà* and Russian *vot/von*" *Journal of Pragmatics*, 25, 819-838.

Greaves, C. and Warren, M. (2010) "What can a corpus tell us about multi-word units?" in O'Keeffe, A. and McCarthy, M. eds., *The Routledge handbook of corpus linguistics,* London: Routledge, 212-226.

Greenbaum, S. (1988) "A proposal for an international computerized corpus of English" *World Englishes,* 7(3), 315.

Greenbaum, S. (1991) 'The development of the international corpus of English" in Aijmer, K. and Altenberg, B. eds., *English Corpus Linguistics*, London: Longman, 83-91.

Halliday, M. (2002) *On Grammar*, London: Continuum.

Harrington, J. (2010) *Phonetic analysis of speech corpora*, Oxford: Wiley-Blackwell.

Hewings, M. and Hewings, A. (2002) "'It is interesting to note that ...': A comparative study of anticipatory *it* in student and published writing" *English for Specific Purposes*, 21, 367-383.

Hickey, R. ed. (2004) *Legacies of colonial English,* Cambridge: Cambridge University Press.

Hickey, R. (2007a) *Irish English: History and present-day forms*, Cambridge: Cambridge University Press.

Hickey, R. (2007b) "Southern Irish English" in Britain, D. ed., *Language in the British Isles,* 2nd ed., Cambridge: Cambridge University Press, 135-151.

Hickery, R. (2012) "English in Ireland" in Hickey, R. *Areal features of the Anglophone world*, Berlin: Mouton de Gruyter, 79-108.

Hickey, R. (2011) *The dialects of Irish. Study of a changing landscape*, Berlin: Mouton de Gruyter.

Hickey, R. (2015) "The pragmatics of Irish English and Irish" in Amador-Moreno, A., McCafferty, K. and Vaughan, E. eds., *Pragmatic markers in Irish English,* Amsterdam: John Benjamins, 17-36.

Hinkel, E. (1995) "The use of modal verbs as a reflection of cultural values" *TESOL Quarterly,* 29(2), 325-343.

Hinkel, E. (2005) "Hedging, inflating, and persuading in L2 academic writing" *Applied Language Learning,* 15, 29-53.

Hirschberg, J. and Litman, D. (1993) "Empirical studies on the disambiguation of cue phrases" *Computational Linguistics*, 19(3), 501-530.

Holmes, J. (1982) "The functions of tag questions" *English Language Research Journal*, 3, 40-65.

Holmes, J. (1985) "Sex differences and miscommunication: Some data from New Zealand" in Pride, J. ed., *Cross-cultural encounters: communication and miscommunication*, Melbourne: River Seine, 24-43.

Holmes, J. (1986) "Functions of *you know* in women's and men's speech" *Language in Society,* 15(1), 1-21.

Holmes, J. (1990) "Hedges and boosters in women's and men's speech" *Language and Communication,* 10(3), 185-205.

Holmes, J. (1993) "New Zealand women are good to talk: An analysis of politeness strategies in interaction" *Journal of Pragmatics,* 20, 91-116.

Holmes, J. (1995) *Women, men and politeness*, White Plains, NY: Longman.

Hosman, L. (1989) "The evaluative consequences of hedges, hesitations, and intensifiers. Powerful and powerless speech styles" *Human Communication Reseach*, 15(3), 383-406.

Hübler, A. (1983) *Understatements and Hedges*, Amsterdam: John Benjamins.

Hunston, S. (1994) "Evaluation and organization in a sample of written academic discourse" in Coulthard, M. ed., *Advances in written text analysis,* London: Routledge, 191-218.

Hunston, S. (2002) *Corpora in applied linguistics*, Cambridge: Cambridge University Press.

Hunston, S. (2011) *Corpus approaches to evaluation: Phraseology and evaluative language*, New York: Routledge.

Hunston, S. (2013) "Systemic functional linguistics, corpus linguistics, and the ideology of science" *Text & Talk,* 33(4-5), 617-640.

Hunston, S. and Sinclair, J. (2000) "A local grammar of evaluation" in Hunston, S. and Thompson, G. eds., *Evaluation in text: Authorial stance and the construction of discourse,* Oxford: Oxford University Press.

Hunston, S., and Thompson, G. eds. (2000) *Evaluation in text,* Oxford: Oxford University Press.

Huspek, M. (1989) "An analysis of 'you know'/'I think' variation in working-class speech" *Journal of Pragmatics*, 13(5), 661-683.

Hyland, K. (1994) "Hedging in academic writing and EAP textbooks" *English for Specific Purposes*, 13, 239-256.

Hyland, K. (1996a) "Talking to the academy: Forms of hedging in science research articles" *Written Communication*, 13, 251-281.

Hyland, K. (1996b) "Writing without conviction? Hedging in science research articles" *Applied Linguistics*, 17(4), 433-454.

Hyland, K. (2000) *Disciplinary Discourses: Social Interaction in Academic Writing*, London: Longman.

Hyland, K. (2016) *Teaching and researching writing,* 3rd ed., New York: Routledge.

Hyland, K. and Guinda, C. eds. (2012) *Stance and voice in written academic genres,* United Kingdom: Palgrave Macmillan.

Hyland, K. and Jiang, F. (2018) "'In this paper we suggest': Changing patterns of disciplinary metadiscourse" *English for Specific Purposes,* 51, 18-30.

Hyland, K. and Milton, J. (1997) "Qualification and certainty in L1 and L2 students' writing" *Journal of Second Language Writing*, 6(2), 183-205.

Hyland, K. and Tse, P. (2005) "Hooking the reader: A corpus study of evaluative that in abstracts" *English for specific purposes*, 24(2), 123-139.

ICE website, (2016) *International Corpus of English* homepage, available: http://www.ucl.ac.uk/english-usage/ice/index.htm.

Ide, S. (1989) "Formal forms and discernment: Two neglected aspects of universals of linguistic politeness" *Multilingua*, 2(3), 223-248.

Janney, R. (2002) "Cotext as context: Vague answers in court" *Language & Communication,* 22(4), 457-475.

Jiang, K. and Hyland, K. (2015) "'The fact that': Stance nouns in disciplinary writing" *Discourse Studies*, 17(5), 529-550.

Jones, C. and Carter, R. (2014) "Teaching spoken discourse markers explicitly: A comparison of III and PPP" *International Journal of English Studies*, 14(1), 37-54.

Jones, C. (2016) "Teaching spoken stance markers: A comparison of receptive and productive practice" *The European Journal of Applied Linguistics and TEFL,* 5, 83-100.

Jucker, A. (2013) "Corpus pragmatics" in Östman, J. and Verschueren, J. eds., *Handbook of pragmatics*, Amsterdam: John Benjamins, 2-17.

Jucker, A., Schneider, K. and Bublitz, W. eds. (2018) *Methods in pragmatics*, Berlin: Mouton de Gruyter.

Kallen, J. (2006) "'*Arrah, like, you know*': The dynamics of discourse marking in ICE-Ireland" Plenary paper, *Sociolinguistics Symposium 16*, Limerick, available: http://hdl.handle.net/2262/50586.

Kallen, J. (2012) "The English language in Ireland: An introduction" *International Journal of Language, Translation and Intercultural Communication,* 1, 25-41.

Kallen, J. and Kirk, J. (2007) "ICE-Ireland: Local variations on global standards" in Beal, J., Corrigan, K. and Moisl, H. eds., *Creating and digitizing language corpora*, London: Palgrave Macmillan, 121-162.

Kallen, J. and Kirk, J. (2008) *ICE-Ireland: A User's Guide: Documentation to accompany the Ireland Component of the International Corpus of English (ICE-Ireland),* Belfast: Cló Ollscoil na Banríona.

Kallen, J. and Kirk, J., (2012) *SPICE-Ireland: A user's guide*, Belfast: Cló Ollscoil na Banríona.

Kasper, G. ed. (1995) *Pragmatics of Chinese as native and target language.* Second Language Teaching & Curriculum Center, University of Hawai'i, Honolulu: University of Hawai'i Press.

Keisanen, T. and Rauniomaa, M. (2012) "The organization of participation and contingency in prebeginnings of request sequences" *Research on Language and Social Interaction*, 45(4), 323-351.

Kendon, A. (1967) "Some functions of gaze-direction in social interaction" *Acta Psychologica*, 26, 22-63.

Kimps, D., Davidse, K. and Cornillie, B. (2014) "A speech function analysis of tag questions in British English spontaneous dialogue" *Journal of Pragmatics,* 66, 64-85.

Kimps, D. (2018) *Tag questions in conversation: A typology of their interactional and stance meanings*, Amsterdam: John Benjamins.

Kirk, J. (2011) "What is Irish Standard English?" *English Today,* 27(2), 32-38.

Kirk, J. and Kallen, J. (2006) "Irish standard English: How Celticised? How standardised?" in Tristram, H. ed., *The Celtic Englishes IV: The interface between English and the Celtic languages,* Potsdam: Potsdamer Universitätsverlag, 88-113.

Kirk, J. and Kallen, J. (2009) "*Just* in Irish Standard English" in Bowen, R., Mobärg, M., and Ohlander, S. eds., *Corpora and discourse and stuff: Papers in honour of Karin Aijmer*, Göteburg: University of Gothenburg, 149-158.

Knott, A. (1993) "Using linguistic phenomena to motivate a set of coherence relations" *Department of Artificial Intelligence*, University of Edinburgh.

Koester, A. (2006) *Investigating workplace discourse*, London: Routledge.

Kroon, C. (1995) *Discourse particles in Latin. A study of nam, enim, autem, vero and at*, Amsterdam Studies in Classical Philogy, 4, Amsterdam: Gieben.

Labov, W. (1972.) *Sociolinguistic patterns*, Philadelphia: University of Pennsylvania Press.

Labov, W. (1984) "Field methods of the project on linguistic change and variation" in Baugh, J. and Sherzer, J. eds., *Language in Use*, Englewood Cliffs: Prentice Hall, 28-53.

Labov, W. and Fanshel, D. (1977) *Therapeutic Discourse: Psychotherapy as Conversation*, New York: Academic Press.

Lakoff, G. (1972) "Hedges: A study in meaning criteria and the logic of fuzzy concepts" *Chicago Linguistic Society Papers*, 8, 183-228.

Lakoff, R. (1975) *Language and woman's place*, New York: Harper & Row.

Lam, W. (2008) "Discourse particles in an intercultural corpus of spoken English", unpublished thesis (Ph.D.), The Hong Kong Polytechnic University.

Lee, D. (1987) "The semantics of 'just'" *Journal of Pragmatics*, 11(3), 377-398.

Leech, G.(1983) *Principles of Pragmatics*, London: Longman.

Leech, G. (1991) "The state-of-the-art in corpus linguistics" in Aijmer, K. and Altenberg, B. eds., *English corpus linguistics: Studies in honour of Jan Svartvik*, London: Longman, 8-29.

Leech, G. (2000) "Grammars of spoken English: New outcomes of corpus-oriented research" *Language Learning*, 50(4), 675-724.

Leech, G. (2004) "Adding linguistic annotation" in Wynne, M. ed., *Developing linguistic corpora: A guide to good practice,* Oxford, Oxbow Books.

Leech, G. (2014) *The pragmatics of politeness*, Oxford: Oxford University Press.

Leech, G., Myers, G. and Thomas, J., eds. (1995) *Spoken English on computer: Transcription, mark-up and application*, London: Longman.

Lenk, U. (1998) "Discourse markers and global coherence in conversations" *Journal of Pragmatics,* 30, 245-257.

Lenk, U. (1998) *Marking discourse coherence: Functions of discourse markers in spoken English*, Tübingen: Gunter Narr Verlag.

Levinson, S. (1983) *Pragmatics*, Cambridge: Cambridge University Press.

Lindemann, S. and Mauranen, A. (2001) "'It's just really messy': The occurrence and function of just in a corpus of academic speech" *English for Specific Purposes* 20, 459-475.

LoCastro, V. (2012) *Pragmatics for Language Educators*, London: Routledge.

Locher, M. and Bousfield, D. (2008) "Introduction: Impoliteness and power in language" in Bousfield, D. and Locher, M. eds., *Impoliteness in language: Studies on its interplay with power in theory and practice*, Berlin: Mouton de Gruyter, 1-13.

Lougheed, W. C. (ed) (1986) *In search of the standard in Canadian English*, Strathy Language Unit, Queen's University, available: http://hdl.handle.net/1974/23743.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. (2017) "The spoken BNC2014: Designing and building a spoken corpus of everyday conversations" *International Journal of Corpus Linguistics*, 22(3), 319-344.

Lucek, S. (2011) "I came up and I seen this haze of smoke, like": How Irish are invariant tags? *Trinity College Dublin Journal of Postgraduate Research*, 10, 95-108.

Lyons, J. (1981) *Language and linguistics,* Cambridge: Cambridge University Press.

Lyons, J. (1993) "Subjecthood and subjectivity" in Yaguello, M. ed., *Subjecthood and subjectivity: Proceedings of the colloquium, the status of the object in linguistic theory,* Paris: Ophrys.

Mac Mathúna, S. (2004) "Remarks on standardization in Irish English, Irish and Welsh" in Tristram, H. ed., *The Celtic Englishes 4*, Potsdam: Potsdam University Press, 114-129.

Macaulay, R. (2002) "Discourse variation" in Chambers, J., Trudgill, P. and Schilling Estes, N. eds., *The handbook of language variation and change*, Malden: Blackwell, 283-306.

Maltz, D. and Borker, R. (1982) "A cultural approach to male- female miscommunication" in Coates, J. ed., *Language and gender: A reader*, Oxford: Blackwell, 417-434.

Markkannen, R. and Schröder, H. (1997) "Hedging: A challenge for pragmatics and discourse analysis" in Markannen, R. and Schröder, H. eds., *Hedging and discourse*, Berlin: Walter de Gruyter.

Marriott, J. (1993) "ESL in mathematics: Why 'easy' is more difficult", unpublished dissertation (MA), Research Centre for English and Applied Linguistics, University of Cambridge.

Martin, J. (2000) "Beyond exchange: Appraisal systems in English" in Hunston, S. and Thmonspon, G. eds., *Evaluation in text,* Oxford: Oxford University Press.

Martin, J. and White, R. (2005) *The language of evaluation: Appraisal in English,* Houndmills: Palgrave Macmillan.

Martin, P. (2003) "A genre analysis of English and Spanish research paper abstracts in experimental social sciences" *English for Specific Purposes,* 22(1), 25-43.

Mauranen, A. (2004) "They're a little bit different… observations on hedges in academic talk" in Aijmer, K. and Stenström, A. eds., *Discourse patterns in spoken and written corpora*, Amsterdam: John Benjamins, 173-197.

Maynard, D. (1989) "Perspective display sequences in conversation" *Western Journal of Speech Communication,* 53, 91-113.

Maynard, D. (1997) "How to tell patients bad news: The strategy of 'forecasting'" *Cleveland Clinic of Journal of Medicine*, 64(4), 181-182.

Maynard, S. (1986) "On back-channel behavior in Japanese and English casual conversation" *Linguistics,* 24, 1079-1108.

Maynard, S. (1997) "Analyzing interactional management in native/non-native English conversation: A case of listener response" *IRAL,* 35, 37-60.

McCarthy, M. (1998) *Spoken language and applied linguistics*, Cambridge: Cambridge University Press.

McCarthy, M. (2002), "Good listenership made plain: British and American non-minimal response tokens in everyday conversation" in Reppen, R., Fitzmaurice, S. and Biber, D. eds., *Using corpora to explore linguistic variation*, Amsterdam: John Benjamins, 49-71.

McCarthy, M. (2010). "Spoken fluency revisited" *English Profile Journal,* 1, Available: http://journals.cambridge.org/action/displayJournal?jid=EPJ.

McCarthy, M. (2015) "'Tis mad, yeah': Turn openers in Irish and British English" in Amador-Moreno, C., McCafferty, K. and Vaughan, E. eds., *Pragmatic markers in Irish English*, Amsterdam: John Benjamins, 156-175.

McCarthy, M. and Carter, R. (2000) "Feeding back: non-minimal response tokens in everyday English conversation" in Heffer, C. and Sauntson, H. eds., *Words in context: A tribute to John Sinclair on his retirement*, Birmingham: ELR Discouse Monography, 18, 263-283.

McCarthy, M., and Handford, M. (2004) "'Invisible to us': A preliminary corpus- based study of spoken business English" *Discourse in the professions: Perspectives from corpus linguistics*, Amsterdam: John Benjamins, 167-201.

McCarthy, M. and O'Keeffe, A. (2012) "Analysing Speech Corpora" in Cobb, T. ed., *The Encyclopedia of Applied Linguistics,* New York: Wiley-Blackwell, 104-112.

McCarthy, M. and O'Keeffe, A. (2014) "Spoken Grammar" in Celce-Murcia, M., Brinton, D., and Snow, M. eds., *Teaching English as a second or foreign language* (4th ed.), Boston: National Geographic/Cengage, 271-287.

McEnery, T. and Gabrielatos, C. (2006) "English corpus linguistics" in Aarts, B. and McMahon, A. eds., *The handbook of English linguistics*, Oxford: Blackwell.

McEnery, T., and Hardie, A. (2012) *Corpus linguistics: Method, theory and practice*, Cambridge: Cambridge University Press.

McEnery, T. and Wilson. A. (2001) *Corpus linguistics*, 2nd ed, Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R. and Tono, Y. (2005) *Corpus-based language studies: An advanced resource book*, in Candlin, C. and Carter, R. eds., *Routledge Applied Linguistic Series,* London: Routledge.

Megyesi, B. (2009) "Corpus usage" *Lecture Notes* Uppsala University, unpublished.

Meyer, C. (2002) *English corpus linguistics: An introduction*, Cambridge: Cambridge University Press.

Migge, B. (2015) "*Now* in the speech of newcomers to Ireland" in Amador-Moreno, C., McCafferty, K., and Vaughan, E. eds. *Pragmatic markers in Irish English*, Amsterdam: John Benjamins, 390-407.

Miller, J. and Weinert, R. (1995) "The function of 'like' in dialogue" *Journal of Pragmatics,* 23, 365-393.

Milroy, L. (1999) "Standard English and language ideology in Britain and the United States" in Bex, T. and Watts, R. eds., *Standard English: The widening debate*, London: Routledge, 173-206.

Mott, H. and Petrie, H. (1995) "Workplace interactions: Women's linguistic behavior" *Journal of Language and Social Psychology,* 14(3), 324-336.

Murphy, B. (2015) "A corpus-based investigation of pragmatic markers and sociolinguistic variation in Irish English" in Amador-Moreno, C., McCafferty, K. and Vaughan, E. eds., *Pragmatic markers in Irish English,* Amsterdam: John Benjamins, 65-88.

Myers, G. (1989) "The pragmatics of politeness in scientific articles" *Applied Linguistics,* 10, 1-35.

Myers, G. (1992) "Textbooks and the sociology of scientific knowledge" *English for Specific Purposes,* 11, 3-17.

Nelson, G. (1996). "The design of the corpus" in Greenbaum, S. ed., *Comparing English worldwide: The international corpus of English,* Oxford: Oxford University Press, 27-35.

Nesselhauf, N. (2005) *Collocations in a learner corpus*, Amsterdam: John Benjamins.

Newman, J. (2008) "Spoken corpora: Rationale and application" *Taiwan Journal of Linguistics*, 6(2), 27-58.

Norrby, C., Wilde, C., Lindström, J. and Nilsson, J. (2012) "Finland Swedish as a non dominant variety of Swedish- extending the scope to pragmatic and interactional aspects" in Muhr, R. ed., *Non-dominant varieties of pluricentric languages. Getting the Picture: In memory of Michael Clyne*, Frankfurt: Peter Lang Publishing, 49-60.

Norris, J. and Ortega, L. (2000) "Effectiveness of L2 instruction: A research synthesis and quantitative Meta-analysis" *Language Learning*, 50(3), 417-528.

Ochs, E., (1979) "Introduction: What child language can contribute to pragmatics" in Ochs, E., and Schieffelen, B. eds., *Developmental pragmatics*, New York: Academic Press, 1-17.

Ochs, E. and Schieffelin, B. (1989) "Language has a heart" *Text*, 9, 7-25.

Ogiermann, E. (2009) "Politeness and indirectness across cultures: A comparison of English, German, Polish and Russian requests" *Journal of Politeness Research,* 5, 189-216.

O'Keeffe, A. (2002) "Exploring indices of national identity in a corpus of radio phone-in data from Irish radio" in Sanchez-Macarro, A. ed., *Windows on the world: Media discourse in English,* Valencia: University of Valencia Press, 91-113.

O'Keeffe, A. (2003) "Strangers on the line: a corpus-based lexico-grammatical analysis of radio phone-in" unpublished thesis (Ph.D.), University of Limerick.

O'Keeffe, A. (2006) *Investigating media discourse*, London: Routledge.

O'Keeffe, A. (2018) "Corpus-based function-to-form approaches" in Jucker, A., Schneider, K., and Bublitz, W. eds., *Methods in pragmatics*, Berlin: Mouton de Gruyter, 587-618.

O'Keeffe, A. (Forthcoming) "Irish English corpus linguistics" in Hickey, R. ed., *The Oxford handbook of Irish English,* Oxford: Oxford University Press.

O'Keeffe, A. and Adolphs, S. (2008) "Response tokens in British and Irish discourse: Corpus, context and variational pragmatics" in Schneider, K. and Barron, A. eds.,

Variational pragmatics: A focus on regional varieties in pluricentric languages, Amsterdam: John Benjamins, 69-98.

O'Keeffe, A. and Breen, M. (2007) "At the hands of the Brothers: A corpus-based lexico grammatical analysis of stance in newspaper reporting of child sexual abuse" in Cotterill, J. ed., *The Language of Sex Crimes,* Basingstoke: Palgrave, 217-236.

O'Keeffe, A., Clancy, B. and Adolphs, S. (2020) *Introducing pragmatics in use,* 2nd volume, Abingdon: Routledge.

O'Keeffe, A. and Farr, F. (2003) "Using language corpora in language teacher education: Pedagogic, linguistic and cultural insights" *TESOL Quarterly,* 37(3), 389-418.

O'Keeffe, A. and Mark, G. (2018) "The grammars of English" in Seargeant, P., Hewings, A. and Pihlaja, S. eds., *The Routledge handbook of English language studies,* London: Routledge, 136-149.

O'Keeffe, A. and McCarthy, M. eds. (2012) *The Routledge handbook of corpus linguistics*, London: Routledge.

O'Keeffe, A., McCarthy, M. and Carter, R. (2007) *From corpus to classroom: Language use and language teaching,* Cambridge: Cambridge University Press.

O'Keeffe, A. and Walsh, S. (2012) "Applying corpus linguistics and conversation analysis in the investigation of small group teaching in higher education" *Corpus Linguistics and Linguistic Theory*, 8(1), 159-181.

Östman, J. (1981) *You know: A discourse functional approach,* Amsterdam: John Benjamins.

Ostman, J. (1983) *You know: A discourse functional approach,* Amsterdam: John Benjamins.

O'Sullivan, J. (2004) "Accommodative phenomena in the speech of Traveller girls: A corpus-based analysis of the convergences to the speech patterns of the settled community" unpublished thesis (MA), University of Limerick.

Palma, T. (2013) "Investigating communicative strategies in novice professional communities of practice: A comparative study of engineering and marketing meetings" unpublished thesis (Ph.D.), University of Limerick.

Palmer, H. (1924) *A grammar of spoken English*, Cambridge: W. Heffer and Sons Ltd.

Palmer, F. (1986) *Mood and modality,* Cambridge: Cambridge University Press.

Pérez-Paredes, P. (2021) *Corpus Linguistics for education: A guide for research.* London: Routledge.

Pichler, H. (2013) *The structure of discourse-pragmatic variation.* Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/silv.13

Plevoets, K., Speelman, D. and Geeraerts, D. (2008) "The distribution of T/V pronouns in Netherlandic and Belgian Dutch" in Schneider, K. and Barron, A. eds., *Variational pragmatics: A focus on regional varieties in pluricentric languages*, Amsterdam: John Benjamins, 181-210.

Precht, K. (2003) "Stance moods in spoken English: Evidentiality and affect in British and American conversation" *Text- Interdisciplinary Journal for the Study of Discourse*, 23(2), 239-257.

Preisler, B. (1986) *Linguistic sex roles in conversation: Social variation in the expression of tentativeness in conversations,* Berlin: Mouton de Gruyter.

Prince, E., Frader, J. and Bosk, C. (1982) "On hedging in physician-physician discourse" in di Pietro, R. ed., *Linguistics and the Professions. Proceedings of the Second Annual Delaware Symposium on Language Studies,* Norwood: Ablex, 83-97.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985) *A comprehensive grammar of the English language*, London: Longman.

Rauniomaa, M and Keisanen, T. (2012) "Two multimodal formats for responding to requests" *Journal of Pragmatics*, 44(6-7), 829-842.

Redeker, G. (1991) "Linguistic markers of discourse structure" *Linguistics*, 29, 1139-1172.

Roger, D., Bull, P. and Smith, S. (1988) "The development of a comprehensive system for classifying interruptions" *Journal of Language and Social Psychology,* 7, 27-34.

Rosch, E. (1973) "Natural Categories" *Cognitive Psychology,* 4, 328-350.

Rosch, E. (1978) "Principles of Categorization" in Rosch, E. and Lloyd, B. eds., *Cognition and Categorization*, New Jersey: Erlbaum Associates, 27-48.

Rounds, P. (1982) *Hedging in academic discourse: Precision and flexibility*, Ann Arbor: The University of Michigan.

Rühlemann, C. (2007) *Conversation in context*, London: Continuum.

Rühlemann, C., and Aijmer, K. (2015) "Corpus pragmatics: Laying the foundations" in Aijmer, K. and Rühlemann, C. eds. *Corpus pragmatics: A handbook*, Cambridge: Cambridge University Press, 1-26.

Ruzaitė, J. (2007) "Vague references to quantities as a face-saving strategy in teacher student interaction" *Lodz Papers in Pragmatics*, 3(1), 157-178.

Sacks, H., Schlegloff, E. and Jefferson G. (1974) "A simplest systematics for the organization of turn-taking for conversations" *Language* 50(4), 696-735.

Salager-Meyer, F. (1994) "Hedges and textual communicative function in medical English written discourse" *English for Specific Purposes,* 13, 149-170.

Sankoff, D. (1988) "Sociolinguistics and syntactic variation" in Newmeyer, F., ed., *Linguistics: The Cambridge survey*, Cambridge: Cambridge University Press, 140-161.

Sankoff, G. (1980) "A quantitative paradigm for the study of communicative competence" in Sankoff, G. ed., *The social life of language*, Philadelphia: University of Pennsylvania Press, 47-79.

Schegloff, E. (1982) "Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences" in Tannen, D. ed., *Analyzing discourse: Text and talk*, Georgetown: Georgetown University Press, 71-93.

Schiffrin, D. (1987) *Discourse markers*, Cambridge: Cambridge University Press.

Schiffrin, D. (2001) "Discourse markers: Language, meaning, and context" in Schiffrin, D., Tannen, D. and Hamilton, H. eds., *The handbook of discourse analysis,* Oxford: Blackwell, 54-75.

Schleppegrell, M. (1992) "Subordination and linguistic complexity" *Discourse Processes,* 15, 117-131.

Schmitt, N. ed. (2004) *Formulaic sequences: Acquisition, processing and use,* Amsterdam: John Benjamins.

Schneider, G. (2013) "Using automatically parsed corpora to discover lexico- grammatical features of English varieties" in Doa, F., ed., *Penser le Lexique-Grammaire, perspectives actuelles*, (Papers from the 3rd International conference on lexis and grammar, Nikosia, Cyprus, 5-8 October, 2011), Paris: Éditions Honoré Champion.

Schneider, K. (1988) *Small talk: Analyzing phatic discourse*, Marburg: Hitzeroth.

Schneider, K. (2001) "Using dictionaries: Studies of dictionary use by language learners and translators" *System*, 29(1).

Schneider, K. (2005) "*No problem, you're welcome, anytime*: Responding to thanks in Ireland, England and the USA" in Barron A. and Schneider, K., eds., *The pragmatics of Irish English*, Berlin: Mouton de Gruyter, 101-140.

Schneider, K. (2010) "Variational pragmatics" in Fried, M., Östman, J. and Verschueren, J. eds., *Variation and Change: Pragmatic Perspectives*, Amsterdam: John Benjamins, 239-267.

Schneider, K. (2014) "Comparability and sameness in variational pragmatics" in Silvia Mergenthal, S. and Nischik, R. eds., *Anglistentag 2013 Konstanz: Proceedings*, Trier: Wissenschaftlicher Verlag Trier, 361-372.

Schneider, K. (2021) "Notes on variational metapragmatics" *Journal of Pragmatics,* 179, 12-18.

Schneider, K (2021) "Variational pragmatics" in Haugh, M., Kádár, D. and Terkourafi, M., eds., *The Cambridge handbook of sociopragmatics,* Cambridge: Cambridge University Press, 663-686.

Schneider, K. and Barron, A., eds. (2008a) *Variational pragmatics: A focus on regional varieties in pluricentric languages*, Amsterdam: John Benjamins.

Schneider, K. and Barron, A. (2008b) "Where pragmatics and dialectology meet: Introducing variational pragmatics" in Schneider, K. and Barron, A., eds. *Variational pragmatics: A focus on regional varieties in pluricentric languages*, Amsterdam: John Benjamins, 1-32.

Schourup, L. (1985) *Common discourse particles in English conversation*, New York: Garland.

Schourup, L. (1999) "Discourse markers: A tutorial overview" *Lingua*, 107, 227-265.

Schröder, H. and Zimmer, D. (1997) "Hedging research in pragmatics: A bibliographical research guide to hedging" in Markkanen, R. and H. Schröder, H., eds., *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts,* Berlin: Walter de Gruyter, 249-271.

Schweinberger, M. (2012) "The discourse marker *like* in Irish English" in Migge, B. and Ní Chiosáin, M. eds. *New Perspectives on Irish English*, Amsterdam: John Benjamins, 179-202.

Schweinberger, M. (2015) "A comparative study of the pragmatic marker like in Irish English and in south-eastern varieties of British English" in Amador-Moreno C., McCafferty, K. and Vaughan, E., eds., *Pragmatic markers in Irish English,* Amsterdam: John Benjamins Publishing, 114-134.

Schweinberger, M. (2020) "Speech-unit final *like* in Irish English" *English World Wide,* 3(4), 327-361.

Schweinberger, M., Siemund, P. and Maier, G. (2009) "Towards a more fine-grained analysis of the areal distributions of non-standard features of English" in Pentitilä, E. and Paulasto, H. eds. *Language Contacts Meet English Dialects: Studies in Honour of Markku Filppula*, Newcastle-upon-Tyne: Cambridge Scholars Publishing, 19-45.

Spada, N. and Tomita, Y. (2010) "Interactions between type of instruction and type of language feature: A meta-analysis" *Language Learning*, 60(2), 263-308.

Scott, M. (2021) *WordSmith Tools,* version 8, Stroud: Lexical analysis software.

Sinclair, J., ed. (2003) *Collins COBUILD advanced dictionary*, 4th ed., London: Collins.

Skelton, J. (1988) "Comments in academic articles" in Grumwell, P. ed., *Applied Linguistics in Society*, 20th Nottingham, England, September 1987, *British Studies in Applied Linguistics,* 3, 98-108.

Spencer-Oatey, H. (2000) "Rapport management: A framework for analysis" in Spencer-Oatey, H. ed., *Culturally speaking: Managing rapport through talk across cultures*, London: Continuum, 11-46.

Staley, L. (2018) *Socioeconomic pragmatic variation: Speech acts and address forms in context*, Amsterdam: John Benjamins.

Stefanowitsh, A. (2018) *Corpus linguistics: A guide to the methodology*, *Textbooks in Language Sciences,* 7, Language Science Press.

Strathy Language Unit (2011), Queen's University, available: https://www.queensu.ca/strathy/.

Swales, J. M. and Burke, A. (2003) "'It's really fascinating work': Differences in evaluative adjectives across academic registers" in Leistyna, P. and Meyer, C., eds., *Corpus analysis: Language structure and language Use,* New York: Rodopi, 1-18.

Sweet, H. (1899) *The practical study of languages. A Guide for teachers and learners,* London: J. M. Dent and Co.

Sweet, H. (1900) *A new English grammar, logical and historical,* Oxford: Oxford University Press.

Szabo, Z. ed. (2005) *Semantics versus pragmatics,* Oxford: Clarendon Press.

Tagliamonte, S. (2005) "*So who? Like how? Just what?* Discourse markers in the conversations of young Canadians" *Journal of Pragmatics*, 37, 1896–1915.

Takimoto, M. (2015) "A corpus-based analysis of hedges and boosters in English academic articles" *Indonesian Journal of Applied Linguistics,* 5(1), 95-105.

Tao, H. and McCarthy, M. (2001) "Understanding non-restrictive *which*-clauses in spoken English, which is not an easy thing" *Language Sciences,* 23, 651-677.

The Irishmen's Prayers (1689) Broadsheet, Pepys collection, Magdalene College Library, Cambridge, Reprinted in Bliss, A. (1979) *Spoken English in Ireland 1600-1740,* Dublin: Cadenus Press, 130-132.

Thomas, J. (1995) *Meaning in interaction: An introduction to pragmatics*, London: Longman.

Thompson, P. and Hunston, S. (2019) *Interdisciplinary research discourse: Corpus investigations into environment journals*, London: Routledge.

Timmis, I. (2009) "'Tails' of linguistic survival" *Applied Linguistics*, available: http://applij.oxfordjournals.org/cgi/search?fulltext=timmis+tails&x=14&y=4. Accessed 23.2.2021.

Timmis, I. (2012) "Spoken language research and ELT: Where are we now?" *ELT Journal* 66(4), 514-22.

Tognini-Bonelli, E. (2001) *Corpus linguistics at work,* Amsterdam: John Benjamins.

Tognini-Bonelli, E. (2010) "Theoretical overview of the evolution of corpus linguistics" in O'Keeffe, A. and McCarthy, M. eds., *The Routledge handbook of corpus linguistics*, London: Routledge, 14-28.

Tottie, G. (1991) "Conversational style in British and American English: The case of backchannels," in Aijmer, K. and Altenberg B., eds., *English corpus linguistics*, London: Longman, 254-271.

Tottie, G. and Hoffmann, S. (2006) "Tag questions in British and American English" *Journal of English Linguistics*, 34(4), 283-311.

Tottie, G. and Hoffmann, S. (2009) "Tag questions in English: The first century" *Journal of English Linguistics*, 37(2), 103-161.

Tsui, A. (1992) "A functional description of questions" in Coulthard, M., ed., *Advances in spoken discourse analysis*, London: Routledge, 89-110.

UAL Dataverse (2015) *Voices of the International Corpus of English (VOICE)- Canada,* University of Alberta.

Urgelles-Coll, M. (2010) *The syntax and semantics of discourse markers*, London: Bloomsbury.

Uwajeh, M. (1996) "Is 'May I ask you a question?' a question?" *International Pragmatics Association (IPrA)*, 6(1), 89-109.

Vaughan, E. (2009) "'Just say something and we can all argue then': Community and identity in the workplace talk of English language teachers" unpublished thesis (Ph.D.), University of Limerick.

Vaughan, E. and Clancy, B. (2011) "The pragmatics of Irish English" *English Today*, 27(2), 47-52.

Vaughan, E. and Clancy, B. (2016) "Sociolinguistic information and Irish English corpora" in Hickey, R., ed., *Sociolinguistics in Ireland*, Basingstoke: Palgrave Macmillan, 365-388.

Vaughan, E. and O'Keeffe, A. (2015) "Corpus analysis" in Tracy, K., Ilie, C. and Sandel, T., eds., *The international encyclopedia of language and social interaction*, Denver: John Wiley & Sons, 252-268.

Villegas, C. (2019) "Discourse markers and pragmatic markers in spoken and written everyday life" Arizona: Arizona State University, available: DOI: 10.13140/RG.2.2.33992.57602.

Walsh, S. (2002) "Construction or obstruction: Teacher talk and learner involvement in the EFL classroom" *Language Teaching Research*, 6, 3-23.

Walsh, S. (2013) "Corpus linguistics and conversation analysis at the interface: Theoretical perspectives, practical outcomes" in Romero-Trillo, J. ed., *Yearbook of Corpus Linguistics and Pragmatics,* 1, Dordrecht: Springer, 37-51.

Walsh, S., Morton, T., and O'Keeffe, A. (2011) "Analysing university spoken interaction: A CL/CA approach" *International Journal of Corpus Linguistics*, 16(3), 325-345.

Walters, K. (1988) "Dialectology" in Newmeyer, F. ed. *Linguistics: The Cambridge Survey*, Cambridge: Cambridge University Press, 119-139.

Wang, S. and Tatiana, K. (2016) "Corpus research on hedges in applied linguistics and EFL journal papers" *International Journal of Education*, 9(1), 44-49.

Watts, R. (2003) *Politeness*, Cambridge: Cambridge University Press.

Weber, E. (1993) *Varieties of questions in English conversation,* Amsterdam: John Benjamins.

Weisser, M. (2015) "Speech act annotation" in Aijmer, K. and Rühlemann, C., eds., *Corpus pragmatics*, Cambridge: Cambridge University Press, 84-112.

Werkhofer, K. (1992) "Traditional and modern views: The social constitution and the power of politeness" in Watts, R., Ide, S. and Ehlich, K. eds., *Politeness in language: Studies in its history, theory and practice*, Berlin: Mouton de Gruyter, 155-159.

White, L. (2003.) *Second language acquisition and universal grammar, Cambridge: Cambridge University Press.*

Wijayanto, A., Malikatul, L., Aryati P. and Susiati S. (2013) "Politeness in interlanguage pragmatics of complaints by Indonesian learners of English" *English Language Teaching,* 6(10), 188-201.

Wooffitt, R. (2005). *Conversation analysis and discourse analysis: A comparative and critical introduction*, California: Sage Publications.

Wynne, M. and Berglund, Y. (2012) "Types of corpora" *Weblearn.ox.ac.uk,* available: http://tinyurl.com/669o4zt.

Yilmaz, E. (2004) "A pragmatic analysis of Turkish discourse particles: Yani, işte and şey" unpublished thesis (Ph.D), Middle East Technical University.

Youmans, M. (2001) "Cross-cultural differences in polite epistemic modal use in American English" *Journal of Multilingual and Multicultural Development*, 22(1), 57-73.

Yngve, V. (1970) "On getting a word in edgewise" in *Papers from the 6th regional meeting, Chicago Linguistic Society*, Chicago: Chicago Linguistic Society.

Zhang, G. and Sabet, P. (2016) "'I don't think' versus 'I think + not'" *Text & Talk,* 37(3)*,* Berlin: Mouton de Gruyter.

Zimmerman, D. and West, C. (1975) "Sex roles, interruptions and silences in conversation" *Language and sex: Difference and dominance,* Stanford: Stanford University Press, 105-129.

APPENDICES

Appendix A

Top-down framework for the analysis of the Forms and Functions of Spoken Grammar

Features of spoken grammar	Types	Functions	Forms
Pragmatic markers "are a class of items which operate outside the structural limits of the clause and which encode speakers' intentions and interpersonal meanings" (Carter & McCarthy 2006, p.208; also see section 2.3)	**Discourse markers:** are "words and phrases which function to link segments of the discourse to one another in ways which reflected choices of monitoring, organisation, and management exercised by the speaker." (Carter & McCarthy 2006, p.208; also see section 2.3.1 and Biber *et al* 1999, p.1086; Schiffrin 1987; Fraser 1990a, 1996, 1999; Brinton 1996)	Opening up and closing down the discourse. This is done by managing the discourse "in terms of launching and concluding topics, opening, concluding or temporarily closing a whole conversation, re-opening previously closed or interrupted conversations". (Carter & McCarthy 2006, p.214)	*Right, so, now, all right, right then, now, good, well, anyway, fine, lovely, okay then, okay, great*
		Sequencing which "indicates explicitly the order in which things occur or how different segments of a discourse are being organised" (Carter & McCarthy 2006, p.216)	*And, and then, finally, first (of all), firstly* (more formal than first), *for a start, going back to, in general, in the end, in the first place, last of all, lastly, next, on top of that, second, secondly*, (more formal than second), *so, there again, third(ly), to sum up, what's more*

359

		Marking boundaries and linking segments of the topic. This is done by "indicating the beginning or end of a topic or a transition from one topic or bit of business to another" (Carter & McCarthy 2006, p.218)	*Okay, so, yeah, and, right*
		"focus the attention of the listener on what the speaker feels is important" (Carter & McCarthy 2006, p.218). This is done by diverting, shifting, resuming the topic of the discourse.	*Hey, listen, look, oh, well, anyway, oh yeah, so*
		Monitoring and managing the ongoing discourse through reformulations and alternative expressions indicating that "the speaker has not selected the most appropriate way of expressing things and is adding to or refining what they say with a more apt word or phrase (Carter & McCarthy 2006, p.220).	*As I was saying, as it were, I mean, if you like, in a manner of speaking, in other words, not to say, or rather, so to speak, strictly speaking, that's to say, to put it another way, to put it bluntly/mildly, well*
		Monitoring and managing the ongoing discourse through "monitoring the state of shared knowledge in the conversation"	*You know, you see, see*

		(Carter & McCarthy 2006, p.221).	
	<u>Response tokens</u> can be defined generally as items in discourse such as *yeah*, *mm*, a head nod etc., which demonstrate engaged listenership without changing the speaker turn. (O'Keeffe & Adolphs 2008; also see section 2.3.2 and Biber *et al* 1999, p.1089; see also Fries 1952; Kendon 1967; Yngve 1970; Maynard, D. 1989, 1997; Maynard, S. 1986; 1997; Tottie 1991; Drummond & Hopper 1993a, 1993b; McCarthy 2002; Gardner 2001; O'Keeffe & Adolphs 2008).	Replying to a request for a service or favour (Carter & McCarthy 2006, p.189; Leech 2014, p.176; Goldschmidt 1998; Keisanen & Rauniomaa 2012; Rauniomaa & Keisanen 2012; Flöck 2016; Aijmer 1996).	*Certainly*, *definitely*, *no problem, sure, you can, course, go for it, fine, okay, that's fine, I'd be glad to, right, all right, yes, yeah*
		Making arrangements and reaching decisions. (Carter & McCarthy 2006, p.189).	*Fine*
		Offering a positive feedback to the speaker and often marking the boundaries of topics where speakers express their satisfaction. (Carter	*excellent, fine, great, good, lovely, right, perfect*

		& McCarthy 2006, p.189).	
		Continuers which hand the floor back to the immediately prior speaker (Gardner 2001).	*Mm, hm, uh, huh, I see*
		Acknowledgments which claim agreement or understanding of the prior turn (Gardner 2001).	*Mm, yeah*
		Change-of-activity tokens which mark a transition to a new activity or a new topic in the talk (Gardner 2001).	*Okay* and *alright*
		Assessments which evaluate the talk of the prior speakers (Gardner 2001).	*Great, how interesting, what a load of rubbish*
		For clarification or other types of repair which seek to clarify mishearing or misunderstandings. This is usually done by brief questions (usually known as follow-up questions) (Gardner 2001).	*Who, Which book do you mean*, or (the very generalised) *Huh?*
		Collaborative completions whereby one speaker finishes a prior speaker's utterance (Gardner 2001).	A: so, he's moved into…. B: commercial interests
		Newsmarkers, and newsmarker-like objects marking the prior speaker's turn as newsworthy in some way (Gardner 2001).	*Really*, the change-of-state token *Oh*, the "idea-connector" *Right*

		Indicating further emphasis in response. This is done through premodification by intensifying adverbs and negation by adding *not*. (Carter & McCarthy 2006, p.190).	*Jolly good, most definitely, absolutely not, definitely not*
		Other functions of response tokens are strongly associated with particular contexts which "refer to a whole preceding utterance rather than their word-class identity as adjectives or adverbs." They often occur in pairs for more emphasis (Carter & McCarthy 2006, p.189).	*Indeed, yeah exactly, possibly, precisely, yeah definitely, absolutely, brilliant, fabulous, cheers, bye, that's excellent, thank you very much, thank you so much, thanks, wonderful, is that so?, by all means, fair enough, not at all, true enough, of course, what a pity!* (Carter & McCarthy 2006, pp.188-192, p.221-222).
	Stance markers are "expressions that mark the speaker's stance or attitude towards the message" (Carter & McCarthy 2006, p.222; also see section 2.3.3 and Biber *et al* 1999, p.965; Hunston & Thompson 2000; Hunston & Sinclair 2000; Hunston 1994; Martin, J. 2000, Martin, P. 2003;	To signal stances, attitudes, or points of view towards some of the segments in the discourse. (Carter & McCarthy 2006, p.222).	*Actually, admittedly, amazingly, basically, certainly, clearly, confidentially, doubtless, essentially, frankly, to be frank, fortunately, honestly, to be honest, hopefully, ideally, if you ask me, I'm afraid, I must admit, I must say, I think, in fact, indeed, literally, naturally, no doubt, obviously, of course, predictably, putting (or, to put) it*

	Martin, J. & White 2005; Biber & Finnegan 1988, 1989; Ochs & Schieffelin 1989; Besnier 1990; Chafe 1986; Chafe & Nichols 1986; Biber *et al* 2002; Jones 2016).		*mildly/bluntly, (quite) rightly, really, sadly, seriously, (I'm) sorry, strictly speaking, surprisingly, thankfully, to tell you the truth, understandably, undoubtedly, unfortunately*
	Hedges are a range of expressions and markers used in everyday spoken language in order for speakers to downtone the assertiveness of a segment of discourse. (Carter & McCarthy 2006, p.223; also see section 2.3.4; also see Biber *et al* 1999, p.557; Hyland 1996a, 1996b; Clemen 1997; Schröder & Zimmer 1997; Crompton 1997, 1998; Lindemann & Maurenen 2001; Maurenen 2004)	To enable speakers to be less assertive and blunt in formulating their message.	*Apparently, arguably, by any chance, I think, just (about), kind of, like* (as a mitigator), *maybe, perhaps, presumably, probably, roughly, sort of, surely, now* (hedging *now* vs presentative *now* (Migge 2015; also see Clancy & Vaughan 2012)) (see Kallen 2006; Schweinberger *et al* 2009; Lucek 2011; Diskin 2017; Schweinberger 2020)
	Interjections "normally refer to exclamative utterances consisting of single words that do not easily fit into the major word classes	They express positive or negative emotional reactions to any segment of discourse.	*bother, crikey, damn, god, goodness (me), gosh, (good) heavens, hooray, jeez, ooh, oh no, oops, ouch, ow, ugh, tut-tut, whoops, wow, yippee, yuk*

	(noun, verb, adjective, adverb)" (Carter & McCarthy 2006, p.224; also see section 2.3.5 and Biber *et al* 1999, p.1083).		
Deixis "refers to the way speakers orient themselves and their listeners in terms of person, time, and space in relation to the immediate situation of speaking" (Carter & McCarthy 2006, p.178; also see Biber *et al* 1999, p.1041).	**Personal deixis.**	Personal reference; some of them can be exclusive and inclusive. (Carter & McCarthy 2006, pp.178-179)	*I, me, he, him, she, her, they, them, we, us*
	Temporal deixis.	Time reference (Carter & McCarthy 2006, pp.178-179).	*Tomorrow, tomorrow afternoon, around, now, today, then, soon, recently*
	Spatial deixis.	Space reference (Carter & McCarthy 2006, pp.178-179)	*Here* and *there*
		Referring to physical closeness and distance (Carter & McCarthy 2006, pp.178-179)	*This, that, these, those*
		Referring to psychological closeness and distance (Carter & McCarthy 2006, pp.178-179)	*This* and *that*
		In narratives, it functions like the definite article *the* to refer to things that are well known to listeners (Carter &	*That*

		McCarthy 2006, p.181).	
		It carries and indicates a feeling of dismissal or rejection of something as problematic (Carter & McCarthy 2006, p.180).	*That*
Questions and Tags	**Question tags (canonical TQs)** which can be combined with rising and falling intonation to generate several types of meaning based on different intonations (falling tones or falling tone plus rising tone) (Carter & McCarthy 2006, p.197, 212; also see Biber *et al* 1999, p. 1079; and Allerton 2009; Axelsson 2011, pp.33-35; Barron *et al* 2015) **Fixed tags (invariant TQs)**	Creating questions used by the speaker in the main clause which gives an expectation of either a *yes* or *no* answer or the potential of both a *yes* or *no* answer. Requests can be expressed with tag questions as interrogatives (Carter & McCarthy 2006, p.198). To check and make sure that something has been understood or to confirm that an action has been agreed upon (Carter & McCarthy 2006, p.198) The general pragmatic functions of QTs are classified under information oriented TQs (which are mainly about the exchange of information between speaker and	A tag after a declarative clause (Carter & McCarthy 2006, p.725; for more explanation and examples, see section 2.4.4). Formed in the pattern of negative clause + affirmative tag with the fall and rise intonation pattern (Carter & McCarthy 2006, p.198). *(all) right, okay, yeah, eh, don't you think, huh, (do) you know, see, I think, you know, ah, is it?, isn't it? no, you see* (cf. Columbus 2010) (Carter & McCarthy 2006, p.198; Barron *et al* 2015)

		addressee, covering such pragmatic functions as: establishing common ground, topic-initiating, surprised reactions, stating a fact or opinion (incl. self-monitors), acknowledging responses, and challenging) and action-oriented TQs (which are used to "give or demand goods and services and include requests, offers and suggestions") (Barron *et al* 2015, p.6; Axelsson 2011; Kimps *et al* 2014, pp.81-82).	
	Echo questions	To request more clarification about the noun phrases or parts of them which may not have been heard by the listener correctly (Carter & McCarthy 2006, p.199; also see Biber *et al* 1999, p.1101).	They typically have declarative word order and include a wh- word such as: *where, which, what, the what, a what, who, the what stuff?* (Carter & McCarthy 2006, p.199, 726).
	Follow-up questions contain different types, such as reduced questions with wh- words, that occur either with a wh- word by itself, or with a substitute word, or with stranded prepositions. (Carter &	as a signal of engagement and attention shown by the listener to keep the conversation going by inviting further responses or to expand the discourse by requesting further specification (Carter & McCarthy 2006, pp.199-201).	*What about?, where to?, who?, which one?, did you?, does she?, oh they haven't, have they* (with falling and rising intonation), *oh he does, does he* (with falling and rising intonation), *how come?, like what?* (Carter &

	McCarthy 2006, pp.199-201; for more explanation and examples, see section 2.4.4).		McCarthy 2006, pp.199-201).
	Two-step questions and responses involve a two-step process occurring in a way that the first question acts like a preface for the upcoming question (Carter & McCarthy 2006, p.201)	It occurs for the speaker to avoid being rude, too direct, or too general (Carter & McCarthy 2006, p.201)	**"A: Are you going to the match tonight?** B: Yeah, I am. **A: Do you mind if I tag along?** B: Sure. We're leaving around seven. **A: Are you in this Sunday afternoon?** B: I expect so. I think we might be going out later. **A: Okay, do you mind if I pop round to pick up the drill?** B: Of course not."** (Carter & McCarthy 2006, p.201)
	Another type of two-step question is called the **pre-question** in which the speaker asks a question in order to have permission (for the next question).	This is done in the discourse to show more respect. Politeness or formality between the speaker and the listener may trigger this type of question.	"A: **I wondered if I might ask you something**? B: Sure. A: **Would you be able to write a reference for me**?" (Carter and McCarthy 2006, p.201)
	Preface questions which occur a lot in the spoken language as a four-word cluster.	These questions have the discourse function of showing what the speaker considers newsworthy or important information to the listener (Carter &	*Do you know what?* (Carter & McCarthy 2006, p.202)

		McCarthy 2006, p.202).	
Ellipsis is often defined as "the absence of elements normally required by the grammar (e.g. a subject before a tensed verb form)" (Carter & McCarthy 2006, p.181; also see Biber *et al* 1999, p.156, 1103-1107).	**Situational ellipsis**	They have various functions in the discourse due to the shared contexts and the relationship between the listener and the speaker (Carter & McCarthy 2006, p.181; also see McCarten 2010, pp.423-424 and McCarthy & O'Keeffe 2014, p.13).	"A: *Don't know what's gone wrong here.* B: *Oh. Need any help?* (situational; understood: **I** don't know… **Do you** need…)"
	Textual ellipsis	They have various functions in the discourse due to the shared contexts and the relationship between the listener and the speaker (Carter & McCarthy 2006, p.181; also see McCarten 2010, pp.423-424 and McCarthy & O'Keeffe 2014, p.13).	"*He applied and got the job.* (textual; understood from previous clause: … and **he** got the job.)"
	Structural ellipsis	They have various functions in the discourse due to the shared contexts and the relationship between the listener and the speaker (Carter & McCarthy 2006, p.181; also see McCarten 2010, pp.423-424 and McCarthy & O'Keeffe 2014, p.13).	"*The car he was driving was stolen*. (structural; optional use of *that*: The car **that** he was driving…)"
Headers commonly refer to	A noun phrase followed by one	To put focus and emphasis on the	"I like David but **Pat** I find rather odd.

the subject, object, object complement or prepositional complement which can be placed within the S-V-X clause structure. In other words, what conventionally comes in writing at or towards the end of the clause (e.g. object, complement, adverbial, question-word clause) will be placed in front of it, and this is called "fronting" (Carter & McCarthy 2006, pp.192-194).	or more pronouns which refer to the noun phrase. This is the most typical type of headers (Carter & McCarthy 2006, p.193).	object, complement etc. that has been fronted.	(fronting of the object Pat)" "**Owen, he's** my favourite nephew." "**Joe**, I've never seen **him** at a single football match this season." (Carter & McCarthy 2006, pp.192,194).
Tails are "items that are placed outside the S-V-X clause structure, but they occur after the clause." (Carter & McCarthy 2006, p.194; also see Biber *et al* 1999, p.1103).	Noun phrases, prepositional phrases or clauses (Carter & McCarthy 2006, p.195).	Providing focus, emphasis, and attention (Carter & McCarthy 2006, p.196).	"**He's** amazingly clever, **that dog of theirs.**" (Carter & McCarthy 2006, p.194).
Social Routines	Greetings and leave-taking routines (Carter & McCarthy 2006, p.227; also see Biber *et al* 1999, p.1085).	To mark different levels of formality in the discourse, or between the speaker and listener. (Carter & McCarthy 2006, p.227).	*Good morning, good evening, hello, morning, evening, hi, hi there, hiya, goodnight, goodbye, bye, bye bye, see you later, see you* (often pronounced *see ya*), *cheers*
Swearing and taboo expressions	Taboo naming expressions and taboo intensifiers.	Swearing often takes the form of interjections with the function of expressing "a variety	*Damn!, Christ!, shit!, oh fuck it!, what the hell, what the fuck, you bitch, bloody hell*

		of strong feeling, in particular, annoyance, frustration and anger" (Carter & McCarthy 2006, p.226; also see Biber *et al* 1999, p.1094).	
Vocatives	Names (including abbreviated names) and titles, terms of kinship and endearment, general plural vocatives, impersonal vocatives, honorifics.	Used for calling, summons, turn management, ritual and sociable contexts, softening and lessening threats to dignity, softening an utterance, topic management, joking and banter. (Carter & McCarthy 2006, pp.228-235; also see Biber *et al* 1999, p.1111).	*Anne, Mr. Roberts, doctor, Pete (Peter), professor, sweetheart, sis, love, dear, folks, children, everybody, someone, sir, madam, mate, minister, waiter*

Appendix B

The wordlist of the top 100 most frequent words of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	the	15764
2	and	8954
3	I	7659
4	to	7569
5	of	7388
6	a	6119
7	you	6036
8	it	6020
9	that	5995
10	in	5777
11	's (is)	5134
12	was	3245
13	is	2973
14	uh	2735
15	yeah	2625
16	't (not)	2615
17	we	2471
18	he	2404
19	on	2286
20	they	2253
21	for	2212
22	there	2142
23	be	1952
24	but	1909
25	have	1878
26	this	1834
27	so	1830
28	know	1754
29	what	1686
30	like	1672
31	uhm	1617
32	with	1587
33	as	1584
34	at	1501
35	one	1428
36	she	1414
37	no	1363
38	are	1308
39	do	1281
40	well	1251

41	now	1234
42	all	1224
43	not	1198
44	or	1187
45	're (are)	1172
46	oh	1108
47	just	1095
48	if	1055
49	from	1021
50	think	992
51	can	956
52	about	953
53	up	922
54	were	913
55	very	912
56	out	896
57	going	892
58	would	890
59	by	888
60	Don (don't)	832
61	sylls	832
62	then	827
63	laughter	821
64	had	782
65	've (have)	771
66	an	762
67	his	738
68	'd (would)	736
69	them	735
70	get	734
71	which	734
72	has	729
73	who	727
74	'm (am)	723
75	two	689
76	been	680
77	when	673
78	'll (will)	668
79	me	666
80	will	666
81	people	652
82	go	650
83	your	650
84	said	639
85	say	638

86	time	625
87	did	622
88	my	621
89	really	614
90	right	612
91	see	563
92	some	544
93	because	539
94	her	520
95	okay	510
96	him	502
97	here	493
98	mean	493
99	more	481
100	down	472

Appendix C

The wordlist of the top 100 most frequent words of Spoken ICE-Canada

#	ICE-Canada	Freq
1	the	27644
2	and	18871
3	I	16760
4	to	16427
5	you	14471
6	that	14125
7	it	13915
8	a	13821
9	's	13230
10	of	13051
11	in	9742
12	uh	7058
13	is	6604
14	they	5837
15	't	5661
16	was	5141
17	know	4874
18	we	4834
19	so	4706
20	this	4604
21	he	4497
22	have	4481
23	for	4479
24	but	4212
25	on	4094
26	there	3886
27	like	3725
28	what	3661
29	ya	3590
30	be	3485
31	with	3452
32	're	3291
33	or	3089
34	at	3062
35	not	3038
36	uhm	3010
37	laugh	2935
38	one	2885
39	well	2879

40	she	2845
41	just	2815
42	do	2809
43	are	2765
44	no	2624
45	as	2533
46	if	2526
47	don	2367
48	about	2357
49	going	2212
50	all	2197
51	can	2185
52	mm	2149
53	okay	2066
54	think	2004
55	hmm	1951
56	two	1942
57	'm	1933
58	out	1929
59	right	1889
60	oh	1857
61	get	1774
62	up	1771
63	from	1705
64	then	1704
65	my	1663
66	really	1617
67	when	1615
68	had	1587
69	yeah	1584
70	would	1566
71	word	1539
72	now	1527
73	very	1508
74	your	1507
75	go	1450
76	them	1450
77	because	1432
78	by	1414
79	people	1411
80	were	1353
81	an	1350
82	'll	1339
83	words	1337
84	has	1329

85	time	1300
86	me	1295
87	some	1291
88	've	1269
89	here	1258
90	who	1231
91	her	1215
92	three	1211
93	his	1207
94	will	1192
95	more	1178
96	want	1159
97	did	1156
98	got	1139
99	how	1134
100	mean	1117

Appendix D

The two-word cluster list of the top 100 most frequent clusters of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	of the	1761
2	It's	1659
3	in the	1465
4	you know	1036
5	That's	848
6	Don't	832
7	to the	760
8	I'm	699
9	and I	683
10	on the	665
11	I think	663
12	going to	586
13	it was	586
14	to be	565
15	and the	541
16	at the	483
17	I don't	479
18	I was	479
19	's a	469
20	There's	459
21	yeah yeah	449
22	kind of	443
23	They're	437
24	for the	429
25	He's	429
26	do you	408
27	and then	393
28	You're	388
29	I mean	384
30	if you	374
31	that the	347
32	't know	322
33	in a	320
34	Didn't	319

35	have to	313
36	you can	310
37	with the	309
38	he was	303
39	and it	297
40	to do	288
41	We're	281
42	I've	280
43	of a	267
44	it is	262
45	I'd	260
46	's the	260
47	the the	257
48	yeah I	257
49	and he	255
50	and that	255
51	from the	255
52	She's	253
53	so I	251
54	and you	250
55	but I	248
56	and uh	247
57	as well	245
58	I I	242
59	all the	239
60	is the	234
61	would be	231
62	want to	230
63	was a	228
64	is a	226
65	I know	225
66	is that	222
67	you have	222
68	like that	220
69	there was	220
70	that I	219
71	and they	218
72	one of	218
73	several sylls	217

74	by the	216
75	's not	215
76	Can't	209
77	Wouldn't	209
78	to get	208
79	I'll	205
80	have a	204
81	that you	204
82	is it	203
83	that was	203
84	and we	202
85	she was	201
86	they were	191
87	for a	190
88	this is	189
89	to go	187
90	Wasn't	187
91	and she	186
92	the first	186
93	the other	182
94	yeah and	182
95	are you	181
96	we have	180
97	the same	178
98	what I	178
99	that we	177
100	did you	176

Appendix E

The three-word cluster list of the top 100 most frequent clusters of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	I don't	479
2	Don't know	276
3	It's a	209
4	It's not	143
5	and it's	127
6	one of the	127
7	yeah yeah yeah	119
8	There's a	117
9	I think it	116
10	going to be	111
11	I didn't	111
12	a lot of	109
13	's going to	103
14	That's the	103
15	do you know	99
16	you don't	98
17	I think that	97
18	're going to	97
19	I'm not	95
20	you know the	95
21	don't think	89
22	fianna fa il	89
23	I can't	89
24	I'd say	85
25	I'm going	85
26	I haven't	84
27	think it's	84
28	and I think	80
29	at the moment	76
30	'm going to	76
31	the end of	74
32	but it's	72
33	you know what	72
34	it was a	71

35	That's what	68
36	I mean I	67
37	you know I	67
38	what do you	66
39	and I was	65
40	I want to	65
41	Isn't it	65
42	and I'm	64
43	you have to	64
44	It's just	63
45	That's a	63
46	there was a	63
47	and that's	62
48	do you think	62
49	it would be	61
50	We're going	61
51	and it was	60
52	It's the	59
53	a little bit	58
54	you know and	58
55	that kind of	57
56	you can't	57
57	and I said	56
58	to be a	56
59	be able to	55
60	have to be	55
61	yeah that's	55
62	That's right	54
63	a bit of	53
64	and so on	53
65	at the end	53
66	It's very	52
67	simultaneous and unclear	52
68	in relation to	51
69	we don't	51
70	I wouldn't	50
71	the fact that	50
72	There's no	50
73	you know that	50

74	you want to	50
75	no it's	49
76	't want to	49
77	was going to	49
78	yeah it's	49
79	Don't want	48
80	it wasn't	48
81	you know it	48
82	Don't have	47
83	I'm sure	47
84	out of the	47
85	some of the	47
86	the attorney general	47
87	is going to	46
88	like you know	46
89	't know I	46
90	What's the	46
91	and then you	45
92	and there's	45
93	it it's	45
94	so it's	45
95	't know what	45
96	the the the	45
97	as well as	44
98	that it's	44
99	wouldn't be	44
100	going to have	43

Appendix F

The four-word cluster list of the top 100 most frequent clusters of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	I don't know	228
2	I don't think	82
3	I'm going to	76
4	I think it's	66
5	We're going to	50
6	's going to be	46
7	yeah yeah yeah yeah	46
8	Don't know I	45
9	at the end of	41
10	Don't know what	40
11	simultaneous and unclear speech	39
12	It's going to	38
13	you know what I	38
14	and I don't	37
15	at the same time	36
16	It's it's	36
17	know what I mean	34
18	the end of the	34
19	of the high court	33
20	I don't want	32
21	you know the way	31
22	Don't want to	28
23	I'd like to	27
24	president of the high	26
25	That's what I	26
26	you know it's	25
27	and I think that	24
28	but I don't	24
29	going to be a	24
30	do you know what	23
31	I think that's	23
32	no I don't	23
33	one point four one	23
34	a bit of a	22

35	so I don't	22
36	They're going to	22
37	a little bit of	21
38	and I think it	21
39	are you going to	21
40	You're going to	21
41	I didn't know	20
42	if you want to	20
43	the attorney general's	20
44	what do you think	20
45	attorney general's office	19
46	I think it was	19
47	It's kind of	19
48	're not going to	19
49	and there's a	18
50	Don't know how	18
51	going to have to	18
52	in relation to the	18
53	in the middle of	18
54	It's a very	18
55	it seems to me	18
56	or something like that	18
57	're going to have	18
58	and I was like	17
59	and it's a	17
60	Don't know if	17
61	I don't I	17
62	I haven't seen	17
63	I'm I'm	17
64	i thought it was	17
65	if you don't	17
66	That's that's	17
67	Don't have to	16
68	I can't remember	16
69	what do you mean	16
70	you don't have	16
71	I'll tell you	15
72	I mean it's	15
73	I was going to	15

74	in the attorney general	15
75	It's not a	15
76	na na na na	15
77	That's all I	15
78	That's right yeah	15
79	well I don't	15
80	yeah that's right	15
81	a lot of people	14
82	an awful lot of	14
83	as president of the	14
84	Didn't want to	14
85	end of the day	14
86	I know I know	14
87	I think there's	14
88	in the case of	14
89	is one of the	14
90	't be able to	14
91	that kind of thing	14
92	the centre of the	14
93	the minister for finance	14
94	the top of the	14
95	There's going to	14
96	They're they're	14
97	yeah I don't	14
98	a lot of the	13
99	and so on and	13
100	do you want to	13

Appendix G

The five-word cluster list of the top 100 most frequent clusters of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	I don't know I	40
2	you know what I mean	33
3	I don't know what	27
4	president of the high court	26
5	yeah yeah yeah yeah yeah	25
6	at the end of the	24
7	It's going to be	24
8	the attorney general's office	19
9	I don't know how	17
10	I don't know if	17
11	I don't want to	17
12	in the attorney general's	15
13	's going to be a	15
14	and I don't know	14
15	na na na na na	14
16	the end of the day	14
17	There's going to be	14
18	as president of the high	13
19	and I think it's	12
20	I don't think I	12
21	I don't think it	12
22	so I don't know	12
23	You're not going to	12
24	I'm going to do	11
25	I'm just going to	11
26	It's a it's	11
27	Don't know I think	10
28	I don't know it	10
29	I don't think that	10
30	I think it's a	10
31	We're going to have	10
32	what do you think of	10
33	and I'm going to	9
34	do you know what I	9

35	no I don't think	9
36	all that kind of stuff	8
37	and all that kind of	8
38	by one point four one	8
39	It's not it's	8
40	it seems to me that	8
41	're going to have to	8
42	That's what I'm	8
43	what I'm going to	8
44	and all this kind of	7
45	and I don't think	7
46	but I don't know	7
47	Don't know it's	7
48	Don't know what to	7
49	for those of you who	7
50	I'd like to thank	7
51	I didn't want to	7
52	I don't think so	7
53	in the middle of the	7
54	like I don't know	7
55	of the high court and	7
56	so I'm going to	7
57	to be honest with you	7
58	well I don't know	7
59	what do you call it	7
60	What's the name of	7
61	Won't be able to	7
62	Wouldn't be able to	7
63	appointment as president of the	6
64	at the centre of the	6
65	at the top of the	6
66	but I don't think	6
67	divided by one point four	6
68	Don't I don't	6
69	Don't know what it	6
70	I don't have to	6
71	I don't I don	6
72	I don't know whether	6
73	I don't think they	6

74	I don't think you	6
75	I'll tell you what	6
76	I'm going to be	6
77	I'm going to have	6
78	I was going to say	6
79	it I don't know	6
80	It's not too bad	6
81	It's uh it's	6
82	know I don't know	6
83	know I know I know	6
84	mean of that data set	6
85	no I don't know	6
86	now I don't know	6
87	one point four one so	6
88	're going to have a	6
89	's a bit of a	6
90	's a it's a	6
91	teaching through the target language	6
92	the mean of that data	6
93	the vast majority of the	6
94	yeah I don't know	6
95	you don't have to	6
96	all this kind of stuff	5
97	and I think that's	5
98	at the same time as	5
99	do you want me to	5
100	do you want to go	5

Appendix H

The six-word cluster list of the top 100 most frequent clusters of Spoken ICE-Ireland

#	ICE-Ireland	Freq
1	at the end of the day	14
2	in the attorney general's office	14
3	as president of the high court	13
4	It's going to be a	13
5	na na na na na na	13
6	yeah yeah yeah yeah yeah yeah	13
7	I don't know I think	9
8	and all that kind of stuff	7
9	I don't know it's	7
10	president of the high court and	7
11	what I'm going to do	7
12	appointment as president of the high	6
13	divided by one point four one	6
14	I don't I don't	6
15	It's a it's a	6
16	the mean of that data set	6
17	Don't know I don't	5
18	I don't know how many	5
19	I don't think it's	5
20	That's what I'm saying	5
21	a couple of weeks ago and	4
22	and all this kind of stuff	4
23	and I think it's a	4
24	at this time of the year	4
25	de de dah de de dah	4
26	do you know what I mean	4
27	Don't know what to do	4
28	ever ever ever ever ever ever	4
29	going to show you how to	4
30	I don't know how much	4
31	I don't know I don	4
32	I don't know I'm	4
33	I don't know if it	4
34	I don't know what to	4

35	I don't want to be	4
36	I don't want to get	4
37	I know I know I know	4
38	I'm going to have to	4
39	if you don't have the	4
40	it's not it's not	4
41	one two three four five six	4
42	p v doyle memorial railway stakes	4
43	's going to be a free	4
44	't know I don't know	4
45	the end of the day I	4
46	we're going to have a	4
47	yeah yeah yeah yeah yeah mm	4
48	you know what I mean yeah	4
49	you're going to have to	4
50	you're not going to be	4
51	a lot of a lot of	3
52	allow anybody to think that fianna	3
53	and I don't know if	3
54	and I don't know what	3
55	and it didn't fall apart	3
56	and so on and so forth	3
57	and unclear speech a leas cheann	3
58	anybody to think that fianna fa	3
59	are the properties of this data	3
60	at the centre of the x	3
61	but you know what I mean	3
62	couple of weeks ago and I	3
63	dah de de dah de de	3
64	didn't fall apart because of	3
65	don't have to be in	3
66	don't know I can't	3
67	don't know I think she	3
68	don't think I've ever	3
69	don't want to get involved	3
70	firearms and falsely imprisoning two people	3
71	for appointment as president of the	3
72	forty three year old man at	3

73	from the point of view of	3
74	going to say I want to	3
75	have been dead for some time	3
76	have the book in front of	3
77	I don't have to be	3
78	I don't know how you	3
79	I don't know I can	3
80	I don't know if there	3
81	I don't know what I	3
82	I don't know what it	3
83	I don't know what she	3
84	I don't know whether I	3
85	I don't think I've	3
86	I don't think I was	3
87	I going to say to you	3
88	I have it I have it	3
89	I haven t seen him in	3
90	I haven't seen you in	3
91	I'm going to be talking	3
92	I'm going to divide each	3
93	I'm going to say I	3
94	I'm going to show you	3
95	I'm just going to show	3
96	I'm very proud of him	3
97	I mean isn't it extraordinary	3
98	I think at the end of	3
99	I think it's a very	3
100	I think it's important that	3

Appendix I

The two-word cluster list of the top 100 most frequent clusters of Spoken ICE-Canada

#	ICE-Canada	Freq
1	It's	4589
2	you know	3140
3	That's	2760
4	of the	2684
5	in the	2338
6	Don't	2287
7	I'm	1775
8	going to	1723
9	mm hmm	1559
10	You're	1427
11	and I	1312
12	I think	1285
13	on the	1273
14	to be	1228
15	it was	1218
16	There's	1210
17	to the	1198
18	's a	1178
19	I don	1159
20	They're	1143
21	have to	1031
22	and the	973
23	I mean	952
24	He's	919
25	gonna going	917
26	at the	888
27	if you	878
28	and then	873
29	and uh	868
30	't know	824
31	to do	806
32	I was	781
33	this is	765
34	want to	760

35	kind of	758
36	Didn't	740
37	and it	701
38	for the	691
39	you have	691
40	and you	673
41	's not	662
42	that you	661
43	a lot	660
44	with the	657
45	but I	629
46	you can	625
47	I I	614
48	We're	614
49	that I	577
50	She's	574
51	and that	572
52	in a	568
53	sort of	567
54	do you	558
55	's the	548
56	and they	546
57	to get	541
58	have a	537
59	and he	533
60	and and	531
61	so I	528
62	is a	511
63	lot of	509
64	of a	508
65	the the	496
66	I've	494
67	it is	494
68	ya ya	486
69	to go	481
70	one of	468
71	that the	462
72	from the	458
73	is that	455

74	he was	453
75	Can't	450
76	well I	447
77	a little	444
78	I'll	443
79	what I	443
80	Doesn't	437
81	was a	435
82	I can	432
83	is the	429
84	to have	425
85	I know	423
86	out of	419
87	know I	418
88	and so	416
89	I have	415
90	all the	412
91	't have	406
92	but it	402
93	's it	401
94	when you	397
95	uh the	394
96	that was	393
97	we have	392
98	ya I	391
99	the same	386
100	as a	385

Appendix J

The three-word cluster list of the top 100 most frequent clusters of Spoken ICE-Canada

#	ICE-Canada	Freq
1	I don't	1152
2	gonna going to	917
3	Don't know	738
4	a lot of	486
5	It's a	483
6	It's not	412
7	going to be	366
8	you don't	362
9	wanna want to	320
10	and it's	319
11	you have to	302
12	There's a	298
13	I didn't	272
14	I'm not	269
15	mm hmm mm	269
16	hmm mm hmm	256
17	That's right	256
18	and that's	255
19	're gonna going	255
20	you know I	245
21	mm hmm and	240
22	That's what	240
23	one of the	239
24	It's just	236
25	It's it	227
26	That's the	223
27	's it's	222
28	Don't think	221
29	Don't have	218
30	two or three	211
31	I think it	200
32	they don't	199
33	're going to	198
34	a little bit	197

35	but it's	197
36	if you're	193
37	or three words	192
38	a few words	187
39	That's a	187
40	I think that	186
41	's gonna going	178
42	you know it	173
43	and you know	165
44	It's like	165
45	kinda kind of	160
46	so it's	159
47	you know what	159
48	going to have	154
49	There's no	153
50	I can't	150
51	you know you	147
52	It's the	145
53	I'm gonna	143
54	That's it	143
55	'm gonna going	142
56	you know and	142
57	know it's	139
58	You're gonna	135
59	you know the	133
60	and I'm	132
61	a couple of	131
62	it doesn't	130
63	They're not	130
64	it was a	129
65	to be a	129
66	going to do	128
67	be able to	127
68	's going to	127
69	so that's	127
70	think it's	127
71	ya it's	127
72	and I think	125
73	I mean I	125

74	and he's	120
75	pause about seconds	119
76	some of the	119
77	uh you know	119
78	you can't	119
79	i have to	118
80	out of the	118
81	you know that	118
82	you know like	117
83	that it's	116
84	three or four	116
85	part of the	114
86	this is a	114
87	uh it's	113
88	c b c	112
89	i mean it	110
90	't know I	110
91	I'm going	109
92	It's very	108
93	You've got	108
94	't know if	107
95	and there's	106
96	I think I	106
97	That's not	106
98	we don't	106
99	You're not	106
100	when you're	104

Appendix K

The four-word cluster list of the top 100 most frequent clusters of Spoken ICE-Canada

#	ICE-Canada	Freq
1	I don't know	561
2	're gonna going to	255
3	mm hmm mm hmm	249
4	It's it's	206
5	I don't think	194
6	two or three words	192
7	's gonna going to	178
8	gonna going to be	177
9	I'm gonna going	142
10	'm gonna going to	142
11	You're gonna going	133
12	gonna going to have	108
13	Don't know I	105
14	Don't know if	103
15	I think it's	102
16	you know it's	99
17	you wanna want to	96
18	three or four words	94
19	I'm going to	86
20	dunno don't know	85
21	I dunno don't	83
22	going to have to	82
23	Don't know what	81
24	That's that's	78
25	I'd like to	77
26	't wanna want to	77
27	We're going to	76
28	gonna going to do	75
29	you don't have	73
30	You're going to	72
31	That's what I	71
32	c b c news	68
33	It's gonna going	67
34	you know what I	67

35	I mean it's	64
36	We're gonna going	64
37	Y know you know	63
38	Don't wanna want	61
39	I think that's	61
40	's going to be	61
41	Don't have to	58
42	gonna going to get	58
43	I wanna want to	58
44	at the same time	57
45	to be able to	57
46	hmm mm hmm mm	55
47	one or two words	55
48	the end of the	55
49	They're gonna going	55
50	a little bit of	54
51	but I don't	51
52	no I don't	51
53	not gonna going to	51
54	Don't know how	50
55	I'm not sure	50
56	or something like that	50
57	and I don't	49
58	I'm I'm	49
59	They're going to	48
60	was gonna going to	48
61	Don't want to	47
62	gotta have got to	47
63	's a lot of	47
64	It's going to	46
65	thank you very much	46
66	the honourable member for	46
67	I don't have	45
68	if you don't	45
69	There's a lot	45
70	you don't know	45
71	c j a d	44
72	four or five words	44
73	know what I mean	44

74	a lot of people	43
75	going to be a	43
76	I don't want	43
77	I I don't	42
78	That's it for	41
79	well I don't	40
80	It's not a	39
81	you know I'm	39
82	Don't think I	38
83	gonna going to go	38
84	mm hmm you know	38
85	're going to be	38
86	you have to do	38
87	and that's what	37
88	just gonna going to	37
89	a lot of the	36
90	and I think that	36
91	at the end of	36
92	know I don't	36
93	mm hmm and I	36
94	That's a good	36
95	uh I don't	36
96	beyond a reasonable doubt	35
97	I would like to	35
98	is gonna going to	35
99	That's right and	34
100	the rest of the	34

Appendix L

The five-word cluster list of the top 100 most frequent clusters of Spoken ICE-Canada

#	ICE-Canada	Freq
1	I'm gonna going to	142
2	you're gonna going to	133
3	I don't know if	92
4	I don't know I	89
5	I dunno don't know	83
6	It's gonna going to	67
7	We're gonna going to	64
8	's gonna going to be	63
9	don't wanna want to	61
10	I don't know what	61
11	gonna going to have to	58
12	're gonna going to have	55
13	they're gonna going to	55
14	mm hmm mm hmm mm	54
15	hmm mm hmm mm hmm	49
16	you know what I mean	42
17	I don't think I	35
18	There's a lot of	34
19	you don't have to	34
20	this week's ecofacts and	33
21	Week's ecofacts and trends	33
22	you don't wanna want	31
23	I was gonna going to	30
24	're gonna going to be	30
25	It's going to be	29
26	you gotta have got to	28
27	Don't know if I	27
28	I don't know how	27
29	's it for this week	26
30	That's it for this	26
31	That's what I'm	26
32	it for this week's	25
33	It's it's a	25
34	well I don't know	25

35	gonna going to be a	24
36	It's a it's	24
37	mm hmm mm hmm and	24
38	That's right that's	24
39	He's gonna going to	23
40	I don't wanna want	23
41	qu eacute b'eacute cois	23
42	That's gonna going to	23
43	are you gonna going to	22
44	c b c news montreal	22
45	I don't know it	22
46	at the end of the	21
47	I don't think so	21
48	I don't think that	21
49	I don't want to	21
50	're not gonna going to	21
51	uh I don't know	21
52	ya I don't know	21
53	a little bit of a	20
54	and it's it's	20
55	audience member asks a question	20
56	ecofacts and trends I'm	20
57	for this week's ecofacts	20
58	I don't know why	20
59	're gonna going to do	20
60	's ecofacts and trends I	20
61	You're listening to the	20
62	I don't think it	19
63	I'm just gonna going	19
64	I'm not going to	19
65	know I don't know	19
66	laugh I don't know	19
67	'm gonna going to do	19
68	'm just gonna going to	19
69	mm hmm mm hmm so	19
70	you know I don't	19
71	if you wanna want to	18
72	is gonna going to be	18
73	It's uh it's	18

74	know it's it's	18
75	're gonna going to get	18
76	's right that's right	18
77	She's gonna going to	18
78	tape stops and starts again	18
79	uhm I don't know	18
80	You're going to be	18
81	do you wanna want to	17
82	I don't know where	17
83	It's it's it	17
84	's a it's a	17
85	the c b c news	17
86	you don't want to	17
87	and I don't know	16
88	but I don't know	16
89	Don't know it's	16
90	dunno don't know I	16
91	f c ocircumflex t'eacute	16
92	It's not it's	16
93	no I don't think	16
94	pierre f c ocircumflex t	16
95	're gonna going to go	16
96	's it's it's	16
97	so on and so forth	16
98	the scott tournament of hearts	16
99	you know it's it	16
100	but at the same time	15

Appendix M

The six-word cluster list of the top 100 most frequent clusters of Spoken ICE-Canada

#	ICE-Canada	Freq
1	mm hmm mm hmm mm hmm	48
2	You're gonna going to have	36
3	It's gonna going to be	34
4	this week's ecofacts and trends	33
5	you don't wanna want to	31
6	're gonna going to have to	27
7	I don't know if I	26
8	That's it for this week	26
9	's it for this week's	25
10	I don't wanna want to	23
11	for this week's ecofacts and	20
12	it for this week's ecofacts	20
13	's ecofacts and trends I'm	20
14	week's ecofacts and trends I	20
15	I'm gonna going to do	19
16	I'm just gonna going to	19
17	That's right that's right	18
18	I dunno don't know I	16
19	it's a it's a	16
20	pierre f c ocircumflex t eacute	16
21	I don't know it's	15
22	I'm gonna going to have	15
23	I'm not gonna going to	15
24	it's it's it's	15
25	're listening to the chevron ecofile	15
26	so I'm gonna going to	15
27	you know it's it's	15
28	you're listening to the chevron	15
29	and so on and so forth	14
30	audience member comments right audience member	14
31	if you're gonna going to	14
32	c b c news for quebec	13
33	don't know I don't	13
34	I don't know I don	13

405

35	it's not it's not	13
36	member comments right audience member comments	13
37	's gonna going to be a	13
38	taped phone interview with guest speaker	13
39	the c b c news for	13
40	what I'm gonna going to	13
41	You're gonna going to be	13
42	I don't I don't	12
43	I don't know if it	12
44	parti qu eacute b eacute cois	12
45	That's gonna going to be	12
46	you're gonna going to get	12
47	you're not gonna going to	12
48	bloc qu eacute b eacute cois	11
49	hmm mm hmm mm hmm mm	11
50	I don't know I think	11
51	I don't know if you	11
52	I'm gonna going to go	11
53	no I don t think so	11
54	We're gonna going to have	11
55	and I'm gonna going to	10
56	and trends I'm Annie Gillis	10
57	don't wanna want to be	10
58	I don't know if they	10
59	I don't think there's	10
60	'm gonna going to have to	10
61	't know I don't know	10
62	the p q parti qu eacute	10
63	there's a there's a	10
64	what are you gonna going to	10
65	with this week's ecofacts and	10
66	afternoon here's the c b	9
67	are you gonna going to do	9
68	c b c news Quebec City	9
69	good afternoon here's the c	9
70	here's the c b c	9
71	I don't know why I	9

72	I mean it's it's	9
73	're gonna going to have a	9
74	's it for our feature report	9
75	that's it for our feature	9
76	well I don't know I	9
77	you don't know how to	9
78	amendments to the young offenders act	8
79	and you're gonna going to	8
80	answer exchange with extra corpus speaker	8
81	ecofacts and trends I'm Annie	8
82	going to have to deal with	8
83	gonna going to have to deal	8
84	I don't know I just	8
85	I don't know what it	8
86	I don't know what the	8
87	I was gonna going to say	8
88	listening to the chevron ecofile now	8
89	p q parti qu eacute b	8
90	q parti qu eacute b eacute	8
91	question answer exchange with extra corpus	8
92	the bloc qu eacute b eacute	8
93	They're gonna going to be	8
94	to the chevron ecofile now here	8
95	we're gonna going to be	8
96	what I'm going to do	8
97	you know I don't know	8
98	you know what I mean like	8
99	you're going to have to	8
100	you're gonna going to do	8

Appendix N

The candidates list of Spoken ICE-Ireland

One word	Two-word cluster	Three-word cluster	Four-word cluster	Five-word cluster	Six-word cluster
that	You know	Yeah yeah yeah	I don't know	You know what I mean	at the end of the day
Uh	That's	You don't	Yeah yeah yeah yeah	So I don't know	I don't know I think
yeah	I think	I think that	At the same time	I'm just going to	and all that kind of stuff
so	Yeah yeah	I'd say	You know the way	What do you think of	That's what I'm saying
like	Kind of	At the moment	I'd like to	All that kind of stuff	at this time of the year
uhm	Do you	You know what	Do you know what	It seems to me that	do you know what I mean
well	And then	It's just	No I don't	I'd like to thank	I know I know I know
now	I mean	That kind of	What do you think	I don't think so	you know what I mean yeah
oh	I know	Yeah that's	it seems to me	to be honest with you	and so on and so forth
just	Like that	That's right	Or something like that	well I don't know	but you know what I mean
which	Are you	And so on	What do you mean	what do you call it	I think it's important that
really	Did you	At the end	That's all I	It's not too bad	
right	What that	I'm sure	That's right yeah	yeah I don't know	
okay	Uh-huh	That's bad	Yeah that's right	What do you think about	
	Oh yeah	That's disgusting	I know I know		
	Like what	That's it	That kind of thing		
	Oh well	That's all	That's fine sure		
	Oh right	Yeah I know	That's right excellent		
	Just not	Is that so	That's all you		

	Just no	By the way	And you know what
	Just too		
	Right right		

Appendix O

The candidates list of Spoken ICE-Canada

One	Two	Three	Four	Five	Six
That	You know	That's right	I don't know	you know what I mean	and so on and so forth
Uh	That's	It's just	I don't think	Well I don't know	I don't know I think
so	Mm hmm	I think that	I'd like to	I don't think so	you know what I mean like
like	I think	And you know	I think that's	Uh I don't know	No I don't think so
ya	I mean	kinda kind of	At the same time	ya I don't know	
uhm	And then	You know what	but I don't	I don't know why	
well	And uh	It's like	No I don't	I'm just gonna going	
just	Kind of	so it's	I'm not sure	You know I don't	
mm	You can	That's it	Or something like that	uhm I don't know	
okay	Sort of	So that's	Thank you very much	But I don't know	
right	Do you	And I think	That's it for	no I don't think	
oh	So I	Uh you know	well I don't	But at the same time	
really	Ya ya	You know that	mm hmm you know		
yeah	I know	you know like	and I think that		
	And so	That's not	That's a good		
	Uh uh	What that is	And I don't		
	Or so	Is that so	and then after that		
	Oh well	Well let's	But oh ya anyways		
	Mm mm				
	Ya right				
	Oh ya				
	Okay okay				
	Right right				

Appendix P

Transcription and markup conventions used in the illustrated examples in this PhD thesis

Symbol	Meaning
+	Overlapping speech
[..]	Editing notes ie. [laughs] [tape cuts out]
~	Speaker cuts off word
^	Self-correction by speaker

Appendix Q

Spoken text category file ID of ICE-Canada and ICE-Ireland (ROI)

Text category	*ICE-Canada*	*ICE-Ireland*
Broadcast discussions	S1B-021 *to* S1B-040	S1B-031 to S1B-040
Broadcast interviews	S1B-041 *to* S1B-050	S1B-046 to S1B-050
Broadcast news	S2B-001 *to* S2B-020	S2B-011 to S2B-020
Broadcast talks	S2B-021 *to* S2B-040	S2B-031 to S2B-040
Business transactions	S1B-071 *to* S1B-080	S1B-076 to S1B-080
Classroom lessons	S1B-001 *to* S1B-020	S1B-011 to S1B-020
Demonstrations	S2A-051 *to* S2A-060	S2A-056 to S2A-060
Face to face conversation (Direct)	S1A-001 *to* S1A-090	S1A-046 to S1A-090
Legal cross-examinations	S1B-061 *to* S1B-070	S1B-066 to S1B-070
Legal presentations	S2A-061 *to* S2A-070	S2A-066 to S2A-070
Parliamentary debates	S1B-051 *to* S1B-060	S1B-056 to S1B-060
Scripted speeches (not broadcast)	S2B-041 *to* S2B-050	S2B-046 to S2B-050
Spontaneous commentaries	S2A-001 *to* S2A-020	S2A-011 to S2A-020
Telephone conversation	S1A-091 *to* S1A-100	S1A-096 to S1A-100
Unscripted speeches	S2A-021 *to* S2A-050	S2A-036 to S2A-050