

A CORPUS BASED STUDY OF PRAGMATIC MARKERS IN SPOKEN STANDARD SRI LANKAN ENGLISH

A DISSERTATION

submitted to the University College London

in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

by

Ranaweera Kaluarachchige Mahishi Chami Ranaweera

University College London

September 2024

I, Ranaweera Kaluarachchige Mahishi Chami Ranaweera, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.



RKMC Ranaweera

Abstract

A Corpus Based Study of Pragmatic Markers in Spoken Standard Sri Lankan English

Pragmatic markers (PMs) add very little contribution to the propositional meaning of an utterance, but they are an essential characteristic of natural speech. The past few decades have seen an extensive interest in research on PMs in all varieties of English. However, there is still little research into PMs in South Asian Englishes. Therefore, this study aims to explore an untrodden area in Sri Lankan English (SLE) by examining the presence and functions of PMs in spoken SLE.

This thesis examines 2949 instances of PMs in a purpose-built corpus of spoken Standard Sri Lankan English (SSLE). The corpus contains 202,557 words, consisting of interviews of 36 men and 36 women participating in several talk show series found in YouTube channels published between 2016-2021. All the speakers were selected on the basis of their use of SSLE, and they represent four different ethnicities in Sri Lanka. Their ages vary from 27 to 81 years. This study examines the distribution and the pragmatic functions of PMs found in the corpus. The results reveal that PMs in SSLE show similar functional patterns to British English, American English and Indian English to varying degrees. There is evidence of nativized PMs and nativized functions for PMs shared with other varieties. The thesis goes on to present a gender-based analysis of PM use. Men in the corpus use marginally more PMs than women, which is a deviation from the picture often presented by genderlectal research. The thesis also reveals a correlation between age and PM use. Some PMs show different patterns of frequency and different functions as speaker age rises. Overall, the study contributes novel data to the pragmatic description of SSLE and to the growing research on PMs in world Englishes.

Impact Statement

Being the first of its kind, my research on Pragmatic Markers in Standard Sri Lankan English adds new knowledge to the body of literature broadly on Sri Lankan English (SLE) and specifically on Standard Sri Lankan English (SSLE). SSLE is still in a state of being defined in research and therefore, this study adds to the body of work that identifies the key characteristics of SSLE.

My research contributes to variational research in world Englishes and specifically South Asian Englishes and complements existing work on the intersection between social factors and pragmatic features. The results of this study can be used in cross-linguistic comparisons of communicative behaviour.

Beyond its contribution to academic research in the disciplines of world Englishes and pragmatics, this study has relevance in other areas of education. My research will contribute to the pedagogy of English as a second language (ESL) by defining SSLE as it is the dialect that is promoted in the ESL classroom. As perception studies reveal, the key features of SSLE are not understood by a great part of the teachers of ESL in Sri Lanka. This is partly because there is a lack of research that specifically address the features of SSLE. Studies based on spoken discourse in SSLE are in particularly short supply. Therefore, my research will inform the teachers of SSLE usage based on authentic data. Additionally, the very important role PMs play in spoken discourse is generally overlooked in ESL materials. My research reveals how SSLE speakers use PMs to ensure a smooth flow of discourse. I am a permanent staff member of the Department of English Language Teaching at the University of Kelaniya, Sri Lanka and as a department we work closely with the textbook writing service of the Ministry of Education. The findings in my study can directly inform how dialogues are presented as samples of natural conversation in the textbooks by showing that even the most formal conversations contain PMs fulfilling various functions.

PMs are often stigmatised, and my research shows that they are widespread and communicatively valuable feature of language despite the stigma. Therefore, the results of this study are relevant even for the Sri Lankan public interested in English to understand that PMs should be held

above the stigma as they facilitate very useful pragmatic strategies to ensure a smooth flow of conversation. Most of these linguistic forms convey subtle messages to the interlocutor that enhance the speaker-listener relationship positively.

Finally, a key contribution of this study is the creation of a new corpus of spoken Sri Lankan English, the Corpus of Standard Sri Lankan English (CSSLE). Currently, only one corpus of specifically spoken Sri Lankan English exists, and this is the International Corpus of English – Sri Lanka (ICE-SL). At the completion of the study, CSSLE will be hosted in the Digital Humanities Lab of University of Colombo, Sri Lanka (dhlab.cmb.ac.lk) so that it can be used by other researchers working on Sri Lankan English.

Dedication

To

SLW, my companion

YRW, my life's work

Acknowledgements

I am greatly indebted to my supervisors, Dr. Rachele de Felice, Dr. Kathryn Allan and Dr. Guyanne Wilson for their guidance and encouragement. This thesis is guided by their tremendous knowledge, insightful comments and advice. I am indebted to the three of them in different ways. I am grateful to Rachele for having faith in me, guiding the design of the study and continuously checking its progress amidst transition of work. I deeply appreciate Kathryn for her wisdom, insightful guidance, kind encouragement and exceptional support extended well beyond her supervisory role. I am immensely thankful to Guyanne for her excellent knowledge, feedback and advice that shaped the thesis in its final stage. I owe any well-founded points in my thesis to them. Any remaining inadequacies are all mine.

In addition to my supervisors, I owe this thesis to many individuals. First, I would like to thank Prof. Doug Biber, Prof. Randi Reppen, Prof. Maithree Wickremesinghe, and Prof. D.M. Semasinghe for their recommendations. I acknowledge Prof. Nilanthi de Silva, the current Vice Chancellor of the University of Kelaniya, with sincere gratitude for all her support. I am also indebted to Prof. Patrick Ratnayake, Prof. Lakshman Senevirathne and Dr. Sudath Senarath for their support in many ways. I am very grateful to Prof. Dilkushi Wettewe and Gevani Prahalathan for their unwavering support throughout the study leave. I wish to thank Prof. Helen Hackett for being a very supportive graduate tutor. I also acknowledge Ms. Wendy Appleby, the Registrar and Head of Student and Registry Services for making special efforts to facilitate my funding.

My sincere thanks are due to the financial assistance provided by the University of Kelaniya at the beginning of the PhD. Additionally, I am thankful to the World Bank AHEAD grant for funding part of the PhD expenses and to the UCL financial assistance fund for providing me relief in unexpected circumstances.

Credit and deep gratitude are due to Prof. Manique Gunsekera posthumously for bewitching me with Sri Lankan English. I gratefully acknowledge Dinali Fernando, Mihiri Jansz, Thilanka Wijesinghe, Prof. Dushy Mendis and Prabha Manuratne for helping me in various ways

regarding the data. I also thank Prof. Kim McDonough and Mansour Al-Sourmi for their valuable comments on my initial proposal. I thank Rev. Thapovanaye Sutadhara, Ms. Kamini Perera, Shyamani Hettiarachchi, Tharupathi Munasinghe, Hasitha Pathirana, Anton Pushparaj, and Thilina Wickramaarachchi for all their encouragement throughout. I especially acknowledge Dr. Lakshmi de Silva for being an inspirational mentor and her unfailing faith in me.

A full time PhD student with family is also an employee. I wish to thank all those who gave me an opportunity in the working world: Prakash and Jenny Gautam, Dr. Sue Sing, Calvin Fell and Eric Stober. Prakash and Jenny hold a very special place in my heart as I owe much of my writing to the time and space freely given to me in their college. I fondly acknowledge the encouragement and support of Sameer, Neelam, and Kiran.

My friends made some of the trying days bearable. I am very grateful to Shani and Navein, Karthy and Debbie, Dinesh, Omesha and Reyhan, Raj and Uma, Sanjeewa, Pushpa and Neili, Gayathri and Evan, Surangi, aunty Anula and Nelly Gillet. They have helped me in numerous ways and I am forever indebted. I wish to thank Daisuke Suzuki for always offering a helping hand when needed. I think fondly of my friends and cousins who sent me good cheer from Sri Lanka constantly.

Whatever pursuits I have achieved rest firmly on my foundation: my parents, my sisters Vidushi and Nilushi, my brothers in law Chandana and Eranda, and my in-laws. I thank Kaveesha, Kavya, Akain and Krishen with affection for their best wishes. Finally, Sanjeewa and Yavin: your love and patience inspire me every day. I quote one of your favourite music groups, *Cold Play*, in awe of you: "Under this pressure, under this weight, you are diamonds". You truly are. Thank you both very much.

Table of Contents

Chapter 1 : Introduction	1
1.1 Aims of the Study	1
1.2 Background	1
1.3 Rationale of the Study.....	6
1.4 Research Questions	7
1.5 Significance of the Study.....	7
1.6 Overview of the Thesis.....	8
Chapter 2 : Sri Lankan English.....	12
2.1 Introduction	12
2.2 English in Sri Lanka.....	12
2.3 The Status of English in Sri Lanka.....	14
2.4 Definitions of Sri Lankan English	19
2.5 The Manner of SLE Acquisition	23
2.6 Standard Sri Lankan English Speaker Identity.....	24
2.7 Features of Sri Lankan English	26
2.7.1 Phonology	27
2.7.2 Morphology and semantics.....	28
2.7.3 Syntax.....	30
2.7.4 Prosodic Features of SLE	31
2.7.5 Pragmatics.....	32
2.8 Primary Dialects of Sri Lankan English	34
2.8.1 Standard Sri Lankan English	35
2.8.2 The Non-Standard Sri Lankan English	40
2.8.3 Burgher English	41
2.9 Summary	42
Chapter 3 : Pragmatic Markers	43
3.1 Overview	43
3.2 Terminology and Definitions of Pragmatic Markers.....	44
3.3 Characteristics of PMs.....	48
3.4 Co-occurring PMs.....	52
3.5 Nativised PMs	57
3.6 The Functions of PMs.....	59
3.7 Labelling Functions	62
3.8 Perceptions about Pragmatic Markers	63

3.9 Sociolinguistic Factors Affecting the Use of PMs	66
3.10 Conclusion	70
Chapter 4 : Chapter 4 Design of the Study	71
4.1 Introduction	71
4.2 Research questions	71
4.3 Data	73
4.4 Research Design	77
4.5 Designing and Creating the Corpus of Spoken Sri Lankan English (CSSLE)	79
4.5.1 The Speakers	86
4.5.2 The Rating Process	87
4.5.3 The Transcription of Data	88
4.5.3.1 British National Corpus 2014	89
4.5.4 Justification for Using BNC 2014 Transcription Conventions	91
4.5.5 Size of the Corpus	92
4.5.6 Providing Identification to the Data	93
4.5.7 Additional Remarks about the Transcriptions	94
4.6 Variationist Approach	97
4.7 The Approach to Analysing the Functions	99
4.8 Approach to Analysing Gender	115
4.9 Approach to Analysing Age	116
4.10 Chapter Summary	117
Chapter 5 : A Functional Analysis of the Pragmatic Markers	118
5.1 Overview	118
5.2 Preliminary Results	119
5.3 The Criteria for Deciding the Pragmatic Functions of the PMs.....	122
5.4 The Analysis of the Functions	143
5.4.1 The Functions of <i>You Know</i>	144
5.4.2 The Functions of <i>I Mean</i>	150
5.4.3 The Functions of <i>Like</i>	155
5.4.4 The Functions of <i>Well</i>	157
5.4.5 The Functions of <i>Right</i>	159
5.4.6 The Functions of <i>Kind of</i> and <i>Sort of</i>	161
5.4.7 The Functions of <i>Yeah</i>	162
5.4.8 The Functions of <i>You see</i>	165
5.5 The Functions of Nativised PMs.....	167
5.5.1 The Functions of <i>Aah</i>	168

5.5.2 The Functions of <i>No</i>	168
5.6 The functions of Co-occurring PMs.....	172
5.7 Conclusions	176
Chapter 6 : Gender and Pragmatic Marker Variation	180
6.1 Introduction	180
6.2 Linguistic Differences Based on Gender: A Broad Survey of Previous Research	182
6.2.1 Gender in World Englishes research.....	185
6.2.2 Gender in Sri Lanka and in Sri Lankan English.....	186
6.3 Research on the Impact of Gender on PMs	190
6.4 Gender-based Results of the Present Study	194
6.4.1 General Results	194
6.4.2 The Gender-based Distribution of Co-occurring PMs.....	199
6.5 Gender-based Data Analysis of Individual PMs	200
6.5.1 A Gender-based Analysis of <i>You Know</i> in the CSSLE.....	200
6.5.2 A Gender-based Analysis <i>I Mean</i> in the CSSLE.....	205
6.5.3 A Gender-based Analysis <i>Like</i> in the CSSLE.....	206
6.5.4 A Gender-based Analysis of <i>Well</i> in the CSSLE	209
6.5.5 A Gender-based Analysis of <i>Right</i> in the CSSLE	210
6.5.6 The Nativized PMs in CSSLE	212
6.5.6.1 <i>No</i>	212
6.5.6.2 <i>Aah</i>	213
6.6 The Sub-corpus of Academics.....	214
6.6.1 <i>You know</i> in the CSSLEAc corpus.	216
6.6.2 Analysis of the other PMs in CSSLEAc.....	221
6.7 A Summary of Findings	222
Chapter 7 : Age and Pragmatic Marker Variation.....	225
7.1 Introduction	225
7.2 Age as a Sociolinguistic Variable	226
7.3 Research on the Impact of Age on PMs.....	227
7.4 Theoretical Framework for Analysing Age and Language in the Current Study.....	229
7.5 Age-based Data Presentation and General Results	232
7.5.1 General Results	232
7.6 Data analysis of Individual PMs	236
7.6.1 <i>You know</i>	237
7.6.2 <i>Like</i>	244
7.6.3 <i>I mean</i>	249

7.6.4 <i>Well</i>	251
7.6.5 <i>Kind of and sort of</i>	253
7.6.6 <i>Right</i>	256
7.6.7 <i>Yeah</i>	258
7.6.8 <i>You see</i>	259
7.6.9 <i>No</i>	260
7.6.10 <i>Aah</i>	261
7.7 A Summary of Findings	261
Chapter 8 : Conclusions	264
8.1 Introduction	264
8.2 PMs use and function: key findings	265
8.3 The impact of gender on PMs: key findings	269
8.4 The impact of age on PMs: key findings	272
8.5 Implications with regard to ESL teaching in Sri Lanka	275
8.6 Limitations	276
8.7 Future directions	277
References	279
APPENDIX A	308
APPENDIX B	315
APPENDIX C	321
APPENDIX D	324
APPENDIX E	327
APPENDIX F	329
APPENDIX G	333
APPENDIX H	341

Table of Tables

Table 2.1: Phonological features of SSLE.....	38
Table 2.2: Standard Sri Lankan English Vowels (Gunesequera, 2005, p. 117)	39
Table 2.3: Standard British English (RP) vowels (Gunesequera, 2005, p. 117)	39
Table 4.1: Sample sources and length of the YouTube videos featuring artists.....	83
Table 4.2: Overview of sub-corpora of CSSLE	92
Table 4.3: A comparison of the auto-generated script with the manually corrected auto-generated script	97
Table 4.4: The data included in the Excel sheet	114
Table 5.1: PMs present in the CSSLE	120
Table 5.2: Distribution of PMs across all speakers (excluding co-occurring PMs)	121
Table 5.3: The list of functions used in the classification	126
Table 5.4: The functions of the co-occurring PMs	174
Table 5.5: the functions of the co-occurring PM like you know in CSSLE.	174
Table 6.1: Tokens of total PMs including co-occurring PMs used in the CSSLE corpus	194
Table 6.2: Percentages of the tokens of the PMs used in the SSLE corpus according to gender and profession.....	197
Table 6.3: Distribution of you know by gender of speaker in CSSLE	200
Table 7.1: Social generations and the distribution of the speakers in the SSLE corpus	232
Table 7.2: The normalised tokens per 100,000 words	232
Table 7.3: Raw data for PM use for the three social generations	234
Table 7.4: The functional distribution of you know according to age.....	240
Table 7.5: The functional distribution of like according to age	245
Table 7.6: The functional distribution of I mean according to age	250
Table 7.7: The functional distribution of kind of and sort of according to age.....	254

Table of Figures

Figure 4.1: AntConc version 4.3.1 concordance lines for well in CSSLE	105
Figure 4.2: AntConc version 4.3.1 concordance lines for you know in CSSLE	108
Figure 5.1: The functional use of you know in the CSSLE	145
Figure 5.2: Functional Use of I mean in the CSSLE	150
Figure 5.3: Functional Use of Like in the CSSLE.....	155
Figure 5.4: Functional Use of Like in the CSSLE.....	157
Figure 5.5: Functional Use of Right in the CSSLE.....	159
Figure 5.6: Functional Use of Yeah in the CSSLE	163
Figure 5.7: Functional Use of You See in the CSSLE	166
Figure 6.1: The overall distribution of PMs according to gender.....	196
Figure 6.2: A gender-based distribution of the co-occurring PMs used in the SSLE corpus	199
Figure 6.3: A gender-based distribution of functions of you know	202
Figure 6.4: A gender-based distribution of functions of I mean	205
Figure 6.5: A gender-based distribution of functions of like.....	208
Figure 6.6: The functional distribution of well.....	209
Figure 6.7: A gender-based frequency of right in the SSLE corpus	211
Figure 6.8: Distribution of PMs among academics	215
Figure 6.9: The distribution of functions of you know according to gender in the SSLEAc sub-corpus	218
Figure 6.10: Functions of right in SSLEAc sub-corpus	221
Figure 7.1: Normalised tokens for the three social generations.....	234
Figure 7.2: A normalised token comparison of the most popular PMs	236
Figure 7.3: The distribution of pragmatic markers (normalised frequencies)	237
Figure 7.4: The distribution of functions of like according to age	246
Figure 7.5: The functional distribution of I mean according to age.....	250
Figure 7.6: The functional distribution of well according to age	251
Figure 7.7: The functional distribution of kind of and sort of according to age	255
Figure 7.8: <i>The functional distribution of right</i>	257
Figure 7.9: The functional distribution of yeah.....	259
Figure 7.10: The functional distribution of you see	259
Figure 7.11: The functional distribution of no	261

List of Abbreviations

AmE	American English
BB	Baby Boomers (born between 1946 – 1964)
BNC2014	British National Corpus 2014
BrE	British English
CoSEM	Corpus of Singaporean English Messages (CoSEM)
CSSLE	Corpus of Standard Sri Lankan English
EFL	English as a Foreign Language
ENL	Native Language
ESL	English as Second Language
Gen X	Generation X (born between 1965 – 1980)
Gen Y	Generation Y (born between 1981 – 1997)
ICE	International Corpus of English
ICE-GB	International Corpus of English -Britain
ICE-GHA	International Corpus of English - Ghana
ICE-IND	International Corpus of English – India
ICE-Ireland	International Corpus of English - Ireland
ICE-NZ	International Corpus of English - New Zealand
ICE- Nigeria	International Corpus of English - Nigeria
ICE-SL	International Corpus of English - Sri Lanka
IE	Indian English
L1	First language
L2	Second language
LINDSEI	Louvain International Database of Spoken English Interlanguage
LTTE	Liberation Tigers of Tamil Eelam
NL	Native Language
NSSLE	Non-standard Sri Lankan English speakers
PM	Pragmatic Markers
RP	Received Pronunciation
SLE	Sri Lankan English

SPICE-Ireland Systems of Pragmatic Annotation in the Spoken Component of ICE-Ireland

SSLE Standard Sri Lankan English

UWL- Role Play University of West London Role Play Corpus

Chapter 1: Introduction

1.1 Aims of the Study

The present research examines the use of pragmatic markers (PMs) among speakers of Standard Sri Lankan English (SSLE) and focuses particularly on their functions. Additionally, I look at how the social factors gender and age interact with the PMs and their functions. PMs are short words and phrases such as *you know*, *I mean*, *well* and *you see* which indicate speaker attitude and assist in discourse management; hence, they are an essential, integral and a widespread linguistic phenomenon associated particularly with spoken language. Despite their common presence, they are syntactically detachable from an utterance, and they do not add anything to its propositional meaning (Beeching, 2016). Interlocutors use PMs to guide the process of interpretation of an utterance as they provide important hints about what has been said or is about to be said (de Klerk, 2005, p.1184). However, they are often overlooked and underestimated in the important role they play in conducting a smooth conversation.

The main aim of this study is to understand how PMs are used in SSLE. My study is at the formal level of pragmatic analysis i.e. it is the starting point of looking at PMs in SSLE. In this endeavour, the empirical data in the study undergoes a form-to-function mapping (Schneider, 2010). The communicative functions of the PMs are analysed to understand the purpose of the PMs in SSLE. From this level of pragmatic analysis, I move on to the second aim of the study, which is an analysis of how PM production reflects the social factors of gender and age of the speakers. This analysis will provide important information about how SSLE speakers show their gender- and age-based identities using PMs.

1.2 Background

PMs are referred to by a myriad of terms such as discourse markers, discourse particles, pragmatic particles, boosters, fumbles and modal particles (Beeching 2002; Edmondson 1981; Hyland, 2000; Schiffrin 1987; Weydt, 2006). The most widely used term, apart from the term PM, is

discourse marker. PMs are considered an overarching term which includes discourse markers as a subset (Crible, 2018; Aijmer and Simon-Vandenberg, 2011; Carter and McCarthy, 2006). Phrases such as *I mean* and *you know* are functional phrases as opposed to content phrases. The present study examines pragmatic meaning and is situated in pragmatics. Therefore, it was decided to use the term pragmatic marker over discourse marker.

PMs are multifunctional (Brinton, 1996), and scholars have compiled long inventories of these functions (e.g. Beeching, 2016; D’Arcy, 2017). There is little agreement on the classification of the PM functions among researchers (Brinton, 1996). Nevertheless, it is agreed that PMs contribute to the discourse-pragmatic level of language by performing textual and interpersonal functions (Scivoletto, 2023). Östman (1981, p. 153) states that their sole function is to “implicitly anchor the propositional meaning”. For instance, they signpost the proposed meaning of the discourse and often direct an emotive interpretation of the message. They show the speaker’s attitude towards the interlocutor or to the topic of the discussion (Wierzbicka, 2003). Hence, without PMs, a discourse would sound discordant, awkward, unnatural and may also lack in interpersonal connection. Although they sit outside the syntax of an utterance, they play a crucial role in the progress of the discourse.

PMs have been studied extensively since the 1960s, especially in Kachru’s inner circle¹ varieties of English such as American English (AmE), British English (BrE), and Canadian English (Beeching, 2016; D’Arcy, 2017; Denis, 2015; Denis & Tagliamonte, 2016). PM functions and positions in an utterance, and the ways they interact with social factors such as gender and age, are well researched in inner circle varieties (Murphy, 2010; Vine & Holmes, 2023; Schweinberger, 2020). Their

¹ Kachru (1985) provided a three-circle model to classify the world Englishes, based on the role English plays in the respective countries. The inner circle variety Englishes provide the norms for other varieties of English. The outer circle varieties such as Indian English and Sri Lankan English emerged as a result of colonialism. They are norm developing countries. They reproduce and further develop the norms provided by inner circle varieties. The expanding circle varieties are norm dependent. They are depended on and fully follow the norms provided by the inner circle varieties.

presence in some outer circle varieties of English such as Indian English, Singaporean English, Philippine English and Trinidadian English are recorded, yet not with the same thorough detail as the inner circle varieties (Lange, 2009; Leimgruber et al., 2021; Westphal, 2024). Currently there is a lack of systematic descriptions of the features of PMs in world Englishes.

The idea of a monolithic conception of English was challenged and its ideological implications considered when the groundwork for addressing the plurality of English was laid in the 1980s with the discussion of world Englishes (Kachru, 1985; Onysko, 2021). In the literature, 'world Englishes' refers mainly to the institutionalised second language varieties of English used around the world (Wolf & Polzenhagen, 2009). Kachru's (1985; 1992) view of world Englishes is important because it deviates from the native versus non-native speaker dichotomy aligned with inner circle and outer circle varieties. It recognises that there are native speakers of world Englishes and gives more agency and ownership to the speakers of the outer circle varieties. According to Kachru, the term world Englishes (1992, p. 2) "symbolises the functional and formal variations, divergent sociolinguistic contexts, ranges and varieties of English in creativity, and various types of acculturation in parts of the Western and non-Western world."

One such variety out of the multi-cansons of English that has its own variations, divergent sociolinguistic contexts and creativity is Sri Lankan English (SLE). Sri Lanka gained English as one of its languages as a result of British colonisation. The British invaded Sri Lanka in 1796 and succeeded in subjugating the entire country in 1815. Ever since, English has been viewed as a language of power and as a means of upward social mobility (Gargesh, 2019, p.90). After centuries of language contact, English in Sri Lanka has developed its own characteristics and emerged as a variety in its own right. In a very open definition, Gunesequera (2005) defines SLE as an umbrella term that refers to varieties of English that is used by Sri Lankans for whatever purpose in Sri Lanka.

English exists in Sri Lanka in plural forms (Mendis & Rambukwella, 2021; Bernaisch, 2015; Gunesequera, 2005). In fact, it is more apt to refer to English in Sri Lanka as *Englishes* in Sri Lanka as there are several dialects in the island (Gunesequera, 2005). Vaish (2005, p. 190) makes a similar point

about Indian English (IE) when she argues that “The term ‘Indian English’ misleads the readers into thinking that this is one monolithic whole”. There are several dialects such as Standard Sri Lankan English, Non-Standard Sri Lankan English, and Burgher English (Fernando, 2010). These dialects will be discussed in more detail in Chapter 2. Among these dialects, Standard Sri Lankan English (SSLE) is the dialect that is desired, envied and despised by those striving to access English because of its perceived prestige (Amarasuriya, 2010; Parakrama, 2010; Samarakkody & Braine, 2014). This dialect is desired as it is perceived to be the most privileged dialect with the highest social value (Gunsekera, 2005). It is envied because accessing it is related to social factors such as social class and education that work as gatekeepers. It is despised as SSLE itself is a gatekeeping tool that keeps people from accessing job opportunities and social acceptance (Amarasuriya, 2010). It is also the dialect that is promoted in the ESL classroom in Sri Lanka (Samarakkody & Braine, 2014). In addition to these complexities, it is a dialect that is shrouded in a cloud of debate and disagreement because its features are not adequately and empirically described.

There is no doubt that SSLE enjoys a hegemonic position among all varieties of English in Sri Lanka due to the advantage the speakers hold over the speakers of other varieties (Parakrama, 1995; 2012; 2016). This view is challenged by linguists such as Parakrama (2015; p. x) who observes that SSLE has an ‘extra-linguistic value’ and argues that as linguists we need to deviate from traditional paradigms which privilege standard varieties and find more inclusive anti-paradigms which assert an equal status for non-standard varieties. Such views raise important questions about the celebration of diversity and the pedagogy of English in postcolonial settings such as Sri Lanka (Kubota, 2015). They also force one to realistically observe the multiple linguistic codes used in Sri Lanka. While I agree with Parakrama with regard to democratising English and recognising non-elite varieties within Sri Lanka, I am also aware of the unequal position of Sri Lankan English(es) in the Englishes of the world arena. From a prescriptive perspective, all world Englishes are placed rather in an unequal position compared to the more dominant varieties such as BrE or AmE (Tupas & Rubdy, 2015). Therefore, in the present context where all of the varieties of Englishes in Sri Lanka lack an in-depth

pragmatic study, exploring pragmatic features of even the hegemonic variety within Sri Lanka is also a way of nuancing perceptions about Englishes in the world. Such an inquiry would help to show that pragmatic differences, if there are any, are not errors or mere deviations from Kachru's inner circle varieties such as BrE or AmE but systematic occurrences but are norms unique to a variety perceived as the standard in Sri Lanka.

Standard Sri Lankan English was also chosen as the focus of this research because of its role in the field of teaching of English as a second language (ESL) in the country; that too is a position that needs questioning, though again recognising SSLE as a standard in its own right represents progress from holding British or American English as a model for Sri Lankan speakers of English. Given that SSLE is still not adequately described, this research on the PMs is intended to further define SSLE and to help to understand SSLE features. The current study concerns the presence of PMs and their patterns of use in Standard Sri Lankan English (SSLE). The data used in the study is spoken data from a corpus compiled from online semi-formal to formal dyad interviews with seventy-two Sri Lankan English speakers, which produced a total of 2949 PMs (a complete description of the data is provided in Section 4.5). Spoken data was selected because PMs are most prevalent in speech. Allwood (1996) states that PMs are among the top ten word forms, based on analysis of corpora of spoken interaction, and Luke (1987) states that a PM is found in continuous talk every 1.5 seconds on average (both facts mentioned in Fung & Carter, 2007, p. 410). Therefore, it is well documented that PMs play a fundamental role in spoken interaction. Semi-formal to formal discourse settings were selected for observation with the rationale that if the PMs are an integral part of speech, they will appear even at high levels of formality. Perception studies show that PMs are generally regarded as characteristic of loose or weak language (Brinton, 2007; Pettersson-Traba, 2018). Therefore, the appearance of these items in language used in a formal public setting would also be indicative of how much this view is adhered to by the speakers of the more privileged dialect of SLE.

An important aspect of this study is its attention to sociolinguistic factors in the use of PMs, and specifically to gender and age. Much research agrees that gender as a social factor affects how

we use language (Coates, 2016; Leimgruber et al., 2021; Li et al., 2023), including features such as PMs (Erman, 1992; Laserna et al., 2014, Tagliamonte, 2006). By understanding how gender interacts with PMs in SSLE, this study can contribute to the understanding of the complex social construction of gender in Sri Lankan society. Age is another key social factor that determines language choice (Murphy, 2010), and it is associated with language change. Language variation in progress can be understood by observing a linguistic structure in individuals from different age groups at a given point in time. Age is observed from the perspective of chronological age in this study as opposed to psychological² or social age³. To understand how PM use is affected by age, the speakers are divided into three social generations according to the year of birth. These generations are Baby Boomers (1946-1964), Generation X (1965-1980) and Generation Y (1981-1997). The results yielded by this analysis reveal how speakers of different generations maintain interpersonal cohesion and text management by using PMs.

1.3 Rationale of the Study

Research into Sri Lankan English has considered a wide range of different features relating to different areas of linguistic study including morphology, syntax, phonology and semantics. However, there are only a handful of studies on pragmatic features (Degenhardt, 2023, 2024; Funke & Bernaisch, 2022; Kithulgodla & Mendis, 2020; Revis & Bernaisch, 2018). There is research on text structure and movements, nativized pauses, intensifiers and downtoners, but none which looks at a repertoire of PMs as used in the present study. However, these previous studies which found nativized pragmatic features in various aspects of SLE suggest that similar results might be found in other pragmatic aspects of language like PMs as well. There is no empirical record of the diverse types of PMs in SSLE to date, or information about their frequency, variation and the impact of social

² Psychological age is based on the logical abilities and emotional developments of an individual. It is a subjective measure of how old someone feels, acts and behaves and can include cognitive capacity and beliefs about how someone feels. For example, someone who is mature and feels older than their actual age may have a psychological age that is higher than their chronological age (Symons, 2011).

³ Social age is the place gained by an individual at a particular point in time within the society to which s/he belongs (Séguy et al., 2019)

factors. The results obtained can be used to compare SSLE with other varieties of English such as British English, which is the input variety for SLE, and Indian English, which has the largest number of speakers in South Asia, using results obtained in previous research for these varieties. Therefore, my findings add new insights to the understanding of SSLE in an area that has, to date, limited research.

Another reason for the interest in the study of PMs is the fact that they have existed under the radar, so to speak, even within research in the pedagogy of English as a Second Language (ESL) in Sri Lanka. As an ESL practitioner with 25 years of experience in Sri Lanka, United Arab Emirates, United States and the United Kingdom, where I have had the privilege of teaching learners at various levels, I have never come across any lesson material that explicitly taught pragmatic markers. This lack of pedagogical and scholarly attention to PMs in both ESL teaching and SLE research in Sri Lanka is striking. It is valuable to understand the role PMs play in conversations in terms of appropriacy so that learners are guided on what is suitable in natural speech, and on how PMs support interpersonal relations. An ESL learner should have an explicit understanding of how PMs are used by fluent speakers in Sri Lanka especially because PMs are a key feature of fluency and ease of conversation.

1.4 Research Questions

This study addresses four research questions:

1. What pragmatic markers are used by SSLE speakers in semi-formal to formal conversations?
2. What are the functions of each of these PMs in the corpus?
3. How does gender interact with PMs in terms of frequency and function?
4. How does age interact with PMs in terms of frequency and function?

1.5 Significance of the Study

While contributing to the study of pragmatic variation, this study adds to the body of work that identifies the key characteristics of SSLE. The fact that research into SSLE is now addressing specific pragmatic features of the language, rather than just syntactic or lexical aspects, suggests that

it is considered to be on a par with other, more established, varieties of English. Studies based on spoken discourse in SSLE are also in short supply. Therefore, on all these fronts, this study contributes to the body of research on Standard Sri Lankan English.

Research has already established that South Asian Englishes have variety-specific nativized PMs, and that some of the PMs that are common to all varieties of South Asian Englishes have variety-specific functions. Moreover, some of these pragmatic features are gender- and age-sensitive. The results of this study contribute to this body of work, and to research on the intersection between social factors and pragmatic features.

Finally, a key contribution of this study is the creation of a new corpus of spoken Sri Lankan English, the Corpus of Standard Sri Lankan English (CSSLE). Currently, only one corpus of specifically spoken Sri Lankan English exists, and this is the International Corpus of English – Sri Lanka (ICE-SL). CSSLE is more uniform than the data available in the ICE-SL. For instance, metadata on the gender and age of the speakers is unavailable for some of the spoken texts in ICE-SL, and especially lacking in those that are from a semi-formal to formal setting. CSSLE is made up of only semi-formal to formal interviews between two people. Further, the corpus is gender-balanced and the number of words from each interview is controlled. Such uniformity allows for a reliable comparison between data sets for each group of speakers. One could also combine the data in ICE-SL and CSSLE and examine linguistic items in a bigger pool of data. At the completion of the study CSSLE will be hosted in the Digital Humanities Lab of University of Colombo, Sri Lanka (dhlab.cmb.ac.lk) so that it can be used by other researchers working on Sri Lankan English.

1.6 Overview of the Thesis

This chapter has provided a basic definition of PMs. It has also presented Standard Sri Lankan English (SSLE) as a dialect of Sri Lankan English (SLE) and introduced the role of gender and age in relation to PM use. The chapter concludes with a brief description of the study's research questions, its aims, and its significance.

Chapter 2 is a literature review of English in Sri Lanka. It starts with a brief description of how English was established in Sri Lanka. It offers a demographic overview of Sri Lanka to show its ethnic diversity and presents statistical information on English speakers in the country. It goes on to discuss the status of English in Sri Lanka, where English plays many roles, viz. as a first language, second language, a third language and a link language. The identity of SSLE speakers is discussed, emphasising the importance of social class, education and the manner of language acquisition in the formation of identity. Survey-based attitudes towards SSLE are also presented to show how much people are aware of a nativized variety of English in Sri Lanka.

Chapter 3 concerns PMs. PMs indicate the speaker's perspective and structure the conversation (Pichler, 2013). This section discusses the diverse definitions of PMs given in the existing literature and the lack of agreement about the terminology. These definitions nevertheless agree that these are items that sit outside the syntax structure, with no input to the propositional meaning of an utterance, and that they are a requisite to maintain interpersonal relationships. The chapter presents the features of PMs identified in previous research. It also discusses the functions of PMs and perceptions about them. The chapter concludes with an overview of sociological factors affecting PMs.

Chapter 4 describes the design of the study, which uses an specially created corpus to answer the research questions. This chapter offers a description of the data and how the corpus was compiled and the relevant data identified within it. This is followed by a discussion of the variationist theory that influenced the analysis of the data. Finally, the approach adopted to analyse functions of PM, and the approaches to analysing gender and age are presented.

In Chapter 5 I present a synchronic analysis of the functions of the 2949 PMs found in the corpus and their frequencies. The analysis of functions reveals speakers' attitudes, such as friendliness, caution, disapproval, and linguistic behaviours, such as discourse structuring, repair, turn-taking, backtracking, and softening or intensifying assertions. Essentially, it reveals the role of PMs in the social interaction of SSLE speakers. This analysis also investigates whether SSLE speakers

use PMs for the same functions as speakers in other varieties of English based on previous research. More importantly it answers the question of whether there are nativized PMs and nativized functions. The results show the influence of both British English (BrE) and American English (AmE) on SSLE, with some patterns of use similar to AmE and others following BrE.

Chapter 6 provides the answer to the third research question by considering the impact of gender on the use of PMs in SSLE. I review genderlectal studies of inner circle varieties and world Englishes. The literature reveals mixed results for the correlation between gender and PMs. Comparing these results with those of other studies requires that the studies share similar data sets, which are very rare among previous research. However, very clearly, the frequency of use of PMs relates to the different communication styles of women and men. The data analysis indicates that there is meaningful genderlectal variation in this data set, but also that other factors are significant: in addition to gender-based variation, the data reveals that the occupation of the speaker has a bearing on PM use.

Chapter 7 presents another kind of sociolinguistic inquiry as its main objective is to understand the patterns of PM use in different age groups of SSLE speakers, thus providing the answer to my final research question. To date, age variation in PM use has been approached relatively less often in research (Andersen, 2001). In this chapter, the data interpretation considers context-based meaning and is largely dependent on inferential processes in utterance interpretations. The interpretation is based on what is implied in an utterance and any supporting evidence. Previous research shows language is highly variable with regard to age stratification. It is often assumed that each generation reflects the language as it existed when that generation learnt this language, though studies on age grading show this is not always the case. Therefore, a synchronic analysis of language according to age reveals how language has been used in different generations, and, in this case, provides an idea about how PMs have evolved through time. The results show that certain PMs are characteristic of certain generations and others have evolved as prolific PMs over time. This implies that there is competition among PMs for survival and that they

continue to change over time. While some PMs fade away, others take up innovative roles or functions, and new PMs emerge.

Finally, in chapter 8, I present a number of conclusions that emerge from the study. The results reveal that while there are trends that are congruent with BrE, AmE and IE, there are also patterns unique to SSLE. The findings imply that PMs are inclined to evolve, not just in terms of frequencies but in terms of functions, semantic uses and meanings. Their frequencies among generations change over time. Therefore, PMs are indicators of how language changes over time. I also discuss some of the limitations of the study and propose directions for future research. The findings point to the possibility of comparisons with other South Asian varieties of English, and with different modes of communication such as informal conversations.

Chapter 2: Sri Lankan English

2.1 Introduction

Variation in the features of different world Englishes challenges the idea of a monolithic concept of English. Varieties have evolved their own features, expressions and norms, and ownership of these norms is claimed by their speakers. Distinctive Englishes are an integral part of the post-colonial identities of some of the multilingual and multicultural nations (Tupas & Rubdy, 2015, p. 1). As such, Sri Lankan English (SLE) is the most salient legacy of colonial inheritance for Sri Lankans. SLE is an umbrella term that refers to any dialect of English used by Sri Lankans in post-colonial Sri Lanka (Bernaisch, 2015; Gunsekera, 2005). The purpose of this chapter is to summarise existing research on Sri Lankan English, with a particular focus on what is relevant to the current study. To this effect, this chapter begins with a description of the origins of Sri Lankan English and its current status, in Section 2.2 and 2.3. In Section 2.4, in a bid to define SLE, it is positioned in the map of world Englishes employing models that describe world Englishes such as Kachru's concentric circles (1985) and Edgar Schneider's Dynamic Model (2007). Section 2.5 discusses the manner in which SLE is acquired in Sri Lanka, as the manner of acquisition determines the dialect of the speaker. In section 2.6, the identity of the Sri Lankan English speaker is discussed to understand the politics of language use. Then, Section 2.7 presents the variety-specific features of SLE. Finally, I present a description of SLE dialects in Section 2.8 and discuss the reasons for the selection of one particular dialect for analysis in the present study.

2.2 English in Sri Lanka

The native languages of Sri Lanka before the advent of British colonialism were Vedda language, Sinhala⁴, Tamil, Portuguese Creole and Malay. Vedda language is the language of the indigenous people of Sri Lanka who have lived in the island since 6th Century BC (Attanapola & Lund, 2013). Sinhala is the language of the majority ethnic group and Tamil is the language of the second

⁴ The terms Sinhala, Sinhalese and Singhalese are used interchangeably to refer to the language and the ethnicity.

largest ethnic group. Sri Lanka Creole Portuguese is a result of Portuguese colonisation of Sri Lanka in 1505, when the Portuguese were the first European invaders of Sri Lanka. The second European invaders of Sri Lanka were the Dutch in 1658, and the Malays arrived in the island mainly as soldiers during the Dutch and British regimes of Sri Lanka (Nordhoff, 2013). These five languages predate English in Sri Lanka.

The British expelled the Dutch and occupied Sri Lanka in 1796. English became the dominant language in terms of its power over the other languages in Sri Lanka when the British declared full control of the island in 1815. Ever since, English has prevailed under many labels such as the language of the colonizers (Goonetilleke, 2003, p. 338), official language under the Crown Colony (Fernando, 1996, p. 485; Gunesequera, 2005, p. 14), first language (Gunesequera, 2005, p. 23; Kandiah, 1979, p. 86), second language (Fernando, 1985, p. 43), third language and link language (Gunesequera, 2005; Walisundara & Hettiarachchi, 2016, p. 309). These labels are given from the point of view of its speakers. Throughout its development in Sri Lanka, English has never lacked limelight or power in this colonial and postcolonial context, because as a language it has always been the tongue of those who wield power and prestige in Sri Lanka, thriving in the socio-economic and political platforms of the country (Ranwala, 2015). For instance, Fernando (2011) presents evidence from post imperial years on how English dominated public life irrespective of its legislative status. Although language provisions were made to incorporate a person's first language use in public offices such as the police, these provisions were not implemented at ground level. Fernando (*ibid*, p. 486) quotes a reference from 1939 which demonstrates that government offices continued to use English as the main language of administration despite such language reforms and shows that a person who did not know English had to find the services of an interpreter or a translator to get assistance during post imperial Ceylon⁵. The power of English in post-imperial times continued in post-colonial Sri Lanka as well. Gunesequera (2005, p. 34) mentions a comment by a former Minister of Education V. J. M.

⁵ Ceylon was the name of the country until 1972 when it repudiated dominion status and became the republic of Sri Lanka.

Lokubandara, when he addressed the academic community and the private sector in 1993, to the effect that the real gap in Sri Lankan society is based on those who know English and those who don't, rather than on religion, ethnicity, money or caste. These examples show that English has held enormous power in Sri Lankan society since post imperial days.

2.3 The Status of English in Sri Lanka

The status of English in Sri Lanka is complex, as it is in other contexts in many parts of the world (Parakrama, 2016). Therefore, a simplistic explanation of its status would create a misleading image. Its status description on paper and its role in the social context are two entirely different narratives. This section describes how English is acknowledged in legal documents such as the Constitution and explores misconceptions of its status by its users and researchers and the role it plays in the Sri Lankan society.

As mentioned above, Sri Lanka, with English, has 6 languages: Sinhala, Tamil, English, Sri Lankan Malay, Sri Lankan Portuguese Creole, and Vedda language. Census and statistics (2012) given on the language distribution by the Department of Census and Statistics declare that Sinhala is the first language of the Sinhalese, and Tamil respectively of the Tamil community. Tamil is also the first language of a majority of Moors⁶ and the Indian Tamils (2012, Census and Statistics). The Vedda language is the mother tongue of the Vedda population which is reported to be 0.0044% of the total population of Sri Lanka (De Silva & Punchihewa, 2011). 86.9% of the population (age 10 years and over) speak Sinhala and 28.8 % speak Tamil as a mother tongue. The statistics report 23.8% as speakers of English; however, it does not specify if it is spoken as a first or second language. Interestingly, the figures are lower for reading and writing for both Sinhala (79.4%) and Tamil (26.5%) than speaking, in contrast to English (30.5%) which has a higher percentage of users of English who

⁶ The Sri Lankan Muslim community of diverse origins are referred to as Moors in Sri Lankan official documents such as Census and Statistics (2012). Bernaisch (2015:3) refers to all groups of Muslims in Sri Lanka as Sri Lankan Muslims. The groups included in the umbrella term are Sri Lanka/ Ceylon Moors, the Coast/Indian Moors, the Malays, the Memons and the Borahs.

can read and write English than speak English. In terms of multilingualism, 22.5% of the total population speak both Sinhala and English, while both Tamil and English are spoken by 8.4% and all three languages are spoken by 7.5% of the population. This is almost a four-fold increase for English from statistics reported in 1960. The Department of Census and Statistics declared in 1960 that English only was spoken by 0.2% of the population, Sinhala and English by 4.2%, Tamil and English by 2.0% and Sinhala, Tamil and English by 3.2% (Lim, 2013, p. 65). Something which distinguished the 2012 Census from previous records is that there was no column for English-only speakers which indicate that all native L1 speakers of SLE could be bilinguals. Gunesekera (2005) notes that being bilingual or multilingual is the norm for L1 speakers of SLE. There was a column for speakers of English only in the 1946 and 1960s statistics, indicating there were monolingual English speakers in Sri Lanka in that time (Lim, 2013, p. 63).

Although six languages are used, only Sinhala and Tamil are official languages of Sri Lanka (Article 18 of the Constitution of Sri Lanka, 2021). Sinhala and Tamil are also recognised as national languages of Sri Lanka under the Article 19 of the Constitution. While this status quo with regard to national languages raises questions about the language rights of some of the minority languages of Sri Lanka, it also dissembles the actual status of English as a prominent language in Sri Lanka by not naming it alongside Sinhala and Tamil.

Instead, the Constitution (amended in 2021) declares and refers to English as a link language. The legislative definition of English as a link language is rather ambiguous as there is no clear explanation of the linking feature of English in Sri Lanka. However, given the history of the 30-year-old civil war from 1983-2009, which erupted because of ethnicity-related issues between the Sri Lankan government and the Liberation Tigers of Tamil Eelam (LTTE) guerilla organisation, it could be surmised that English is a neutral language for discussion between the conflicted parties. Language has been a bone of contention for the ethnic minorities as Sinhala, the language of the majority, is more widely used for state affairs than Tamil. In that light, English might be considered a neutral language for conflict resolution in theory. However, how practical it is to use English as a neutral

language between the two ethnic groups is debatable for several reasons. For instance, the conflicted parties comprise of the majority ethnic group, the Sinhalese, and the minority Tamils. Fluency in English relates to socio-economic and class-associated factors common for both these groups. English is not an easily accessible language for persons of lower socio-economic class irrespective of ethnicity (Gunesequera, 2005, p. 13). This is explained further when the identity of the SLE speaker is examined in detail in Section 2.6. Against this social backdrop, English can act as a link language only between the English-fluent upper middle-class population of both ethnicities (Fernando, 1996). Given the way English is acquired in Sri Lanka (discussed in Section 2.5), it is hard to imagine that the ground-level social classes who actually experienced the war would use English to communicate with each other. Those who faced the war as an everyday reality were the members of the LTTE guerilla organisation, the soldiers in the Sri Lanka army, and the villagers from both Tamil and Sinhala border villages ('border villages' is a term referring to villages that were bordering the no man's land in the war zone in the North and North Central provinces of Sri Lanka). The border villages are situated far from urban cities in Sri Lanka where people can easily access English. The only English that is used in border villages might just be the 45 minutes of English period per day at school, and this provided the school has an English teacher at all. De Silva et al. (2019) report in a study done with a combination of qualitative and participatory methodologies that education was often disrupted in these areas during the war mainly due to displacement. A case study of a border village in this report states that there were no English teachers in school during the war and even in post-war 2012 (ibid, pp. 35-36). In this context, English is not a lingua franca in these areas. Speakers use rudimentary Sinhala and Tamil to communicate instead. Canagarajah (1999) states that English can act as a link language between government offices of education, commerce and communication and Tamil or any non-Sinhala speaking citizen. Whether this is the reality is highly doubtful given the fact that English is the language of a privileged minority.

While Sinhala and Tamil are stated as national languages of Sri Lanka under Article 19 of the Constitution, Article 22 of the Constitution declares that Sinhala and Tamil shall be the languages of

administration in Sri Lanka. Complicating these statements, Article 22 (2a) declares that a person is entitled to receive communication in Tamil, English or Sinhala. This gives the impression that English is on par with Sinhala and Tamil in some capacity. This ambiguous legislative status misrepresents the social reality of English especially given the presence and the power English holds over other languages of Sri Lanka in terms of upward social mobility and employability in the job market (Ariyawansa, 2013; Ranwala, 2015). However, the ambiguous status of English in the Constitution leads many to believe that English has an official capacity. An example is Mukherjee's (2012) assertion that English is an official language in Sri Lanka. This is addressed by Mendis and Rambukwella (2020) who state that this confusion may arise from the prominent presence that English has in Sri Lanka despite its ambiguous link status. To quote Mendis and Rambukwella:

English is still pervasive in many areas of officialdom; it is the language used in Sri Lanka's Supreme Court; it has a strong presence in the media and in advertising; it is making a comeback in the country's education system and it is the undisputed language of choice in the private business and commercial sectors. In other words, its hegemonic grip on the country is still very evident. (Mendis & Rambukwell, 2020, p. 179)

The 2021 amendment of the Constitution says that the languages of the Courts are Sinhala and Tamil but the Minister of Justice may issue directions permitting the use of English in or in relation to the records and proceedings. In other words, English is very much permitted in the courts despite the fact that only Sinhala and Tamil are the declared languages of the Court. The language of legislation is both Sinhala and Tamil together with a translation in English.

SLE is often referred to as a second language in Sri Lanka. However, it should be noted that no official document gives English the status of a second language. Fernando (1996) states that in 1946, English ceased to be the medium of instruction in schools and it was taught in schools as a second language from grade 3 upwards. English can be considered a de facto second language given the fact that it is a mandatory subject in all government schools (De Silva, 1996 cited in Abeywickrama, 2007). It continues to be taught in institutes of tertiary education by departments of

English Language Teaching (DELTA), and a pass in English is compulsory to complete any undergraduate degree course.

In a bid to democratise English use, the Government in 2009 initiated a campaign that aimed at attitudinal change towards the ownership of English. The initiative was termed 'Speak English Our Way' evoking patriotism and conscious change of attitude towards SLE. The slang symbol for English is *kaduwa*, 'sword' in Sinhala, and on the one hand this is a weapon that can be used to destroy those who are not armed with it. On the other hand, if the user is not skilled in using the *kaduwa* or sword it can harm them. This symbol was replaced with a *manne*, a machete-like knife that is a useful tool in any rural household. The message was not to use English as a weapon of oppression but a tool or a skill that is useful for upward mobility. This movement was helpful for discourse in the ESL classroom about the politics of English and established a specifically Sri Lankan variety of English as part of the Sri Lankan identity. However, after a short period of about one and a half years, the programme ceased. This movement was important in the sense that, had it continued, it would encourage people to speak English without any inhibitions especially with regard to pronunciation and grammar. One inhibition to speaking English in Sri Lanka is the fear of incorrect pronunciation. The manner of articulation defines the difference between prestigious dialect of SLE and others (Gunsekera, 2005). More information on dialects is provided in Section 2.8.

The President of Sri Lanka referred to *kaduwa* recently in 2024 in a speech made at a state university⁷. He urged the undergraduates to take the *kaduwa* in their hand without any fear. In sum, this reference to *kaduwa* and his advice is a great example to show that the perceptions and the oppressive fear about English in Sri Lanka is still unchanged and its status among the hierarchy of local languages is still privileged.

⁷ Speech made in Sinhala on 20 July 2024 at the opening of the Faculty of Allied Health Sciences, University of Ruhuna, Sri Lanka. <https://youtu.be/ApUmQ8AKVe4?si=KlxBlOkLBarECKiX> (refer 2.29 min)

2.4 Definitions of Sri Lankan English

A very simplistic definition of Sri Lankan English would be any variety of English spoken in Sri Lanka by Sri Lankans as a first language, a second language or as a third language. First language speakers of SLE are considered native speakers (NS) and are explored in Section 2.6. Gunesekera (2005), who was an eminent authority on SLE, states that “Sri Lankan English is the language used by Sri Lankans who choose to use English for whatever purpose in Sri Lanka” (p.11). She further characterises SLE as “Sri Lankan English, with its borrowings and influences from Sinhalese and Tamil, as used in Sri Lanka” (p.20). Bernaisch (2015) states that Sri Lankan English is a term used most frequently to refer neutrally to any Sri Lankan variety of English. It is an umbrella term for Englishes used in Sri Lanka which includes a standard dialect, a non-standard dialect and other dialects such as Bugher English and Tamil English. These dialects are discussed in Section 2.8 of this chapter.

We can define Sri Lankan English in terms of the models of world Englishes as well. Various models and approaches have been proposed to explain the evolution of world Englishes and to describe the status of varieties of English. Most of the models are based on the method of English acquisition by respective speakers (Kachru, 1985). Therefore, the varieties take names such as English as a Native Language (ENL), English as a Second Language (ESL) and English as a Foreign Language (EFL). Kachru’s Three Circle model (1985), the most widely discussed model thus far, classifies Englishes according to the role they play in the relevant societies. He uses three broad categories: a norm-providing inner circle, a norm-developing outer circle, and a norm-dependent expanding circle. In this model, SLE can be categorised as a norm-developing outer circle variety. One feature of norm-developing varieties is that English is only one code in the linguistic repertoire of bilinguals or multilinguals in a country as in the situation of Sri Lanka. English is generally an additional language to a vernacular mother tongue. Alternatively, English can be simultaneously acquired with another vernacular language in Sri Lanka leading to simultaneous bilingualism. Another condition of outer circle varieties is that English has acquired an important status in the language policies. This is the case with Sri Lanka as well. English is a major language of government, education and the legal

system. Further, Sri Lanka falls within the norm-developing outer circle variety because there is a divergence between the way Sri Lankan speakers use English and the way it is used in the input variety, viz. British English. There are phonological, morphological, syntactical, semantic and pragmatic features in Sri Lankan English (SLE) that are different to British English (BrE). For example, a phonological deviation from BrE is the absence of aspiration of word initial voiceless plosives such as /p/ and /k/ in SLE (Fernando, 1985). Another example is the word stock that has been added to SLE to explain Sri Lanka experiences such as *pirith chanting* (chanting Buddhist scriptures with special prayers), *wedding house* (the house where a wedding is held), *funeral house* (a house where a funeral is held), *podering* (drizzling) and *rice pullers* (side dishes or accompaniments to a rich curry meal) (Gunesekera, 2005). The variety-specific features of SLE are discussed in Section 2.7 and give ample evidence to show that Sri Lanka fits within the norm-developing outer circle variety in Kachru's Concentric Circles.

Schneider's Dynamic Model (2007, 2014) too can be used to explain the variety status of SLE. The Dynamic Model is based on the idea that a traceable unvaried process underlies the origin of post-colonial Englishes. The process has five consecutive stages: foundation, exonormative stabilization, nativization, endonormative stabilization, and differentiation. Schneider has revised the model since its original inception and a brief condensed description of the revised Dynamic model (2014) is given below.

Foundation is the initial stage in which English is established in a country through colonization, and with time some members of the indigenous community use a limited amount of English. In this stage the colonisers adjust their pronunciation and lexical usage to facilitate the understanding of the locals. This leads to the beginning of a complex language contact situation, which is evident in the use of toponymic borrowing of typically local place names into English discourse among other exchange of language usage.

Phase two is exonormative stabilization, where the settlers stabilize politically and English is regularly spoken by the indigenous population. The standard variety is the variety of the colonial

masters while a variety of English as a second language emerges. The settlers as well as the indigenous population begin to gradually adopt local vocabulary to describe mainly fauna and flora of the territory. Early phonological and syntactic transfer phenomena are increasingly found.

The third phase, which is the central and the most vibrant phase, is nativization. This involves both cultural and linguistic transformation. With the country gaining independence, the social gap between the colonizers and the colonized is narrowing, paving more opportunities for interaction between these parties. The result is the increasing emergence of structures distinctive to the newly evolving variety on the levels of lexis, phonology and grammar via second language acquisition, L1 transfer and innovation. A heavy load of lexical borrowing from the native languages takes place markedly restructuring the English language. The settlers' descendants adopt these local features and indigenized people will start showing a distinguishable local accent and structures of the newly evolving variety.

The fourth phase, which is called endonormative stabilization, is associated with nation building, and is usually followed by total political independence by the settler's mother country giving way to the birth of a new nation. This phase is marked by cultural and self-independence and a conscious effort for identity separation from the former 'mother country'. Linguistic and cultural homogeneity is emphasized and the codification of the new variety of English is initiated.

The fifth phase is differentiation. By this time the emergence of a new variety is a thing of the past, the new variety is already established and this is the stage of dialect and sociolect birth. According to Schneider (2007) this is also the phase when the citizens of the independent country no longer define themselves as a 'single social entity in relation to the former colonial power but rather as a composite of subgroups, each being marked by an identity of its own' (Schneider, 2007 p. 53). The subgroups are defined by gender, age, ethnicity, geographic location, social group or stratum. Schneider further states that in this phase, the selection of specific language forms marks group membership. This phenomenon leads to increased variation within the formally new variety.

Mukherjee (2008, p. 361) applies this model to SLE, and he assumes SLE is on its way to endonormative stabilisation based on criteria relating to history and politics, language contact, language use, the sociolinguistic attitudes of the speakers and codification. An attitudinal survey of 122 informants from academic backgrounds conducted by Künstler et al. (2009) shows that Sri Lankan English speakers are to a certain extent aware of the dialect that they speak. All the informants in this study were university lecturers, undergraduates and secondary school teachers, and data was collected using a survey in 2007 and 2008. In this survey, 60% of the informants were aware of the existence of SLE but only 40% admitted that they accepted SLE as their target norm. Half of the informants preferred Received Pronunciation (RP) as a teaching target in schools, indicating a preference for BrE over SLE. The findings of this study support Mukherjee's (2008, p. 361) assumption with regard to SLE because the criteria of acceptance of local norms are only partially met. Bernaisch (2012) endorses the same view in his survey which is almost identical to Künstler et al. (2009) in terms of informant profile and results. Both studies include participants from an academic background within the age range of 21 to 76. In Bernaisch's (2012) survey, 65.6% out of a total of 169 informants were students, academic staff at university or teachers from secondary schools. Similar to Künstler et al. (2009), the results conclude that SLE is an endonormatively stabilized variety of English as a majority of SLE speakers rated BrE as the most valued variety of English. The study also reported that the informants were aware of SLE and had a positive attitude towards it.

However, this placing of SLE at phase 4 can be challenged, as some characteristics of the fifth phase are also evident and are confirmed by studies. Stronger evidence that SLE is in the fifth phase of the Dynamic Model is the evidence of dialectal birth⁸. Gunsekera (2005) states that internal variation or dialectal birth could be ignored or rejected by its speakers due to unawareness. The awareness and evidence of dialectal birth is largely limited to the few studies and the little

⁸ The emergence of distinctive, new language variety following the migration of people speaking mutually intelligible dialects to what, to all intents and purposes, in linguistically 'virgin' territory. New dialects are also formed by a process of split and subsequent dialect divergence. (Kerswill & Trudgill, 2009, p. 196)

community of researchers on SLE. This knowledge is not publicly accessible as the research is published only in academic journals. For instance, Fernando (1985), Gunesequera (2005), Mayler (2007), Bernaisch (2015), Parakrama (2016) and Rajapakse (2024) discuss the dialects within SLE in their respective work. It can be argued that SLE is in phase 5 and has completed its emergence as new variety. Dialectal variation is discussed further in Section 2.8.

In sum, SLE is the variety of English spoken in Sri Lanka. It is a post-colonial variety of English used as a native language, a second language and a third language. There are emerging dialects in SLE. Given the linguistic criteria considered for classifying varieties, SLE is a well-established variety of English in the Outer circle. Therefore, strengthening its research on its pragmatics to be on par with other linguistic aspects is appropriate at this stage.

2.5 The Manner of SLE Acquisition

SLE is acquired as a first language (L1) (Fernando, 1985; Gunesequera, 2005; Jansz, 2019), as a second language (L2) (Fernando, 1985) or as a third language. It is a first language and native language (NL) for those whose home language is English, with either Sinhala or Tamil as a second language (Fernando, 1985; Gunesequera, 2005). The variety of English that is spoken in Sri Lanka as a first language is considered Standard Sri Lankan English (SSLE). Its phonological features are better documented than other linguistic aspects. These features are discussed in Section 2.8. Most L1 speakers of SLE are simultaneous bilinguals (Gunesequera, 2005; Meyler, 2007). They have access to English before they are formally introduced to the language in a classroom setting.

A majority of speakers learn English as a second language. Second language learners of SLE learn it at school at primary, secondary or tertiary level. Those who have mastered Sri Lankan English as a second language are associated with access to higher levels of education. Moreover, Siromi Fernando (1985) groups L2 speakers of SLE who show a native-like fluency on par with L1 speakers as opposed to those who are not very fluent. This indicates that native speakers of SLE and L2 speakers of SLE with a high proficiency are considered Standard Sri Lankan English speakers.

Some others learn SLE as a third language mainly at school while using two languages as home languages. For example, you can be a Muslim, whose first language is Tamil, medium of education is Sinhala, learning English as a third language. Speakers who do not acquire English fluently are considered Non-Standard Sri Lankan English (NSSLE) speakers. This variety is also marked mainly by its phonological features as explained in Section 2.8.2.

2.6 Standard Sri Lankan English Speaker Identity

Language is part of a speaker's social identity. The definition of the identity of a SSLE speaker is riddled with contentions (Mendis & Rambukwella, 2020). Descriptions presented by researchers are variously based on the method by which SLE is acquired, the level of proficiency, the class and social background of the speaker, and the education level of the speaker. Speaker identity is linked to who has access to acquire English in Sri Lanka. In 1948, when Sri Lankan English was first documented, attempts were made to describe the Sri Lankan English speaker in terms of education level and proficiency. For instance, Passé (1948) described the Standard Sri Lankan English speaker as an educated Ceylonese⁹ who knew English and "for the most part they know it well" (1979, p.16). There is no literature to suggest that dialectal variation within SLE was recognised by 1948 and therefore it can be assumed that Passé's reference to SLE is the variety spoken by fluent and educated speakers which in today's terminology might be considered Standard Sri Lankan English (SSLE), though presumably some of its features have changed. Manique Guneseckara (2005) provides a class-based definition: the dialect of Standard Sri Lankan English is the "English used by the Sri Lankan elite" (p.24). A standard English speaker in Sri Lanka is also referred to as a habitual English speaker. Herat (2005) refers in a footnote to the habitual speaker of Sri Lankan English:

"By habitual users, I mean, people who use English a lot, but for whom it is neither a first nor native language. These people frequently use English in different domains but have another language as their "best" and/or "native" language". (p. 206)

⁹ Inhabitant of Ceylon

In Herat's analysis, a habitual speaker is anyone who uses English frequently irrespective of the level of proficiency. Ratwatte (2012) too states that speakers can be divided into two groups as habitual and non-habitual speakers (as stated in Jansz, 2019). However, Ratwatte's description of the habitual speaker is very different to Herat's, and in fact the two can be considered contradictory. Similar to Gunsekera (2005), Ratwatte (2012) classifies Standard Sri Lankan English (SSLE) speakers as the habitual speakers or anglophile elite as opposed to the non-habitual speakers who have a knowledge of English but do not speak it habitually. Ratwatte (ibid) does not refer to class directly like Gunsekera (ibid), yet when she mentions the anglophile elite invariably it refers to the middle or upper middle class of the Sri Lankan society. By using the phrase anglophile elite, she implies a high level of proficiency in English in her habitual speakers.

All these definitions indicate that Sri Lankan English speaker identity is closely linked to the manner the language is acquired, as well as class. Chitra Fernando (1976) refers to native speakers of SLE in her article on English and Sinhala bilingualism in Sri Lanka. She defines native speakers as highly educated people with a very anglicized lifestyle who live in the top or middle strata of social class. Commenting on L1 users of SLE, Kandiah (1979) states that "to begin with, Lankan English is by no means foreign or second language, in any real sense of these terms, to a considerable number of its users who determine its distinctive nature." Similarly, Gunsekera's (2005) definition of the native speaker too is a westernized bilingual upper-class individual who is a bi- or trilingual. The native speaker for Gunsekera is:

the westernized elite, which uses English as a home language. In this context, a native speaker of Sri Lankan English is either bilingual or trilingual.....the native speaker of Sri Lankan English is fluent in English and in either Sinhalese or Tamil. The level of fluency in the other language may vary, with English being the dominant language. (Gunsekera, 2005, P.23)

Social class, education and the manner of language acquisition therefore play key roles in determining the SLE speaker identity, i.e. whether you are a first language speaker, second language

speaker or a third language speaker. Additionally, the manner of acquisition also determines if you are considered a standard speaker of Sri Lankan English (SSLE) or a Non-Standard Sri Lankan English speaker (NSSLE). Exceptionally, Sri Lankan Burghers, whether educated or not, are considered as first language speakers of English (Rajapakse, 2018; Jansz, 2019). It is generally accepted that Burghers use English as a home language as they are the descendants of British, Portuguese or Dutch. Usually, their second language is Sinhala. In one of his novels, Carl Muller, a Sri Lankan Burgher fiction writer, presents the socio-political reality for Burghers in the 60s when Sri Lanka adopted a Sinhala Only language policy making Sinhala the medium of instruction in schools (Muller, 1994). This displaced the Burghers who were not fluent in Sinhala and many migrated to western countries such as the UK and Australia as they felt their language rights were being violated. All in all, due to the privilege associated in being a first language speaker of SLE or a highly proficient L2 speaker of SLE, the SSLE speakers are a minority (Ekanayake, 2020). The current study focuses on the standard SLE speaker. The reasons for this selection are elaborated further in chapter 4.

2.7 Features of Sri Lankan English

English has been in contact with the vernacular languages of Sri Lanka for over 200 years and this language contact has led to the distinct characteristics of SLE (Passé 1948; Guensekera, 2005). SLE has plenty of evidence of influence and borrowing from Sinhala, Tamil, Malay, Portuguese, Dutch and various other languages the island had come into contact with (Gunesekera, 2005). Comprehensive studies on the phonology, morphology, structure and on its description have established Sri Lankan English features and SLE as a variety in its own right. These features set SLE apart from its input variety, British English. Kandiah (1979) suggests that Sri Lankan English should not be looked at from the point of view of its former masters but of its actual users, who give SLE its identity and the unique features. In his study on attitudes to SLE, Bernaisch (2012) gives a brief list of these unique features at various structural levels such as phonetics and phonology, morphology, lexicogrammar, and syntax. An overview of research on Sri Lankan English is presented below. This gives a brief overview of SLE features of phonology, morphology, semantics and syntax, but presents

a more complete survey of research pertaining to pragmatics in Section 2.7.5 as it is more relevant to the current study.

2.7.1 Phonology

Phonology is an area where SLE is very clearly distinct from its input variety, and there is substantial research into its unique phonological features. The phonology of SLE is mainly a result of language contact between the vernacular languages of Sri Lanka, Sinhala and Tamil (Meyler, 2007). Phonological features play a crucial role in distinguishing an SSLE speaker from an NSSLE speaker. In this section, I concentrate on the phonological features of Standard Sri Lankan English (SSLE) as it is the dialect that I focus on in my study. Some of the phonology-based research attempts to distinguish SSLE features by comparing it with features of BrE. For example, in early research, Fernando (1985) contrasts SSLE features with BrE. In later research, Meyler (2007) also describes SSLE in relation to BrE in his introduction to the dictionary of Sri Lankan English. Mendis and Rambukwella (2021) note that the similarity between Fernando's (1985) and Meyler's (2007) observations of phonological features over two decades with regard to phonology signals stability in the development of SLE phonological features.

An important feature of SSLE is the absence of aspiration of word initial voiceless plosives such as /p/ and /k/ where BrE word initial voiceless plosives are aspirated. For example, SSLE pronunciation for *pot* as /pɔ̃t/ and BrE pronunciation is /p^hɔt/. Retroflex sounds are used in SSLE where BrE would use alveolar plosives, such as /t/ and /d/. SSLE devoices /z/ in word initial, word final and intervocalic positions. For example, an SSLE speaker would pronounce *zoo* as /su:/ and not as /zu:/ in BrE. Another example is that SSLE uses the dental plosives /t̪/ and /d̪/ in place of BrE /θ/ and /ð/. Both Fernando (1985) and Ekanayake (2020) note the use of /o:/ in place of /ɔ:/ as a feature of SSLE. For example, *door* is pronounced as /do:/ in SSLE whereas it is pronounced as /dɔ:/ in BrE. SSLE uses the diphthongs /ea/ where BrE uses /eə/ as in *there*. Given the distinctiveness of phonological features of SSLE, they were a particularly significant criterion for identifying speakers of

SSLE to be included in this study. A more detailed description is therefore given in the section below on SSLE in Section 2.8.1.

2.7.2 Morphology and semantics

There is substantial evidence from empirical research on the word stock that is unique to SLE. SLE is broadly described and understood as a variety in its own right on the strength of its vocabulary. In fact, Fernando (2012) suggests that the SLE vocabulary and other language aspects was produced under pressure to assert Sri Lankan identity during pre- and post- independence Sri Lanka. Gunsekera (2005), Fernando (2012), Fernando (2015), Senaratne (2009) among many others comment on the morphological features of SLE. Borrowing, compounding and affixation are the most productive morphological processes in SSLE (Gunsekera, 2005, p.143). In her analyses of word formation processes in SLE, Gunsekera (ibid) provides a list of 576 vocabulary items that are unique to SSLE. Ranaweera (2007) analyses a corpus of 1.6 million words of newspaper language and presents 68 lexical items unique to SLE. Among them are vocabulary connected to kinship terms, fauna and flora, traditions, festivals and political satire unique to the Sri Lankan context. Some examples are *anusasana* (Buddhist sermon), *pirivena* (a place of formal learning for Buddhist monks), *nilame* (a historical chief of the central part of Sri Lanka), *perehara* (procession) and *machan* (buddy or close friend).

A study with a different perspective is Fernando's (2015) analysis of SLE borrowings (SLEBs) in learner language. Generally, teachers resist SLEBs in formal learning, especially in writing, because they regard borrowings as a deviation from the 'standard' (Medawattegedera and Devendra, 2006). However, borrowing or the adoption of loan words are characteristic of varieties that exist in a multilingual context. Fernando (2015) takes the theoretical standpoint that "appropriate use of SLEBs in the local context is an indication of the sociocultural competence of the speaker" and "the avoidance of SLE borrowings can demonstrate a lack of sociocultural competence in a user" (p. 44). This indicates that a Standard Sri Lankan English speaker would use SLEBs appropriately. For example, borrowings are used when equivalent terms are lacking in English to express a local experience. The

data reveal direct borrowings, indirect borrowings, and hybrid compounds. Most of the borrowings in this learner data belong to the category of food and local experiences such as *kevum*, *kokis*, and *aluva* (all traditional sweets prepared during Sinhala and Hindu new year in Sri Lanka) which is similar to the findings in formal writing in SSLE. However, some of the borrowings deviated from the SSLE borrowing. For example, during the Sinhala and Tamil new year in Sri Lanka there are many village-level games contests and one essential item is the beauty queen contest. The beauty queen title in SSLE terms is *Avurudu Kumari* (Queen of the new year). In Fernando's (ibid) study, the NSSLE speakers use the literal English translation *New Year Princess* instead of the SSLE borrowing. This indicates that NSSLE speakers lack the linguistic knowledge to use borrowings appropriately.

A crucial contribution to the morphology of SLE was the codification of lexical items in Meyler's (2005) publication of a dictionary of SLE. Although there were several studies which discussed morphological processes and record new SLE vocabulary prior to Meyler's publication, this was the first formal dictionary with an impressive entry count of over 2500 words. It records many noun compounds that are recognisably SLE forms. Some of these are made up of two English words: for example, *marriage broker* refers to a professional match maker. Others are formed from one English word combined with a word from a vernacular language: for example, a *made-up Kandyan* is a type of saree called Kandyan which is cut and tailored to make it easier to wear. In terms of morphology, there is a difference between spoken and written SLE. The dictionary includes code mixing terms, informal usages and colloquial SLE terms which are all found more often in speech. Clearly, SLE borrowings are chosen by speakers or writers depending on context, with attention to function and the level of formality. Illangakoon et al. (2021) analyse the SSLE borrowings in ICE-SL. They first use a binary classification found in the literature which classifies the borrowings into core borrowings and cultural borrowings. However, they find that this classification is inadequate. A more nuanced classification of 4 categories accommodates the SSLE borrowings. The first category is the

local-cultural borrowings, which refer to concepts and items alien to BrE such as *veddas*¹⁰ (a borrowing from Sinhala). The second is regional-cultural borrowings which refer to items alien to BrE but found in South Asian languages such as *kapha* meaning ‘phlegm’ (borrowed from Sanskrit). The third is core borrowings which denote the exact meaning available in BrE like *amma* (‘mother’). The fourth category is core-extension borrowings which refer to an additional meaning to the translated word in BrE such as *maha* in *maha sangha* (referring to Buddhist monks). All these studies demonstrate that SSLE morphology is dynamic and ever-growing to accommodate cultural experiences.

In regard to semantics, there are words that have gained new meanings or interpretations in SLE which are not found in BrE, its source language. For instance, the phrasal verb *come down* in SLE refers to failing a subject at an exam (Meyler, 2007). *A basket woman* refers to a loud-mouthed woman (Kandiah, 1981; Gunesequera, 2005). The words look and sound BrE, but they have a meaning that is understood only within SLE.

2.7.3 Syntax

A few studies report on distinctive syntax features of SLE. Mendis and Rambukwella (2020) assume that these features could be found in other South Asian Englishes as well. For example, topicalization is viewed as a feature of SSLE (Ekanayake, 2020, p.344) and it is also a feature of Indian English (IE) as investigated by Lange (2012, p.122). Kandiah (1981) categorizes ellipsis, focalization and topicalization as characteristic of SLE speech. An example of topicalization is “For one hour, I was waiting” (Ekanayake, 2020, p. 344). Non-specification of certain redundant features as in the example “Are they teachers? I thought students and gave them forms (I thought they were students ...)” (ibid, p. 344) is another characteristic. The words *they were* are omitted because they are considered redundant. Rajapakse’s (2008) study on Burgher L1 speakers of SLE confirms Kandiah’s observations with regard to syntactic structures.

¹⁰ The indigenous people of Sri Lanka.

Another key feature of syntax unique to spoken SLE is the use of the question tag *no* (Gunasekera, 2005). Ekanayaka (2020, p. 343) lists the *no* at the end of a sentence as an operator that affirms or asserts, as in “She couldn’t have been in Colombo. She was here *no*”. This is a nativised feature that has been influenced by Sinhala, since Sinhala has a PM that functions as a tag question and functions as an emphasis marker. Chandralal (2010, p. 231) mentions in a footnote with regard to Sinhala that “*ne* is an interactional particle used in the clause final position to elicit addressee’s consent or attention or to add emphasis to a statement”. The SSLE *no* that is used for the same two purposes of tag question and emphasis appears to be an example of language transfer from Sinhala. *No* is a marker that straddles syntax and pragmatics. It features syntax when it operates as question tag and it is a PM when it is used for emphasis. Therefore, *no* is a pragmatic feature that is investigated in my study.

Another such syntax feature in speech is the use of *here* to draw attention of a speaker as in *Here, where is that shop? Here* in this instance shows the influence of Sinhala in which the word *me* [mee] is used to get attention. *Me* is equivalent to *here*, and is similar to *oi* in BrE. The non-interrogative tag element *will you* expressing coercion is also a feature of SSLE speech. An example is “Go and collect the parcels, will you, without complaining” (Ekanayaka, 2020, p. 344). The use of phrasal verbs too is different from use in BrE. For instance, Meyler (2007, p. 544) notes *bear up* as an SLE use as in the example, “She couldn’t bear up the pain”. *Cope up with* and *vote at* are also phrasal verbs unique to SLE (Ranaweera, 2007, p. 70).

It is commonly believed that areas such as phonology and morphology are the areas that display SLE specific features and that syntax is close to that of BrE. However, a close examination shows that SLE has distinctive syntax and writing style. More research in analysing the features of written SLE is needed to give a more complete picture.

2.7.4 Prosodic Features of SLE

The prosodic features of SLE are also little-studied. Prosodic features define the identity of a language sometimes even more prominently than so-called essential features such as morphology

and phonology. Sometimes, simply by listening to the rhythm of a variety of English, one is able to place the geographical location of the variety. Therefore, ideally, we need more research on SLE prosodic features to validate its status as a distinctive variety.

The little research that exists declares that the stress, rhythm and intonation of Sri Lankan English is different from other varieties of English in the world. For example, in yes/ no questions like ‘You have a class now?’, the pitch of the voice rises to its highest at the beginning of the last syllable. Fernando (1985) notes a flatter intonation curves in SLE than in BrE. Fernando (1985) notes that all stress in SLE can be viewed as weak. Passé (1948) attributes stress on second syllable of a word for L1 influence from Sinhala. Meyler (2007) notes that primary stress is placed on the first syllable of the word in SLE, whereas stress would typically be placed on the second syllable in BrE. Given that six decades passed between these two studies, a possible explanation for this contradictory stress pattern is that perhaps during Passé’s era, SSLE resembled BrE more closely than in 2007. By Meyler’s time, SSLE is more influenced by the vernacular languages of the country than BrE. Once again, it is evident that L1 influence plays a major role in forming the identity of SLE.

2.7.5 Pragmatics

Compared to the other areas of language research, research on the pragmatics of SSLE is scarce. The available research investigates the structural features of text types or genres and frequencies of forms that express pragmatic functions such as apologies and requests. For example, Mendis (2006) researches SMS language, Herat (2014) observes the structure and language that display cultural and religious elements in obituary notices in Sri Lankan English newspapers, and Kitulgoda & Mendis (2020) examine the rhetorical structure of welcome addresses (WA) in Sri Lankan English. It is customary in Sri Lanka to have a formal address at formal functions to welcome guests. These addresses follow a general pattern, and the study reveals a pattern of 3 obligatory moves and 1 optional move in 20 WAs.

With the completion of several corpora in the International Corpus of English, some research on the pragmatic aspects of world Englishes has been produced to complement the body of research

on structural features. There is new research on linguistic aspects such as pauses, backchannelling, requests and apologies that indicate variety-specific usage patterns and language variation sensitive to social factors such as gender and age. ICE Sri Lanka (ICE-SL) (Bernaisch et al., 2019) has been used to explore many of these features in SLE as part of comparative studies. For instance, one study of nativisation of pauses in Asian Englishes in the ICE corpora includes a description of filled and unfilled pauses in SSLE, IE and BrE (Revis & Bernaisch, 2020). The study concludes that while the pauses are systematic, they are pragmatically nativised in SSLE. Kraaz and Bernaisch (2020) examine backchanneling in BrE, IE and SLE in the ICE corpora. They find indications of pragmatic nativisation of backchannels in Indian and Sri Lankan English. The study reports that SLE speakers use backchanneling more than the speakers of BrE and IE. SLE and IE use lexical echo backchannels exclusively, where the last nominal head used by the preceding interlocutor is replicated and are therefore regarded as nativized features of SLE and IE. Degenhardt (2022) investigates apologies in BrE, Indian English (IE) and SLE in the ICE corpora. She finds that BrE speakers employ a larger set of apology forms than IE and SLE speakers. She also looks at gender and age variation and finds that irrespective of age, male speakers do not apologise as much as women in all three varieties. Degenhardt (2023) examines requests in SLE, IE and BrE in the ICE corpora and reveals differences between South Asian Englishes and BrE. Among many other findings the study concludes that speakers with a South Asian background value positive face wants more than speakers of BrE with more Western backgrounds. All this research indicates that there are variety specific pragmatic features in SLE, and such pragmatic features are also sensitive to sociolinguistic factors such as age and gender of the speaker. The findings of these studies reinforce the importance of including these factors in my study as relevant for pragmatic features.

Revis and Bernaisch (2020) among others imply that quantitative research into variation in the use of pragmatic markers is uncommon especially in postcolonial varieties of English. There is no research specifically into PMs, and my study addresses this gap. However, there is one existing study on Sinhala discourse markers. As Sinhala is likely to influence PM use in SLE, especially with regard to

nativised PMs, this study is also pertinent to this literature review. Perera & Strauss (2015) investigates Sinhala demonstrative *me* ('this') and temporal adverb *dæn* ('now') in their functions as discourse markers. The study reports that these discourse markers are used for repairs, hesitation and contrast by their users.

2.8 Primary Dialects of Sri Lankan English

Although there are features common to different SLE varieties, there are also dialect-specific features. The two primary dialects are Standard Sri Lankan English (SSLE) and Non-Standard Sri Lankan English (NSSLE) (Fernando, 1985; Gunsekera, 2005). There are also dialects based on ethnicity, Burgher English and Tamil English (Rajapakse, 2008; Gunsekera, 2005; Meyler, 2007). Meyler (2007, p. x) acknowledges that there is a variety spoken by the elite in Colombo (the commercial capital city of Sri Lanka) which is different to that of the wider community who are more likely to speak English as an L2. Parakrama (1995, 2016) discusses the dilemma of English usage of the non-elite users of English. Meyler (2007) and Parakrama (1995, 2016) refer to the SSLE and NSSLE speakers respectively when they mention elite users and non-elite users. Gunsekera (2005) mentions Tamil English in addition to SSLE, NSSLE and Burgher English, noting that "While Tamil English shows the influence of Tamil on English, Burgher English is largely a mixture of some elements of Portuguese Creole and English" (p. 37). However, there is no literature on the features of Tamil English.

Meyler (ibid.) comments that there is a difference between the language use of older and younger generations. Perera & Weerasooriya (2016) confirm Meyler's observation in their study of generational change observable in SLE vocabulary. The study finds that there are generation-specific words. Yet, there is no large-scale study that has presented differences in a range of linguistic aspects to show variation between generations in SLE.

2.8.1 Standard Sri Lankan English

As discussed earlier in this chapter, Standard Sri Lankan English (SSLE) can be regarded as a dialect of Sri Lankan English (SLE) that is spoken by the native speakers of SLE (Mendis and Rambukwella, 2020) and by very fluent L2 learners of SLE (Fernando, 1985). SSLE, similar to any norm referencing variety, is the most prestigious dialect of SLE. It is perceived as a class marker and a gatekeeping dialect (Parakrama, 1995, 2012, 2016). Thus, Quirk et al.'s (1985, p. 197) definition of the standard of language can be aptly extended to SSLE speakers in Sri Lanka as well; "The standard language is inevitably the prerogative of a rather special minority."

SSLE has a few prominent phonological features that immediately classify the speaker as an SSLE speaker. One of them is the ability to use [o] and [ɔ] differently. [o] and [ɔ] are not interchangeable phonemes in SSLE. It is perceived by SSLE speakers that this single phonological feature is a prime characteristic of BrE and an attestation of speaking similarly to BrE. Using data collected through questionnaire and interviews, Gunsekera (2005) shows that most speakers of Standard Sri Lankan English claimed to speak BrE without hesitation; "it is only now, in the 21st century, that at least some users of English are prepared to say that they speak or use Sri Lankan English" (Gunsekera, 2005, p.11). Even in 2005, acknowledging that there is a variety of English specific to Sri Lanka was rare. This hesitancy to accept the actual variety of English that the Sri Lankans speak as opposed to the preferred target language BrE, links with what Kandiah (1981) explains as 'schizoglossia' of Sri Lankan English. Schizoglossia refers to the breakdown between emotion and reality in terms of speech production. It also shows linguistic insecurities as a result of the post-colonial mentality in SSLE speakers. Attitudinal surveys other than the 2005 Gunsekera study also show evidence that this schizoglossia continues, as about 30% SLE speakers still hesitate to accept that they are speaking a variety different to BrE and recommended RP as the teaching norms in Sri Lanka (Künstler, Mendis & Mukerjee, 2009).

This study investigates the PM use of SSLE speakers in Sri Lanka. As the speech samples were selected based on the presence of features of SSLE, this section describes the SSLE phonological

features in detail. SSLE distinguishes itself from the other dialects mainly on the phonological features in addition to grammatical accuracy. Therefore, the most salient features of SSLE that are attested in previous research are presented in Table 2.1 as a summary. When selecting the speakers for my study I used the features given in this table to guide this judgement. The selection of the speakers is explained in the Sections 4.5.1 and 4.5.2.

Feature	Attested by
The absence of aspiration of word initial voiceless plosives such as [p] and [k]	Fernando, 1985
Retroflex sounds where BrE would use alveolar plosives -eg: [t] and [d]	Fernando, 1985
Devoicing of [z] in word initial, word final and intervocalic positions.	Fernando, 1985
Use of dental plosives [t̪] and [d̪] in place of BrE [θ] and [ð]	Fernando, 1985
Use of a flap articulation of [r] than the frictionless continuant of BrE for initial [r]	Fernando, 1985
The use of a clear [l] in the final position where BrE would use a dark [ɫ]	Fernando, 1985
The use of an approximant or a labio-dental frictionless continuant [ʋ] for both BrE [v] and [w] initially	Fernando, 1985
A degree of lip rounding in the labial sounds generally, but particularly in [f] and [w]. The degree is dependent on the scale of formality, although these consonants are never accompanied in SSLE with as much labialisation as in BrE.	Fernando, 1985

<p>Doubling of a final consonant in a stressed syllable when it is followed by an initial vowel in the next syllable, accompanied by an absence of juncture over word boundaries. This is usually found in informal, friendly conversations. Eg: Come up- [kʌmmʌp]</p>	Fernando, 1985
<p>Long vowels shorter than those heard in Received Pronunciation</p>	Passe, 1948 (cited in Fernando, 1985)
<p>Long vowels in place of diphthongs in BrE</p>	Fernando, 1985
<p>The use of [o:] in place of [ɔ:]</p>	Fernando, 1985; Ekanayaka, 2020
<p>The use of short back rounded half-close vowel [ɔ] in words like <i>omit</i> where BrE would use [ou]</p>	Fernando, 1985
<p>Difference in the quality of diphthongs: BrE uses failing diphthongs, whereas in SSLE the first element is only slightly more prominent than the second element. SLE diphthongs are usually shorter than the corresponding BrE sounds.</p>	Fernando, 1985
<p>The use of the diphthongs [ea] where Std. E. uses [eə] as in <i>there</i></p>	Fernando, 1985
<p>Differences in the diphthongization of the triphthong [aue] in comparison to BrE. For example <i>flour</i> is pronounced as [flauə] in BrE and it is pronounced as [flaa] in SSLE.</p>	Fernando, 1985
<p>In the diphthongs [ai], [ɔi], [au] the final element comes fairly close to the frictionless continuants [j] and [w] in causal colloquial speech, i.e. on the furthest point of the scale of formality. E.g. <i>so how how?</i>¹¹</p>	Fernando, 1985

¹¹ *So how how?* is a common greeting in SSLE which is equivalent to *so, how are you?* in BrE.

A slight tendency to replace the middle element of triphthongs [aue] and [aiə] with bilabial or palatal frictionless continuants in casual colloquial style e.g. <i>power</i> - [pavə]	Fernando, 1985
The use of [a] for final a, ah in unstressed syllables where BrE uses [e]	Fernando, 1985
The non-use of the neutral vowel in weak forms of words like at, for, of, to, do etc.	Fernando, 1985
The tendency to use the neutral vowel [ə], or a sound intermediate between the full vowel [e] and [ə] in all unstressed vowels in final syllables	Fernando, 1985
Primary stress on the first syllable of the word. Stress would typically be placed on the second syllable in BrE.	Meyler, 2007
Use of the diphthong [ai] in SLE in lieu of [i] or [I] in BrE	Meyler, 2007

Table 2.1: Phonological features of SSLE

Gunesequera (2005, p. 117) details the difference between BrE and SSLE vowels in the following Table 2.2 and 2.3. As shown in Table 2.2 the short vowels [i] and [e] in SSLE are not used in BrE. Most of the long vowels in SSLE such as [ee], [II], [εε], [ææ], [əə], [ʌʌ], [uu], [ʊʊ], [oo] and [ɔɔ] are not used in BrE. The most prominent difference is that SSLE uses long monophthongs to match each vowel and this feature is absent in BrE. The central vowels are more open in SSLE than in BrE. The back vowel [ɑɑ] in BrE is realised as a central long and short vowel in SSLE.

	Front	Central	Back
Close	[i], [ii]		[u], [uu]
	[I], [II]		[ʊ], [ʊʊ]
Half close	[e], [ee]		[o], [oo]
Between Half close & Half open/ Mid	[ε], [εε]	[ə], [əə]	[ɔ], [ɔɔ]

Between Half open and open	[æ], [ææ]	[ʌ], [ʌʌ]	
Open		[ɑ], [ɑɑ]	

Table 2.2: Standard Sri Lankan English Vowels (Gunesequera, 2005, p. 117)

	Front	Central	Back
Close	[i]		[u]
	[ɪ]		[ʊ]
Half close			
Between Half close & Half open/ Mid	[ɛ]	[ɜ]	[ɔ]
Between Half open and open	[æ]	[ʌ]	[ɒ]
Open			[ɑ]

Table 2.3: Standard British English (RP) vowels (Gunesequera, 2005, p. 117)

The literature on SLE confirms that Standard Sri Lankan English phonological features are not influenced by the L1 background of the speaker (Gunesequera, 2005). Commenting on Standard Sri Lankan English, Gunesequera (2005) notes that:

This variety is used by the Sri Lankan elite consisting of members of the Sinhalese, Tamil, Moor, Malay, Burgher, Parsi, Borah, Sindhi, Bharatha, Colombo Chetty, and Eurasian ethnic groups. (p. 35)

She further mentions that the members of this group “share togetherness in their use of the language”. Other previous literature too confirms this statement. Kandiah (1981, p. 63) mentions that SLE (the variety that is referred to as SSLE in later times) is spoken by “habitual users” and is used as an L1. He mentions that the habitual Sri Lankan users of English are bilingual and he does not relate it

to any one particular ethnicity. It is clear that the SSLE speaker is not defined by his/her ethnicity, which is part of the reason why other linguistic features are used for identification instead.

SSLE is the variety that is modelled in the ESL classroom in Sri Lanka. In their article on teaching English in Sri Lanka, Samarakkody and Braine (2014) state:

“In terms of which variety of English should be used as the model in the English language teaching classroom, the integration of audiocassettes based on standard Sri Lankan English pronunciation into the curriculum is evidence of the growing awareness and acceptance of this variety as the norm.” (p. 152)

As mentioned in chapter 4, my being an ESL practitioner influenced me to explore the features of SSLE further. It is the norm aimed for in ESL contexts and the prestige norm, and so it is important to understand how it works on all linguistic levels. This is also the variety that offers opportunities in obtaining employment in the competitive private sector job market (Amarasuriya, 2010). Job interviewees in the private sector have been humiliated by overt reference to their lack of English proficiency and features of SSLE (Amarasuriya, 2010). Therefore, it is clear that just speaking English is not enough to empower the youth of Sri Lanka. They should be exposed to the SSLE variety of English so that they can access equal opportunities in employment and not be socially excluded. In order to expand the teaching of SSLE, the feature of SSLE should be common knowledge and clear to those learning and teaching English in Sri Lanka. If this study can provide an explicit description of its pragmatic features, this can assist in the explicit instruction of these features in the ESL classroom.

2.8.2 The Non-Standard Sri Lankan English

Literature on SLE acknowledges a non-standard variety of SLE (NSSLE) spoken by members of the lower strata of society who speak SLE as an L2 in a non-proficient manner (Fernando, 1985; Parakrama 1995, 2012, 2016). Derogatively, this dialect is referred to as *not pot* English (Gunasekera, 2005, p. 36) or as Singlish (Parakrama, 1995, p. 103) by SSLE speakers. Gunasekera (ibid) and Meyler (2007) comment that this is the best known and the most widely spoken dialect in SLE. The name derives from a phonological feature of SSLE that is used in free variation by the NSSLE speakers, viz.

the pronunciation of the open [ɔ] and closed [o] as in *lot* [lɔt] and *gold* [gold]. Therefore, the *not* and *pot* are pronounced as [not] and [pot] instead of [nɔt] and [pɔt]. It is this “confusion” between these sounds that gave the derogatory term *not pot English* for NSSLE (Gunesequera, 2005, p. 113).

Gunesequera (2005) attributes this mistaken free variation or inability to pronounce the open vowel [ɔ] to the influence of Sinhalese which does not have two types of back vowel [o]s. Sinhala only uses the back vowel [o] in its phonology.

Another noticeable characteristic of *not pot English* is confusion between [p] and [f]. For example, the word *perfect* could be pronounced as *ferfect*. The insertion of [i] in initial position of consonant clusters such as *sch*, *st*, and *sl* (in words such as *school*, *stop* and *sleep*) so that the words may sound [isku:l], [istop] and [isli:p] is another common characteristic. In a perception study done on the relevance of pronunciation as standard markers, Dinali Fernando (2014) concludes that the [o] and [ɔ] difference will not matter to the gatekeepers of English language in future because the future gatekeepers themselves will be unaware of the difference between [o] and [ɔ].

2.8.3 Burgher English

Rajapakse (2008) reports on a study which provides evidence of the features of Burgher English. Burghers are a minority group among Sri Lankans, with Dutch, Portuguese and British ancestry. In Rajapakse’s (2008) qualitative study that interviews twelve Burghers across three generations (1920s to circa 2008), each claim their first language is English. Rajapakse states that compared to SSLE, the overuse of the progressive tense in the data, in instances such as “because once you die, you are not getting anything no” or “They are scolding the Burghers but they’re following the Burghers’ side” leads to a usage considered ‘ungrammatical’ Standard Sri Lankan English. Parakrama (1995) states that the overuse of the progressive tense is not a characteristic of habitual users of Standard Sri Lankan English speakers. Rajapakse’s study reports features of English unique to Sri Lankan Burghers in terms of ellipsis, focalization and topicalization. Rajapakse also gives the tag question *no* as a feature of Burgher English but it is a feature that is also common in Standard Sri Lankan English (Gunesequera, 2005). Many examples of Burgher English can be observed in the

writing of the Sri Lankan Burgher writer Carl Muller who is best known for his trilogy on continuity of a Burgher family: *The Jam Fruit Tree, Yakada Yaka and Once Upon a Tender Time*.

2.9 Summary

This chapter has presented a review of the literature pertaining to Sri Lankan English. It is clear that English is a de facto official language as evident from its functions in the country. SLE is an umbrella term for various dialects of English used in Sri Lanka. On paper, SLE is perceived as a neutral language and labelled as a link language but this is dissembling the real status of SLE. It is perceived as a more privileged and powerful language compared to the other languages in Sri Lanka. SSLE is the most prestigious dialect of SLE and it is also a gatekeeper. With regard to other dialects and other languages, it has assumed a role akin to the colonial English which overpowered vernacular languages. Yet, since it is the dialect that is promoted in the ESL classroom, it is necessary to understand its features and create awareness among its users and learners. Therefore, the next chapter explores the role of PMs in SSLE, an area that has not yet been uncovered. This information underpins the necessary knowledge for the ESL learners in Sri Lanka to speak English in a more natural manner as PMs add a degree of spontaneity to speech (Beeching, 2016).

Chapter 3: Pragmatic Markers

3.1 Overview

Pragmatic markers (PMs) are an essential part of spoken language that aid fluency and the smooth flow of conversation. Speakers use these elements to establish social and interpersonal cohesion and discourse management (Degand et al., 2013). These elements direct the conversation and help set the emotions, mood and attitude between a speaker and a hearer. The term pragmatic marker is often used to refer to items like *you know*, *well*, *I mean*, *like*, *right* and *you see* although there is no consensus as to which items exactly can be regarded as PMs. They carry a non-truth-conditional value and can be removed from the utterance they are part of without harming its propositional meaning (Fedriani & Sansò, 2017, p. 3). Example 3.1 illustrates this point¹². In example 3.1, *you know* does not enter the truth conditions of the sentence. *You know* has no impact on the truthfulness or the falsity of this utterance. *You know* can also be detached from the utterance without affecting the propositional meaning.

Example 3.1

<05AcF> er which pr=but it is not only for women to in the Domestic Violence Act anybody
within a family could complain if there is violence within the domestic sphere
<07HM> mm
<05AcF> within certain **you know** types of relationships

As a starting point I will adopt the straightforward definition that PMs are a functional class of linguistic items that do not typically change the propositional meaning of an utterance but are essential for the organisation and structure of discourse, for marking the speaker's attitudes to the proposition being expressed, and for facilitating processes of pragmatic inferences. (Furkó, 2017, p. 2)

¹² Please note <05AcF> in example 3.1 is a code that refers to the individual speakers in the corpus created for my study. All the examples given in this thesis have such codes to refer to the speaker. These codes are explained in Section 4.5.6.

In this chapter, first the forms and definitions of the PMs as stated in previous literature are described. The discussion touches on the PMs that combine to create co-occurring PMs as noted in literature. This is followed by delving into the concept of nativised PMs in outer and expanding circle varieties of English. Then, the functions of PMs are discussed. Perceptions about PMs are also explored as the study investigates the impact of social factors which generally affect perceptions. Finally, the literature on the social variables of gender and age with regard to the use of PMs is explored.

3.2 Terminology and Definitions of Pragmatic Markers

There is no universal agreement on the terminology for PMs among researchers. In fact, the terminology also considerably depends on the researchers' categorization of functions of the PMs. Jucker and Ziv (1998, p. 1) state that "the terminological diversity reflects both the wide range of linguistic approaches that have been employed for their study, and the multiplicity of functions which these elements are said to fulfil." Beeching (2016, p. 5) selects the term PM "to highlight their interpersonal rather than textual usages, though recognising that pragmatic markers have procedural meanings". The words such as *you know*, *well* and *I mean* are functors or function words as opposed to content words or lexical words. Hence, considering the wide range of pragmatic functions these elements play in an utterance, I prefer the term pragmatic marker over discourse marker, following Aijmer (2013) and Beeching (ibid). Mostly, PM has been used interchangeably with discourse marker. For some scholars, PM is considered an umbrella term which includes discourse markers as a subset (Aijmer and Simon-Vandenberg, 2011; Carter and McCarthy, 2006; Crible, 2018).

The vast array of different terms used in the literature highlights the diverse appeal to scholars in terms of the role these items play in language. In the literature, these items have been called discourse markers (Lenk, 1998; Müller, 2005; Schiffrin, 1987; Schourup, 1999), discourse particles (Aijmer 1996; Barnes, 1995; Fischer, 2006; Schourup, 1985), pragmatic markers (Aijmer, 2013; Aijmer & Simon-Vandenberg, 2006; Andersen, 1998; Brinton, 1996; Caron-Prague and Caron 1991; Denke, 2009; Watts, 1998), pragmatic particles (Beeching, 2002), modal particles (Weydt;

2006), theticals (Heine, 2013), and connectives (Bazzanella 1990; Degand, 2000; Fraser 1988; Lamiroy 1994; Unger 1996). They are also called hedges (Coates, 2013; Holmes, 1995; Hyland, 1998), boosters (Beeching, 2009; Holmes 1995; Hyland, 1998, 2000), fumbles (Edmondson, 1981) and punctors (Vincent & Sankoff, 1992). The rationale for dismissing PMs as trivial items or “gambits”, as some of these terms suggest, is because they are simply considered items that “lubricate” conversations and used by speakers in part to gain time (Edmondson, 1981, p. 81). Although researchers have not reached a consensus about the most appropriate label to represent the function of the elements, all these terms hint that these elements are extra-clausal.

Similarly, there is no consensus about the definition of a PM. This lack of a comprehensive definition makes it challenging to know which forms can be accepted as a PM. The selection of the definition controls the selection of the linguistic expression (Schourup, 1999). Therefore, one of the greatest needs in the field to date is an inclusive model of identification and selection (Bolly et al., 2017, p. 71). In their research Bolly et al. (2017) suggest a model for discourse marker annotations (MDMA) to overcome this issue. Even this model depends on the subjective viewpoint of expert coders and manual identification of potential instances of DMs or PMs. Therefore, it is difficult to imagine an agreement on the definitions and items of PM in the foreseeable future.

However, there are some delineating qualities that are commonly attributed to PMs in all definitions, i.e. they are difficult to place in a traditional word class, syntactically optional and polyfunctional (Aijmer and Simon-Vandenberg; 2011). There are a few definitions which are based on these features. For instance, Schourup (1999, p. 242) defines PMs as “a more or less open class of syntactically optional, non-truth-conditional connective expressions”. Shiffrin’s definition too touches on one of the above qualities; they are syntactically optional. Shiffrin (1987, p. 31) defines PMs in relation to units of talk as “sequentially dependent elements which bracket units of talk”. She explains that it was deliberate to leave the phrase *units of talk* vague without defining it in relation to structure, cohesion or interactional relations because all these types of *units of talk* influence the markers. A precise unit would limit the function of PMs. Such limitations are however not evident

with PMs as they can be removed from the unit of talk any time with the structure and the propositional meaning of the unit intact. There are two PMs in the utterance given in example 3.2, *I mean* and *like*. *I mean* is connected and its relevance is limited to the phrase “dance was always a part” and *like*, similarly is connected and relevant to “my teachers would know that”. However, this utterance retains its propositional meaning even without *I mean* and *like*. These PMs can be removed with the structure and meaning intact irrespective of the earlier limitations that was noticed in each unit of utterance.

Example 3.2

<07AF> no no but erm **I mean** dance was always a part even in school **like** my teachers would know that oh she’s not in class she must be in the dancing room

Another quality of PMs that is highlighted in definitions is that they facilitate interpersonal cohesion and can convey the underlying emotions of the speaker. They provide an opportunity for the hearer to interpret the verbal message in a way that the speaker’s intention is fulfilled. In this way, the PMs construct the mood and the attitude of an utterance. This is why PMs are regarded as metalinguistic indicators as well. This quality is touched on by Bazzanella (2006, p. 449) when she defines that PMs “are related to the speech situation [and] (...) express attitude and emotion”. This reference to attitude and emotion is evident in example 3.3 from the CSSLE, the corpus compiled for the current study.

Example 3.3

<02AF> and I think that’s not all that healthy erm other skills maybe the skill to stay with something for a extended period of time erm er er that also I feel is er **kind of** I’m I’m I’m losing that capacity so all of these things I’m I’m trying to er regain er my abilities to do that

Speaker <02AF> is an artist who is being interviewed during the pandemic. She is explaining that she felt that time was being wasted during the pandemic and that she was losing her skills

because she was not using them. Instead of directly stating that she felt that her capacity and skills were being lost, she hedges the statement with the PM *kind of*. It would be a very blunt statement had she stated that she is losing the capacity to do things. The use of the PM highlights that she felt that a direct expression is too strong. This is a reference to her attitude and emotion with regard to this idea she expresses.

Aijmer (2013) observes PMs to be surface phenomena that on a deeper level reflect a speaker's mental processes as relevant to the on-going conversation. Combining all above features Heine et al. (2021) define PM/DM in a more comprehensive definition than Furko's (2017). Their definition defines the PM specifically with regard to its relevance to the grammatical structure as well as its impact on the prosodic structure as well:

“(a) invariable expressions which are (b) semantically and syntactically independent from their environment, (c) set off prosodically from the rest of the utterance in some way, and (d) their function is metatextual, being anchored in the situation of discourse and serving the organization of texts, the attitudes of the speaker, and/or speaker–hearer interaction.” (p. 6)

All of the definitions given above either use or revise the definitions proposed by earlier scholars or describe the functions PMs perform in a particular data set. This is the main reason for either inadequate or vague definitions. Another difficulty in identifying PMs is their diverse forms. PMs come from diverse syntactic categories including adverbs, conjunctions, verbs, etc. and even include non-verbal expressions such as *er* and *arh*, so that many expressions are PMs in some contexts but not others. The invented examples below show the diverse syntactic categories of *like*: example 3.4 shows *like* as a verb, example 3.5 illustrates *like* as a conjunction and example 3.6 shows *like* as a comparative complementizer.

Example 3.4

I **like** ice cream.

Example 3.5

If you want to show bravery, **like** a man face the problem.

Example 3.6

It seems **like** the health system has changed.

This multifunctionality means that PMs can be difficult to identify. The following characteristics of PMs could guide a researcher to recognise and categorise PMs.

3.3 Characteristics of PMs

Firstly, PMs mostly occur in spoken discourse but there are instances found in written genres as well. Previous research show that PMs are mostly present in spoken language (Beeching, 2022; Holmes, 1995; Stenström et al. 2002). Louwerse and Mitchell (2003) report that PMs occur ten times more in spoken discourse than in written genres. However, Shakir and Deuber (2024), in their study of code-switching in Computer Mediated Communication such as tweets and memes reveal that indigenised PMs are widely used in IE, SLE, Pakistani English and Bangladeshi English. This finding challenges the claim that PMs are mostly widely used in spoken discourse rather than written genres, as tweets and memes can be considered modern written genres. They contain features similar to chats and interactive speech such as face to face conversations (Shakir & Deuber, 2024).

PMs operate on three levels of conversation: the interactional, the interpersonal and textual levels (Carter & McCarthy; 2006). Brinton (2017), in her very thorough and useful account of functions of PMs, mentions that they are present in written genres but as an alternative set of forms. Secondly, PMs are usually short phrases or individual words whether they are used in speech or in writing. For example, there are one-word PMs such as *right* and *well*. Generally, the short phrases are limited to a combination of two words as in *you know* and *sort of*. In a sentence structure, they occupy a flexible place. For instance, they are generally not restricted to sentence-initial, medial or final position. Given below are examples of *you know* in these three positions from the data collected for this study. In 3.7 *you know* appears in sentence initial position, and 3.8 shows *you know* in sentence medial position. Example 3.9 shows *you know* in sentence final position.

Example 3.7

Sentence initial position

<02AF> **you know** not to be afraid of boredom to sit to look out to just be um I think that's if I'm to be honest with myself that's something that I struggle with

Example 3.8

Sentence medial position

<02AF> real **you know** the longer pieces I just like to know that I can do both

Example 3.9

Sentence final position

<02AF> I think that really made my sort of er sort of grasps the bull by the horns **you know**

However, there are a few PMs which are restricted to a certain position in a sentence when they are used for a specific pragmatic purpose. An example would be *well* which occurs only in the sentence-initial position as in example 3.10. In this instance, *well* is used to mark taking a turn. In this function, *well* is always used sentence initially.

Example 3.10

<08AcF> **well** the diaspora at least the word itself is from a Greek er two words to give meaning to scatter from or place of origin to scatter or like seeds you scatter the seeds so the origin of the words is closely associated with the Jewish diaspora

PMs are generally perceived to have no propositional meaning. They do not affect the main message of the sentence. 3.11 is from a conversation between an interviewer and a musician.

<05AF>, the musician explains the people who shaped her success. *I mean* is used to introduce the phrase "they moulded us" as a justification to show that her school teachers trained her well. "Visaka Vidyalaya" is the name of her school. The main message in this utterance is that her teachers at Visaka Vidyalaya moulded her. Hence, even if *I mean* is removed from this utterance this main meaning is still intact.

Example 3.11

<05AF> Kumar also my teachers at Visaka Vidyalaya **I mean** they moulded us

Thirdly, PMs are regarded as optional to the syntactic structure. In 3.11, the syntactic structure is intact even if *I mean* is removed. The idea would read as “my teachers at Visaka Vidyalaya (they, as pronoun reference to my teachers) moulded us”. Therefore, the PMs have no propositional meaning and are optional to the syntactical structure.

However, PMs are no longer considered completely devoid of meaning (Brinton, 2017). In fact, they are considered essential for the pragmatic meaning. This is illustrated in 3.12 where *like* precedes “a year” to indicate an estimation. It is the approximative use of *like*. The presence of *like* shows that the time period <04AF> worked in the hotel could be less than a year. If *like* is removed, the utterance would indicate that <04AF> worked for exactly one year, which is not the proposition the speaker intended to convey. Therefore, *like* is not entirely devoid of propositional meaning in this instance.

Example 3.12

<04AF> I worked at a hotel for **like** a year

3.13 is another example of this essential pragmatic meaning. In 3.13, without the PM, the hearer may not know that a reference is made to common knowledge between the speaker and the hearer. The speaker is referring to the living room that is familiar to the hearer as well by using *you know*. We can remove *you know* and the meaning of the utterance would be intact. However, the interpersonal reference and the effect it would have on the speaker-hearer relationship would be lost. By referring to common knowledge the speaker is strengthening her relationship with the hearer as it has a sense of inclusion. The distance of the relationship between the hearer and the speaker is shortened.

Example 3.13

<02AcF> well they are always there Kumar that’s the thing I don’t separate I you can’t compartmentalize life so when I do my PhD now I’m studying in the middle of **you know** the living room upstairs and they’ve seen me studying so it’s part of that

As noted above, most of the PMs are polyfunctional. Beeching (2016) reports five functions for *like* and nine for *well* in her study of British English. For example, *like* is used for exemplifying, approximation, to quote, to focus and to hedge. Each function depends on the role that the PM plays in each discourse situation. The characteristics such as PMs being short phrases, optional to the syntactic structure, their polyfunctionality help a researcher to discern PMs in an utterance from other linguistic items to a great extent when vague definitions do not help.

We also need background knowledge or inferential schemata surrounding an utterance to understand the role of PMs (Aijmer, 2013). All utterances are part of a socio-cultural context. They construct their meaning from this background information. Their meaning also depends on the relationship between the interlocutors. The interlocutors interpret the meaning of the PMs simultaneously although the meaning has never been extrinsically discussed. In other words, this context bound meaning is co-constructed by the speakers. This co-construction is successful only if the speakers have “access to a large number of linguistic and extra-linguistic conventions in order to use pragmatic markers” (Aijmer, 2013, p. 15). This can be seen in examples 3.14 and 3.15 from the corpus of the present study.

Example 3.14

<01HM> and it was a fairly controversial film and yours your role as well arh one what how do you handle sex nudity suggestive sex on stage

<01AF> it's er I don't think that I was handling sex or nudity on I was playing a character er er it's the er need of the role the whole script wanted that character to behave in such a way that's the way certain peo= people behave

<01HM> sure

<01AF> you know

<01AF> that's her natural life so I have to play the role of that er the life of this role Gothami er I was not doing any sex or **you know** and in that role to er for the role to develop and for

us to er come to the er mmmeme to bring out the meaning and the issue behind the character ha= certain situations have to be er shown in the

Example 3.14 is a conversation between the interviewer <04AM> and the female actress <01AF>. *You know* in this discourse refers to sex-related action in a movie scene. <01AF> does not explicitly describe what entails such action but uses the PM *you know* so that the hearer, with his access to the common knowledge about the field, can understand what was meant. 3.15 is an example where the speaker taps common knowledge with the combined PM *you know*.

Example 3.15

<06AcF> **you know** they either sit in front of televisions or

<08HM> the computer

<06AcF> phones or computers and everything information everything is fed to them

To provide some context, example 3.15 is a conversation between an academic and an interviewer. The academic mentions that she conducts workshops for children to improve their theatre and creative skills. She mentions that today's children are always either in front of the television and that is the reason that such workshops are needed. She uses the PM *you know* to refer to the common knowledge of how children spend their time in today's world. Her invite for the common knowledge is successful as her interlocutor <08HM> too suggests yet another device that the children use adding to her idea.

In sum, PMs are short phrases with no propositional meaning yet performing poly functions in contexts co-constructed by speakers emanating context dependent meaning.

3.4 Co-occurring PMs

Pragmatic marker combinations or co-occurring PM are a relatively new field in PM research in all varieties of English and even other languages (Lohmann & Koops, 2022). Examples of co-occurring PMs in English are *well you know*, *well I mean*, and *I mean you know*. In a study based on the British component of the International Corpus of English (ICE-GB), they are described as "a combination of two or more non-identical, immediately adjacent" PMs (Haselow, 2019, p. 4). This

phenomenon of co-occurrence is also known as clustering (Maschler, 1994), combining (Fraser, 2013) or sequencing (Lohmann & Koops, 2022). Important aspects of these co-occurring PMs such as the choice of combination or the factors that determine their linear order are not fully understood yet. For instance, we know that certain PMs can combine with one another to create a more emphatic utterance than can be conveyed with one individual PM. 3.16 illustrates the emphatic force created by the combination of two PMs. In this context, <04AF> claims to be a vegan. She describes that there are non-vegan food cravings sometimes. Both *you know* and *like* can be used individually on their own in this context. Yet by using them together there is an appeal to the interlocutor for agreement and at the same time *like* acts as an exemplifier. It introduces an example situation that happens to <04AF> with regard to non-vegan food craving. This appeal to the interlocutor together with exemplification further attracts attention to the idea of what happens with a craving for non-vegan food craving.

Example 3.16

<03HM> like ribbon cake can make me go non vegan in the first sight
 <04AF> oh my God yes but like I've kind of trained myself to just taste it and not binge eat it
 <03HM> okay
 <04AF> **you know like** sometimes when you see something you're like you keep
 <03HM> that's true
 <04AF> going for it and you'll just end up eating the whole thing

The principles that direct such combinations and the choice of combinations inform us about the thinking patterns of the user and underlying grammar patterns of language. Therefore, my research investigates the PM combinations and their functions in my data in addition to other PMs. Haselow (2019)'s analysis of combined PMs will guide the analysis of the co-occurring PMs in my data.

The literature on co-occurring PMs has revealed possible reasons for co-occurrences. For instance, Aijmer (2002) argues that co-occurring PMs are functionally similar. Lohmann & Koops (2016; 2022) likewise assume that similar functional categories or the functional complementarity of PMs might prompt them to form combinations. In a theoretical investigation that synthesised the results of previous research on BrE, AmE and French, Lohmann & Koops (2022) state that the tendency for functionally similar markers to co-occur is very high. An example of two PMs with similar functions combining with each other for a function similar to their individual functions is given in 3.17.

Example 3.17

<08EM> er well my er personal er yeah my personal er thing is **like you know** er we we are an island

One of the functions of the PM *like* is to show hesitation. A function of *you know* is word search. The utterance in 3.16 has many filled pauses that indicate hesitation. Therefore, it can be assumed that the speaker is hesitating and searching for words in this utterance. Therefore, he has combined *like* and *you know* both for the purpose of hesitation and word search. Haselow (2019) analyses such combinations as a communicative strategy to overcome disfluencies in real time speech.

Crible (2018) examines co-occurring PMs in a data set called DisFrEn. It is a collection of spoken English and spoken French. The texts for the spoken English comes from ICE-GB and the Backbone project (Kohn, 2012) which also features BrE. In this corpus analysis of both English and French, he finds that 20% of all PMs in the corpus are co-occurring PMs. French co-occurrences were more common than English co-occurrences. He also states that the most frequent patterns of co-occurring PMs were to express similar or complementary functions.

There is research on the flexibility of sequencing of PMs as well. The little research available shows that ordering is constrained by the position of the PM (Haselow, 2019; Crible & Degand,

2021). Crible (2018) claims that co-occurrences are more frequent in sentence-initial position than other places. Haselow (2019) investigates the sequencing behaviour of turn initial and final PMs. His markers include a combination of PMs such as *you know* and *well* and DMs such as *then* and *because*. He identifies three functional domains in which the PMs and DMs typically operate, which are interaction, discourse structure, and cognition. Interaction domain refers to all functions related to the management of the interlocutor interaction, such as organising turn-taking and setting up the conditions for successful uptake of an upcoming message. An example of such a PM sequence from the PMs that the present study concerns would be *well I mean*. The second domain, discourse structure, refers to the relationship between the discourse units. It indicates the relationship between discourse units and implied meanings or inferences. An example would be *you know like* as given in 3.16. *You know like* in 3.16 connects the reference to non-vegan food craving to an example of what happens when the speaker gets an opportunity to eat non-vegan food. Therefore, it connects one discourse unit with another. The third domain pertains to the interlocutor's cognitive states. An example of a general function of cognition domain is providing a cue for an utterance interpretation. An example from the present study is *I mean you know* where nuancing and appeal to common knowledge is combined. 3.18 is an example of a co-occurring PMs in the third domain providing a cue for an utterance interpretation.

Example 3.18

<04AM> that's what company should adopt not sack and reduce salary and what is this I

mean you know can you imagine a person in this situation going home and saying you know children your father or mother has lost my job I don't have a job from tomorrow

<04AM> is an actor who is also a CEO of a company. He mentions that some of the CEOs in some Sri Lankan companies fired their employees to manage their businesses during the pandemic. With the use of *I mean you know* he prepares the listener to an opposing idea. Preceded by what is this he implies that he questions this decision by the CEOs, and he appeals to the interlocutors to

accept his idea. The speaker implies that he believes the listener too agrees with his opinion, thus provides a cue for utterance interpretation.

For his research, Haselow (2019) uses a spoken section of the ICE-GB. Then he selects all individual two-part PM combinations and rates the likelihood of a given word to be followed by another word such as *well I think*. The results reveal that the linearisation order corresponds to a particular temporal logic. The functional motivation that is underlying a speaker's thought process is the orientation in discourse processing. Therefore, PMs that ranked higher within a function preceded those that ranked low. This is one possible order that PMs co-occur. In other words, there is a logical and predictable order for the PMs to combine with one another that is linear, temporal and function driven.

Another explanation of an order is provided by Crible and Degand (2021) who explore the variables that govern the ordering of the DMs in a study that looks at DMs more broadly but includes PMs. They analyse all DMs from LOCAS-F, a multi genre corpus of monologues and dialogues in French. This corpus of 36, 912 words includes a range of speeches from political speeches to radio interviews and casual conversations. They finalise three ordering tendencies in their data: the syntactic category of the DM, domains and functional strength. They suggest that weak PMs or sequential domain PMs come first in a combination. Nevertheless, they declare that it is still not conclusive how the DMs combine with each other as more than one tendency is possible for a single sequence. Interestingly, 18 sequences such as *bon ben (well)*, *alors donc (well so)* and *ben en fait (well actually)* remain unexplained by any of these factors. Crible and Degand (2021) assume that these sequences cannot be categorised as they are not restricted by a rigid syntactic rule like the other DMs.

In sum, the literature on co-occurring PMs shows that there is an underlying systematicity to these combinations and their order of combinations. Also, what we understand from the comparisons between English and French co-occurring PMs is that there are cross-linguistic or language-specific PM behaviour. Not all sequences can be explained by the existing ordering

tendencies or generalisations. Therefore, these should be further investigated, as should the effect of combinations on the individual meanings of the PMs. Whether such combinations result in new idiomatic expressions or remain compositional is a question to investigate.

3.5 Nativised PMs

When a country is colonised, the language of the coloniser is spread among the colonised. In the case of Sri Lanka, English, the language of the coloniser was learnt as a second language by the natives of the country. With time, as generations passed, some of the natives of country acquired English as a first language in addition to those learning it as a second language. These new speakers of the coloniser's language added their own phonological, morphological, syntactic, semantic and pragmatic features into the language. For instance, with time British English in Sri Lanka lost the original linguistic features that characterised it as BrE because the natives of the country spoke English with the influence of the vernacular languages of the country. This process creates linguistic features which are new. The new native speakers of this variety create the new norms for the language. This process is called nativisation (Lowenberg, 1986; Schneider, 2007). Nativisation is influenced mainly by the L1 or the vernacular languages of the speaker (Li, Lorenz Siemund, 2023). Nativisation occurs in all aspects of a language like phonology, morphology, semantics, syntax and pragmatics. SSLE has nativised features in all aspects of language as well. These are as presented in Section 2.7. For example, with regard to pragmatics, Lim (2007) records the influence of Bazaar Malay and/or Hokkien for the nativised PMs *ah* and *lah* in colloquial Singaporean English.

The nativisation of a language happens through various strategies. Early research classified the strategies of nativisation to be similar to those found in the process of second language acquisition (Jain, 1974; Selinker, 1972). Later research shows that these strategies are extensions of innovative processes that are similarly productive in established varieties such as BrE or AmE. The varieties of world Englishes use overgeneralisation as a strategy of nativisation; however, this is not the erroneous overgeneralisation made by ESL learners (Lowenberg, 1986). One nativisation strategy

for PMs seem to be an attempt to match the intonation pattern of the first language of the speaker (Lowenberg, 1986, p.8). This is evident in the use of *nah* in Indian English and *no* in Sri Lankan English. As previously mentioned in 2.7.3 *ne* is an interactional particle and an emphasis marker in Sinhala (Chandralal, 2010) that has been transferred to SSLE as *no*. *Neh* also occurs as a variant of *ne*. Therefore, *no* and *neh* both can be considered nativized features of SSLE. *Ne* or *neh* becoming *no* appears to show anglicisation so that the form is recognisable in English. This lexical transfer from Sinhala has resulted in an intonation transfer from Sinhala as well. The utterance final addition *ne* creates an intonation pattern similar to Sinhala. In Sinhala, the emphasis marker *ne* ends with rising intonation. The intonation pattern is similar to that of a tag question. The same intonation pattern is adopted when *no* is used in SSLE for the same purpose of emphasis or a question tag. 3.19 is an example of such use.

Example 3.19

<02AcF> you should also have something inside like a balance and the sense of enjoying a tree as much as a book as much as a party so I wish I could teach them that balance so that where ever they are they would

<01HM> you wish you could or have you already

<02AcF> well I hope I have I never know **no** we are happy together

In this discourse sequence, <02AcF> is responding to the question whether she is ambitious for her children. She is explaining that material comforts are important but emotional intelligence is more important. When <01HM> asks if she has instilled this in the children, her response is that she hopes she has, so but she doesn't know for sure. *I never know no* implies she would not know it. *No* in this instance is an utterance final emphasis marker that is similar to *ne*, a PM in vernacular Sinhala language used for the same purpose.

Recent literature on world Englishes gives plenty of evidence that many post-colonial varieties of English have nativised PMs. For instance, Wilson et al. (2017) examines the presence of Trinidadian tag questions *eh* and *nah* in the International Corpus of English Trinidad and Tobago (ICE

T&T). Oladippu & Unuabonah (2020) discuss how *now* has a new function as a PM in Nigerian English apart from its role as a temporal adverb and as an attention pragmatic marker. Mohr (2021) finds that *you know* and *I think* fulfil personal and subjective functions in Zanzibar English which is typical of world Englishes. Closer to home, there is evidence of nativised PMs in Asian Englishes such as Indian English, Pakistani English and Singaporean English (Columbus, 2010; Leimgruber et al., 2021; Parviainen, 2015; Shakir & Deuber, 2023). For example, Noor (2021) presents twelve nativised discourse markers in her study on Pakistani English. Further, Anwar, Rasool and Kamran (2020, p. 70) report about the use of nativized PMs *na* and *haan* in Pakistani English. Similar to *no* and *na* in IE, these Pakistani English nativized PMs occur generally in sentence-final position to mark emphasis, convenience and assurance. An example is “She said, so tell me no” and “Why marking noise haan?” Shakir and Deuber (2024) report the use of many nativized PMs in code switching instances in their study on computer-mediated communication from Sri Lanka, Pakistan, Bangladesh and India. The results include the nativised PM *neh* used in lieu of *right* in English. All this research promises the possibility of nativised PMs in world Englishes. Therefore, the present study explores the presence of nativized PMs in the spoken data.

3.6 The Functions of PMs

Similar to the concept of definitions, there is no uniformity in identifying the functions which are independent of the propositional meaning. However, across the literature, functions are listed under three broad categories: tools of social and interpersonal cohesion, text-structuring devices and as indicators of how an utterance should be processed by the hearer (Fedriani & Sansò; 2017; Jucker & Ziv, 1998). Similarly, Erman (2001) classifies the functions into sets of three major functions. They are text monitors (editing signs, repairs, hesitation markers), social monitors (interactive markers, comprehension solicitors), and metalinguistic monitors (approximators, hedges, emphasizees). Biber et al. (1999) suggest that PMs facilitate an interactive relationship between interlocutors and the message. Many words which represent a grammatical category such as an adverb (*well, right*) or an

interjection (*eh, arh*) will assume a functional role and renounce their grammatical role when they become a PM. PMs are multifunctional and have a range of pragmatic functions (Brinton, 2017).

Determining the function of a PM is dependent on the pragmatic intention of a speaker (Denis and Tagliamonte, 2016). The functions of PMs stem from the role they play in communication. To understand the function, the context of the occurrence needs to be analysed (Aijmer, 2013). First, in terms of interpersonal cohesion, PM functions range from expressing interpersonal stance to reflecting social stance (Fischer, 1998; Lutzky, 2012). The role is often described as one that affects the personal relationship between two interlocutors and is co-constructed by them. For instance, Beeching (2016, p.1) attributes PMs to “oil the wheels of conversational social interaction” indicating that they are essential for a smooth relationship between two interlocutors. They seem to contribute significantly to the flow of conversation. For example, PMs like *well, yeah* and *you know* get the attention of the hearer by initiating discourse or closing discourse. The PM *right* is a short form when a speaker is aligning epistemic stances with the hearer as in *do you agree that I am right* (Lutzky, 2012, p.254). It is also used for the function of progression check. 3.20 illustrates this function.

Example 3.20

<01AcF> erm there has to be some good in the world **right**

<05HF> mm yeah

Right as a PM in this instance is functioning as a question tag which may or may not request a verbal answer from the hearer. The intention of the speaker is only to establish that there is agreement for what is being discussed between the two interlocutors. PMs promote a cordial relationship between speakers. When the speaker uses *you know* to appeal to common knowledge, the relationship between the two interlocutors immediately improves by a certain degree of cordiality. PMs also control the degree of politeness. *You know* is used as a means of inclusiveness in a conversation so that it saves face and promotes positive politeness. Beeching (2016) suggests that PMs mark friendliness and warmth. Another point is that PMs mark the identity of a particular age or

social group. This is the case with the use of *like* which is associated with suburban white female teenagers (D'Arcy, 2017, p.119).

Secondly, PMs are also text structuring devices. The PMs structure the discourse so that the message is easily and effectively perceived by the hearer. PMs mark a boundary in discourse by indicating a new topic, new information or a reference to an older topic (Brinton; 1996; Dér & Marko; 2010). They also mark the sequence between two clauses. Vincent (2005, p.189) argues that PMs function as connectors and modal elements. They highlight the relevance from one clause to another. They help repair one's own or the hearer's discourse. For instance, *I mean* is generally used to rephrase or repair a word or an idea in an utterance as exemplified in 3.21. <05AcF> first says "sixty percent of women". Then she wants to change the word *women* to *students*. So she introduces this repair with *I mean*.

Example 3.21

<05AcF> now when you look at university enrollment about sixty percent of women I mean
fif=sixty percent of students are women

PMs are essential for discourse structuring. They also work as hesitation markers (Beeching, 2016; Erman, 2001). Beeching (2016, p.4) states that PMs allow for hesitation in spontaneous speech.

Thirdly, PMs mark attitude and indicate how the hearer should interpret the utterance. Attitude is the activation of positive or negative positioning (White 2011, p.16). The attitudinal mode is characterised by three features: affect, judgement and appreciation. Affect refers to fundamental attitudinal meaning associated with emotions (disappointed/ashamed); judgement refers to a view indicated of the social acceptability by reference to social norms. Appreciation is when assessments are made of semiotic and natural phenomena. An utterance that contains one PM in isolation can imply the speaker's attitude. For instance, if an interlocutor asks the opinion about something and the other interlocutor simply states *you know* or *I mean* and ends the utterance, it can be read in

many ways like joy, cynicism, disapproval or disappointment. The tone and intonation coupled with perhaps facial expressions will enhance the understanding of the emotion. It is a subjective response or a reaction to the discourse situation. Example 3.22 illustrates this point.

Example 3.22

<01HM> Peter how did your mother manage fourteen people

<04AM> **well you know** [laughter] some sacrifice

<01HM> fourteen mouths to feed

<04AM> is an actor whose father had passed when he was very young as he mentions at the interview. There were 14 members in his family. So, the interviewer is asking how the mother had managed this big family after the father passed and <04AM> responds with just a PM. He simply says “well you know” and laughs. This implies that the mother somehow managed but it had not been easy. The listener is guided by the PM and Blakemore (1992) states that PMs guide the hearer by constraining the number of possible interpretations. In the case of 3.22, it is implied that the speaker is inviting the listener to imagine what it may have been like for the widowed mother by referring to common social knowledge. Markers select a meaning in relation to the surrounding discourse for the listener (Schiffrin, 1987, p. 318).

In sum, common features are identified across the literature, but with different explanations of functions. PMs operate mainly in the domains of social and interpersonal cohesion (Fedriani & Sansò; 2017). They also function in the textual domain as a discourse structuring device (Brinton; 1996). Their function is intended as it gives direction to the hearer to follow the speaker’s intended meaning.

3.7 Labelling Functions

There is no consensus as to the labels or terms associated with describing the functions. There is also much overlap and contrast of opinion about the functions assigned to the PMs among researchers (Fox Tree & Schrock, 2002). For example, *you know* has been described as a marker of

presumption of shared knowledge (Andersen, 2001), and of pretence of shared knowledge (Östman, 1981), an invitation to fill out the speaker's meaning (Jucker & Smith, 1998), a device that stalls time (Holmes, 1986) and a device to decrease social distance (Stubbe & Holmes, 1995). The classifications in just three studies on *you know* are analysed to show the diversity of the labels. The first study is by Beeching (2016) on British English, the second study based on American English is by Koczogh & Furkó (2011) and the third, is Östman's (1981) study of American English.

In the three studies which are reported in this section alone there are 14 different functions and only three functions are shared by all three studies. The three functions that are common to all three studies are clarification and appeal to common knowledge, reference to shared knowledge and word search. However, the three studies have used three different terminologies to refer to one function: *appeal to shared common knowledge*. Beeching (2016) uses the term *clarification and appeal to common knowledge*. Östman (1981) uses the label *direct plea for cooperation* and Koczogh & Furkó (2011) list this function as *reference to shared knowledge*. In sum, there are some core overlapping functions that are recognised as functions of a PM but there is no agreement among researchers on the terminology used to refer to them.

3.8 Perceptions about Pragmatic Markers

Although PMs are quite frequent items in speech, a heavy use of PMs is discouraged by prescriptivists (Miskovic-Lukovic, 2009). As the objective of the present study is to investigate PM use among Standard Sri Lankan English speakers which is the dialect aimed in the ESL classroom, it is relevant to understand how they are perceived, since these perceptions reveal why a speaker would aim to either use or avoid them and why particular PMs are chosen. It is helpful to understand what is perceived to be a fluent and a disfluent use of PMs as well.

The literature about perceptions of PMs shows mixed results. These perceptions are entwined with the idea of fluency and disfluency of a speaker. Therefore, it is important to first understand the difference between fluency and disfluency as one can be regarded as fluent but not

accurate. A comprehensive distinction between these two terms is provided by Crible (2018). Italics are as used in the original text.

“Fluent characterizes perceptively unmarked talk, which can be plain, eloquent or creative, albeit not necessarily flawless; *disfluent* applies to major breaks in the speech flow or in the syntax, leading to some sort of disruption; a disfluency (or disfluencies) refers to an actual occurrence of phenomena such as pauses, repetitions or truncations.” (p. 9)

In this light, pauses, repetitions and truncations are considered disfluencies, and these are some functions and characteristics of PMs as well. The previous literature considers PMs to be stylistically stigmatised especially in formal speech. They are considered to be characteristic of careless speech or loose use of language as they are used as fillers (Anderson, 1998; Schiffrin, 1987; Brinton, 2007; Pettersson-Traba, 2018). They are even considered as “speech irregularities” (Strassel, 2004). O’Donnel and Todd (1991) claim that PMs are a feature of “unskillful speakers” (p.69). In more recent work, Miskovic-Lukovic (2009) states that the use of *kind of* and *sort of* is bad usage of English. They are viewed as distracting to the listener, particularly when used repetitively, and a mark of disfluency. Perhaps the fact that they do not directly contribute to the propositional meaning of an utterance could be igniting this negativity. An attitude survey study done with undergraduate students in the US reports that a majority of the participants would not use PMs in a professional setting because for them, these items indicated being nervous, insecure or embarrassed (Fox Tree, 2007). PMs, especially *like*, are regarded as an indicator of problems with word choice (Siegel, 2002). It is clear that PMs have a bad reputation. However, not all PMs are equally regarded as a feature of a non-standard or bad language. For example, *you know* is considered part of standard English while *like* does not enjoy the same regard or value in the dimensions of prestige, solidarity and dynamism (Dailey-O’Cain, 2000; Buchstaller, 2006).

Contrastingly, PMs are also associated with fluency (Neary Sundquist, 2014). The presence of PMs is a mark of proficiency as seen in proficiency measurements in ESL based studies. The

frequency of PMs increases with each advancing proficiency level (Neary Sundquist, 2014, p. 638). As discussed in Section 3.6, the functions of PMs include structuring the discourse so that the linguistic message is directed more effectively. In my view, part of the structuring is to maintain coherence. Coherence is a characteristic of focused and fluent speech and is ensured by PMs. Therefore, PMs are part of fluent speech as well. Secondly, PMs, as defined by Beeching (2016) help oil the conversation between two speakers which implies that a smoother conversation is brought out by the PMs. Some PMs indicate a closer relationship between two speakers as they refer to common knowledge between speakers. *You know* is an example of a PM that refers to shared knowledge between two speakers. Sometimes, this reference can mean a presumption of a shared knowledge, yet it helps to establish an emotional closeness between the two speakers. This is beneficial for speakers. Thirdly, it is hard to imagine a spontaneous conversation between two fluent speakers of English without the use of PMs. A mark of fluent speech is fast and effortless automatic processing (Erman, 2001). In her study of adolescent speech, Erman (2001) states that adolescents hardly show that they are at a loss of words or taking time to plan their speech because they use PMs to overcome these issues. Adolescents especially use *you know* to show quick processing and fluency.

Overall, there are contrasting perceptions about the use of PMs. As a solution to mitigate subjectivity in these perceptions, Crible (2018) suggests that there should be a subset of *Potentially Disfluent Functions* and the presence of these functions will determine the level of disfluency. For instance, she suggests using functional and positional features of discourse markers to interpret the relative fluency of the clusters they appear in. She believes that both fluent and disfluent users of PMs would form a continuum of fluency rather than a dichotomy.

In sum, this discussion reveals that there are both positive and negative perceptions about PM use. Additionally, there is a demarcation between fluent and disfluent PM use. A disfluent PM use includes PMs for the functions of pausing and repetitions which indicate hesitation. This discussion is important for the present research as it observed PM use among speakers of English in Sri Lanka considered fluent. The frequencies and functional range of the PMs observed in the data

will provide further evidence of how they are used by these fluent speakers, and how they contribute to create both positive (fluent) and less positive (disfluent) instances of speech.

3.9 Sociolinguistic Factors Affecting the Use of PMs

Variationist sociolinguists have shown that social factors such as age, gender, and social class affect language use, and the current study explores the roles of gender and age in PM use in SSLE. It is generally accepted that gender has an impact on language, but the precise nature of that impact is difficult to establish; in particular, there is still argument about whether gender is a factor in itself or whether gender differences relate to other factors such as register, social status, education and age (Macaulay, 1977; Johnson, 1993; Coates, 2015). All these variables have some effect on language and drawing conclusions about their significance is very complicated. Yet, ideas about gender, based both on research and on popular belief, have primed our understanding of language. Popular beliefs emphasize gender differences as presented in mainstream psychology books such as John Gray's *Men Are from Mars, Women Are from Venus*. In addition, there is prolific research attesting that speakers are, consciously or subconsciously, deeply aware of what is gender-appropriate behaviour in their speech (Holmes, 2003; Bucholtz, 2003; Sauntson, 2018; Slobe, 2018). This is manifested in the choice of language structures they use (Newman et al, 2008; Labov, 1990; Lakoff, 1975; Vanda and Péter, 2011; Holmes, 1995, 2006; Laserna et al, 2014). In many sociolinguistic studies on language and gender, there is evidence of men and women using language differently. Most commonly, research reveals that women often use more standardised language than men. For example, Labov (1966) showed that women were more inclined to use structures perceived as grammatically more accurate and pronunciation patterns associated with higher social status. He interpreted this use as an attempt to achieve prestige. Similar to Labov, in his research on speakers from Norwich, Trudgill (1972) showed that women used more 'standard English' to signal social status. Milroy and Gordon (2003) claim that women prefer supra-local variants that may or may not be prestigious, but men prefer localised variants which are also stigmatised. Moving away from

phonology and syntax, even the latest digital data such as texts, emojis and emoticon usage show gender variation. Koch et al. (2022) investigated age and gender-linked language variation in instant messages. One finding among many gender-linked studies of language variation is that females used more emojis and a broader range of different emojis. There is gender difference manifested even in communication strategies. Recent research shows that women use inclusive in-group discursive strategies than men, and that correspondingly these strategies are perceived to be typically feminine. In an analysis of tweets sent by Canadian Chief Medical Officers during the pandemic, Vessey (2023) observes that women medical officers use stereotypically female ways of speaking as a strategy to navigate online and offline environments. Overall, these studies show that gender-based variation is present in all areas of human communication.

PMs are an aspect of language use that is traditionally associated with genderlectal and age variation. Some PMs are associated more with speakers from one age group or one gender. For instance, *like* is generally associated with young females in American English (D'Arcy, 2017). Hedges, intensifiers and quotatives are linguistic phenomena that show a certain degree of genderlectal and age variation and PMs fulfil these functions among the array of roles they assume in discourse. For instance, Holmes (1990) report that women and men use *you know* as a hedge distinctively in New Zealand English. Men use *you know* as a hedge to register degrees of verbal hesitation and uncertainty. Women, on the other hand, use *you know* to create and maintain interpersonal solidarity.

Like is regarded as showing innovative use among young women in Canadian English and Tagliamonte and D'Arcy (2007) observe that men have retreated from the use of *be like* quotative as its use became associated with women. *(Be) like* is generally not considered a PM; however, this study includes *(be) like* in its analysis as it is considered to have a PM function by Beeching (2016, p. 131). Further clarification is provided in Section 5.4.3. Schleeff (2005) examined the frequencies of *like* in the MICASE (Michigan Corpus of Academic Spoken English), which contains a wide range of speech events from academia. Schleeff's (2005) study observed 18 instructors and 50 students taking

part in 8 hours of lectures and 10 hours of seminars from an equal number of male and female instructors. The results show that female students use *like* more than male students. As well as this, students use *like* more than professors, indicating that *like* is used more by younger members of society. Laserna et al. (2014) conclude that PMs are more common among women and younger participants in their study of 263 speakers of American English. These are only a few examples which illustrate the impact of social variables on the production of PMs.

Research argues for the considerable influence of gender and age specifically on PM use in world Englishes (Botha, 2018; Lange, 2012; Lange & Leuckert, 2020). Lange and Leuckert (2020) investigated the tag questions used in the International Corpus of English India (ICE-India). Tag questions are indicative of interactive speech among interlocutors. The tag questions in the study included the nativised tag question *no* as well. The study reveals that females use more tag questions than men, and that female speakers make use of tag questions more often in all female conversations than male speakers in all male conversations. The study concludes that Indian women are more inclusive of their interlocutors than Indian men. It also showed that women used the nativised tag question *no* more frequently than men. Lange (2009) shows that women initiate the use of *yaar* in Indian English. Such results lead to the assumption that women in world Englishes contexts innovate nativized PM use. The current study investigates whether a similar pattern can be found among the speakers of SSLE.

Some Englishes like Singaporean and Indian English also display the impact of ethnicity on the selection and use of PMs (Leimgruber et al., 2020; Lange, *ibid*). Leimgruber et al. (2020) investigates 45,617 instances of pragmatic markers in the Corpus of Singaporean English Messages (CoSEM). CoSEM was analysed in terms of gender and ethnicity. Overall, men use more instances of PMs than women. The results reveal that speakers of Indian ethnicity use significantly fewer PMs. Chinese ethnicity as well as male gender are positive predictors for most PMs.

In terms of age, the literature shows that there is a clear connection between age and PM use. In some studies, the language innovators for PM use are identifiable by age. For example, Denis

& Tagliamonte's (2016) study investigates utterance final tags (UFT) in Canadian English and shows that males born between 1933 and 1943 are the first to use *right* as an UFT. Men under thirty-five have not adopted this trend to the same degree as their elders. In other research, looking at varieties of English, mostly adolescents are regarded as the innovators (Anderson 2001; Stenström, 2014; Tagliamonte, 2005). For instance, Andersen (2001) report that *like* is mostly used by adolescents in BrE.

The research on the influence of age in world Englishes also reveals age as a deciding factor for PM choice. For instance, in addition to her findings on gender, Lange (2009) reports that the nativized PMs *yaar* and *na* in Indian English are more used by the 18-25 age group compared to the age group from 42 onwards. The results revealed by the diverse research on gender and age influence the current research to find out how these social variables influence the use of PMs in SSLE.

It should also be noted that some PMs have received more attention on how they are impacted by gender and age than some others. For instance, there is much research on how *you know*, *eh* and *like* are used by women and men compared to the use of *sort of* or *kind of* (Holmes: 1986; D'Arcy, 2017; Vine & Holmes; 2023). It has been argued that the functions that some of the PMs play are more relevant to certain age groups than others (Stenström, 2014). It could be assumed that some PMs are more sensitive to the influence of gender or age or both factors than others. This could be investigated in research on PMs across a wide age range and gender such as the current study.

Suárez-Gómez and Seoane (2021) state that there is little research on gender and age with regard to PMs in world Englishes due to the limited metadata available about the speakers. Therefore, my study explores how the PM use in one variety of world Englishes is influenced by the social variables of gender and age, adding new knowledge to the body of work on world Englishes. This investigation into the sociolinguistic factors that affect PM use in SSLE considers who could be

the innovators of PM use in SSLE, speaker preferences of PMs and whether there are gender-specific functions performed by PMs.

3.10 Conclusion

This chapter has looked at the definitions of PMs and their formal and functional characteristics. It has examined the issues that arise from the lack of agreement on the definitions, terminology and functions of PMs. It also examined what has been uncovered by previous research with regard to co-occurring patterns of PMs and nativised PMs. Previous research has shown how sociolinguistic factors such as gender and age interact with PMs, and this has been introduced, and will be discussed further in later chapters. The background knowledge that has been set out in this chapter will help to situate this study in the large canvas of research on PMs and on PMs in world Englishes, and has outlined some issues and gaps in the literature on PMs. Existing studies show very clearly that world Englishes have PMs that are unique to them and that the functions of PMs behave in original and unexpected ways in world Englishes. However, there are still unanswered questions with regard to definitions, functions and identifying items of PMs.

Chapter 4: Chapter 4 Design of the Study

4.1 Introduction

This chapter presents the procedure followed to answer the research questions of this study. A considerable part of the chapter is devoted to describing the 202,557-word Corpus of Standard Sri Lankan English (CSSLE), which was created specifically for the purposes of this study. The CSSLE includes metadata about speaker gender, age, and occupation, allowing for the creation of virtual sub-corpora to investigate specific variables. The sub-corpora consist of a corpus of male speech, a corpus of female speech, 3 separate corpora for three separate social generations and three corpora for the three main occupations of the speakers. The main purpose of the study is to account for PM use in Standard Sri Lankan English. Additionally, PM use is analysed according to gender and age using the sub corpora. This chapter presents a description of the data needed to achieve the purpose of the study and how the ad hoc corpus was created. Then, it describes the variationist approach that is used to analyse and interpret the data. Approaches to comparing the different speaker groups based on gender and age are also discussed.

4.2 Research questions

There is much research on the types of PMs used in many inner circle varieties of English such as British English, Australian English, American English, Canadian English etc. (Beeching, 2016; D'Arcy, 2017; Denis & Taglianmone, 2016; Diskin-Holdaway, 2022). The research on PMs in postcolonial varieties of English has increased considerably over the past two decades as well (Parviainen, 2016; Shakir & Deuber, 2023; Tay et al., 2016). However, there is no reported research on PMs in Sri Lankan English (SLE). In fact, as noted in Section 1.3, any research on the pragmatic aspects of SLE is sparse (Revis & Bernaisch, 2020). As a starting point, this study explores the forms of PMs used in Standard Sri Lankan English (SSLE). It is well-established that PMs are more commonly used in speech situations than in writing (Schiffrin, 1987, p. 31; Gonzalez, 2004, p.1). It almost seems that no natural conversation is complete without the presence of PMs. If PMs are an integral part of speech, it can be hypothesised that PMs would appear even in higher levels of formality. Therefore,

the current study explores the use of PMs in speech used in YouTube videos limited to a formal to semi-formal public setting. The first research question is: **which PMs are used by SSLE speakers in semi-formal to formal conversations?** This study aims to provide an initial inventory of the range of SSLE PMs and their relative frequency, using the methodology detailed in this chapter.

A second research question concerns identifying the functions fulfilled by these PMs. Although the PMs are seen as non-semantic elements, a close analysis carried out in previous research confirms the important role they play in assuring a smooth verbal transaction and the subtle important manipulation they perform on the meaning of the utterance (Beeching, 2016). For instance, by hedging or toning it down, they mitigate an utterance that could be perceived as a strong opinion. Each PM found in the data was analysed for the specific function it may play in the context of that particular utterance. This leads to the second research question: **what are the functions of each of these PMs in the corpus?**

A third research question concerns the role the macro-social factor gender plays in PM use in SSLE. There is much research that suggests that there is a significant gender-based difference in the way speakers use PMs (Leimgruber et al., 2021; Zhang, 2024). Much of this research attributes a higher frequency use of PMs to female speakers than male speakers (Beeching, 2016, p.126; Coates, 2016, p.89). In a gender fluid world, although it is relatively difficult to pin such assertions to speakers based on their biological sex, it is indeed intriguing to read research that continues to make similar claims. Therefore, this study uses the biological sex as a variable and tests the distribution of PMs in SSLE according to gender. Thus, the third research question: **how does gender interact with PMs in terms of frequency and functions?**

A fourth research question investigates the relationship between age and PM use. Earlier research has suggested that age has no impact on the use of PMs (Holmes, 1986) but there is more recent work that found that PM use could indicate the age of the speaker (Reichelt, 2021). In light of these contrasting findings, it is worthwhile to investigate the impact of age on the production of PMs in SSLE and determine whether, as a dialect of Sri Lankan English which is a post-colonial variety, it

exhibits different behaviour to the inner circle varieties. Accordingly, the final research question: **how does age interact with PMs in terms of frequency and functions?**

4.3 Data

Many of the previous PM studies have been conducted on corpora built for general language analysis purposes. The research questions of a study are very important in deciding the corpus that should be used. As PMs are primarily found in spoken contexts rather than in written contexts (Brinton, 1996), it is more suitable to carry out investigations of PMs in SSLE on spoken language corpora to ensure rich results. The research questions also concerned observing PM use in a high prestige dialect of SLE as that is the dialect that is followed in the ESL classroom. Additionally, as perception studies on PMs reveal, PMs are discouraged by prescriptivist views of use (Miskovic-Lukovic, 2009). Prescriptivist views are mostly adhered to by speakers when they use language in the public sphere. Therefore, my study also explores the presence and use of PMs in spoken SSLE in public domains. In other words, my study investigates the public face of standard SSLE in terms of the PM use. Those PMs which emerge even in this more constrained formal environment are likely to be indicative of PMs that are particularly entrenched and widely used in SSLE. Therefore, ideally, the data should be from speakers using SSLE in a formal public platform.

There are only limited options for corpora of Sri Lankan English broadly. There are seven locally compiled Sri Lankan English corpora of written texts, such as Sri Lanka English Newspapers Corpus (SLENC) with 31.8M words, Arts Academic Corpus with 100,000 words and South Asian Varieties of English (SAVE2020) with 18M words hosted in the Digital Humanities Lab of the University of Colombo¹³. The only existing corpus of Sri Lankan English spoken language which is widely known is the International Corpus of English-Sri Lanka (ICE-SL), consisting of 1 million words.

¹³ A list of corpora of Sri Lankan Englishes available at the Digital Humanities lab at the University of Colombo is provided in Karunanayake et al. (2022). Pedagogical and research potential of a digital humanities laboratory: Language corpora as open educational resources (OER), *University of Colombo Reviews (Series III)*, 3(2), 5-18. The website of the lab: [AHEAD-DOR: DH Lab | Department of English \(cmb.ac.lk\)](https://www.cmb.ac.lk/AHEAD-DOR/DH-Lab/)

ICE-SL has a written component as well as a spoken component. The spoken section of the ICE-SL corpus consists of 600,000 words divided into three hundred 2000-word texts, collected from diverse contexts (Bernaisch, Mendis & Mukherjee, 2019). There are 180 dialogues and 120 monologues. The 180 dialogues in the spoken component contain 100 private conversations and 80 public speech samples. The private domain speech samples include 90 face-to-face interviews and 10 telephone conversations. The 80 public domain tests include classroom lessons (20 texts), broadcast discussions (20 texts), broadcast interviews (10 texts), parliamentary debates (10 texts), legal cross-examinations (10 texts), and business transactions (10 texts). The 120 monologues contain spontaneous commentaries (20 texts), unscripted speeches (30 texts), demonstrations (10 texts), legal presentations (10 texts), broadcast news (20 texts), broadcast talks (10 texts), and non-broadcast talks (10 texts). This corpus includes speakers who are above 18 using SSLE. Most of the speakers have a bachelor's degree or above education qualifications and those who don't have a degree were selected based on the features of SSLE as identified by the corpus compilers.

This corpus has not been used for the present study for several reasons. Although the ICE-SL corpus represents texts from one dialect, from speakers with a similar education level and has texts from a variety of contexts, it is not suitable to answer the research questions of my study. There are only 30 texts that are suitable in ICE-SL for my study given the criteria it requires. My study requires SSLE samples in dialogues from a formal public domain to show that PMs are an integral part of language use in SSLE even in higher levels of formality, and to explore how PMs are used in the public domain. The monologues in ICE-SL are omitted as the study explores the functions of PMs between interlocutors, which leaves the 80 public speech samples as potential candidates for inclusion in the study. However, the metadata¹⁴ for gender and age is not available for most of the public domain data. For example, the 20 texts of broadcast discussions (S1B-021 - 040 texts in ICE-SL), 10 texts of

¹⁴ Metadata is the sociobiographical information about the contributors in ICE-SL. Information includes data such as the year of recording, gender, year of birth, occupation, date of moving to Sri Lanka, stays abroad (> six months), ethnicity, educational background, home language, additional language, and information pertaining to linguistic habits and surroundings such as exposure to BrE media and AmE media.

broadcast interviews (S1B-041 - 050 texts in ICE-SL), 10 texts of parliamentary debates (S1B-051 - 060 in ICE-SL), and the 10 texts of legal cross-examinations (S1B-061 – 070) do not have data on the age of the speakers. Out of the 80 public speech texts, metadata for gender and age is available only for the 20 texts of classroom lessons and 10 business transactions. These 30 texts amount to 60,000 words only, which, despite the fact that PMs are a relatively high-frequency item in speech, is not a sufficient amount to guarantee good coverage and avoid issues of data sparseness. Therefore, ICE-SL proved to be unsuitable to answer the research questions in my study.

My study needed controlled data that is as uniform as possible in a number of respects, such as the dialect of SLE, the formality level of the discourse situation, the number of speakers, equal gender groups and comparable age groups to get a clear insight into the use of PMs in the public face of SSLE. Considering all these factors, a bespoke corpus was created for this research. Therefore, I collected data from an equal number of female and male speakers who used the dialect of SSLE in semi-formal to formal interviews created for public dissemination, available on YouTube at the point of data collection. Several interviews were exclusively created as content for YouTube such as C03AF, C05AF, C07AM and C08AM. There are also videos of interviews recorded for television broadcasting at an earlier date but later uploaded to YouTube, including C02AcM, C07AcM, C08AcM and C09AcM. None of the videos had any live audience watching the interviews as they were being recorded.

Although the speaker numbers were gender balanced, the range of speakers invited for interview made it impossible to precisely balance for age as well. However, the data collected fits broadly into three comparable social generations: the Baby Boomers (1946-196), Generation X (1965-1980) and Generation Y (1981-1996). Moreover, conversations between only two speakers were selected so that it would be easier to understand and interpret the functions of PMs clearly in addition to relatively simplifying the task of transcription. Interpersonal cohesion is a functional area that is fulfilled by PMs and therefore, conversation between two speakers rather than a larger group

of 5 or 10 speakers would provide suitable data on how PMs manipulate or affect the interlocutor relationship.

As discussed in Section 2.8, there are several dialects of SLE spoken in Sri Lanka. The dialects recognised in the research literature range from the SSLE to a non-standard variety of SLE, derogatorily known as *not pot English* (Gunsekera, 2005, p. 113; Wijetunga, 2008, p. 3; see further Section 2.8.2). This study elected to explore the dialect of SSLE since it is the dialect that is advocated and taught in ESL classrooms in Sri Lanka irrespective of the elusiveness of the definition and the contentious form of the variety (Samarakkody & Braine, 2024, p. 152). Being an ESL practitioner myself, it is important that insights into SSLE are revealed through empirical data so that the features of the dialect are better understood. It has also been noted in my two decades of ESL teaching experience that no PMs are overtly taught despite their prevalence in natural speech.

When the dialect was selected, the source of data was considered. PMs are mostly seen in naturally occurring dialogues although they are present to a lesser degree and with limited functions in monologues as well (Bogdanova-Beglarian et al. 2020; Han, 2011; Östman, 1982). Hansen (1998, p.52) describes several dimensions that define natural conversations. The dimensions refer to the extent to which the communication is made public, the degree of intimacy between interlocutors, the degree of emotional involvement on the part of the interlocutors, the extent to which the communicative process depends on the situational context, referential immediacy versus referential distance, physical closeness versus distance (in both space and time) between interlocutors, degree of cooperation between interlocutors, degree of dialogicity or how much one is connected to others through dialogue, degree of spontaneity and the degree of topic fixation.

We can assume that the recorded data will differ from non-recorded, spontaneous speech in several respects according to these dimensions. For example, the recorded communication is more public than a conversation that is not recorded. The degree of intimacy is less than that between friends or acquaintances who are not being recorded. Even the emotional involvement will be less

than in an unrecorded conversation. The topic fixation too would restrict the spontaneity of the conversation. All these changes from a natural conversation between familiar speakers that is not recorded will affect the PM production in recorded data as well. For instance, there might be less use of *you know* or *like* for word search or as fillers because of the speakers' awareness of the recording. Speakers are generally advised to use fewer PM as fillers in public speech: for example, Zandon¹⁵ (2018) refers to PMs as crutch words and mentions that they lower the speaker's impact in public speech. With regard to the use of *kind of* and *sort of*, Miskovic-Lukovic (2009, p. 602) mentions that "more often than not these forms have borne the stigma of bad usage through prescriptivism and bias". Given the information from such perception studies, it is expected that the production of PMs will be affected negatively in the data considered for my study. Therefore, in this case, any presence of PMs in the data in my study also confirms how much of an integral part PMs play in SSLE.

4.4 Research Design

A collection of speech samples was required to fulfil the aim of the study which is to understand the use of PMs in spoken English by SSLE speakers. A survey of recent work on PMs shows that corpus exploration is a very common approach to researching PMs (Beeching, 2016; Unuabonah & Muro, 2022; Westphal, 2024) as this allows access to authentic data, on a scale that is likely to ensure a reasonable degree of thoroughness. Therefore, my study too benefits from a corpus approach. Further, corpus data allows the investigation of many aspects such as pragmatic marker functions, functional counterparts between languages, universality and the dialogistic positions, i.e. the way the speaker expresses an attitude towards the propositional meaning through PMs (Fraser, 2006; Furkó, 2017; Stenström, 2006; Varga, 2020). A corpus-based study provides quantitative information about a set of data. Additionally, the context surrounding each instance supports the interpretation of the data to understand if PMs fulfil the same functions described in Section 3.6, such as solidarity building and

¹⁵ N. Zandan is CEO and co-founder of Quantified Communications, a firm that combines data and behavioural analytics to help people change the way they speak. He claims in Harvard Business Review (2018) that his company works globally with leaders of corporations, government organisations, higher education institutes and hundreds of TED speakers

finding common ground, allowing turn taking in interactional conversations, achieving politeness by hedging and downtoning, or working as social markers of a particular social group. Using corpus methodology, the researcher has access to both quantitative and qualitative perspectives on data through the combination of large datasets and broad contextual information.

The current study is underpinned by an initial core quantitative phase, during which the presence of each PM was identified and its frequency recorded. These instances were then further categorised and quantified according to the macro-social categories of gender and age for each PM. To achieve this, first, all the transcripts were read manually. Then, the study used manual sifting to survey PMs in Sri Lankan English. Manual sifting in my study involved reading all the transcriptions in their entirety on paper and marking instances of PMs manually in the transcripts while listening to the videos before the coding process began. Listening to the videos and understanding the context helped to decide if an expression such as *you know* or *I mean* is used as a PM or in its canonical use. Then, the corpus was uploaded to AntConc software (Anthony, 2024) to verify the quantitative manual analysis. Every PM was searched using the concordancing tool. AntConc provided a list of all searched PMs irrespective of their PM or canonical use. However, this electronic search verified that the manual search had not overlooked any instances of PM use. This manual and electronic quantitative analysis of the data was used to understand basic information on forms, frequency and patterns of use. The basic information includes how many PMs are used by the speakers in the data, how much is the use as a percentage of the total language used in the transcript, how many PMs are used by speakers of a certain gender or a certain age group.

This quantitative approach was then combined with a variationist approach to interpret the numerical data qualitatively. The variationist approach mainly refers to the methods developed by Labov (1966) which helped in investigating urban dialects in the United States. It assumes that particular language use is embodied in societies, communities and individuals. A variable, PMs in the case of this study, shows a relationship with some social factor such as class, age or gender. More

details about the variationist approach are provided in Section 4.7. The analysis also draws on my own knowledge of the community and its norms to interpret the data.

A study of this nature requires a combination of methods to gain a more complete picture of the data and results. Such a mixed method approach is suitable to see the relevance of sociolinguistic factors such as gender and age on the stylistic preferences of the users of PMs and how different groups or individuals attain their communication goals.

4.5 Designing and Creating the Corpus of Spoken Sri Lankan English (CSSLE)

A research-specific corpus can be very convenient to observe the aspect of language under study. Davies and Fuchs (2015) mention that proprietary or ad-hoc corpora provide extremely useful and insightful data while noting the possible downside that they may not be as useful for a wider range of researchers. What is important is that the corpora should meet the research needs of a particular study in terms of data, size, genre and accuracy (Davies & Fuchs, 2015).

The optimal size for the ad hoc corpus of this study was carefully considered from the outset. Lourerio-Porto (2017) cautions the researchers against selecting corpora based on size but based on the research aims of a study and stresses the point that carefully built small corpora could be more reliable than big-data-based corpora to research certain genres or lexical items. The key point is that it is not the size of a corpus that matters but how much the data is representative of the research object of a study. The idea of representativeness has been debated for decades and there is little consensus around it. In fact, it is claimed that a representative corpus is never possible (Egbert, Biber & Gray; 2022). Egbert, Biber & Gray (2022) state that in direct contrast to the beliefs of some of the corpus linguists, “statisticians have often noted that the term “representative” is not a precise statistical concept and thus it needs to be operationally specified” (p 29). For instance, although PMs are a phenomenon prevalent in dialogues, they cannot be reliably observed in dialogues in fictional literature even though these dialogues may appear to be natural and life-like, because research has shown that such texts differ considerably from natural everyday speech (Jucker, 2021). Considering

the point made by Lourerio-Porto (2017) with regard to size and representativeness, and Biber & Gray (2022)'s point regarding operationally specified representativeness, the corpus for the current study consists of naturally occurring unscripted speech by speakers of SSLE produced in a conversation between two speakers performed at a public platform.

In addition to the factor of representativeness, the size of a corpus is decided by practicality as well (Reppen, 2010). The data collection process was scheduled to begin when the Covid-19 pandemic started, making it impossible to gain access to naturally occurring data recorded in Sri Lanka especially for this study. The most appropriate replacement was found to be data already present online. Schneider (2016, p. 278) comments that YouTube videos are a "treasure trove" of great resources to study world Englishes. They are a text-type of their own and they present people's perceptions and what people do with language. He further mentions that they are a valuable resource for linguists because they mostly contain "community oriented", "fairly vernacular and localised speech" (Schneider; 2016, p. 259). They contain great examples of how language is used within a community. However, there are limitations to using online data, especially from a sociolinguistic perspective, given the lack of metalinguistic information about the speakers. For instance, there is generally no metadata about the speaker age, social class or any other demographic detail overtly available in online videos. These limitations notwithstanding, the mere accessibility of online videos and their sheer range makes them a great choice for language exploration. They can also be easily downloaded and stored.

Thus, the study explored YouTube online spoken data featuring Sri Lankans speaking English. The following criteria were used in selecting the online videos and in selecting the data for analysis.

- The content is interviews between two participants

- The interviews are face to face interactions or Zoom meetings where speakers had recorded the meetings¹⁶
- The structure of the interview may range from semi-formal to formal
- The interviewee should be either an artist, an entrepreneur or an academic. If an interviewee could be classified as having two occupations (for example, one participant was both an actor and an entrepreneur), the occupation which he/she was most prominently associated with was selected.
- All videos are on publicly accessible online platforms such as YouTube and posted as public videos with standard license which allows fair use.

Although the corpus was designed to capture a variety of speakers, to minimize the variability of demographic differences that may affect the data, it was decided to select speakers from three professional areas only, viz. the arts, the academia and business. Videos from these three categories of profession were readily available online at the point of data collection. However, I do not consider profession to be an index that directly affects the use of SSLE, although occupations are generally linked to socio economic indices which determine language use (Chambers, 2009, p.46). If I may apply the knowledge of the Sri Lankan context as an insider, I would comment that it is difficult to gauge if your profession shapes your language or if you secured a certain job partly because of the variety of English you speak. The variety of SSLE could be a factor that gives an edge at job interviews. In fact, youth employability is closely linked to speaking English with a certain accent and Amarasuriya (2010) records this variety as 'British accent'. What Amarasuriya refers as 'British accent' is the variety spoken by the affluent members of the society, which is SSLE (Gunesekera, 2005). Amarasuriya herself refers to Gunesekera (2005) to make this distinction. However, occupation may not be an estimator of the variety of SLE they use, because in reality, individuals within one profession can have varying levels of SLE proficiency.

¹⁶ The Zoom interviews uploaded to YouTube had been recorded during the pandemic in 2020 and 2021. For example, the script C04AM is from a programme titled *Lockdown Diaries*.

To minimise any unnecessary variability in the data, speakers were restricted to three occupations: academics, artists and entrepreneurs. There were two newly popular online channels named *pulse.lk* and *hi!! online* which contained series of interviews with famous artists, academics and entrepreneurs. One example of a series is *Anything but with Kumar de Silva*, broadcast by *pulse.lk* YouTube channel in which mainly artists were interviewed, and another was called *The Founder* in which only entrepreneurs were hosted. As mentioned in Section 4.3, most of the videos that were selected had been produced specifically to be aired on YouTube, while there are others that had been produced as TV shows and later uploaded onto YouTube.

The videos that were selected had been published between 2012 – 2021. Ethical clearance for using the videos in this research was obtained from the Research Ethics Service, at University College London. This study was exempt by the UCL Research Ethics Committee, and the approval is attached as Appendix A. All the videos selected for the study had been produced under YouTube Standard License except for one. The exception was a video produced for a Sri Lankan TV channel called YA TV and it had been later uploaded to the Vimeo website. This video does not need the copyright holder's permission as the website allows videos to be used for fair use that includes nonprofit educational uses. A list of the videos with weblinks and license type is given in Appendix B. Additionally, Appendix B includes information about the year of publication and the publisher for each video.

The interviews featured in these videos had been recorded at various venues to be exclusively broadcast on YouTube. For instance, the videos for *The Founder* series had been recorded in an office space dedicated to the show. Some of the interviews of artists such as <03AM>, <05AM> and <09AM> on *pulse.lk* were recorded in the residences of the interviewees. Six YouTube videos for the sub corpus of academics were selected from a series called *Sri Lankan Understanding on hi!! online* YouTube channel and they were recorded in a studio. Videos such as C04AcM, C07AcM and C08AcM, recorded in a professional studio, were first broadcast on a private media television channel and later uploaded to the YouTube by the same private media company.

Table 4.1 provides information about the sources of the YouTube videos. It gives details of the YouTube channel publisher and the number of videos selected to compile data for one of the sub corpora viz. the artists. The length of the videos given in Table 4.1 is a calculation of all videos featuring artists. The total length of videos included for this data set is 5 hours 54 minutes and 37 seconds long. The relevant video from the corpus of artists is stated in the fourth column. The video titles are coded, with each letter and number denoting information. For example, the breakdown of C04AF is C = conversation, 04 = fourth video in the corpus of artists, A = artist, and F = female speaker. Similarly, for example in C04AcF and C04EF, AcF refers to academics and E refers to entrepreneurs. In videos where male speakers are featured, the title is coded as C04AcM. M stands for male. All the links for the videos are given in Appendix C.

Name of the YouTube Channel	Number of videos selected	Length of the videos in total (in hr: min: sec)	The artist video
Hi!! Online	01	00:15:31	C04AF
Mind Adventures Theatre Co.	01	00:18:51	C02AF
Pulse.lk	22	05:20:15	01AF, C03AF, C05AF, C06AF, C07AF, C08AF, C09AF, C10AF, C11AF, C12AF and C01AM to C12AM
Total	24	05:54:37	

Table 4.1: Sample sources and length of the YouTube videos featuring artists

The speech of the interviewer/host was excluded from the analysis to avoid over-representation of one speaker. For instance, several YouTube videos from one talk show series had the same interviewer. If the interviewer's language was selected in all those interviews, the data would be skewed. Moreover, it would affect the gender balance of the corpus.

The size of the corpus had to strike a balance between ensuring enough data for meaningful analysis, representing both genders equally, and the manageability of having a single researcher building it and transcribing the data. This was achieved by including 24 speakers for each of the three occupation groups, equally split between 12 male and 12 female speakers, for a total of 72 speakers. Speakers between 25 to 81 years were selected to facilitate the analysis with regard to how age interacts with the PMs. The study observed speakers from three social generations but, as explained in Section 4.3, the criterion to have equal numbers of speakers in each social generation could not be met in this data set.

I consciously tried to maintain ethnic diversity in the data to make the corpus representative of the population of Sri Lanka. I included speakers from the country's four main ethnic groups, which are the Sinhalese, the Tamils, the Muslims and the Burghers. However, it was not possible to find a perfect balance in the representation of the four ethnic groups in the available data. The data is dominated by the Sinhalese (51 speakers-71% of the total participants). There are nine (12.5%) Tamil speakers, eight (11.1%) Muslim speakers, and four (5.5%) Burgher speakers. This could be considered a close approximation of the demographics of Sri Lanka which reports in 2011 census a Sinhala population of 74.9%, Tamil 11.1%, Muslim and Malay 9.52% and 0.19% for Burghers. A table with a breakdown of the speakers according to their ethnicity is given in Appendix D. As stated in Section 2.8.1, the phonological features of SSLE are not influenced by ethnicity. Therefore, the imbalanced representation of ethnicity is not a limitation for the results of the current study as it observes a dialect of SSLE that does not depend on ethnicity.

The YouTube videos selected for the data are all professional interviews but they range in a continuum of formal to semi-formal. Mainly, the greetings and introductions between the interlocutors mark formality. For instance, in the interview C08Ac the interviewer <04HM> greets the interviewee <08AcF> by stating that he is honoured to have her in the discussion. Additionally, <04HM> mentions a list of professional and educational qualifications as in “the Head of the Department of International Relations at the University of Colombo the university from which she earned her first degree her master's is from the Notre Dame University and her PhD in Conflict Analysis and Resolution is from the George Mason University she's also a published author”. This could be partially to inform the audience that she is qualified to comment on the subject of discussion. Mentioning formal qualifications to introduce a person signals a formal interview. In semi-formal interviews like C03AF, the interviewer greets the interviewee with a more personal introduction. For instance, the interviewer <01HM> states “my guest today on this special aluth avurudu¹⁷ edition is a little girl I knew as a baby” and this introduction is followed by the guest’s name and “hi”. This introduction is casual, friendly and semi-formal. The reference to knowing the guest as a baby indicates informality and the casual greeting *hi* confirms that it is an informal introduction. Therefore, the introductions and greetings are more personal in semi-formal interviews than formal interviews. Secondly, the type of questions that are asked mark formality. The formal interviews consist of subject specific questions such as “why do we talk about power sharing and also attach controversy to it” (C05AcM) or “how do you assess the legacy of colonialism and how has it impacted on the debates that we have about state reform today” (C03AcM). Semi-formal interviews contain more personal questions such as “what was life like growing up in Nawala with your sister Umi and your parents er Janaki and Anu” (07AF) or “take us back to your childhood what was life like here in Wellawaththa in the good old days” (in 06EM). Thirdly, this demarcation is visible in the venue that the video is recorded as well. The highly formal interviews are recorded in a

¹⁷ Sri Lankan English morphological borrowing meaning *new year*

dedicated studio, while semi-formal interviews were recorded either in the interviewee's residence or in a café. This variation could be levelled off perhaps by the idea that the speakers were conscious of being recorded and thus were perhaps more mindful of their speech irrespective of the nature of the interview (Sarangi, 2007). It also could be argued that in any interview, the participants are likely to relax to some extent and to become accustomed to the recording environment.

Several decisions with regard to what should be included in the analysis of data were taken. As explained above, although the language of the host is included in the transcription, it was excluded from the analysis to maintain the gender balance and avoid over-representation of particular speakers; 24 videos in the data have the same person as the host/interviewer and had he been included in the data, his idiosyncratic use of PMs may have influenced the results and overall conclusions. Another decision was the length of the text. The online videos of Standard Sri Lankan English speakers were of different lengths. A pre-determined number of words was selected for each text to maintain uniformity across the corpus. All scripts, with the exception of seven, had an average of 2750 words. Therefore, 2750 words was selected as the length of a text in the corpora. The number of words in the scripts included language from both the interlocutors. Seven scripts fell slightly below this threshold, with an average of 2512 words, but were nonetheless included as they represented speakers whose ethnicity was vital to the study or because they represented an occupation with limited online interviews.

4.5.1 The Speakers

As previously explained, the study focuses on SSLE speakers as that is the variety of SLE that has perceived privileges and is disseminated in the ELS classrooms, especially at universities. My sample is a judgement sample: a non-random sample selected from predetermined demography or practically with any predetermined social attribute. The fundamental sampling criterion for this judgement sample was that the participants had to be speakers of Standard Sri Lankan English (SSLE).

The twenty-four artists include eight actors, two dramatists, two dancers, eight singers, two pianists, one designer and one author. Their ages range from 25 to 81 years. The twenty-four

academics are attached to state universities or semi-government/ independent research agencies. Their ages range from 34 to 74 years. The twenty-four entrepreneurs are all either owners of established companies, start-up companies or in higher management in an enterprise. Their ages range from 27 to 75 years. They all are speakers of Standard Sri Lankan English based on SSLE phonological features. The videos of all 72 speakers were checked by two acrolectal speakers of Sri Lankan English, both with research experience in Sri Lankan English, to ensure that the data represented SSLE, as outlined in the next section.

4.5.2 The Rating Process

The objective of the rating process was to ensure that all the speakers in the corpus could be described as speaking one dialect. ICE-SL followed similar criteria at the time of its compilation, although the ICE-SL team avoided the term standard. Instead, the manual to ICE-SL uses phrases such as “highly representative of and/or influential for Sri Lankan English”, “acrolectal speaker of English”, and the texts were “checked against the judgement of native speakers of Sri Lankan English” (Bernaisch et al. 2019, p.7 & 8). All these phrases more or less point to the idea that the texts come from a particular uniform accent which in certain literature is termed as the standard (Gunesequera, 2005).

The two raters of the speakers in the corpus are senior lecturers in Departments of English in two state universities of Sri Lanka. They have substantial experience in teaching and researching Sri Lankan English, one with 17 years of service in academia and the other 27. They were provided a list of SSLE phonological features as reported in empirical research to guide the selection of the speakers. This checklist of phonological features is given in Appendix E. For example, the speaker should not use the phonemes [o] and [ɔ] in free variation, so that *gold* is pronounced as [go:ld] and not [gɔ:ld]. Similarly, *gone* is pronounced as [gɔ:n] and not [go:n]. This phonetic feature is a primary factor in distinguishing between SSLE and Non-Standard Sri Lankan English (NSSLE) (Gunesequera, 2005). The raters were asked to watch a ten-minute clip from any starting point in the selected 72 videos and confirm whether the speakers can be classified as SSLE speakers. Where a speaker

showed an inconsistent and/or minimal deviation from any of these features, I asked the raters to use their professional judgement to either accept the speaker as an SSLE speaker or not. They agreed on all videos except one, which was then replaced by another qualifying video. The raters confirmed that all 72 speakers in the corpus are SSLE speakers.

4.5.3 The Transcription of Data

As the data collected only existed in video form, it had to be transcribed before it could be subjected to analysis. The 72 transcriptions were made by me over a period of one year. A transcription had to recreate the context of the spoken words. There are many problems with transcribing spoken data. On the one hand, a transcription cannot realistically capture the whole picture of a spoken interaction. There are the words primarily, but there are also prosodic markers, gestures and even unspoken understanding that goes between two interlocutors that is merely conveyed by a look. On the other, one can include unlimited information in a transcription. Cook (1990) summarises this issue when he says that what can be included in a transcription is potentially infinite. He further mentions that any amount of contextual data can be included in a transcript. One can mention information from the intonation of the speaker to what he or she was wearing. Hence, a very difficult negotiation of information has to be made in the process of transcribing. At the same time, a difficult reality is that the analysis of this kind of data is only as good as the transcription. Bearing this dilemma in mind, I had to make a few practical decisions that would impact my study. Mainly, the transcription should include only information that is required to provide answers to the research questions of the study. Therefore, it includes pertinent features that help in analysing the functions of PMs. For example, features such as disfluencies or false starts are important as PMs are generally used to repair them. Another important decision made is selecting the transcription norm. It was important to bear in mind that the required information and the transcription norm should not require an extremely detailed or time-consuming process which will affect the timelines of the study but should result in a transcription that will not compromise a thorough analysis of the PMs.

Three transcription conventions used in three corpora were studied: British National Corpus 2014 (BNC2014), Louvain International Database of Spoken English Interlanguage (LINDSEI), and Systems of Pragmatic Annotation in the Spoken Component of ICE-Ireland (SPICE Ireland). BNC 2014 is an 11.5-million-word corpus of conversations among members of UK public recorded between 2012- 2016 (Love et al, 2018). The content of the present study also entails conversations and therefore it was useful to closely study the transcription conventions of BNC2014. Louvain International Database of Spoken English Interlanguage (LINDSEI) contained spoken data collected from advanced learners of English with different first language backgrounds (Granger et al., 2020). The present study similarly examines speakers with multilingual backgrounds. For example, as Gunsekera (2005) states, all SSLE speakers are bilinguals or multilinguals. Therefore, it was judged that LINDSEI shares similar speaker profiles with the current study. SPICE Ireland contains both formal and non-formal conversations similar to the current study (Kirk et al., 2007). Therefore, SPICE Ireland scheme for transcription was helpful for the present study. After a close examination of all three transcription conventions, it was decided that BNC 2014 transcription convention was the most appropriate for my study.

4.5.3.1 British National Corpus 2014

BNC2014 is a spoken corpus compiled by the Economic and Social Research Council (ESRC) funded Centre for Corpus Approaches to Social Science (CASS) at Lancaster University together with Cambridge University Press. It is a 11.5-million-word corpus of informal conversations of L1 British speakers recorded between 2012 - 2016. It features 668 speakers in 1,251 recordings. BNC2014 is a further development of BNC 1994 which is considered the most widely used corpus to date (Love et al., 2017).

The transcription norms are shaped by the recommendations made by Atkins et al. (1992) where they state that script type transcriptions are sufficient for linguistic analysis. BNC2014 follows Atkins et al.'s (1992) suggestions in terms of speaker change, inaudible segments, alternative written

forms for numbers and abbreviations, and permissible forms for spelling deviations representing dialects. Another notable recommendation is to create a large set of orthographic representations to transcribe functional and non-functional sounds to minimize inter-transcriber variability. Appendix F of the BNC2014 manual (Love et al., 2017) gives a complete list of all transcription norms. The general guidelines instruct that the transcription has to be a direct copy of the recording. For example, the transcriber should not correct or paraphrase any grammatical inaccuracies in the speech samples.

The most salient transcription norms used in the transcripts are thus presented. Each speaker is given a unique number such as <022>. If the ID of the speaker is ambiguous, it is marked with ? as in <022?>. Many people speaking together is marked as <MANY>. All speakers were anonymised but had a gender tag. For instance, Dave becomes Name <M>. Similarly, place names become <place>. However, places which are so general like France or London, and that would not lead to identifying the speaker are not changed. Names of famous people like David Cameron are also not changed. The speech is transcribed as utterances rather than sentences. Therefore, units neither start with a capital letter nor end with a full stop. If the speaker asks a question, a question mark is used but it is not followed by a capital letter. No punctuation marks are used with the two exceptions of question marks and hyphens for proper names and numbers. Only the pauses that occur within an utterance are recorded. Any pause at the beginning of an utterance is ignored. A short pause between 1 second and 5 seconds is marked as (.) and a long pause longer than 5 seconds is marked as (..). In unfinished sentences or false starts an equal sign (=) is used to mark where a word is finished. An example is <001> yes he is a ba=bachelor. Overlaps are marked as . When a speaker is unintelligible, the transcriber guesses the words and indicate thus: <u=GUESSEDWORDS> (where GUESSEDWORDS should be replaced by the guess). If a guess cannot be made, <u=?> is used. Filled pauses such as *ums* and *errs* are given set spellings. Although there is dialect and accent variation among speakers, only standard spelling is used with the exception of the word *innit*. Non-verbal vocalisations such as coughing, laughter etc. are marked with square brackets as in [laugh]. If

there is any event is audible and relevant to the recording but that is not produced by the speakers, they are represented by [e=SOMETHING]. An example is [e=sound of phone].

4.5.4 Justification for Using BNC 2014 Transcription Conventions

A comprehensive yet uncomplicated transcription scheme is required for the present study. LINDSEI, SPICE Ireland and BNC 2014 use comprehensive transcription conventions suitable for the purpose of the respective corpora. However, BNC 2014 stood out as the most appropriate. For instance, it is orthographically transcribed which allows searching for patterns in concordance form (Crowdy, 1995). Another reason for selecting BNC 2014 conventions was that it minimised transcriber inference. For example, the BNC 2014 transcription did not include any non-verbal signals which would take more time to transcribe. Such features would challenge the consistency of the transcription as well. In order to check which transcription method is the least time consuming as time was a constraint of the current study, a voice clip of 4 mins and 26 seconds was transcribed in all three conventions. BNC 2014 was the fastest transcription at one hour and 54 minutes.

The main reason for a speedy transcription was the reduced specificities in BNC 2014. For example, LINDSEI tailored its convention to suit certain pronunciation variations, such as phonetic features like syllable lengthening and the articles mentioned above. In SPICE Ireland conventions, a short pause is termed as “any perceptible break in phonation equal in length to one syllable, uttered at the speaker’s tempo” and a long pause is “any longer break in phonation” (Nelson, Wallis, Aarts, 2002, p. 12; Kallen & Kirk, 2012). BNC 2014 does not have such specificities. As BNC 2014 follows an orthographic transcription with fewer details on transcription norms for phonetic features and pauses it is ideal for research that does not need to examine pronunciation closely. With these features, the BNC 2014 transcription conventions were sufficient to analyse the data to answer the research questions. Appendix G is a sample transcript of the present study based on the transcription norms adapted for the current study.

4.5.5 Size of the Corpus

The corpus of Standard Sri Lankan English (CSSLE) created for the purposes of this study has 202,557 words. Sub-corpora can be derived from the overall corpus based on the variables of occupation, gender and age, as summarised in Table 4.2. There are three sub-corpora of similar sizes, for the three occupations considered: artists (68,297 words), academics (67,546 words) and entrepreneurs (66,714). There are two gender-based sub corpora, which make it possible to analyse the effect of gender on the use of PMs. The corpus of female speech has 100,837 words and the corpus of male speech has 101,720 words. Finally, sub-corpora based on age were created to analyse the patterns of PM use according to age. These are based on the three social generations that correspond with the speakers' age; in this case, it was not possible to precisely balance the number of speakers belonging to each generation. The Baby Boomers corpus has 79,293 words, the Generation X corpus has 73,024 words and the Generation Y corpus has 50,240 words.

Name of the corpus	Number of texts	Word count
Corpus of Artists	24	68,297
Corpus of Academics	24	67,546
Corpus of Entrepreneurs	24	66,714
Corpus of Females	36	100,837
Corpus of Males	36	101720
Generation BB	28	79,293
Generation X	26	73,024
Generation Y	18	50,240

Table 4.2: Overview of sub-corpora of CSSLE

Spoken corpora are generally smaller than written corpora due to the difficulties in collecting and transcribing data (McCarthy, 1998). Murphy (2010) mentions that smaller corpora are better than larger corpora for research on high frequency items. High frequency data such as PMs can be

captured even in a small corpus. Less frequent items can only be researched using large corpora. A small corpus has more manageable data for a researcher than handling a million-word corpus. Therefore, the size of the current corpus allows me to analyse the data deeply and provide a reasonable picture of the PM use among SSLE speakers in Sri Lanka. Moreover, the corpus size allows me to manually sift through data and look at individual contexts to see how PMs are used. The corpus therefore suited the aims of the study, and facilitated a deeper and detailed qualitative analysis of how the social factors gender and age affect the use of PMs.

4.5.6 Providing Identification to the Data

The 72 transcriptions in the corpus are each given a unique five-digit code, denoting information about the identity of the transcription. Three examples from the three main occupation groups are provided to understand the code used for the speakers.

C05AF – C (conversation) / 05 (the sequence of the video among 12 videos) / A (the occupation of the speaker- in this case artist) / F (female)

C02AcF - C (conversation) / 02 (the sequence of the video among 12 videos) / AcF (the occupation of the speaker- in this case an academic) / F (Female)

C05EM – C (conversation) / 05 (the sequence of the video among 12 videos) / E (the occupation of the speaker- in this case an entrepreneur) / M (Male)

In sum, C05AF is the transcript of the 5th video which contains a conversation with a female artist. Information with regard to age was not included in the coding as at coding stage, age-related information was not available for all the speakers. In addition to the text ID, the speakers too are assigned an ID. The speaker ID is simply a four- digit number which corresponds with the transcript ID. For example, the interviewee's speaker ID number is the same as the text sequence number. All anonymized speakers are gender tagged. Therefore, at the end of the numerals, either M or F is added to indicated male or female gender. An example would be <01EF>. This is the interviewee in video 1 of female entrepreneurs. Speaker ID is consistent and unique to each speaker and in the

instances where the same speaker appears in more than one video, for example, the interviewer who may appear in a series of online videos, one speaker ID is used in all instances to ensure consistency.

The corpus is stored as plain text files. After coding the transcripts and the speakers IDs, a header was provided for every transcript. For example, a header contains the transcript ID, then the two names of the interlocutors with their speaker ID. As the videos are published in a public platform for public viewing, there is no ethical demand to anonymise the names of the participants. Yet, the actual names of the interlocutors were retained only in the header for easy reference. The remaining text carries only the speaker ID at every utterance turn and not their actual names. The speaker IDs facilitate easy access to some of the meta data about each speaker. This is useful in interpreting the data. For example, immediately it is evident if it is a female or a male delivering a certain utterance. This corpus does not contain mark up, except for the socio demographic details such as occupation and gender given in the header.

4.5.7 Additional Remarks about the Transcriptions

Highly detailed transcription conventions can lead to lower inter-transcriber reliability. A good transcription depends as little as possible on the inference of the transcriber. The transcriber shouldn't have to be guessing as to which symbol should be used from the convention system. For instance, when marking filled pauses, if only a few spelling choices are given, it is easy for the transcriber to select one choice. Therefore, a simple closed set of conventions for backchanneling and pausing such as the one adopted by BNC 2014 is sufficient and suitable for the purposes of the current study. As noted in Section 4.5.4, the BNC 2014 was considered the best transcription convention system to use in the present study. However, time constraints compelled further restrictions to be applied on the number of details to be transcribed. For example, unfilled pauses are not marked in the transcriptions in this study. Nevertheless, since the corpus size allowed me to manually look at each occurrence of a PM when I had to code them for their functions, I could also notice the importance of a pause for a PM, if it was the case. I revisited the video when in doubt.

Filled pauses were marked with a set of permissible forms. For example, all instances of fillers such as *erm*, *arm* and *urm* were assigned *erm* while *ah*, *er*, *ur* were all assigned *er*. The BNC 2014 uses minimal punctuation marks with the exception of question marks and hyphens. The present corpus excludes all punctuation, even these two marks, since not all questions are direct questions and deciding when exactly a question mark should be included required subjective inference. The orthography of the transcripts follows British English. In instances when the speaker pronounces words which had two distinct pronunciations without a spelling variation such as *director*, [daɪ'rektə] or [dɪ'rektə], the spellings in the transcript did not reflect the pronunciation.

Additionally, I decided to remove all that may be extraneous to the text, such as any sound outside what the speakers produce, like vehicles passing, other people talking in the background, music, and so on. I only included the sound of laughter as [laugh]. I also did not label any prosodic features such as pitch, phonological emphasis or accents. It should be restated that accent and pronunciation are largely uniform (apart from unavoidable idiosyncrasies specific to speakers) as the speakers use only SSLE. Transcribing data required much more than merely writing the spoken word (Cencini, 2002). It required listening intently for overlaps of speech between speakers, false starts, intelligible utterances and specific idiolects. The transcripts in the CSSLE included these features. In some cases, the transcription process was supported by the availability of auto-generated transcripts on the videos' YouTube pages. Fifty-eight videos out of the 72 had automated transcripts.

Transcribing was relatively faster for the videos which had auto generated transcripts, although these still required considerable editing, such as removing timestamps assigned to each line, marking speaker turns and applying the BNC2014 transcription norms. The automated text also had to be compared to the live conversation for accuracy, and there were many instances of inaccurate transcribing of place names and expressions due to unfamiliarity of place names and accent. For example, Table 4.3 is a comparison of one minute of auto-generated script with corrected auto-generated script. It is from transcript C04AcM which discusses power plants in Sri Lanka. The SSLE pronunciation is transcribed wrongly. Some of these are highlighted in bold. For example, the

words *that strategy* is auto-generated as *tongue and Shi*. Local place names such as *Trincomalee*, *Mawella*, and *Norochcholai* are transcribed as *trink to Mar to Nori*. All references to the personal pronoun *I* were generated as *i*. These are some examples of the adjustments that were manually made to the auto-generated scripts.

Auto-generated text	Manually corrected auto-generated text
<p>6:00 then it attracts a lot of protest so you can you know some of these protests are protests I was about to say that to you</p> <p>6:07 but Average Joe public doesn't really know about the costs and so on they</p> <p>6:14 unless news first has done an expose on it but no that's tongue and Shi but what</p> <p>6:20 I mean is that Joe public doesn't really know the cost they will look at things</p> <p>6:25 like oh the sound and it's going to be dirty and there's going to be do of dust and there their plantations their homes</p> <p>6:31 their daily lives will be disruptive so who's fueling these protests I can't say</p> <p>6:36 very clearly but definitely there are genuine protests and some some NOS file court cases yeah so it is up to the CB</p> <p>6:45 and the politici behind the the government structure to address these</p> <p>6:50 concerns rather than run away and cancel the power plants because cancelling the</p> <p>6:55 power plants is in fact the habit of various governments if you consider nor</p> <p>7:00 it has been the site has been moved from trink to Mar to Nori over the period 1987 to 1992 so</p>	<p><04AcM> then it attracts a lot of protest so you can you know some of these protesters</p> <p><14HM> but these protests are I was about to say that to you but average Joe Public doesn't really know about the costs and so on they unless News First has done an expose on it but the that strategy but what I mean is that Joe Public doesn't really know the cost they will look at things like oh the sound and it's going to be dirty there's gonna be lot of dust and their their plantation their homes their daily lives would be disruptive so who's fueling these protests</p> <p><04AcM> I can't say very clearly</p> <p><14HM> yeah</p> <p><04AcM> but definitely there are genuine protests</p> <p><14HM> mhm</p> <p><04AcM> and some some NGOs file court cases</p> <p><14HM> yeah</p> <p><04AcM> so it is up to the CEB and the politician behind the the government structure</p> <p><14HM> yeah</p> <p><04AcM> to address these concerns rather than run away and cancel the power plants</p> <p><14HM> mhm</p> <p><04AcM> because cancelling the power plants is in fact the habit of various governments if you consider Norechcholai it has been the site</p>

	has been moved from Trincomalee to Mawella to Norochcholai er over the period nineteen eighty seven to nineteen ninety two
--	---

Table 4.3: A comparison of the auto-generated script with the manually corrected auto-generated script

Three random transcripts out of the 72 were given to two raters to assess the consistency of transcribing. The transcripts that were checked are C05AM, C05AcM and C05EF. The raters are two university academics with a background in linguistics. They manually checked the transcript that I created against the video recording and the transcription convention reference. For example, they had marked any words that had been excluded by mistake in my transcription and any instances where the word order had been mixed. These errors were counted, and then Word Error Rate (WER) was calculated. The rate was less than 1% for all three transcriptions.

4.6 Variationist Approach

This study looks at the heterogeneity of the use of PMs among Standard Sri Lankan English speakers and how gender and age contribute to this diversity. The method of analysis uses a variationist approach. I selected a variationist approach because it combines language, social aspects and a quantitative and qualitative analysis. This combination allows data to be explored and the investigation of possible variation within a dialect due to the influence of social factors. The variationist approach is empirical and is based on authentic data. The data should be sufficient in terms of “quality, quantity, and representativity” (Poplack and Tagliamonte 2001, p. 88).

This method assumes that “the variation is inherent in the individual, the group, the community and beyond” (Tagliamonte, 2012, p. 21). It also recognises that there is no ideal speaker or a homogenous variety of language but a structured variability (Cameron, 1990). The word ‘structured’ refers to systematic variation that arises because of a social condition. For example, there are always two or more ways of expressing an idea. Weinreich, Labov, and Herzog (1968, p. 100-101), as mentioned in Tagliamonte (2012), state that “the key to a rational conception of

language change is the possibility of describing orderly differentiation.” This suggests that it is important to observe the differences between two language data sets to understand language change. For example, one group of speakers might use PMs differently to another group of speakers in the corpus, and this difference could depend on a social factor. By defining the use of a language feature among groups of speakers from a variety of social backgrounds, the language itself is defined. Therefore, the patterns of PM use that I observe in my study from the different social groups are characteristic of SSLE. Therefore, the method confirms that the characteristic of PMs that I may find can be part of the definition of SSLE.

However, this assumption is not without complications. Cameron (1990, p. 57) criticises the variationist theory for its dependence on a ‘naïve and simplistic’ social theory. Cameron points out that the variationist theory depends on concepts like ‘norm’, ‘identity’ and ‘gender’ as a bottom line to explain language variation. She argues that these concepts themselves need explication. Cameron also raises the important question of how we relate the idea of ‘social’ to the linguistic. The variationist theory seems to regard that the social structures existed before language and therefore, language reflects such structures. Cameron challenges the idea that a linguistic pattern in data manifested in frequency reflects a social condition such as the effect of age, gender or class. She rejects the idea that language reflects society and calls this kind of frequency-based descriptions a correlational fallacy (ibid. p. 59). The aspects of society such as gender, age or class can be very complex and cannot be explained fully. Therefore, these categories are too broad to be represented by one set of data. Further, rather than a social aspect such as age or gender, interaction of network with gender for example, could be the reason for language variation (Milroy, 1980). A speaker’s interaction with others leads to construct social aspects such as gender or class. Then they become cultural constructs. Therefore, these concepts are far more complex entities than envisaged in variationist theory. Further, in language situations where more than one language interacts, such as the case in SSLE, when people use different variants of a linguistic feature, variation could well result from the influence of another language as much as gender, class or age (Winford, 1993). Therefore,

correlations between language traits and social factors could be fuzzy. Even though the current study does depend largely on frequency patterns to make generalisations, it is done with the knowledge that these descriptions might not represent wider trends. They could be limited to one particular setting and language situation. For example, the results of this study cannot be generalised to represent all female SSLE speakers or all younger SSLE speakers but rather they are limited to the particular setting and language situation observed in the current data.

An inventory of all the items under observation has been made. Using the variationist approach, I look first at the frequency data for each PM in each cohort of speakers. The proportions of use provide a point of analysis. Then, in line with traditional variationist studies, these quantitative findings are explained and interpreted using minimal statistical interpretation. Minimal descriptive statistics help to review frequencies and understand general trends in the data (Gomez, 2002). I go on to use a more qualitative approach to understand the findings. I look at how the contextual factors such as discourse situation, speaker age, gender and the intention of the speaker influence the use of one PM over another. This allows me to observe how PMs are used in distinct ways by these social groups to fulfil their intentions. In addition, I also tried to consider other factors such as the influence of other languages and profession since the data suggests that these factors might be relevant in accounting for some differences in the use of pragmatic markers.

4.7 The Approach to Analysing the Functions

Although PMs are generally considered not to contribute to the main proposition of a sentence, they do actively manipulate interpersonal messages. They clearly have a function in a conversation. By analysing their functions, details about how people subtly manipulate the thought process of their interlocutor is revealed. The PMs work in the favour of the speaker's intention. These functions are well documented in the research on PMs. The relevant previous research on functions is discussed in Chapters 2 and 4. The functions depend on the relationship between the interlocutors, the context of the utterance and even the mode that is used, for example whether it is a face to face or remote communication. The functions also rely on social factors such as age and

gender. Each PM has an array of functions, and functions vary among different types of PMs as well. Functional variation is well documented in the inner circle varieties and a few of the outer circle varieties such as Singaporean English and Indian English. This study adds new knowledge to the body of research on PMs broadly and on SSLE particularly by analysing the functions of PMs in SSLE.

All 72 transcripts were read manually to identify the different types of PMs that were used. There are 11 different types of PMs in the current data: *aah*, *I mean*, *kind of*, *like*, *no*, *right*, *sort of*, *well*, *yeah*, *you know* and *you see*. This selection was governed by the definition of Heine et al. (2021). The PMs *aah* and *no* are nativized PMs of Standard Sri Lankan English. In addition to the PMs mentioned, the study also found 11 different co-occurring PMs: *I mean like*, *I mean you know*, *like I mean*, *like you know*, *no I mean*, *you know I mean*, *you know like*, *you know sort of*, *well I mean*, *well yeah you know*, and *well you know*. They too were assigned functions based on the role they performed in the speech context. These functions are discussed in Section 5.6.

I omitted *I think*, *actually* and *now* although these had been called PMs in previous research (Farahani & Ghane, 2022; Novotana, 2016) due to descriptive challenges. The challenge is that their boundary with regard to meaning and syntax is vague, so that it is difficult to discern their use in context. For example, the randomly selected transcript C08AM has 09 instances of *actually* used by the interviewee. Examples 4.1, 4.2 and 4.3 are the first four instances in the transcript.

Example 4.1

<08AM> yeah well I I think er it's it's highly connected with my art because I I live for theater
I I **actually** I think theatre saved my life as well

Example 4.2

<08AM> er well it's not hallucinations those **actually** some of those things
<01HM> <u=?>

<08AM> I can't talk about some of these things because they were non-disclosure [laughter]
agreements and things like that [laughter]

Example 4.3

<08AM> everyone knows about my ribs most restaurants because of my blog posts they **actually** invite me to the their restaurants to **actually** try out their versions of the ribs so I get to critique and blog about these things people send their their cupcakes to me and things like that [laughter]

In all these examples, it is unclear if *actually* is used as to mean the truth or fact of the situation or for merely emphasis. In example 4.1, it could well be that theatre actually saved his life. In example 4.2, he could mean that he is actually prohibited from speaking about some things by law. Example 4.3 could be instances that <08AM> wants the listener to truly believe that the restaurants invite him to try their food because of his blog posts.

I think can also be observed in the same transcript. There are a total of 19 instances of *I think* in the transcript C08AM. Similar to the examples of *actually*, the first four occurrences in sequence are presented as examples 4.4, 4.5 and 4.6.

Example 4.4

<08AM> yeah well I I **think** er it's it's highly connected with my art because I I live for theater
I I actually I **think** theatre saved my life as well

Example 4.5

<08AM> er but I **think** the real person is known by very few a a very small circle I try to keep
my circle small and very real

Example 4.6

<08AM> er but but Kumar it's bee=it's just been crazy because I I do theatre it's absolutely
impossible to get time to do these things but I am really glad that I'm on your show and I
have been watching all your episodes and I **think** you are doing an amazing job

It is difficult to distinguish *I think* as a PM in these four instances as well. It is syntactically detachable from the utterance; however, it is unclear if it does not add any meaning to the propositional content of the utterance. The speaker intention is not clear. On the one hand, *I think* could be regarded in its canonical sense that the speaker is consciously thinking or introducing an opinion. On the other hand, *I think* could be a PM which is used for emphasis. This is the reason for omitting *I think* from the analysis.

Next, *now* as a PM is examined. There are no instances of *now* used by the interviewee in the same script. There are five instances used by the host but these are excluded from the search. Therefore, another random transcript is selected. For example, C08AcF is selected. There is a total of three instances of *now*, but only one instance is used by the interviewee <08AcF>, shown in example 4.7.

Example 4.7

<08AcF> now first of all we don't have full statistical information on how many Sri Lankan that is Sinhala Tamil Muslim er Burgher communities are residing outside of Sri Lanka

In this example, *now* does not refer to time in the actual instance of speaking as a temporal marker. However, it does have an extended temporal meaning such as *broadly at the time of speaking* or *currently*. Similar to *I think*, it can be removed from the utterance without harming the syntax of the utterance. However, it is not clear if *now* refers broadly to the period of speaking as in *currently* or if it is merely a turn taker such as *well* or an emphasis marker such as *right* or *yeah*. Due to these ambiguities, I omitted *actually*, *I think* and *now* from my search.

The manual search of transcriptions allowed me to eliminate instances of false positives such as the canonical use of items: *like* used as a verb, *you know* used as a subject pronoun and a verb in a sentence, *well* used as a conjunction. After that the data was also analysed by AntConc version 3.5.9 through concordance lines to double check the PM instances.

All instances of false positives for PMs were omitted from the analysis. For example, *I mean* is used for both canonical and pragmatic usage. *I mean* is used as part of the expressions *you know what I mean*, *get what I mean*, or *what I mean* is in its grammatical or canonical usage. Example 4.8 from CSSLE illustrates this canonical function. In this example, *what I mean* cannot be detached from the utterance as it will harm the grammaticality and the meaning. Such instances were omitted from the analysis.

Example 4.8

<01HM> sorry

<12AF> **what I mean is** <u=?>

<01HM> no I didn't say anything I'm just

The PM *kind of* and *sort of* are discussed together as they both have very similar functions. Similar to *you know* and *I mean*, *kind of* and *sort of* have a canonical meaning and a pragmatic meaning. For example, *kind of* and *sort of* can be replaced by *type of* in its grammatical meaning. Some invented examples are *what kind of a thing*, *that kind of a thing*, *that kind of music* and *that sort of a thing*. Reichelt (2021, p. 565) calls this the “type construction”. Example 4.9 and example 4.10 from CSSLE illustrate this type of false positive for *kind of* and *sort of*.

Example 4.9

<12AM> and you know just get a shirt and wear that **kind of** thing so we have seen a very shy er side to him

Example 4.10

<12AcM> he's in some **sort of** bubble

There were 1105 instances of *like* in CSSLE. However, *like* is a PM that has many false positives, and only 337 instances showed pragmatic usage. D'Arcy (2017) reports 6 grammatical uses for *like*, as a verb, a noun, a preposition, a conjunction, a comparative complementiser, and as a denominal adjectivalising suffix. All such instances are false positives for *like* in this study. Some of the grammatical uses that D'Arcy mentions are found in CSSLE. *Like* as an adjectivalising suffix (e.g.

like-minded people) does not occur in this corpus. Example 4.11 illustrates *like* as a verb. *Like* as a noun occurs in one instance in CSSLE and it is given in example 4.12. *Like* functioning as a preposition is seen in example 4.13. Example 4.14 is an instance of *like* used as a conjunction. Conjunctive *like* can be alternated with *as*. In example 4.15, *like* is used as a comparative complementiser. It introduces a finite subordinate clause (D'Arcy, 2017, p. 7). In such instances, *like* can be replaced with *as if*, *as though*, and *that*. Example 4.16 is an example of *like* occurring as a general extender as in *stuff like that*, *things like that* and *something like that* (Beeching, 2016).

Example 4.11: *like* as a verb

<02AcF> the it has to be there but it depends finally even the Booker even the Pulitzer Kumar on those three who read it and that matters and if I I've been a judge in in these things if the three writers who were chosen to be judges don't **like** that particular work and not all of us are the same then someone else will win it so it's also yes merit yes skill but also luck

Example 4.12: *Like* as a noun

<01AcM> if I'm correct on Southeast Asia that means we introduced mobile before the **likes** of Singapore and places who you know everybody tends to think that they were the benchmark

Example 4.13: *Like* as a preposition

<08EM> you know it sounds **like** a good name you know to go with

Example 4.14: *like* as a conjunction

<06EF> no but on a serious note erm I know I might sound **like** a cliché

Example 4.15: *like* as a comparative complementiser

<01AM> it's it's almost **like** they just walk into the center of your life and they you know they take over

Example 4.16: *like* as a general extender

<02EF> **like** the kind of quality I want and detailing **like** piping and things **like** that so

All such uses of *like* were treated as false positives with regard to *like*.

Another PM that has both canonical and pragmatic meaning is *right*. In terms of non-pragmatic use, *right* acts as an adjective (example 4.17), an adverb (example 4.18), verb (invented e.g. *we righted the fallen box.*), noun (example 4.19) and as an exclamation (invented e.g. *oh right!*). These instances were treated as false positives with regard to *right*.

Example 4.17

<03EM> the basic principles in life where you do the **right** thing at whatever the cost

Example 4.18

<10AcM> yeah comprador class yes that's **right**

Example 4.19

<07AcF> it doesn't mean that the that the Tamils car=don't have the **right** to return

An AntConc word search showed that *well* occurs a total of 559 instances in the CSSLE corpus. A screen shot of the AntConc search of *well* is given in Figure 4.1.

File	Left Context	Hit	Right Con
C01AcF Milani Salpitrikonala.txt	involved in the fashion industry er what what exactly <08EM>	well	I think yeah I met her when she was
C01AcM Dr. Shami Appathurai.txt	and I do a a clinic once a week <09AcF>	well	I think you look younger than ever touch wood [
C01AF Kaushalya Fernando.txt	in Ladies' College and what not <03EF> ahem [slight laughter]	well	I had a very free erm free and er
C01AM Ferocze Kamardeen.txt	this idea and bring out your first product <02EF> erm	well	I had obviously worked in fashion before I had
C01EF Kasturi Chellarajah Wilsor	many times were you pinched in your rear <08AF> oh	well	I had to get out of the place [laughter] <01
C01EM Bilal Yusuf.txt	HM- you wish you could or have you already <02AcF>	well	I hope I have I never know no we
C02AcF Madhubhashini Disanay...	s taking a lot of my time as well but	well	I hope Soul Sounds there's a lot of
C02AcM Dr. Ajantha Dharmasiri...	is having a biased attitude isn't that right <08AcM>	well	I think that the last government and this
C02AF Saundarie David Rodrigo.txt	AM- er <01HM> you have another life also <08AM> yeah	well	I think er it's it's highly
C02EM Ariyaseela Wickramanay...	now you do <03AF> now I have no choice er	well	I like it I mean not in a bad
C03AcF Shamara Wettimuny.txt	that's that's that's you to me <09AM>	well	I like that compliment because I'm a real
C03AcM Prof. Jayadeva Uyangoc	start off with what you mean by management education <02AcM>	well	I would say every manager er can be a
C03AcF Aseka Wijewardena.txt	or those of a higher recruitment or employment band <02AcM>	well	I would say management erm in a way it'
C03AM Seneka De Silva.txt	the cabinet and the the inflated size of the cabinet	well	I also have one er along those lines and
C03EF Obara Gunewardena.txt	which do you like more the acting or directing <11AM>	well	I always always always er preferred acting over directing
C02EM Mitra Dhammika Ganeg...	there he said it is necessary so many people <09AcM>	well	I cannot answer exactly er to your question
C04AcF Dr. Asha De Vos.txt	incident or moment in your career as an actor <11AM>	well	I can't forget my past you know the
C04AcM Dr. Ajantha Dharmasiri.txt	I I mean I know that this isn't your	well	I don't know if I'm just making

Figure 4.1: AntConc version 4.3.1 concordance lines for *well* in CSSLE

When one manually visits the context of each use of *well* and looks at its syntactic and pragmatic role, only 205 instances qualify as PM use. The following examples 4.20 and 4.21 from the CSSLE corpus show false positives of PM use for *well*.

Example 4.20

<12EF> and we try to ask=make it as=make them truly understand the subject as **well** I hate where they just memorise

Example 4.21

<01AF> more stuff in the world through FB through internet through so many things so er I er it was not be= no big deal I mean nobody nobody said anything no not there sc= I mean were teachers who had watched this movie of their schools and er b I mean nobody commented on they always what they said was that my I acted **well** so I mean so

Example 4.22

<03AM> **well** er I run a very hectic er life style you know rehearsals shows this that training so actually er the real Senaka other er than that virago that you all er know as to be

Well is a conjunction in example 4.20. It carries the meaning of *in addition to* in this example.

Example 4.21 shows *well* functioning as an adverb. It qualifies the verb *act*. *Well* in example 4.22 functions as a PM that marks turn taking. Examples 4.20 and 4.21 are false positives that are irrelevant to the current study.

Yes or *yeah* occur grammatically as an affirmative response to a question or as an exclamation (e.g. *yes!*). Instances of *yeah* where the use is syntactically dependent and the absence affects the meaning were regarded as the grammatical use of *yeah*. These were counted as false positives. Example 4.23 illustrates the grammatical use of *yeah*.

Example 4.23

<06AF> well done Kumar

<01HM> I believe I pitched correctly

<06AF> yeah

<01HM> Soul Sounds happened quite by accident you were telling me

<06AF> yes it did I

You know had many false positives. There were 1687 occurrences of *you know* but only 1405 occurrences functioned as a PM. One type of false positive is *you know* occurring as a pronoun and the corresponding verb. Example 4.24 illustrates this.

Example 4.24

<03AF> so many dancers in the country now and er it's not easy at all but as long as you love what you do and **you know** your goal

Another two functions that incur false positives are *you know* as a declarative as in, *as you know* and *you know* as an interrogative as in, *don't you know* (Östman, 1981, p. 21). There are 22 instances of *as you know* in CSSLE but there are no occurrences of *don't you know*. Figure 4.2 shows occurrences of *you know* as a declarative.

The screenshot displays the AntConc software interface. The main window shows a list of concordance lines for the search query "as you know". The interface includes a menu bar (File, Edit, Settings, Help), a toolbar with various analysis options (KWIC, Plot, File View, Cluster, N-Gram, Collocate, Word, Keyword, Wordcloud), and a search bar. The search results are displayed in a table with columns for File, Left Context, Hit, and Right Context. The search query is "as you know" and the context size is 10 tokens. The results show various instances of "as you know" used in different contexts, such as "m keeping my hand in in the thyroid research group" and "advice for sure and the show is called The Founder".

Figure 4.2: AntConc version 4.3.1 concordance lines for *you know* in CSSLE

The use of *as you know* denotes an assumption on the part of the speaker for a presumed certainty of knowledge as illustrated in example 4.25.

Example 4.25

<09EM> so this is the future so **as you know** Narendra Modi the Prime Minister of India has given a firm pledge that by twenty thirty all the i=cars in India are going to be electric cars

The role of *you know* would be different if this utterance contained *you know* rather than *as you know*. If the conjunction *as* is removed from this utterance, the use of *you know* could mean a reference to shared knowledge, but with the conjunction *as* its use is unclear. In such occurrences, *you know* can be detached from the utterance and the propositional meaning will be intact. Then it becomes an instance of *you know* functioning as a PM. Therefore, I did not include instances where *you know* was combined with *as*. I also eliminated instances of *you know what I mean* from the PM analysis as it an indirect question which evokes the canonical meaning of *you know*. There were

three instances of *you know what I mean* in CSSLE. This kind of an elimination process was conducted for all PMs in the CSSLE.

Similar to *you know*, *you see* is used grammatically in a pronoun and verb combination. Example 4.26 from CSSLE illustrates this grammatical use. Such instances are treated as false positives.

Example 4.26

<01AcF> on a daily basis erm so it's extremely difficult and then when **you see** a child and the child is so pure

All false positives were eliminated for all the forms and the occurrences of PM use were noted. These PMs were assigned functions drawn from a list compiled based on previous literature. The study by Beeching (2016) mainly guided the list of functions. Additionally, when Beeching's categorization did not satisfy the meaning denoted in a context, other classifications were considered (Denis & Tagliamonte, 2016). Some of the functions identified in previous research were revised to better align with the patterns found in the corpus.

As for instances of nativized PMs, the function was decided based on the context as there is no previous PM descriptions in the literature on SLE. In the analysis of functions, the current study did not include instances of *no* as a question tag that required an answer although there were such instances in the data. Example 4.27 is such an instance. In this interview, the interviewer <01HM> is asking some personal questions from the interviewee <06AM>. The interviewee is surprised about the questions and he says "doing your research no". In this instance, *no* is a question tag and it can be replaced by *aren't you* as in "doing your research, *aren't you*". Such instances are part of the syntactical structure and are not disconnected to the propositional meaning. Therefore, such instances were excluded.

Example 4.27

<06AM> it's all about what you want at the end of the day right [laughter]

<01HM> right you erm

<06AM> doing your research **no**

<01HM> sorry I I I take my show very seriously I do a lot of homework and I do it very very well

Section 5.3 presents the criteria for deciding the PM functions. An initial search for the PMs in CSSLE showed that there are 11 PMs in the corpus. After that a list of functions for each of the 11 PMs was prepared based on previous research. I decided to follow the classifications given by Aijmer (2013), Beeching (2016), Denis & Tagliamonte (2016), Ekanayake (2020), Farahani & Ghane (2022), Reichelt (2021) and Tao (2003) for these offered comprehensive and clear functions for the same PMs found in the CSSLE.

Beeching (2016) offers a detailed classification of functions for *like*, *sort of*, *well*, *I mean*, *just* and *you know*. Since my definition of PMs did not consider *just*, it was omitted from the list I prepared. Beeching lists 5 functions for *like*: exemplifying *like*, approximative *like*, quotative (be) *like*, discourse marking/focusing/ and hedging *like*. All these five functions were included in the initial checklist I created. Beeching listed 7 functions for *you know*, which are hesitation and appeal to common knowledge, word search and appeal to the interlocutor to fill in the gap, clarification and appeal to common knowledge, attention getting/ launching a new piece of information, direct appeal to shared knowledge/initiating a topic, repair, you know in final position: pointing out a self-evident truth/ impositional. She assigns 7 functions to *well*, which are hesitation, transitional *well*, changing the topic, taking a turn/ polite interruption, other-correctness, self-correction, and 'quotative' *well*. In addition to these 7, I also added the function word search and self-repair as noted by Aijmer (2013). Beeching (2016) listed 6 functions for *I mean*. They are self-repair, hesitation, clarification, exemplification, hedging. The sixth function she listed had four elements bundled together: elaboration, reformulation, justification, concession and nuancing. There were 5 functions listed for

sort of. They are metacommenting, hedging and qualifying, mitigating face threats, pause-filling, the general extender *sort of thing*. The list of functions for *right* was based on the eight functions listed in Denis & Tagliamonte (2016) study. These are accusation, command, exclamation, fact reporting, narrative fact, question, quotative delimiter, strict narrative and expressing a presumption of the interlocutor's common ground. The functions for *kind of* is based on the list of functions in Reichelt (2021)'s study of *sort of* and *kind of* in BrE. She lists four functions. The first one is metacommenting, hedging and qualifying, and the others are mitigating face threats, pause filling, and general extender. The functions for *you see* are based on the classification by Farahani & Ghane (2022). They list seven functions. They are introducing new topic or information, indications of objects and places, hesitation markers, pauses and restarts, explanations, justifications and conclusions, exemplification, emphatic lexis, and checking comprehension of the audience. Literature on SSLE mentions *no* as an operator that affirms/asserts (Ekanayake, 2020). Although there is no literature specifically on PMs in SSLE, this use of *no* too was added to the initial checklist of functions. Tao (2003) notes *yeah* as a turn initiator and as a marker of an assessment or an agreement/ affirmation or disagreement with regard to prior talk. These two functions were added to the checklist as well.

After creating the initial list of functions, the PM instances in the CSSLE were individually checked to see if their functions matched with the list. Functions that were found in the data were retained and the redundant functions, i.e. functions which are not reported in the CSSLE data, were removed from the list. In instances when the list was insufficient to describe a function found in CSSLE, a new description of a function was introduced. For example, *no* has the function of turn taker and *yeah* marks the end of a turn.

In the process of assigning the functions, it was also discovered that some of these functions serve as an umbrella classification for functions that differ from each other slightly. For example, the PMs *well*, and *yeah* share the common function of turn taking (Beeching, 2016; Tao, 2003). While *well* is exclusively used at turn initial to indicate taking a turn, *yeah* is used for both taking a turn and ending a turn in CSSLE. Therefore, in my individual analysis of functions for each PM, I assign two

separate turn taking functions to *yeah*: taking a turn and marking the end of a turn. This is because they are clearly two functions performed by *yeah* in a discourse situation. Examples 4.28 and 4.29 illustrate the functions of turn taking and marking the end of a turn respectively.

Example 4.28

<06AF> favourite composer is Rachmaninoff

<01HM> why are you off the beaten track if I may ask you with all due respect to your choice

<06AF> I mean I do respect other er composers the greats er Beethoven Baroque Bach but Rachmaninoff has that I mean he's a romantic composer and he has that extra bit of passion for me passion is important and **yeah**

<01HM> how passionate a person are you when you are performing accompanying is also as you said it's a performance

In this example, <06AF>, who is a musician, explains why she has selected a particular musician as her favourite composer. Then, she explains the reasons for her choice. When she completes her answer, she signals her interlocutor the ending of the explanation by using *yeah*. *Yeah* in this instance is not an affirmative response or an exclamation. The interlocutor proceeds to the next question because he acknowledges that <06AF> has concluded the response. Similarly, *yeah* also marks turn taking as in example 4.29.

Example 4.29

<01HM> what about your role in that

<01AF> **yeah** that was also that was a very I think er very er er what would you say it was a small role er

<01HM> harsh role

<01AF> not a harsh role not a harsh role er it was also face of this same character in Bora Diya Pokuna another face er but it was er more I think I don't know whether to you this word

is correct more it was more abstract I think the whole movie was whereas Bora Diya Pokuna from A to Z it the story went A to Z

In example 4.29, <01AF>, an actor, begins her answer with *yeah* although the interlocutor's question does not require a *yes/no* answer. This is the pragmatic use of *yeah* where it is used to mark turn taking. Therefore, *yeah* at the beginning of an utterance was assigned a separate function to *yeah* at the ending of an utterance in my classification.

After the list of functions for each PM was finalised, each PM instance in the corpus, embedded in its context, was saved on an Excel sheet. This step involved a manual survey of all the transcripts. A sample extract from the data recorded in the Excel sheet is given in Table 4.5. The information includes the utterance containing the PM (A), the PM (B), the function (C), speaker ID (D), the transcript ID (E), and finally the justification for deciding the function (F).

A (the line containing the PM)	B The PM	C The function	D The speaker ID	E The conversation ID	F The justification for deciding the function
because every= otherwise it's er you know it's a real common	you know	Clarification and appeal to common knowledge	<04EF>	C04EF	otherwise it would be a common design- she is clarifying and appealing to common knowledge

Table 4.4: The data included in the Excel sheet

The criterion for selecting a function for a PM is to identify the role a PM plays in a discourse context. The rationale for selecting the function is provided for example 4.30.

Example 4.30

<17HF> and what what are your products like basically
 <04EF> er it's as I mentioned it's Christmas décor
 <17HF> Christmas décor
 <04EF> and it is different like you know I don't have the same theme every year
 <17HF> okay
 <04EF> you know it's a unique thing
 <17HF> okay
 <04EF> and it's customized
 <17HF> right
 <04EF> anybody who wants something you know I make it different
 <17HF> okay
 <04EF> because every= otherwise it's er **you know** it's a real common
 <17HF> right

The speech context is relevant to deciding the function for *you know* (in bold). <04EF> is an entrepreneur and she is explaining her business making Christmas décor to her interlocutor. She is establishing the fact that her décor is uncommon. She emphasises that she creates custom tailored products because she likes them to be unique. She says if the Christmas décor is not unique, the designs tend to be very common. In the use of the bold *you know* she clarifies this point and implies that it is common knowledge. In fact, the interlocutor agrees by saying *right*. After understanding the role the PM plays, the function is assigned as clarification and appeal to common knowledge. Crucially, the speech context is examined to assign the function of a PM.

A total of 2949 PMs found in the corpus underwent this manual classification process. This kind of pragmatic annotation can be very challenging as it is not always straightforward to determine the pragmatic intention of the speaker. Contextual clues surrounding the PM are a powerful aid in this process of assigning the functions. However, there is no assurance of 100% accuracy as to whether what I interpreted as the function is actually the meaning the speaker wanted to convey. In order to assure reliability, after my own coding, I asked two further coders to code a sample of the data. The two raters were selected based on their academic experience in the field of linguistics. Neither of them have any specific expertise on PMs, but both are Sri Lankans. One is an ESL practitioner at a state university and the other is an academic in the field of speech and language therapy. I gave them an introduction about PMs and how they are defined. Then I presented the final list of PMs that I compiled based on previous research for them to code for functions of the PMs. I did not share my coding with either of them. Eight scripts were blindly coded by one coder and 4 of the same scripts were blindly coded by another coder, amounting to 11% of the scripts and 300 PM instances, equal to 10% of the total PMs. There was a disagreement in 4 instances between my coding and the two coders' and after considering their rationales, the instances were amended in the coding sheet. Since the two coders and myself agreed on 296 instances, which is an inter-coder agreement of 98%, my annotation can be considered reliable.

4.8 Approach to Analysing Gender

The correlation between gender and PM use was understood by analysing the frequencies of occurrences in keeping with the variationist approach. Gleason (1961, p. 392) describes the variationist approach as studying variants in a corpus "by seeking correlations with non-linguistic factors, commonly the speaker and the circumstances" and analyses the variants in their language use. Variationists examine correlations and attribute the variation to social factors or conditions. To answer the third research question, which looks at the relationship between gender and PM use, PM use by the female speakers and the male speakers is compared.

The corpus was divided into two sub-corpora of female and male speakers, consisting of 100,837 and 101,720 words respectively. Then, raw frequencies and associated percentages were manually calculated for each data set and used to determine the proportion of PM use in the conversations. These figures were then analysed and compared. Additionally, all PM frequencies were also calculated for women and men in each of the three different occupation groups. This allowed me to see if there are any genderlectal behaviour differences within one occupation. The distribution patterns underwent a qualitative analysis supported by previous literature and by my knowledge of being a practitioner of research and teaching in SSLE. Certain uses of PMs have been regarded as evident of personal qualities in previous research. For example, research shows that hesitation is a feminine quality, and women show this by using *you know* more frequently than men (Koczogh & Furkó, 2011). This type of attributes attested in previous research were used in the qualitative interpretations in the present study.

4.9 Approach to Analysing Age

Murphy (2010) lists four contexts of talk that are found in research on age and ageing: standardised tests, interviews, conversations and real-life interactions. She notes that context plays an important role in the type of results yielded in research on age. The current research uses data from real life interactions. The advantage of such contexts is that there is no direct influence of the researcher on the language produced by the speakers and the observer's paradox is minimised as well, as the interactions are not being produced with a researcher present. A central extralinguistic concern in analysing language according to age is time. A researcher has the choice of observing language change synchronically or diachronically as well as in real time or in apparent time. The data that is collected for this research allows only an apparent time study of language, as all the data has been collected synchronically. Differences of language use among different generations are taken to manifest diachronic variation in language (Murphy, 2010). Murphy (2010) claims that most of the adult age-related research observes the language of middle aged speakers, and advises that a good research study should avoid this misrepresentation. The current study includes both relatively young

speakers and the elderly speakers, with speakers in the age range 27-81. This study follows an etic approach to grouping the speakers. An etic approach groups the speakers arbitrarily according to equal spans of time such as decades or social generations. The CSSLE corpus was divided into three sub corpora to represent the three social generations: Baby Boomers (1946-1964), consisting of 28 speakers and 79,293 words, Generation X (1965-1980), consisting of 26 speakers and 73,024 words, and Generation Y (1981-1996), consisting of 18 speakers and 50,240 words. As in other components of this research, the PMs used in each sub corpus were analysed both quantitatively and qualitatively.

4.10 Chapter Summary

In this chapter I have described the data and the process of the analysis I followed to address the four research questions. Additionally, I shared the demographic details of my participants, choice of terminology, justifications for some of my decisions, how the data was analysed, and challenges faced in completing the study. In the next chapter, I will answer the first and the second research questions, i.e., the repertoire of PMs in my data and the functions of the PMs of SSLE.

Chapter 5: A Functional Analysis of the Pragmatic Markers

5.1 Overview

This chapter discusses the results of the corpus analysis pertaining to the functions of PMs. Pragmatic markers do have a function, although they do not contribute to propositional meaning as discussed in Section 3.6. The functions of PMs subtly manoeuvre a conversation. One main objective of this study is to investigate how the SSLE speakers use PMs to navigate conversations. I intend to answer two research questions in this chapter: first, what are the PMs in SSLE, and secondly, what are their functions? To answer these questions, I analyse the tokens of PMs present in the Corpus of Standard Sri Lanka English (CSSLE) that was compiled specifically for this study. The functions of each of the PMs are then presented.

A clear function is performed by each PM, though whether or not the function is obvious to the interlocutor is not clear. If the function is obvious, the response from the interlocutor would be favourable to the speaker's intentions. Determining a list of functions of PMs for the analysis is challenging for two reasons. One is that the taxonomies available for PMs based on functions greatly differ from each other (Brinton, 1996). Secondly, even studies of individual markers provide a wide range of meanings to a single marker; some even have overlapping meanings. Therefore, it is difficult to understand the difference between two functions in certain classifications. For example, Farahani and Ghane (2022) lists six functions for the PM *I mean*: among them, are 1) clarification and explanation and 2) elaboration. They give the following two examples for these two functions:

Example 5.1

if you just follow it blindly yeah i mean it would be worse if we were all like having different ideas and all arguing about it. (p. 62)

Example 5.2

and radar does anyone know what it is i mean you might not you might think that radars measure rainfall. (p.63)

One can argue that both these examples could be classified as explanation. Example 5.1 is an explanation of what would happen if you blindly follow (whatever is mentioned in that particular context). Example 5.2 is an explanation of what a radar is not. In addition to this kind of overlap, previous research and literature recognize the element of subjectivity in classifications on function. In fact, there is no agreement on the list of functions or the number of functions for each PM among linguists (Aijmer, 2014). Each description of function in previous studies broadly reflects the particular context observed in the study. Summing up the context, the attitude and the intention of the speaker, Schiffrin (1987, p. 326) claims that pragmatic markers “function as contextual coordinates of talk in that they are used to index an utterance to the local context in which the utterance was produced and interpreted”. The function of the PMs is an integral part of its context.

To understand the functions of the PMs in this study, first the PMs in the data were identified. Then a probable list of functions for each of these PMs was created based on previous research. A clear set of criteria to identify the functions was identified. Next, the function for each PM was determined based on the criteria and the context. The terminology to describe the function was selected from a list compiled from previous studies as explained in Section 4.7. Where a function was discovered which was not described in previous studies, study-specific terminology was created.

5.2 Preliminary Results

The first analysis was conducted to report the inventory of PMs in the data. As mentioned in Chapter 4, I followed Heine et al. (2021)'s definition of a PM to identify the PMs. Any marker that I was certain was part of the non-propositional part of the sentence, which could be removed unambiguously without affecting the propositional meaning, was selected as a PM in the corpus. The procedure that was followed is described in Section 4.8.

The following 11 PMs were observed in the corpus (arranged in alphabetical order). They are *aah*, *I mean*, *kind of*, *like*, *no*, *right*, *sort of*, *well*, *yeah*, *you know*, and *you see*. In addition to these PMs, the data also had evidence of 11 co-occurring PMs. The combinations are listed in table 5.1.

As evident in Table 5.1, the PMs and the co-occurring PMs identified are similar to the PMs identified in other studies based on British English, American English and Indian English with the exception of the two nativized PMs, viz. *no* and *aah*. *No* is noted in Indian English as a question tag (Lange & Leuckert, 2021).

PMs	Co-occurring PMs
Aah	I mean like
I mean	I mean you know
kind of	like I mean
like	like you know
no	no I mean
right	you know I mean
sort of	you know like
well	you know sort of
yeah/so yeah/yes	well I mean
you know	well yeah you know
you see/see	well you know

Table 5.1: PMs present in the CSSLE

There is a total of 2949 PMs in this corpus of 202,557 words. As a percentage of the total words in the CSSLE, PMs comprise 1.4%. 2949 PMs are produced by 72 speakers and although not every speaker produces an equal number of PMs, this roughly translates as an average of 50 words per speaker in an exchange of 2750 words. As a percentage, 1.8% of an individual's speech in this data consists of PMs.

There are 83 co-occurring PMs among the 2949 PM instances. The individual PMs are calculated as part of the total of PMs excluding co-occurring PMs. The total of PMs without the co-occurring PMs is 2866. Given below is a list of the PMs arranged in descending order of frequency in

the CSSLE. As the frequencies for each PMs reveal, *you know* is the most used PM in the data. Almost half the data consists of this single PM. Macaulay (2002) investigated *you know* in a similar word count in Scottish English. He used two corpora consisting of 120,000 and 125,000 words, the first containing interviews between adults, and the second including speech samples between adolescents. He reports a frequency of 3.41 *you know* per thousand words. My study reports a frequency of almost double this figure for SSLE, 6.9 per thousand words. This high occurrence suggests that a high use of *you know* is a characteristic of SSLE. According to the frequency figures, the use of nativised PMs is less than 0.5% of total PM production. A more detailed analysis is provided in Section 5.4 onwards for each PM.

PM	Raw Frequency	As a percentage of the total production
You know	1,405	49%
I mean	362	13%
Like	337	12%
Well	205	7%
Right	184	6%
Kind of	170	6%
Sort of	121	4%
Yeah/ so yeah/ yes	46	2%
You see/see	21	0.7%
No	13	0.4%
Aah	02	0.06%
Total	2866	100%

Table 5.2: Distribution of PMs across all speakers (excluding co-occurring PMs)

5.3 The Criteria for Deciding the Pragmatic Functions of the PMs

This study reports 23 functions performed by the 11 PMs. In this section, the criteria for coding the functions of the PMs that are attested in my study are explained, supported with examples from CSSLE for each function. These functions were selected and attributed from a broad pool of functions that were included in an initial checklist based on previous research as mentioned in Section 4.7. Then, instances in the corpus were checked against the initial pool of functions. Only the functions that are performed by the PMs in the data of the present study were selected from the pool of functions. The finalised functions of each PM in the present study and the source of the research that previously attested these functions are given in the following bulleted list. The functions that are not attributed to a source are specific to the present study.

Aah

- Emphatic lexis

I mean

- Self-repair (Beeching, 2016)
- Hesitation (Beeching, 2016)
- Clarification, exemplification, elaboration, reformulation (Beeching, 2016)
- Justification (Beeching, 2016)
- Concession and nuancing (Beeching, 2016)
- Hedging (Beeching, 2016)

Kind of and sort of

- Metacommenting, hedging and qualifying (Beeching, 2016)

Like

- Exemplifying like (Beeching, 2016)
- Approximative like (Beeching, 2016)
- Quotative (be) like (Beeching, 2016)
- Discourse marking/hedging (Beeching, 2016)

- Focuser (Beeching, 2016)

- Hesitation

No

- Emphatic lexis

- Taking a turn

Right

- To indicate shared knowledge/common ground between hearer and speaker (Denis & Tagliamonte, 2016)

- Shared knowledge presumed by the speaker

- Progression check question tag (Othman, 2010)

- Narrative fact: the overlap of narration and fact reporting (Denis & Tagliamonte, 2016)

- Explanations, justifications, and conclusions (can be replaced by *you see*)

- Emphatic lexis

Well

- Transitional well (Beeching, 2016)

- Taking a turn/polite interruptions (Beeching, 2016)

- Self-correction (Beeching, 2016)

- 'Quotative' well (Beeching, 2016)

- Prefacing a dispreferred response (Beeching, 2016)

- Hesitation (Beeching, 2016)

- Emphatic lexis

- Changing the subject (Beeching, 2016)

Yeah

- Topic shift

- Taking a turn (Tao, 2003)

- Marking the end of a turn

- Explanations, justifications, and conclusions

You know

- Attention-getting/ launching a new piece of information (similar to you know what?) (Beeching, 2016)
- Clarification and appeal to common knowledge (Beeching, 2016)
- Direct appeal to shared knowledge
- Filler
- Hesitation and appeal to common knowledge (Beeching, 2016)
- Introducing a quotation
- Repair (Beeching, 2016)
- Word search
- Word search and appeal to the interlocutor to fill in the gap (Beeching, 2016)
- You know in final position: pointing out a self-evident truth/ impositional (Beeching, 2016)

You see

- Hesitation markers, pauses and restarts (Farahani & Ghane, 2022)
- Emphatic lexis (Farahani & Ghane, 2022)
- Exemplifications (Farahani & Ghane, 2022)
- Explanations, justifications, and conclusions (Farahani & Ghane, 2022)
- Indications of objects and places (Farahani & Ghane, 2022)
- Shared knowledge presumed by the speaker (Farahani & Ghane, 2022)
- Checking comprehension (Farahani & Ghane, 2022)

A summary of these overlapping functions is presented in Table 5.3 below. One PM has many functions and the functions of different PMs can be overlapping. For example, explanation and elaboration is a function shared by the PMs *I mean*, *yeah* and *you see*. The section following Table 5.3 discusses the criteria that were used to define the functions that were attested in the present study.

PM function	Aah	I mean	kind of	Like	No	Right	sort of	Well	Yeah	you know	you see
(1) Explanation and elaboration		√				√			√		√
(2) Emphatic lexis	√				√	√		√			√
(3) Exemplification		√		√							√
(4) Concession and nuancing		√									
(5) Hedging		√	√	√			√				
(6) Hesitation		√		√				√		√	
(7) Justification		√							√		
(8) Self-repair		√						√		√	
(9) Approximative				√							
(10) Focuser				√							
(11) Quotative				√				√		√	
(12) Taking a turn					√			√	√		
(13) Shared knowledge/common ground						√				√	
(14) Narrative fact						√					
(15) Progress check						√					
(16) Topic shift								√	√		
(17) Prefacing a dispreferred response								√			

(18) Attention getting										√	
(19) Clarification and appeal to common knowledge										√	
(20) Filler										√	
(21) Word search										√	
(22) You know in final position: pointing out a self-evident truth/impositional										√	
(23) metacommenting, qualifying			√				√				

Table 5.3: The list of functions used in the classification

A brief description of each function is given here to explain how PMs were assigned to a function. Explanation and elaboration (1) is typically performed by *I mean, right, yeah* and *you see*. Example 5.3 illustrates this function. Generally, the speaker first makes a point and then elaborates by either giving additional information or an example. The speaker uses one of the above PMs to introduce the additional information or the example. The examples or additional information is generally preceded by *I mean*, followed by *right* and *yeah* and either preceded or followed by *you see*. In example 5.3, the speaker <05AcF> makes a general statement to the effect that it is important to create a society that recognises men and women equally. Then, she further elaborates that “if men can so can women”. To elaborate this point, she uses an example introduced by *I mean*. The PM *you know* is connected to her previous utterance which is “if men can so can women”. *I mean*

introduces the elaborative example that most celebrated cooks are men. Such instances are coded as explanation and elaboration. It is not a clarification of her statement. It is merely an example to elaborate her point.

Example 5.3

<05AcF> but the most important thing is creating a societal environment a social environment

<07HM> mm

<05AcF> that recognises the equal worth of men and women

<07HM> mm

<05AcF> a society that says yes if men can so can women

<07HM> mm

<05AcF> if women can so can men

<07HM> mm

<05AcF> you know **I mean** some of our most celebrated cooks are men

<07HM> mm

The second function is emphatic lexis (2), performed by the PMs *aah*, *no*, *right*, *well* and *you see* in the CSSLE data. These help the speaker emphasise an idea or a key point in the narrative by directly drawing attention to it. Usually in the case of *well*, it is placed before an important idea, word of phrase (Huang, 2019). *See* is used as a shortened form of *you see*, used interchangeably by speakers in this corpus. *You see* introduces the most important point of the discourse situation. In other words, *you see* introduces the topic sentence in a unit of talk. Example 5.4 illustrates this function.

Example 5.4

<09AM> anyways I was there I was in the US army for six years and I served in the California National Guard I was stationed at Fort Ord in California it was a wonderful experience

wonderful experience and er I er am very that also **see** I I've been very lucky good college good me=mentors and then the army with discipline I can tell you stories about the army take too long

<9AM> emphasises that he has been very lucky in life especially in terms of a good education and job opportunities. This is the point he wants to highlight, and summarises his earlier descriptions. The speaker uses *you see* to emphasise this important point he makes. As *see* is actually *you see*, it immediately and directly requests the interlocutor to pay attention to what is being said. Such uses are categorised as emphatic lexis.

The third function of the PMs explored in CSSLE is exemplification (3). Exemplification refers to providing examples. This function is usually performed by *I mean*, *like* and *you see*. These PMs can be paraphrased by 'such as' or 'for example' when they are used for exemplification. Example 5.5 is an illustration of this function performed by *like*. Preceded by *like*, the speaker gives an example of an unforgettable incident.

Example 5.5

<09AF> you cannot fight either so it's not a very good thing and erm hopefully all this will change erm unforgettable lots of unforgettable incidents **like** er I suppose the most important one would be the fact that I almost lost my life twice while

According to Beeching (2016), *like* as an exemplifier serves as a face-saving device as well. This is achieved by using an example to validate an earlier statement. The interlocutor, at least in the eyes of the speaker, is convinced by the example. This is illustrated in 5.5.

The fourth function of PMs in my data is concession and nuancing (4), and this is performed exclusively by *I mean*. Example 5.6 exemplifies this function. Concession and nuancing refer to providing a subtle shade of meaning by modifying or adjusting an utterance. *I mean* implies that the speaker is covertly questioning an idea or is gently urging the listener to think more deeply about what the speaker is saying. This subtle adjustment to meaning is usually preceded by *I mean*.

Example 5.6

<04AM> well you know er now that you asked the question I might as well you know talk about it straight away I mean you know five top CEOs among the top ten brands in Sri Lanka were interviewed recently and er they spoke about the crisis and all that and then er you know they they they asked them the interviewer asked them if there is one thing that the government can do for you in this crisis what would it be shockingly but not surprisingly every single one of them said we want flexibility in labour laws meaning we should be able to hire and fire people at will and reduce their salaries

<01HM> what

<04AM> yeah

<01HM> ouff

<04AM> every one of them **I mean** is this leadership Kumar **I mean** what kind of leader would sacrifice your team at the very first sign of trouble

In this example, the speaker <04AM> explains that top CEOs in five companies requested the government for the autonomy to hire and fire employees during the pandemic to overcome the crisis at the time. <04AM> does not agree with this request and in his opinion these CEOs are not true leaders. Therefore, he introduces this concession with the first use of *I mean* and nuances his stance on true leadership. Both uses of *I mean* in this example are used for concession and nuancing an argument. This function is characterized by a situation which is followed by a hint of criticism introduced by *I mean*.

The fifth function is hedging (5). Hedging refers to avoiding a definite statement, softening an argument or presenting an idea cautiously. Hedging facilitates a speaker to save face in a speech situation. We lose face when we cannot uphold the truthfulness of a statement and hedging saves us from this by indicating a sense of probability in the truth condition of an utterance. It allows the speaker not to be held accountable for a statement made. This function is typically carried out by *I mean*, *kind of*, *like* and *sort of* in my data. Example 5.7 illustrates this function.

Example 5.7

<09EF> in our own home for the last year and a half and if we can manage that then with just a set of bags and I think that's **kind of** how it should be in every home

Generally, a PM used for hedging precedes a strong or a definite statement. The speaker in this example is an entrepreneur who produces environmentally friendly bags. She has experimented with her product in her own home and feels that it was successful. She advises that every home should use such bags. Advising *every* home can be seen as a very strong recommendation. In order to soften her suggestion, she uses *kind of* as a hedging device. Such softening or downtoning instances introduced by any PM is recognized as hedging function.

The next function noticed in the CSSLE corpus is hesitation (6). Speakers most often use *I mean, like, well, and you know* to indicate hesitation. Hesitation is marked with many filled or unfilled pauses. In the present corpus, which was marked only for filled pauses, one indication of hesitation was a PM used in between these filled pauses. Another sign of hesitation is repetitive and almost aimless use of several PMs indicating that a speaker is either searching for the right word or ideas. Kirk (2018, p. 142) interprets *well* as an indicator that the speaker is thinking about what to say while maintaining the conversational engagement. This interpretation is similar to hesitation in my classification of *well*. *You know* and *like* are PMs that are used in combination to mark hesitation in this corpus. When *you know* is used for hesitation, it is an appeal to common knowledge as well. A speaker may use *you know* several times consecutively in one utterance trying to plan the speech or recall a word, but the addressee feels that s/he has an idea about what is being expressed. Hesitation is also marked by repetitive use of one PM before several false starts. In example 5.8, *like* is such an example of a hesitation marker, and precedes three false starts.

Example 5.8

<01HM> what the hell does Uguj mean

<07AF> I have no idea **like** it just **like**

<01HM> <u=?>

<07AF> one day it just **like**

<01HM> <u=?>

<07AF> yeah

In this conversation, the interviewer <01HM> asks the interviewee <07AF> what is meant by a word, the meaning of which is not clear. In this speech context, it is a pet name used for the interviewee's sister. <07AF> says she has no idea and then she is at a difficulty to find the right words, thus hesitating to provide a useful answer. All three instances of *like* in this example show hesitation and this is marked by the false starts or unfinished utterances that have no connection to each other. False starts and disconnected words or utterances joined or preceded by a PM were regarded as instances of hesitation. *Well* is also used for hesitation. Kirk (2018, p. 143) mentions function of *well* as a staller or a delaying device, referring to the work of Svartvik (1980, p. 171). An example is given in example 5.9:

Example 5.9

<01HM> er any place that you haven't been to which you are dying to visit you want to go

<07AM> yeah er **well** er

<01HM> mine <u=?> one day

In example 5.9, the speaker is stalling for time to give a well-thought-out answer to the question. In this instance, he is using *well* for hesitation. The filled pause *er* precedes *well* to get even more time.

Justification (7) is another function played by PMs. A justification is a type of an explanation but with a clear provision of a reason: a speaker presents an idea and then this idea is justified by another statement. Justification is indicated by *I mean* and *yeah*. The PM *I mean* that preceded a justification can be replaced by *because* if needed. Example 5.10 illustrates this use.

Example 5.10

<09AF> I mean it it's a very pleasant feeling because **I mean** to know that you are still loved there for all these years fifty-two years in industry now and you are loved and and you are still remembered and that people sort of

There are two uses of *I mean* in this example. The speaker <09AF> is a famous singer in Sri Lanka and the interlocutor asks her if she feels uncomfortable when unknown people reach out to her because of her stardom. The first *I mean* is prefacing an explanation to this question. However, the second *I mean* is a justification for her response that it is a pleasant feeling because you feel still loved even after many years. Instances where the comment is a justification to a statement mentioned earlier and the PM can be replaced by *because* are assigned the function of justification.

Example 5.11 illustrates the function of self-repair (8), typically performed by the PMs *I mean*, *well* and *you know*. Self-repair occurs when a word or a few words are mentioned by mistake in an utterance, or a false start is made. To correct this, the repaired speech is introduced by *I mean*. The self-repair generally leads to a completion of an idea so that it is more appropriate to the discourse situation.

Example 5.11

<12EM> so we are interested in

<17HF> to learn martial arts and all

<12EM> yes we have people **I mean** kids starting from three years old

In this example, the speaker <12EM> mentions *people* wrongly. Then he corrects himself and changes it to what he actually wanted to say: *kids*. He introduces the self-repair with *I mean*.

Beeching (2016) mentions this function as one of the least ambiguous functions of *I mean* because it is obvious when someone halts their speech and changes a word or phrase to produce an appropriate utterance.

Another function carried out by PMs is approximation (9). Approximation implies something that is similar or nearly similar to another thing but not exactly. For instance, with numerical examples, it refers to an approximate number, or a number close to what is mentioned. The speaker

uses an approximation when s/he does not know an exact detail or a numerical detail or does not want to share an exact value or a detail with the interlocutor. *Like* can be replaced by *about* or *approximately* in examples that mention a number, but not in non-numerical approximations.

Examples 5.12 and 5.13 show this use.

Example 5.12

<04AF> so I was like okay I will do this just to please them and I studied er hospitality and event event management and I worked at a hotel for **like** a year

Example 5.13

we keep taking new people so now they get from junior to **like** a mid level

Like is used to show an approximation of time. The speaker <04AF> mentions that she worked in a hotel for a period of approximately a year. *Like* is used for approximation in numerical contexts as in example 5.12 and in non-numerical contexts as in 5.13 as well.

The next function is PMs acting as a focuser (10). *Like* is used to focus on new information in an utterance (Underhill, 1988, p.234) and also to focus on given information (Miller & Weinert, 1995, p.376). Giving new information is the point of the sentence. Given information is either information that the speaker assumes that the interlocutor can infer from the conversation or information that had been already introduced in the discussion. *Like* precedes new or given information, introducing the most significant information in the utterance. *Like* generally focuses on a one-to-five-word phrase of information in an utterance (Underhill, 1988). It can also appear at or near the end of new information (Underhill, 1988). D'Arcy mentions (2007) that in such instances, *like* can be replaced by *I mean* or *you know*. Example 5.14 illustrates this use.

Example 5.14

<06AM> I'm in tune with reality Kumar and and the other thing is I just got **like** a new pair of teeth fixed before that I wasn't really smiling a lot so people thought I have a bitch face

The speaker draws the attention to the *new pair of teeth* by using *like* to introduce it. The *new pair of teeth* is new information in this utterance.

Another function is PMs acting as quotatives (11). The PMs *like*, *well* and *you know* fulfil the function of introducing a quote in my data. In a strict sense, quotative *like* as a quotative cannot be considered as a PM use, since it cannot be omitted from an utterance as it would challenge the syntactic structure of the utterance. D’Arcy (2017, p.16) points out that although it is not wholly pragmatic quotative *like* “straddles the interface between discourse-pragmatics and lexical encoding”. This ambiguity is recognized by Buchstaller (2002, p. 4) as well. She mentions that when a quotation is introduced by *(be) like*, it is rendered only approximately in terms of its emotional and contextual message. Therefore, quotative *(be) like* signals “the possible non-equivalence of what is reported and the actual utterance” (p. 4). This allows the speaker a reduced responsibility in being accountable for what is reported. She concluded that *like* then functions at an interpersonal-pragmatic level. Beeching (2016) also considers quotative *(be) like* as a PM on a similar argument. She says that quotative *(be) like* is different from other quotatives such as *say*, *think* and *go*. *(Be) like* as a quotative reports speech as something said “along the lines of” (Beeching, 2016, p.131). It gives the idea that the words that are reported by quotative *(be) like* may not be the exact words. In this sense, Beeching argues that quotative *(be) like* is a PM. About 15% of the total production of *like* in the present data is used for the function of quotative *(be) like*. Quotative *(be) like* is also considered a very fast-growing linguistic innovation. I follow Buchstaller’s (2002) and Beeching’s (2016) argument and consider quotative *(be) like* as a PM despite its deviation from the definition I follow for PMs. Example 5.15 illustrates this function.

Example 5.15

<06AF> yeah I’m not myself so he=yeah so we won this competition and I **was like** what next
and there was this conductor called Gregory Rosewood

In 5.15 the speaker does not introduce an exact direct quotation. She introduces the general direction of the thought process that occurred in her mind preceded by the quotative *(be) like*. It is

an approximation of reported speech. It is in this sense that the current study codes *(be) like* as a pragmatic function.

The next function under discussion is PMs that indicate taking a turn (12). The PMs *no*, *well* and *yeah* can function to indicate either taking or ending a turn. A turn initiator is the very first form with which a speaker begins a new turn in a speech exchange (Tao, 2003). Taking a turn can be a polite interruption when the flow of a conversation is interrupted by using *well*, which acts as a marker of politeness. As this function is linked to discourse structure, its identification is more straightforward because it relies on initial position. Example 5.16 elaborates this use. <8EM> begins his turn with *well*.

Example 5.16

<16HM> so tell me starting off to what do you owe your success

<08EM> **well** er I I believe er it's my commitment

Another function of PMs in this data is to indicate shared knowledge or to share common ground (13). The speaker uses a PM to either precede or follow a piece of information that is incontestable mutual knowledge (Holmes, 1986). Shared knowledge or sharing common ground is evident in explicit interpersonal references. Example 5.17 is from a conversation between an interviewer and an artist. The artist <02AF> talks about the pressure for her to produce artistic material quickly in this modern age where news items and information is produced instantly. To illustrate her point she refers to the general everyday routine of a person in today's world. As this scenario is common knowledge between her interlocutor and herself, she precedes these sections of incontestably mutual knowledge with *you know*. Reference to shared knowledge is indicated by *right* and *you know* in CSSLE.

Example 5. 17

<02AF> so in a way you get up and you every morning and you you know you open your twitter feed or you you know it's all social media but that's how we all communicate

Example 5.18 elaborates yet another function of the PMs, narrative fact (14). This is only reported for the utterance final tag *right* in the present data. What is meant by narrative fact is the overlap of narration and fact reporting (Denis & Tagliamonte, 2016). A fact is presented followed by *right* in the middle of a narration as in 5.18.

Example 5.18

<01AM> and I know Harshitha has control of my Instagram and my Twitter accounts **right** er
and for a while I think the Puswedilla twitter account was linked to the Puswedilla Facebook
page something like that

In this example, <01AM> is a dramatist and the interviewer asks him jokingly if he is technologically challenged. <01AM> is narrating how he handles social media and then reveals factual information about who handles his social media. He uses the utterance final tag *right* in this overlap of narration and fact reporting.

Example 5.19 is an example of the function of progress check (15). The PM *right* is used in this function in the current data. *Right* is used only to ensure that the listener is following the narration. The speaker does not stop for any response from the interlocutor although *right* is projected as a question tag. The interlocutor may respond with backchanneling or even use *yes* or *no*. However, the use of *right* at the end of an utterance does not demand a response, and can be used without pausing or giving a chance to yield the floor to the interlocutor. The speaker simply continues the narration moving on to new ideas. This is illustrated in example 5.19. The speaker <01AcF> is narrating her experiences of handling child abuse professionally in this context. She checks if the listener is following her narration in both instances of *right* in this context. The listener acknowledges the progress check by the backchannel response *mm*. Such progression checks were recorded for this function irrespective of evidence of backchanneling.

Example 5.19

<01AcF> on a daily basis erm so it's extremely difficult and then when you see a child and the
child is so pure

<05HF> mm

<01AcF> **right** erm there has to be some good in the world **right**

<05HF> mm yeah

Another function evidenced in the PMs of this study is topic shift (16), conveyed by both *well* and *yeah* in my data. Example 5.20 illustrates this function. The speaker is addressing one topic and almost abruptly s/he wraps up the topic with *yeah* and moves into a new topic. There is no direct reason for the speaker to use an affirmative *yes* or *yeah* in the discourse situation. Therefore, the use of *yeah* merely signals the end of one topic and the beginning of another. This use of *yeah* is usually preceded by *and*. The use of *yeah* at the end of one narration allows the speaker to move to another topic in the discourse unit without any formal introduction to the new topic or without yielding the floor to the interlocutor.

Example 5.20

<02AF> I think that really made my sort of er sort of grasps the bull by the horns you know er had to sit with myself er for long hours um discovered that I wasn't such a nice person quite boring to spend my entire time [laughter] time with er and **yeah** er that was that was how it was started kids came back it was a whole different set of challenges

In this instance, the speaker <02AF> narrates her experience of self-isolation during the lockdown as she was being quarantined. She describes in length how her experience was and then she wants to focus on the next part of the narration, and she indicates that she is changing the topic by using *yeah*. Then she moves on to describe how the lockdown changed for her when her kids came to live with her.

Example 5.21 illustrates a PM use in my data that is confined to *well*. It is prefacing a dispreferred response (17). The criterion to mark this function is to see that a negative response or a refusal is preceded by *well*. In this example, <05AF> wants to give a negative answer to the interviewer's question and she prefaces it with *well* to delay her negative response.

Example 5.21

<01HM> yeah Menaka people's sweat in their palms has that ever happened and their fingers slip off the keys

<05AF> **well** that hasn't happened to me but er if er somebody with sweaty palms plays before me on a keyboard and I have to play on the piano afterwards then it's a problem

Example 5.22 illustrates the function of attention getting (18) by launching a new piece of information performed by *you know*. The attention is sought by highlighting one piece of information presented as new information. It is preceded by *you know*. It is not a reference to shared knowledge. *You know* is generally in the initial position in such instances (Beeching, 2016). For example, in 5.22, even the idea that the speaker's two sisters and brother had to sacrifice their university education is new information for the interlocutor. However, the next piece of information is highlighted by introducing it with *you know*. *You know* is in utterance initial position. *You know*, is a direct demand for the listener to give attention as it has the personal pronoun 'you'.

Example 5.22

<04AM> but you know God can't drop er food from the through our roof no so two of my sisters and my brother had to sacrifice their university life and go and work and **you know** even when I went to work Kumar I went to work much later

The next function is clarification and appeal to common knowledge (19) performed by *you know*. In example 5.23, <05AM> mentions that he loves nature and clarifies this with a simile that it was as if he and nature were hand in hand. With the use of the idiomatic expression *hand in hand* the speaker appeals to the common knowledge of the listener to understand this special affection. It can be regarded as an explanation or even an elaboration. It is extremely difficult to distinguish between a clarification, an elaboration and an explanation. These terms were not used as my classification followed Beeching's (2016) terminology. There are three movements for this function. First one is that the speaker makes a statement and secondly follows it up with a further clarification for the statement. Third movement is the actual clarification preceded by *you know*.

Example 5.23

<05AM> loved nature I was very much in touch almost **you know** it was hand in hand kind of thing I don't know why I was so attached to it

Another function in the data is the use of PMs as a filler (20). When *you know* is used as a filler there is no appeal for common knowledge or cooperation from the addressee. This is different from its use for hesitation, when there is an appeal to the interlocutor to understand what the speaker is trying to say. As a filler, *you know* gives the speaker time for planning the speech and the speaker will construct the rest of the utterance on his/her own without any support from the addressee. When *you know* is used without a clear aim and without any appeal to the interlocutor it is categorised as a filler. Example 5.24 illustrates this function.

Example 5.24

<04EF> he helps me even to blend the colours and

<17HF> nice

<04EF> you know so all those things so you know that's how **you know** I

<17HF> yeah

<04EF> is an entrepreneur who makes Christmas décor. This utterance is from a section where she describes how she started her business. She mentions that at the beginning her husband helped her to blend the colours in her creations. There are three instances of *you know* in this context. The first two instances are attempts for clarifications and appeal to common knowledge. “(Y)ou know so all those things” refer to the many activities she had to do at the beginning of her business. The second *you know* is the beginning of another clarification. “(Y)ou know that's how” is the beginning of another clarification about how things were at the beginning of her business. However, she does not finish her idea and uses another third *you know*. This third *you know* has no particular function, and such aimless uses of *you know* are counted as fillers.

The function of word search (21) is performed only by *you know* in my data, and can be separated into two slightly different uses. The first is word search and the second is word search and appeal to the interlocutor to fill in the gap. First, *you know* can be used to hold the floor until the

speaker finds the appropriate word to utter. This function is slightly different to word search and appeal to the interlocutor to fill the gap, where the speaker is searching for a verb and appealing to the listener to fill the gap with an appropriate word. In such instances, the speaker would hold the floor with *you know* until the interlocutor says an appropriate word to complete the utterance. Example 5.25 is an example of the first type. The speakers appear to find the appropriate word themselves than appeal the interlocutor to fill in the gap.

Example 5.25

<10EF> so erm yeah so that's that's where my struggles are because trying to do everything being a housewife

<17HF> mhm

<10EF> erm checking the home front and try to **you know** erm what do you say like meet the erm deadlines

In example 5.25, <10EF> is an entrepreneur, and she says that one of her problems is that she is a sole operator and does not like to ask for help. She explains the number of responsibilities she has to shoulder, and one of them is to meet deadlines. She is looking for the word “deadlines”. She rephrases this as “what do you say like meet the” and this search for words is introduced by *you know*. An instance of the second type is given in example 5.26.

Example 5.26

<01AcF> and she introduced me to a lawyer called Sajeewa Samaranakayeke who passed away few years ago erm so he was also my mentor and you know this complete erm he used to call himself a social working lawyer

<05AF> right

<01AcF> so not a lawyer in the **you know**

<05HF> traditional sense

<01AcF> traditional sense

<05HF> yeah

In example 5.26, <01AcF> is the interviewee. She is trying to find a word to explain the sort of mentor she has. She starts her utterance by saying “so not a lawyer in the” and then she is looking for the right word. She uses *you know* to hold the floor until her interlocutor says the appropriate word. Then, she repeats the word “traditional sense” confirming that is the word she was looking for. Word search and word search and appeal to the interlocutor to fill in the gap are treated as two different functions because of this very clear difference.

Another function of *you know* in my finalised list is *you know* in final position: pointing out a self-evident truth or impositional (22). *Impositional* refers to an idea being imposed upon the addressee. Beeching (2016, p. 103) describes it as “no argument can be raised against it”. Instances in which a strong fact or a relatively truthful statement is presented with *you know* at the end of the proposition, are classified under this function. Example 5.27 illustrates this function.

Example 5.27

<05AcF> but the most important thing is creating a societal environment a social environment

<07HM> mm

<05AcF> that recognises the equal worth of men and women

<07HM> mm

<05AcF> a society that says yes if men can so can women

<07HM> mm

<05AcF> if women can so can men

<07HM> mm

<05AcF> **you know** I mean some of our most celebrated cooks are men

<07HM> mm

In this example, *you know* is a continuation of the utterance that says “if women can so can men”.

The interviewer <07HM> in this context is a lawyer and the interviewee <05AcF> is an academic in

the legal field. <05AcF> presents a self-evident truth to her listener and she confirms this by using *you know* at the end of the utterance. This is not a direct appeal to shared knowledge. <05AcF> is building her argument step by step. First, she says that we need a society that recognizes the equal worth of men and women. Then she goes onto her first supporting point that if men can so can women. This can be viewed as a common argument in society at large. She then subverts this argument and brings forth a strong argument that is not so common: “if women can so can men”. Even if both interlocutors are educated and in the legal field, <05AcF> does not use *you know* to indicate shared knowledge. *You know* is used to introduce this impositional strong argument.

The last of the list of functions is metacommenting, hedging, and qualifying (23). A metacomment is a comment made about another comment. Hedging has been explained earlier in this section, and refers to diminishing the impact a word would have on the listener. Qualifying refers to making a statement or an assertion less absolute. Both *kind of* and *sort of* fulfil this function and therefore they are discussed together in this section. The use of these two PMs implies that “the description is only approximate” (Davidse et al., 2008, p.157). This function was adopted from the classification by Beeching (2017). She uses metacommenting as one function while hedging and qualifying is classified as another function, although she herself admits that the boundary between metacommenting and hedging and qualifying is hard to draw (Beeching, 2016, p. 158). As mentioned in the Section 3.7, given the challenges in precisely identifying the dividing line between metacommenting and hedging and qualifying, all three categories are grouped as a single function fulfilled by both PMs. Given below are examples 5.28 and 5.29 to elaborate this function fulfilled by *kind of* and *sort of* respectively.

Example 5.28

<05EF> er when I graduated from college and came back I was **kind of** er trying to figure out what I want to do with life so it was like more like trial and error process for me so it was this

you know experimenting with er loads of things actually and then I came across erm this company who's also into lighting it was actually Interior House

Example 5.29

<08AF> I have a I feel a very er **sort of** I feel emotional to the song I feel it has done some magical you know thing for me and er it's er it's what I call my signature tune

In these examples, the speakers use *kind of* and *sort of* to metacomment, hedge or qualify the statement followed by the PMs. In the example 5.28, *kind of* is linked to the verb *trying* and in example 5.29, *sort of* is linked to *feel*. The effect or the impact of these two verbs are diminished or softened by the use of *kind of* and *sort of*.

The subjectivity associated with determining the functions of PMs in the data was minimised as I employed two coders to assess the classification. Wiltschko et al. (2018, p. 569) argue that the different ways of encoding the same meaning or function define the object of the study, and therefore they are necessarily subjective. However, accountability should be maintained in interpretations as far as possible. Therefore, I took into account the linguistic and extra-linguistic context as far as possible along with conventions of the situation in interpreting the functions of the PMs. The interpretation given below shows the thought procedure I adopted when I decided on the function of each PM.

5.4 The Analysis of the Functions

I applied the criteria outlined in the previous section to code all the data to identify what function each PM was fulfilling. The function that is used most frequently is clarification and appeal to common knowledge. This was performed solely by *you know*: 68% of the total functions of *you know* is used for clarification and appeal to common knowledge. This particular function is used 32% as a percentage of all functions performed by all PMs. The least used function is for hesitation or as a filler. As a percentage of all functions performed by all PMs, hesitation and filler function is used only 0.4%. Hesitation and filler functions are performed by *you know*, *I mean*, *like* and *well*. These two

functions are used less than 2% of the total functions performed by each relevant individual PM. Beeching (2016), who analysed British English claims that the greatest usage of *you know* in most genres in BrE is for hesitation and appeal to common knowledge. She mentions that 30% of all occurrences among adults are used for this purpose. This finding is a stark contrast to my study.

The next sections analyse each PM individually in terms of their functions so that the role of each PM is understood deeply.

5.4.1 The Functions of *You Know*

As listed by Janet Holmes (1986) in a review of different functions listed in previous studies, *you know* has been labelled as a verbal filler (Brown, 1977: 102), a softening connective (Crystal & Davy, 1975), cajoler (Edmondson, 1981:155), compromiser (James, 1983), hedge (Lakoff 1975, Brown & Levinson 1978), and a conversational greaser (Wong-Fillmore, 1976). The many labels are a manifestation of the many roles *you know* plays in interaction.

You know has a very flexible syntactic positioning: it can occur initially, medially or finally in an utterance. It is one of the most productive PMs, and accounts for 49% of the total frequency of PMs. This high use of *you know* is in line with Zheng's (2015) investigation about *you know* used in BrE. The 1405 tokens of *you know* present in this corpus perform 10 functions, which are listed below.

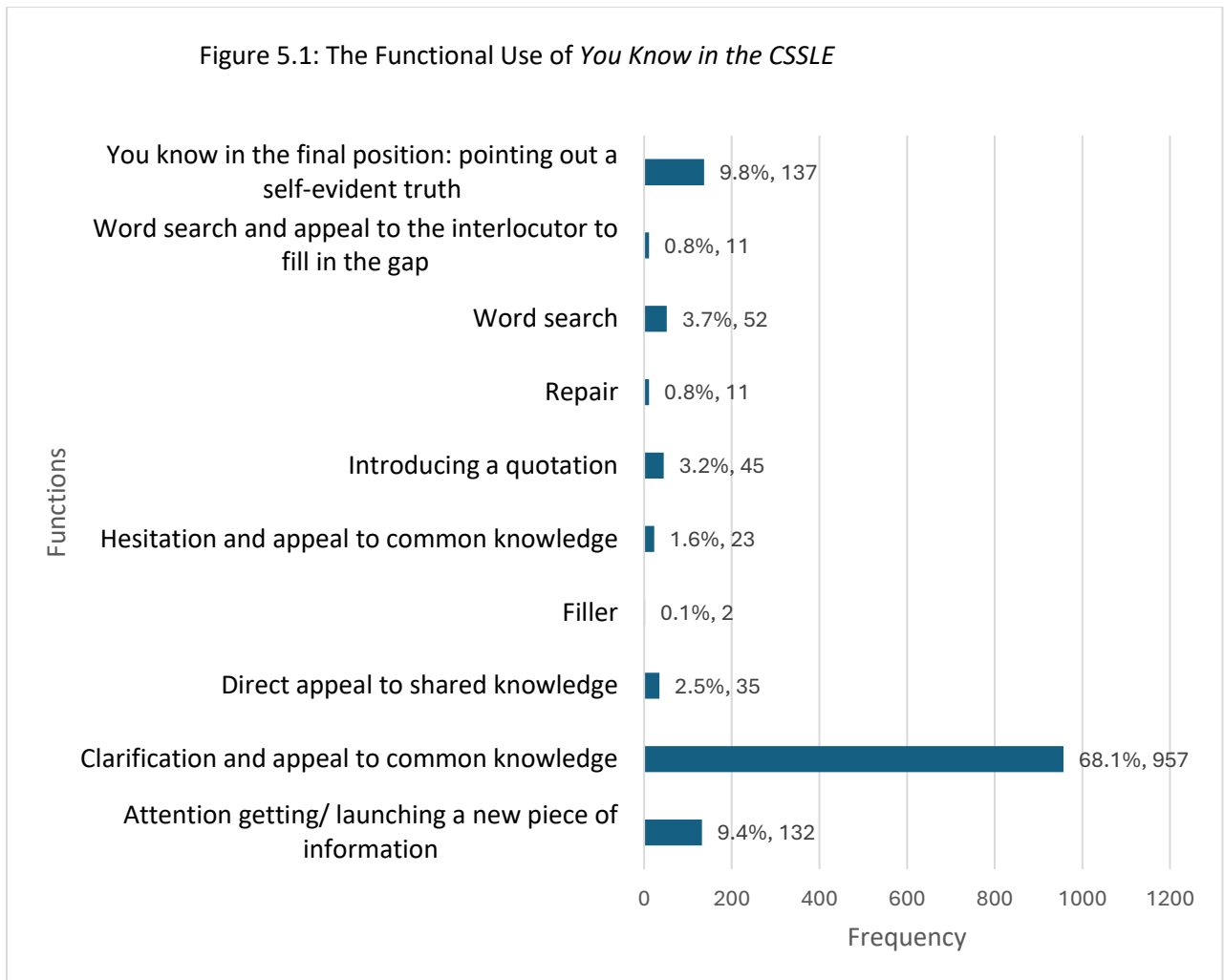


Figure 5.1: The functional use of *you know* in the CSSLE

Often, speech which includes frequent use of PMs is perceived as disfluent or careless, on the basis that PMs are most frequently used as fillers or hesitation markers as explained in Section 3.8. The list of functions in figure 5.1 very clearly shows that in the case of *you know* in SSLE, there is a specific and intended purpose for PMs and they are not just used for hesitation. The highest use of *You know* by far is for clarification and appeal to common knowledge. Next, although with a much lower frequency, *you know* is used in the final position pointing a self-evident truth and attention getting/launching a new piece of information respectively. The fourth most frequent use is word search. The least used function of *you know* is as a filler.

The most frequent function used by the SSLE speakers is clarification and appeal to common knowledge, with a total of 957 instances or 68% of the total use of *you know*. This reference to

common knowledge as a primary function is attested in much of previous research on *you know* (Östman, 1981; Fox Tree & Schrock, 2002). This prolific use of *you know* for this function can be argued to be a sign that, in this interactional context, SSLE speakers make a lot of effort to ensure that the interlocutor is following the discussion. It also implies an attempt to be polite to the listener by creating common ground and drawing attention to the immediate topic of discussion. 9.8% of the total use of *you know* is in the final position, pointing out a self-evident truth. Given below is example 5.30 to from CSSLE to illustrate this point.

Example 5.30

<02AF> I tell my kids all the time you know not to be afraid of boredom to sit to look out to just be um I think that's if I'm to be honest with myself that's something that I struggle with I'd at least want to watch a movie

<02HM> yeah

<02AF> cos I feel its I'm wasting time otherwise **you know**

<02HM> correct

Example 5.30 is from an interview with a female artist. There are two instances of *you know* in this example. The first introduces a quotation, but the second is relevant in this context. <02AF> says that she tells her children that it is alright to do nothing and feel bored. Then, she admits that although she says that to her children, it is a difficult advice to follow herself. She says that it feels like wasting time. This self-evident truth is presented with *you know* in utterance final position. This self-evident truth implies an honesty that the speaker wants to share with the addressee. This exchange of honesty can be interpreted as being comfortable in the company of the addressee. This can even lead to further cooperation between the two speakers.

In a similar percentage of cases (9%), *you know* is used to get attention by launching a new piece of information. Example 5.31 elaborates this use.

Example 5.31

<18HF> I can see erm after twenty nineteen the wall is er quite kept for some upcoming

<10EM> yeah

<18HF> stuff and we can also talk about

<10EM> **you know** after the easter tragedy

<18HF> yes

<10EM> last year we didn't do a campaign because the situation was not cordial

This context illustrates a conversation between an entrepreneur who sells jewellery and an interviewer. The interviewer is in the jewellery shop and in this context, they are referring to a wall where each marketing campaign for the jewellery company is recorded with a photograph. There is no marketing campaign after 2019 and the interviewer quizzes the gap in the wall. Immediately, the entrepreneur highlights the Easter bomb attack in 2019 in Sri Lanka which shocked the country and brought the industries to a halt. He introduces this new turn of events with *you know*.

You know as a word search device is only used in 3% of cases, and it is used even less frequently as a device of word search and appeal to the interlocutor to fill in the gap, in only 0.7% of cases. Similarly, *you know* as a repair device too is not frequently used in my corpus. 3% of total use of *you know* introduces a quote. It should be noted that the quote that is introduced does not have to necessarily be a verbal quote. Even a thought that is produced as an utterance is presented using *you know*.

Example 5.32

<08EF> so I thought **you know** Styled by Lily would give justice to what I do

Example 5.33

<02AcF> yes and I told my daughters seriously **you know** you don't always have to be like if you say something is wrong or feel something is wrong you know you can voice it and say no I don't think that's quite the way

In example 5.32, the entrepreneur is explaining how she came up with the name for her company and she introduces her thought that leads to the name with *you know*. The example 5.33 introduces a quote that was actually said as opposed to a thought. Both these instances follow the subject + verb + PM pattern of *I thought you know* and *I told you know*.

The function of repair is infrequent compared to the other functions. There are only 11 instances which is 0.8% of the total number of instances of *you know*. Example 5.34 illustrates this use. The speaker begins an utterance with “we it” and then reformulates the sentence to “we send”. This repair is preceded by *you know*.

Example 5.34

<8AcF> and we we it=**you know** we send invitations and we inform the Sri Lankans living there

Pettersson-Traba (2018) in her study of three corpora of American English and Beeching (2016) in her survey of British English report similarly low frequencies for repair with *you know*.

In only 2% of instances *you know* is used as a direct appeal to shared knowledge. *You know* is always a direct appeal to the hearer to establish the common ground with the speaker. However, for *you know* to be meaningful as a direct appeal to shared knowledge, the two speakers should have much overlapping personal experiences. The interviews that are in CSSLE are from formal to semi-formal situations where the interviewer may not necessarily be familiar with the interviewee. This could be a reason for this low occurrence. Macaulay (2002) also notes that there is no strong evidence of this function in his analysis of 846 tokens in Scottish English. He says that reference to shared knowledge is a basic meaning and function of *you know* and therefore, this result is surprising. His explanation is that the original meaning and function of *you know* may be partly bleached out. *You* in *you know* is an indefinite pronoun and it is a direct reference to the addressee. Therefore, it is awkward to repeat *you know* in an

utterance, causing this bleaching. Given below is 5.35, an example of reference to direct shared experience:

Example 5.35

<03AcM> based on individual rights you see and if you read the Donoughmore commission report **you know** that chapter on communal representation you know there was a famous phrase used by them

In the interview in 5.35, the interviewer and the interviewee are both academics and therefore, the reference to the Donoughmore commission report¹⁸ and the particular chapter is common knowledge for both of them. This is not the canonical use of *you know*, where the speaker would use *you* and *know* as a subject and a verb. In this instance, *you know* is uttered as one unit, as a PM. The context of the data in the present corpus does not provide such shared common ground much. The reference to “that chapter” also demonstrates through deixis that it is a chapter that the addressee is aware of.

You know is the least used as a filler. In fact, there are only two instances by one female entrepreneur. A reason could be the nature of the data. The data is from recorded interviews for public consumption. Generally, people plan their speech more consciously when they speak publicly. There is no information which suggests that the guests in the shows whose speech is the data for this study had been given the questions before hand to have prepared answers. However, even if it is completely spontaneous responses, the speakers would still organise their speech so that there are only a minimum number of fillers or hesitations. This is one explanation for the very minimal use of *you know* as a filler. Secondly, PMs are used with a high frequency as a filler to hide disfluencies. This low frequency could mean that fluent speakers do not overuse PMs, or in this case, *you know*, as a filler but for a sincere purpose that directs the conversation in favour of the speaker.

¹⁸ The Report of the Committee on Ministers' powers in Ceylon published in April 1932.

5.4.2 The Functions of *I Mean*

The second most frequently used PM is *I mean* which accounts for 362 instances. *I mean* is used for a total of 6 functions. As shown in figure 5.2 below, the function with the highest frequency, by a very large margin, is clarification, exemplification, elaboration, reformulation; the other five functions are concession and nuancing, hedging, hesitation, justification and self-repair. The high proportion of use for clarification, elaboration, reformulation is a finding similar to Pettersson-Traba (2018), who explored three American English corpora and reports a frequency of 62.15% of *I mean* for clarifications. Beeching (2016) too reports an overwhelming use of *I mean* for the same function in BrE, although accounting for just 30% of uses. The current research reports 71% of the total use of *I mean*. This suggests that the broad usage patterns of *I mean* in SSLE may be similar to those observed in other varieties of English. As Pettersson-Traba (2018) suggests, this could be because it is the closest in meaning to the canonical *I mean*.

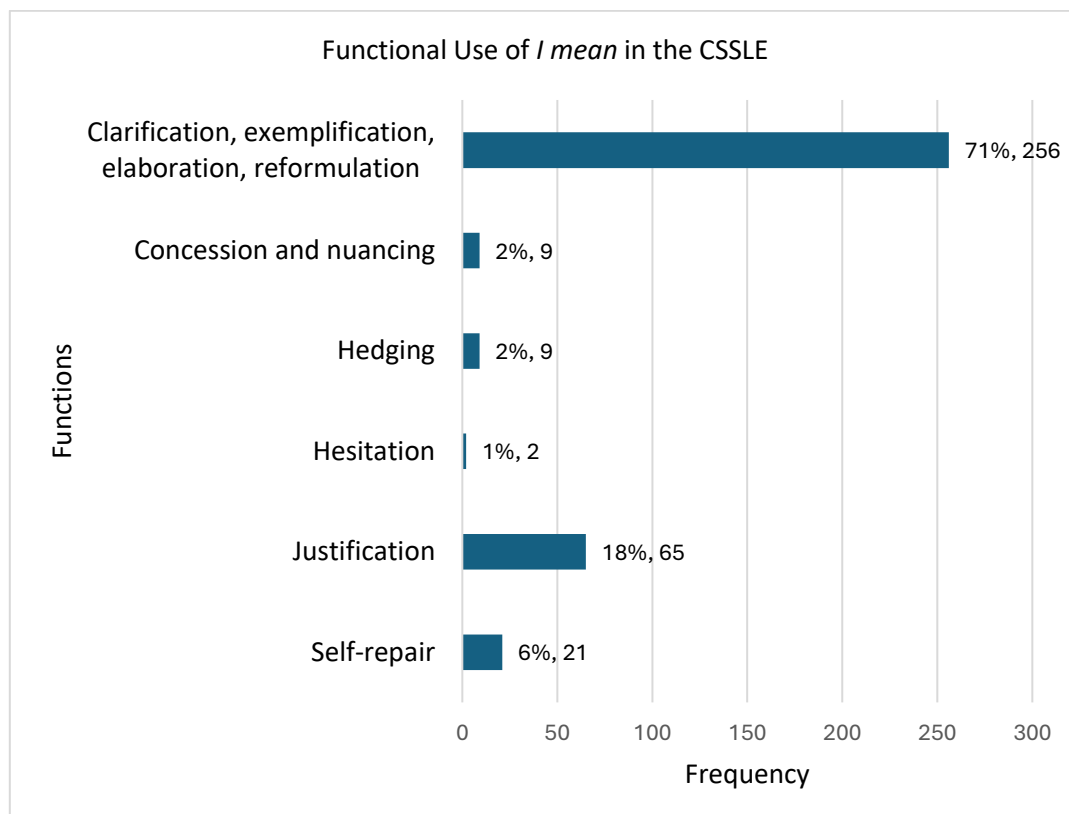


Figure 5.2: Functional Use of *I mean* in the CSSLE

A typical example of this use is provided in example 5.36. The female entrepreneur <03EF> uses *I mean* to preface the clarify and elaborate as to how she started her animal charity.

Example 5.36

<16HM> so tell me a bit about how you got into the er Embark er how you came up with rather the Embark

<03EF> er well I mean over the years even throughout the business in Odel I always erm always erm had a close combination with environment and erm and other animal related projects always supporting them and the **I mean** other people projects as well but always erm you know

Here *I mean* is used to further clarify and elaborate the point the speaker is trying to make. For example, in the second *I mean*, the speaker senses that she must explicitly state that she has always been interested in animal projects and not only hers alone, but even other people's projects. She introduces her supporting utterance with *I mean* to make her idea clearer to the interlocutor. This is not a justification to how she started her project or why she started her animal rescue project. It is merely an elaboration about her background to starting an animal shelter.

Justification is another function performed by *I mean*. Sixty-five of the 362 instances of *I mean* in the corpus (18%) function for justification. Beeching (2016) reports 22% use of *I mean* for the same function in her study of BrE. As a comparison, this is a close resemblance to BrE. This is another example to show that the broad usage patterns of *I mean* in SSLE are similar to inner circle varieties of English, and in this case to BrE. In this function, the speaker uses *I mean* to support or justify an idea s/he mentioned earlier. In example 5.37, the speaker <08AcM> provides justification to the comment he made to the effect that he believes the issue under discussion is serious:

Example 5.37

<14HM> erm now then I just wanted to ask you erm is it very significant or is it just

nothing that Sri Lanka=this er resolution sort of asking Sri Lanka I don't want to use the word against Sri Lanka but asking Sri Lanka to er you know up its anti and be better at whatever undertakings it's given in the past is it nothing or is it serious

<08AcM> I think it is very serious **I mean** the UN Human Rights Council is where forty seven countries of the world meet these forty seven countries are elected by the UN General Assembly of hundred and ninety odd countries they are given the mandate of ensuring that human rights in this world are better protected so the fact that Sri Lanka has gone is=before it and that these forty seven countries sat in judgment over Sri Lanka is very s=is very serious

In this conversation, the interviewer <14HM> is asking the <08AcM> how serious is the fact that Sri Lanka was issued a UN resolution for violation of human rights. <08AcM> being a human rights researcher mentions that it is very serious and gives a justification/reason for his comment preceded by *I mean*. *I mean* in this instance does not introduce an elaboration of his point. It introduces a justification to why he thinks the UN resolution is serious. It is serious because it is the decision with regard to Sri Lanka from 47 countries represented by the UN Human Rights Council. The speaker uses *I mean* to bring to attention the large number of countries that are part of the UN and that all these countries now have judged Sri Lanka as a justification to his response.

There are 21 (6%) occurrences of self-repair using *I mean*. This is almost double the percentage of both Beeching (2016) and Pettersson-Traba's (2018) studies: they both report 3.2% and 3.02% respectively. Roughly, it could be argued that SSLE uses *I mean* for self-repair twice more than both AmE and BE at least in this particular setting. Beeching (2016) investigated speech files of the British National Corpus (BNC) and University of West London Role Play Corpus (UWL- Role Play), which included spontaneous everyday conversations. Pettersson-Traba (2018) observed informal face to face speech interactions selected from the Corpus of Contemporary American English, Corpus of American Soap Operas and Corpus of Spoken, Professional American English. These text types differ on the scale of formality and spontaneity. Broadly, both Beeching (2016) and Pettersson-Traba (2018) explored informal speech. It can be assumed that the formal interview setting of my corpus makes

the speakers more likely to be aware of their speech and therefore, more likely to repair. Example 5.38 illustrates one such occurrence. The speaker begins to say something with “fu” and then repairs that utterance to “we go”. He indicates this change or repair with *I mean*.

Example 5.38

<05EM> I used to do a lot of stuff for Rotary clubs Rotaract Clubs Lion’s club we used to raise
fu=**I mean** we to go and play music

The least prominent functions are concession and nuancing, hedging and hesitation. There are 9 instances of hedging, 9 occurrences of concession and nuancing and merely 2 occurrences of hesitation. These functions add up to no more than 5% of its use in CSSLE. Similar results are reported for AmE for self-repair and concession and nuancing (Pettersson-Traba, 2018). Only 2% out of the total production of *I mean* is used for hedging. However, it is worth highlighting that these results are similar to BrE. Beeching’s (2016) study of BrE too reports a low percentage for hedging, 6.4%, but records more frequent occurrences for concession and nuancing than my study, 10% of total use. Another big different is seen in uses of hesitation, which account for 20% in Beeching’s data, whereas the CSSLE reports only 1%. Pattersson-Traba (2018) states that the corpus which contains the most formal AmE speech reports more hesitation than the other corpora. She suggests that the type of data is a reason for this high use, as hesitation is generally expected to be low with formal speech. Similar to *you know*, this result points towards the idea that the PMs are not typically used for hesitation or as fillers in SSLE in this context.

Instances of hedging and hesitation are given in examples 5.39 and 5.40, and 5.6 in Section 5.3 provides an example of concession and nuancing for *I mean*.

Example 5.39

<08AcM> er ge= fortunately the government has said they they will not implement it right
now they're going to consider it more which is
<14HM> oh so
<08AcM> very important

<14HM> that too is a proposal

<08AcM> yeah act=actually

<14HM> there seems to be lots of proposals

<08AcM> yeah **I mean** and in fact it's very sad that the government came up with this these laws which in a way targeting the Muslim community just before the UN human rights council took its vote because I'm sure

Example 5.38 illustrates *I mean* used for hedging. <08AcM> is an academic and heads the National Peace Council. The interviewer is asking him about a draconian law called deradicalisation that the government at the time was considering implementing. Since <08AcM> is a human rights activist, the interviewer asks him whether he wrote to the Minister of Justice with regard to this law. <08AcM> explains that several issues pertaining to this law had been highlighted by his organisation several times and the government had informed him in each occasion that it is only a proposal. The interviewer mockingly states that "there seems to be lots of proposals". <08AcM>, being an academic and a head of an organisation, does not want to overtly agree or disagree with this sarcastic suggestion. So, he says "yeah" but follows it with *I mean* to show hedging. He softens his evaluative comment. He continues to hedge his stance with regard to the government as he starts his next point with the term "it's very sad".

Example 5.40 provide the only two instances of *I mean* in the corpus for hesitation. In this speech situation, <12AM> is explaining how his household was when his father, who was a famous actor, was at home. The first *I mean* (not highlighted) could be counted as introducing a clarification of what happened when his renowned father was at home. However, the two *I mean* instances that follow seem to help him to gather his thoughts. Therefore, they mark hesitation while he is thinking about how to express himself.

Example 5.40

<12AM> er when he was here er I mean known people were **I mean** open to just **I mean** there was we had security

5.4.3 The Functions of Like

Schweinberger (2014) reports frequencies of *like* as a PM in several varieties of Englishes. The frequency for British English is 0.49 per one thousand words (ptw), 1.51 ptw for Indian English, 2.18 ptw for New Zealand English, 2.23 ptw for Philippines English and 4.38 ptw for Canadian English. Diskin (2017) reports a staggering number of 10-12 ptw for Irish English. If such a calculation is to be made based on my corpus, which shows 337 instances for 198,000 words, a frequency of 1.7 ptw can be reported, close to the figure reported for Indian English. The 337 occurrences of *like* in the entire CSSLE corpus are used for 5 functions: approximative *like*, discourse marking/hedging, exemplifying *like*, focuser, hesitation, quotative (be) *like*. Figure 5.3 shows their frequency of use.

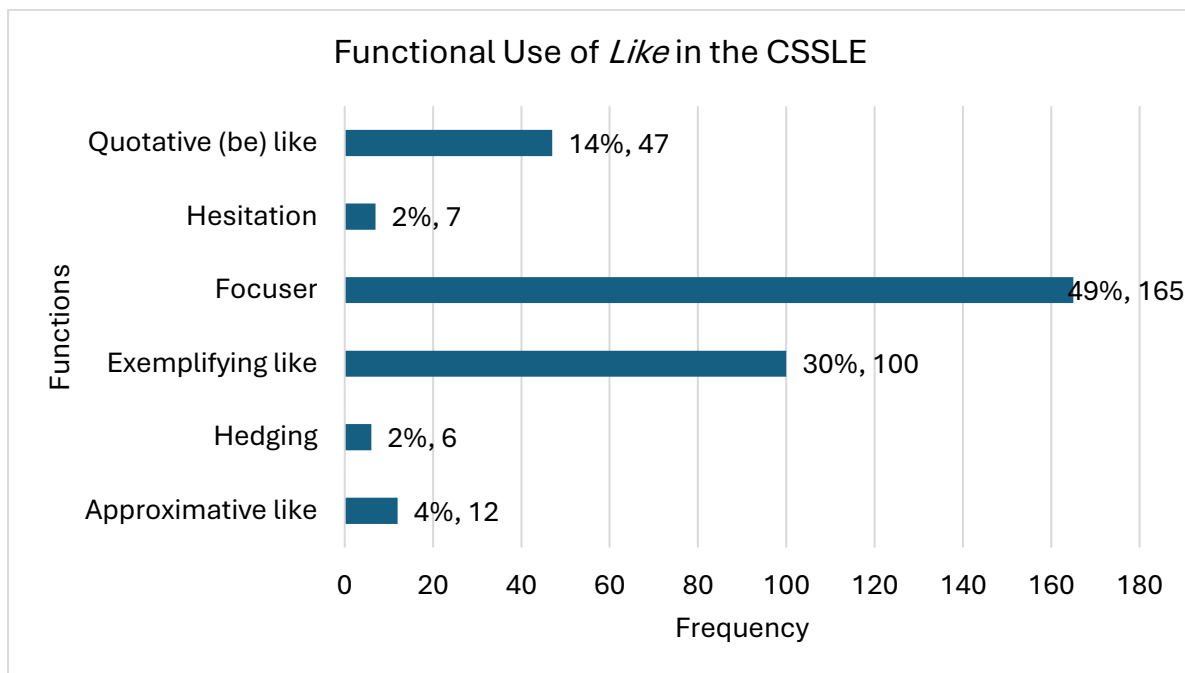


Figure 5.3: Functional Use of Like in the CSSLE

The most prominent function is *like* as a focuser which accounts for 49% of all *like* usage. This is a noticeable contrast to BrE as according to Beeching (2016), there are no occurrences of *like* as a focuser in her data, which consist of speech samples from BNC and UWL Role-Play. Beeching's (2016) data contains informal conversations. It could be assumed then that *like* as a focuser is unusual in BrE or that this function depends on specific text type. It is also evident in my data that all instances of

like as a focuser precede the phrase or word they focus which confirms Underhill's (1988) assertion that *like* is rule governed. In fact, *like* precedes the phrase or utterance that is subject to pragmatic marking in every function. A typical example of this use is presented in Section 5.3 under example 5.14.

Thirty percent of the total use of *like* in my corpus is for exemplification. Beeching (2016) reports 26% for her data. These two figures can be regarded as showing a similar pattern. Example 5.41 illustrates this use.

Example 5.41

<12EM> in terms of a hard decision I'd say I had to step down from fighting for few reasons
like I had a few injuries coming in

Example 5.41 is from the speaker <12EM>, who says that he used to be a boxer. In this utterance, he is providing one reason among many for giving up boxing. He uses *like* to introduce this reason.

The trend that is noticeable with regard to *like* in other Englishes is that it is more frequently used to introduce a quote (D'Arcy, 2017, p.16). In my analysis, for instance, this function accounts for 14% of the data, while it is non-existent in Beeching's (2016) BrE data.

Similar to *you know* and *I mean*, *like* used as a filler or for hesitation is marginal, with only 7 instances or 2% of the total use. The example 5.42 illustrates this function.

Example 5.42

<07AF> I have no idea **like** it just **like**

<01HM> <u=?>

<07AF> one day it just **like**

In example 5.42 *like* is used in three instances without a clear formation of an idea in the sentence. It seems that the speaker is trying to form a thought and is using *like* to hold the floor until the appropriate words are connected to create an utterance. The seven instances of hesitation in the CSSLE corpus relate to four speakers: four of the instances are produced by speakers born between 1981-1996. It is the youngest age group in the data. D'Arcy (2017) says that this pattern could be part

of discursive competency among young speakers. The very limited numbers make it difficult to analyse this use of hesitation meaningfully. However, the lack of the use for this function point to the assumption that it is not a feature of SSLE speakers.

5.4.4 The Functions of *Well*

Well comprises 7% of the total production of PMs in the current data. The corpus includes 205 instances of *well*, showing 8 functions. These functions are changing the topic, emphatic lexis, hesitation, prefacing a dispreferred response, self-correction, taking a turn/polite interruptions, transitional *well*, and quotative *well*. Biber et al. (1999: 1096) report that *well* is more common than *you know* and *I mean* in British and American English. However, this is not true for SSLE as *you know* and *I mean* are the two most frequent PMs in the CSSLE.

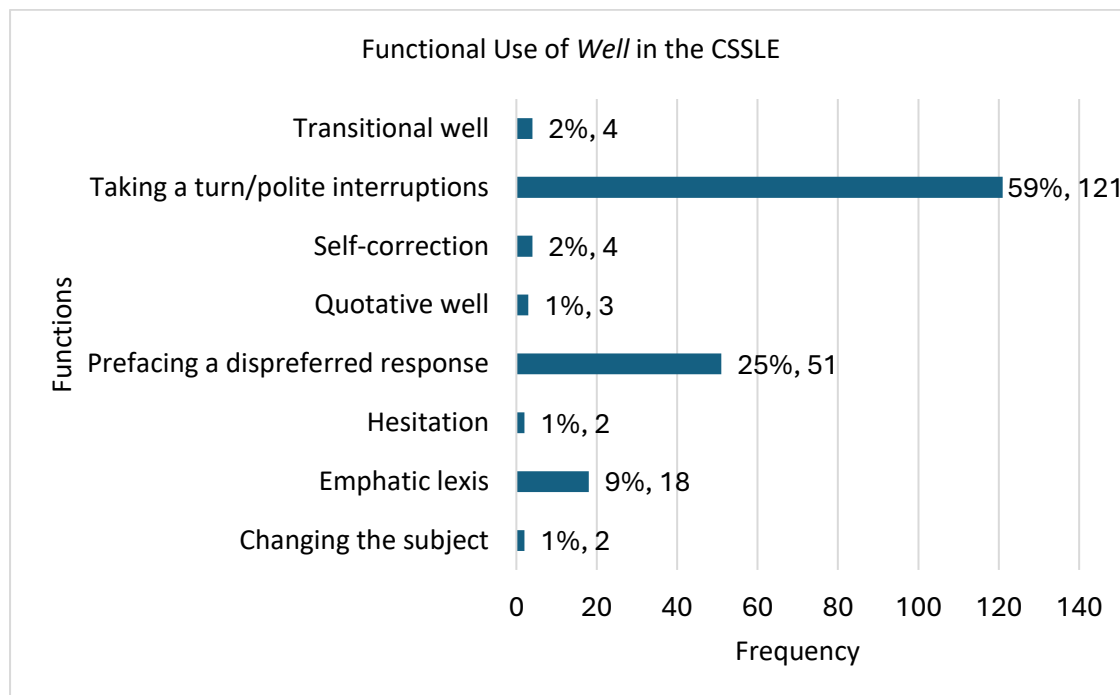


Figure 5.4: Functional Use of *Like* in the CSSLE

Well is by far most frequently used in taking a turn/taking a turn by a polite interruption. There are 121 such instances, which is 59% of total use of *well* in the CSSLE. It can be argued that *well* provides the speaker an easy entry into a conversation as it is established as a turn taker. *Well*, as an attitude marker could perhaps indicate confidence (Hale, 1999). When one starts an utterance

with *well*, it gives the idea that what will be said next is well thought out. In this sense, *well* indicates confidence. Therefore, a high use of turn taking with *well* could indicate a strong personality. The data in CSSLE is from prominent and accomplished persons in the Sri Lankan society. Therefore, the high use of *well* for turn taking could be an indication of their confident personalities. 25% of the total occurrences of *well* in this corpus are used to preface a dispreferred response. A defensive argument is presented in a polite manner by the use of *well*. The example 5.43 below from the CSSLE shows this kind of use.

Example 5.43

<01HM> <u=?> it irks you

<02AcF> **well** no it's just that women have been defined by normally whose father they are whose wife they are and that kind of as a woman not irks me but I wish the system was different

In example 5.43, <01HM> is asking <02AcF> whether she gets annoyed when she is referred to as her father's daughter, the father being a famous professor in Sri Lanka. She briefly says "no" and <01HM> persists asking if it irks her for which she gives the dispreferred response preceded by *well* to imply politeness. *Well* makes her response more polite though she is disagreeing. In contrast, *well* is used at the start of a turn to show approval as well as evident in example 5.44.

Example 5.44

<01HM> actually you always struck as being a character out of a Clint Eastwood movie don't ask me why but that's that's that's you to me

<09AM> **well** I like that compliment because I'm a real major fan of Clint

Here the interviewer <01HM> tells the interviewee <09AM>, a Sri Lankan film director that he looks like a character out of a Clint Eastwood movie. This comment is happily embraced by <09AM> as a positive comment. His approval is highlighted by the use of *well* in this instance. When *well* functions

as a turn taker, it introduces both positive and negative replies. Similar to the previous PMs *I mean* and *like, well* is infrequent as a hesitation marker in this context in SSLE. This is a contrasting finding to BrE, in which Beeching (2016, p.53) reports that *well* is used frequently to mark pausing and searching for words.

5.4.5 The Functions of *Right*

The fifth most frequent PM in the corpus is *right*, with a total of 184 instances (6% of total PM production). The Figure 5.5 illustrates its respective functions and frequency of use.

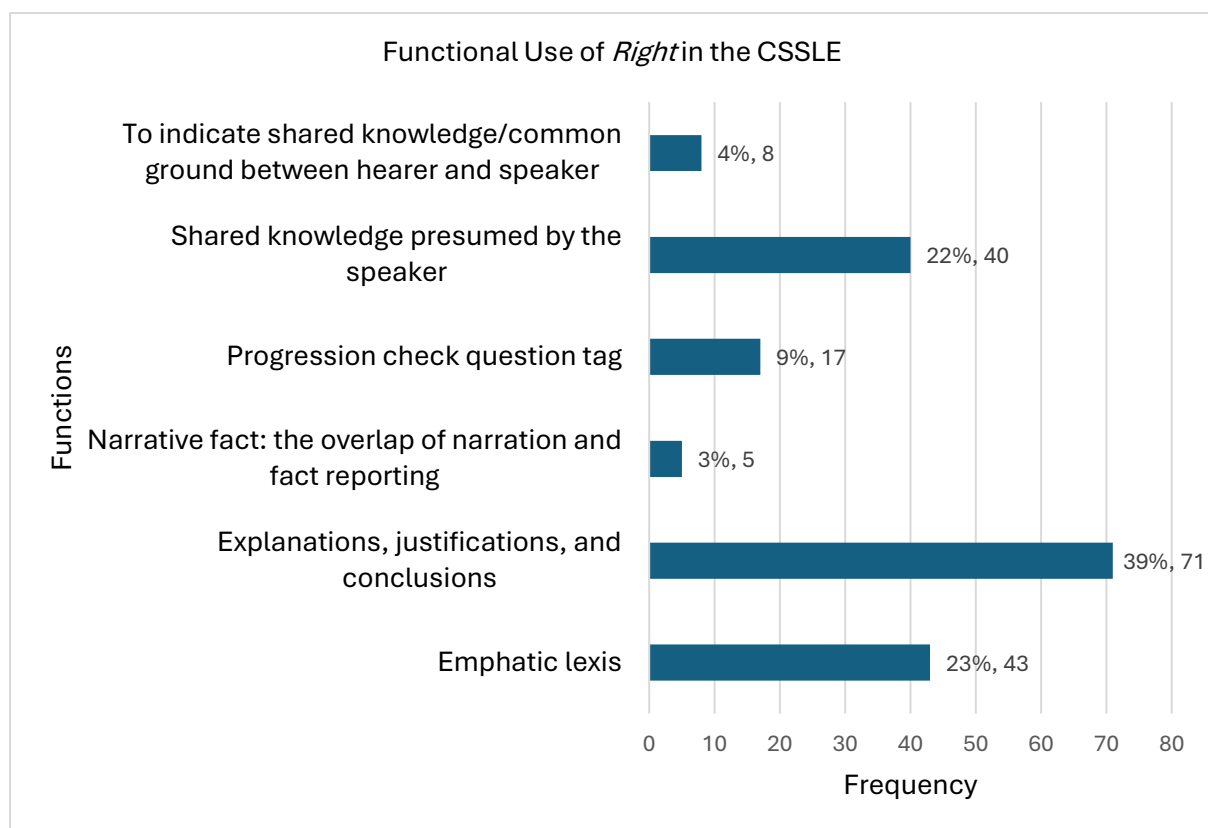


Figure 5.5: Functional Use of *Right* in the CSSLE

Unlike other PMs discussed so far, no single function dominates, although over a third of instances (39%) of *right* are used for explanations, justifications and conclusions. It is used second and third most frequently (just over a fifth of the total each) for emphasis and shared knowledge presumed by the speaker. *Right* is used as a progression check question in 9% of the total instances. *Right* has an equivalent in the vernacular Sinhala and in Tamil, the two main languages of Sri Lanka.

In both Sinhala and Tamil, the word *right*, in Sinhala *hari /ha'ri/* and in Tamil *sari /sa'ri/*, is used as a progression check question tag, for emphasis and for explanations, justifications and conclusions.

Therefore, it could be assumed that this native language use has influenced the speakers use of English as well. Example 5.45 from CSSLE illustrates its use:

Example 5.45

<17HF> small or big and they would like to know you know what your first work space was like

<04EF> yeah I was a secretary **right**

In example 5.45, <04EF> is asked what her first workspace was like and in response, as part of the explanation she says that she was a secretary. *Right* in this context can be replaced by *you see* which is also mainly used for explanations. In SSLE *right* seems to at once prepare the listener for a careful explanation. This is generally followed by further explanations.

In Trinidadian English, Wilson et al. (2017, p. 730) identifies one of the functional uses of *right* as a question tag. Nevertheless, the PM *right* could be differentiated from the invariant question tag *right*. Biber et al. (1999, p. 1089) state that the question tag *right*, which functions in an affirmative or confirmatory manner, generally requires a verbal response from the listener. My study regards question tags which do not require a verbal response from the listener as PMs and *right* serves as a progression check question tag which does not necessarily require a verbal or non-verbal response from the listener. A progression check monitors the listeners' understanding of a previous utterance or a proposition, not to excite a question, but perhaps to generate backchannelling (Schleef, 2009). The progression check question tags are used to indicate the end of a section or some important information. The speaker merely uses *right* as a progression check question tag for her/his satisfaction, to establish that the listener is following what is being said. Example 5.46 presents such an occurrence:

Example 5.46

<2AM> okay we'll create a song with Sinhala Tamil Tamil English in all these things embedded talking about peace here the project was we chose this er beautiful rhythm **right** a lot of you know er with a with a with a with a with a with a up tempo rhythm where the youth kind of loved it they embraced the rhythm we brought a melody where they embrace the melody then the subject

Right in this context is used to merely ensure that the listener is following the conversation. It is not used to elicit a verbal agreement. The speaker does not stop for any response from the listener and continues with the explanation after using *right* almost as an act of ingratiation. Then the speaker goes on to build on it by initiating new ideas. *Right* is least used to present narrative fact, the overlap of narration and fact reporting. There are only 5 instances, i.e. 3%. Example 5.47 illustrates this function.

Example 5.47

<04AcM> according to the Electricity Act

<14HM> okay

<04AcM> it's it is stated in the act **right** so plans are available and you know we can discuss about the plan in detail but the plan is professionally prepared

The two speakers in this conversation are discussing the inefficiency of the Ceylon Electricity Board (CEB) in Sri Lanka. The interviewer <14HM> is asking if the CEB has long term plans for which <04AcM> is stating that there are plans professionally prepared. The plans are stated in the Act which is a fact. In this occasion there is an overlap of narration and fact reporting. This is not an instance of simple emphasis as the narrative is backed by facts.

5.4.6 The Functions of Kind of and Sort of

Both *kind of* and *sort of* are discussed together in this section as they are used only for one pragmatic purpose, i.e., metacommenting, hedging, and qualifying in the CSSLE. Examples of these two PMs for this function are provided in Section 5.3, in examples 5.28 and 5.29. There were 170 and

121 occurrences of *kind of* and *sort of* respectively in the CSSLE. *Kind of* accounts for 6% of the total production of PMs while *sort of* accounts for 4.2% of all PMs. Taken together, metacommenting, hedging and qualifying fulfils 10% of all functions performed by the PMs. This distribution is a contrast to Reichelt (2021) study of *kind of* and *sort of* in BrE. It shows a higher use of *sort of* than *kind of*.

A pattern that is noticeable in my study is the filled pause followed by both *kind of* and *sort of*. There were 28 instances of *kind of* and 14 instances of *sort of* followed by the pause filler *er* or *erm*. This pattern accounts for 16% and 11% of the total use of the two PMs. The study did not mark unfilled pauses in the transcripts and therefore, whether there are unfilled pauses made after these PMs cannot be stated accurately. However, PM use coupled with filled pauses denotes that the speaker rather consciously takes a moment to metacomment, hedge or qualify a strong opinion or an idea. Although both the PMs are used for similar functions, SSLE speakers appear to prefer *kind of* to *sort of*. This is similar to AmE where *kind of* is more frequently used than in BrE (Biber et al. 1999, p. 870). Another pattern is the very infrequent use of what Beeching calls the general extender use of *sort of* and *kind of*, where they are coupled with *thing*. There are only three instances of *sort of thing* and 13 instances of *kind of thing*. Although these two phrases are also used to downplay or hedge an idea, they were not counted as PMs in my study. I regarded these two phrases as noun phrases which are part of the direct propositional meaning. An example would be:

Example 5.48

<12AM> I was pushed into it by other people so I didn't want that **kind of thing** to happen I just wanted to just take time off and do my own thing

In example 5.48, *kind of thing* refers to being pushed into something by other people. There is a direct relationship to an NP and therefore a direct reference.

5.4.7 The Functions of *Yeah*

There are 45 instances of *Yeah* used as a PM. *Yeah/so yeah/yes* are three variations of *yeah* noticed in the corpus and they are equivalent to each other. Since these are three variants of one

PM, I will refer to them collectively using *yeah*. Unlike other PMs discussed so far, which display a wide range of functions, *yeah* performs only four pragmatic functions, and is mainly used to mark the end of a turn. It is also used for taking a turn, topic shift and explanations, justifications, and conclusions. Figure 5.6 shows the distribution of the functions.

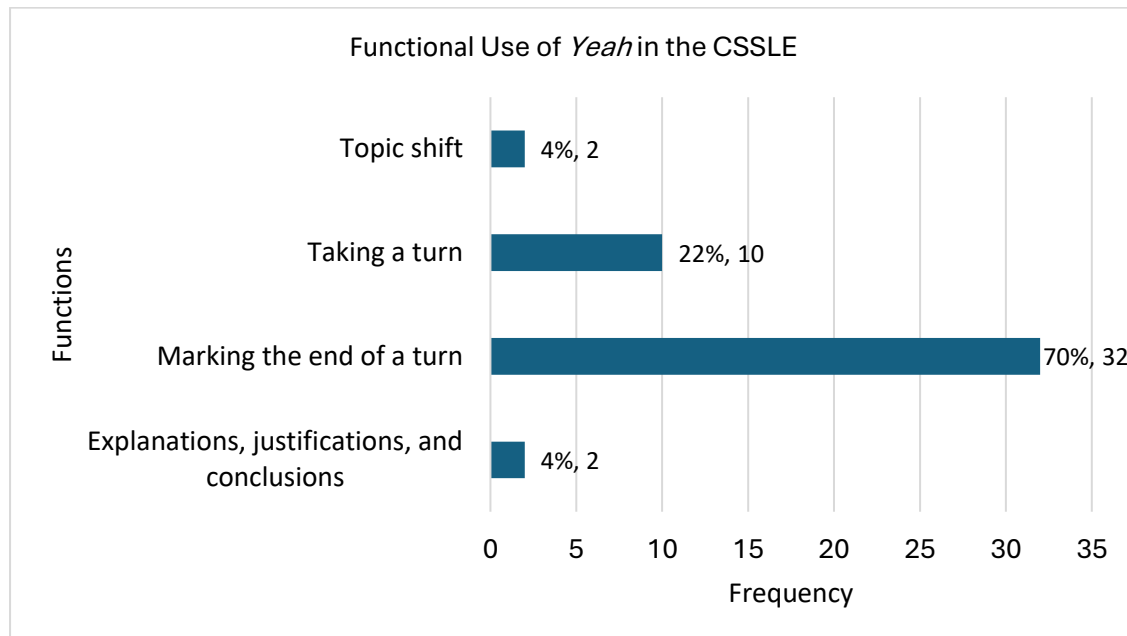


Figure 5.6: Functional Use of *Yeah* in the CSSLE

Much of earlier research on *yeah* report that it is used mostly as a turn initiator (Tao, 2003 - with regard to AmE). In fact, Drummond and Hopper (1993) mention that it is quite rare for *yeah* to occur outside a turn in initial position with regard to AmE (p. 205). 22% of *yeah* use in CSSLE is used for turn initiating. They do not record any turn ending function for *yeah*. With regard to BrE, Othman (2010) in a corpus study of academic lectures notes that *yeah* does have a pragmatic function when it is used to confirm shared knowledge. Her study too does not record *yeah* as a turn ending PM. She concludes that *yeah* is used to partition thoughts in speech. However, the striking finding from the CSSLE is that 70% of *yeah* used by SSLE speakers mark the end of a turn. So *yeah* or *yeah* seem to have a wrapping up or concluding effect in a conversation in SSLE. With reference to Othman (2010) it could be considered a partitioning of thoughts as it technically ends a topic. When a speaker is

elaborating or describing something and when the speaker has said enough, s/he tails off the comment by a *yeah*.

In example 5.49 from the sub corpora of female entrepreneurs, the speaker is asked what the frustrating moments in her start-up company have been. She describes her frustrations by giving a list, and then she continues her elaboration by admiring the staff and the clients. When she wants to stop the elaboration, she uses *so yeah* as an indicator of a stopping point.

Example 5.49

<06EF> yeah mm apart from that there is always frustrations like erm handling staff you know erm clients who don't who can't really erm communicate what they want so I have to kind of imagine you know but erm if I look at the big picture again I couldn't do without my staff I've a great team and or from my clients I learn so much **so yeah**

The turn initiator or turn taking function is the second most frequently used function for *yeah* in CSSLE, accounting for 22% of instances. It must be noted, however, that the total number of instances of this PM is relatively small overall (46), so any findings must be taken with caution. Tao (2003) reports on a study on turn initiators which studied more than 3000 turns in two corpora of AmE, where *yeah* showed the highest frequency, while *well* was placed sixth in order. This too is a contrasting result to CSSLE in which *well* had been used twice more than *yeah* as a turn taker, i.e. 59% vs. 22% of instances.

There are only two instances of *yeah* used for explanations, justifications, and conclusions, as illustrated in example 5.50.

Examples 5.50

<12AcM> the JR Jayewardena presidency of nineteen seventy eight **yeah** and parliament in this respect is purely a rubber stamp

<14HM> mhm

Yeah marks the explanation about the presidency of a former President of Sri Lanka. It is not used as a progression check, but rather draws attention on the shared knowledge with the listener, as *yeah*

can be replaced by *you see*. It seems to be used in lieu of *Do you recall?* Or *Do you agree?* which is clearly a reference to a shared knowledge.

Another function of *yeah* in this corpus is topic shift. Again, instances are limited to two occurrences. Example 5.51 illustrates this function.

Example 5.51

<02AF> that I wasn't such a nice person quite boring to spend my entire time [laughter] time with er and **yeah** er that was that was how it was started kids came back it was a whole different set of challenges er from online schooling you know everything else erm and as an artist

Here, an artist is asked how the lockdown and her Covid experience affected her and she gives a long explanation of how her break into the lockdown was intense and what she did during that time. Then she shifts her attention to the challenges of the lockdown with kids. The shift in focus is marked by *yeah*. It also should be noted that the speakers preferred to use *yeah*, which is a feature of AmE than *yes* in BrE in most of the uses. Perhaps, *yeah* is the more universal or globalized version of *yes* currently.

5.4.8 The Functions of *You see/See*

There were only 0.7% of *you see/see* in the CSSLE. *You see* is used in the hope that your interlocutor will understand what you are saying or asking. As a PM, it shows only three distinct functions. They are emphatic lexis, exemplifications and explanations, justifications and conclusions. While noting that, with such a small sample, findings need to be treated with caution, the data shows that more than half (57%) of the occurrences of *you see* are used for emphasis, while the remainder are used for explanations, justifications, and conclusions (24%) and exemplifications (19%).

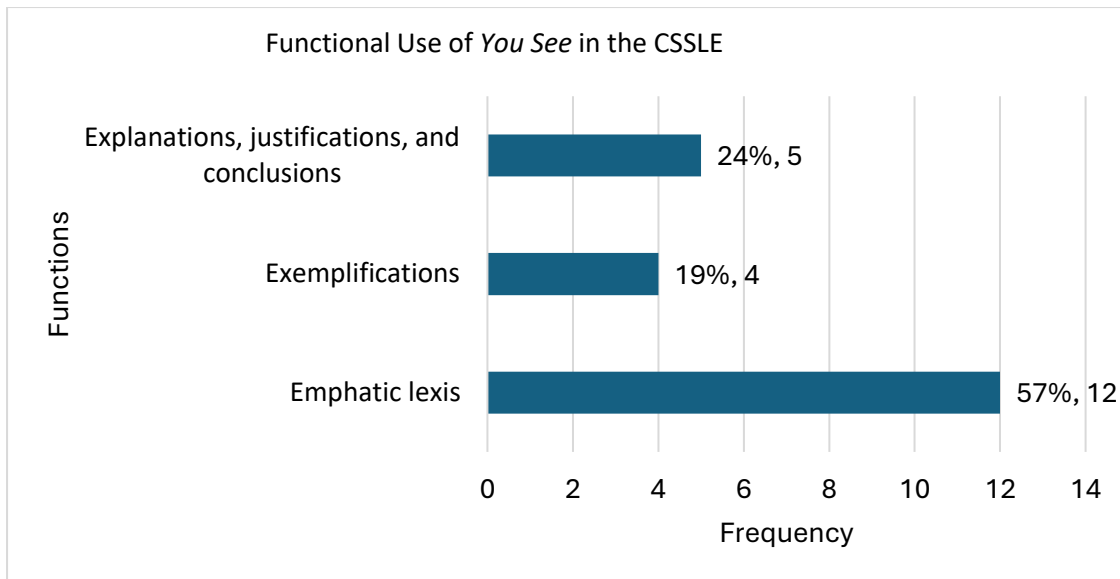


Figure 5.7: Functional Use of *You See* in the CSSLE

You see suggests that the speaker wishes the listener to understand and accept the speaker's ideas, as the example 5.52 below shows:

Example 5.52

<04AM> you know Kumar and that is what is lacking I think **you see** people think being a CEO or you are being a leader not necessarily true because most of them are managers they are not leaders so what does it take to lead

Here *you see* urges the listener to pay attention to the information that is about to be revealed. It introduces an example of what is lacking as well. In the above example, <04AM> is explaining what is lacking in the leadership in current companies. He is giving an example of this deficiency preceded by *you see*. There is a sense of sharing information that might be agreeable to the listener. Therefore, *you see* can be identified as a very strong seeker of agreement with the interlocutor.

In example 5.53, *you see* precedes an explanation, justification and conclusion.

Example 5.53

<03AcM> you see the critique of er the colonial boundaries in Sri Lanka comes from a variety of perspectives er the mostly nationalist and there are geographers also who critiqued that their arguments these are very arbitrary boundaries which is probably true

This turn initial *you see* presents an explanation of the colonial demarcation of boundaries. Because *you see* directly addresses the listener in the second person, instant attention is demanded and the listener feels as if s/he is co-constructing the response. The speakers sometimes drop the second person *you* and simply use *see*, and there are nine instances of this in my corpus. Example 5.54 illustrates this use.

Example 5.54

<01HM> like you know that adi pudi fight scenes

<11AM> **see** he's at a different league you know not adi pudi fights [laughter]

5.5 The Functions of Nativised PMs

There are two nativised PMs reported in this data, *no* and *aah*, although they occur in very limited frequency. There are 15 instances (0.5% of the total PMs) of these nativised PMs in the total corpus: two instances of *aah* and 13 instances of *no*. Despite the small number of occurrences, these still merit attention as they are signs of nativisation in PMs which support unique features of Standard Sri Lankan English. Nativised features are important evidence of forming the identity of new Englishes such as Sri Lankan English. *No* is a nativised feature of Indian and Pakistani (and even possibly Bangladeshi) Englishes (Shakir and Deuber, 2023). However, as mentioned in Section 2.7.3, the *no* in SSLE could be the influence of the interactional particle in Sinhala used clause finally for emphasis.

The functions of the two nativised PMs are analysed in this section. Although their instances are extremely limited, it is important to note them as they appear in public formal to semi-formal style video recordings published on YouTube. The fact that the speakers feel comfortable using nativised features may show that they are beginning to be perceived as the norm, since usually speakers avoid non-anglicised or less prestigious features of SSLE in public use of the language (Gunasekera, 1989).

5.5.1 The Functions of *Aah*

There were 2 instances of *aah* in the corpus. *Aah* can be replaced by *you see* in these instances and it is used for emphasis.

Example 5.55

<12AcM> who go on and do whatever they want particularly with regard to violating people's rights **aah** so we don't really have a government in the sense of a cohesive coherent government with a cohesive coherent plan to govern

Example 5.56

<09AM> yeah taste it Ceylon tea **aah** okay how is it Kumar

<01HM> not bad I'm sure Nihara makes better cup of tea

In example 5.55, <12AcM> is discussing the nature of politicians in Sri Lanka and he mentions that they particularly do what they want with people's rights. Since the interviewer is familiar with the context and the politics, he emphatically refers to the situation with *aah* indicating agreement with the listener. In example 5.56, <09AM> wants to state that he served his guest the best tea which is Ceylon tea. He uses *aah* to emphasise the tea variety.

5.5.2 The Functions of *No*

There are 13 instances of *no* in CSSLE. This corpus shows examples of two functions: emphatic lexis and taking a turn, which occur in almost equal numbers: six instances are used as emphatic lexis and seven are used for taking a turn. In Sri Lankan English *no* is identified as a question tag, an emphasis marker (Gunsekera, 2005) and as an operator that affirms or asserts (Ekanayake, 2020).

In addition to these functions, *no* used for taking a turn has been attested in previous research on other varieties of English. Lee-Goldman (2011) claims that that turn-initial *no* adds seriousness to a comment in AmE. He investigates the International Computer Science Institute (ICSI) meeting corpus, Berkely, California, USA and Fisher English Training Corpus of University of

Pennsylvania, USA. In this data *no* brackets previous talk as non-serious and presents an idea with a sense of seriousness.

The example 5.57 is from the transcript C03AM. It is a conversation between an artist and an interviewer. There are many occasions in this transcript that the interviewer is asking questions which are not of a serious nature although the interview in its entirety falls in the category of semi-formal. For example, the interviewer asks the artist <03AM> about his school life and the artist replies that he reluctantly went for cricket practice. The next question is one of light banter: “what happened you chased the ball”. This background is provided to show that the conversation in this transcript shifts between a light sense of humour and seriousness. In the following example 5.64, the host asks an artist in a light tone what he really learnt at Heywood art school and the artist adds seriousness to the whole answer by selecting to tell what he actually learnt; this move to seriousness is preceded by *no*. This is the function that Lee-Goldman (2011) attests in his study on AmE as well.

Example 5.57

<01HM> what did you really learn at Heywood that you have brought with you throughout the years to today

<03AM> **no** formally I mean you're taught from life drawing to object drawing to colour to all the the crafts

No as a turn taker is an established PM, as these examples from outside the CSSLE data shows. In an interview¹⁹, the President of Sri Lanka, Ranil Wickremesinghe who is considered as a fluent Standard Sri Lankan English speaker (Gunsekera, 2005) uses *no* to indicate turn taking. I refer to minute 1.26, minute 3.50, minute 9.49 and minute 21.50 in the interview, where he starts the turn with *no*. Example 5.58 begins with a direct question. The interviewer asks if the previous government should face a punishment as the courts gave a verdict that the former President and the finance

¹⁹ Firstpost. (2023, November 22). Exclusive: Sri Lanka caught in power battle between India and China? Ranil Wickemesinghe Answers [Video]. YouTube. <https://www.youtube.com/watch?v=A28AIP-Mk-c>

minister are responsible for the economic crisis in Sri Lanka in 2021. It might seem that he is giving a negative response. However, *no* is not a direct answer to the interviewer's question because his answer *no* is not connected to the rest of the utterance. *No* can be regarded as a turn taker that adds seriousness and weight to the President's answer in all these instances.

Example 5.58 (9.49 min)

Interviewer: what is the most difficult part of your job sir as President of Sri Lanka

RW: **no** first was to make the country realise the problems we are facing

Example 5.59 (3.50 min)

Interviewer: but the court has not announced any punishment do you think that they should face the consequences for their action or the very least apologise

RW: **no** the er crisis took place during their time no the uh the crisis took place during their time and then when Mahinda Rajapakse the Prime Minister gave up the president then called on the opposition leader to form a government which is the normal practice and in this case the opposition leader should have taken over because he had the backing of the government

Example 5.60 (9.49 min)

Interviewer: one of your biggest expenses is servicing the debt but Sri Lanka has not been able to generate the kind of revenue they need to

RW: no

Interviewer: do that how do you plan to meet your debt obligations then

RW: **no** we have to er depend more on our revenue we must we have to get the economy grow much faster which means er restructure the economy removing away the remaining fetters to growth

Example 5.61 (21.50 min)

Interviewer: but yes about the connectivity what were the big takeaways from that meeting

RW: **no** we decided to go into close economic relations

In all these instances, *no* as a marker of turn taking precedes an explanation. *No* adds seriousness to the utterance or statement that is presented and can be recognised as a well-established SSLE turn taker similar to *well*. A diachronic study could reveal whether *no* is replacing *well*.

The second function of *no* is its use as emphatic lexis. Given below is example 5.62 from CSSLE. <04EF> is a female entrepreneur and she explains that her products have a market in Sri Lanka. The reason for the demand for her products, she explains is the fact that they are hand-made. To emphasise that they are hand-made, she uses the PM *no* which can be replaced by *you see*. A direct translation to Sinhala would carry the Sinhala vernacular PM *ne* at the end of “handmade stuff”.

Example 5.62

<04EF> there is a market for it

<17HF> that’s true over here in Sri Lanka

<04EF> yes

<17HF> cos

<04EF> because hand

<17HF> yeah

<04EF> made stuff **no** it is very difficult sometimes you know

Here is another example:

Example 5.63

<04AM> I became a grandfather at sixty two last year but you see my father was a became

became a father at sixty so Kumar he is a very capable man and he’s a catholic **no** so he lived

up to the biblical saying go forth and multiply

<01HM> [laughter]

In example 5.63, <04AM> uses *no* to emphasize humorously that his father was a strong catholic and was influenced by the Bible. In this instance too, *no* can be replaced by *you see* and retain a similar meaning.

The use of *no* in such contexts is very common among Sri Lankan English speakers especially in informal speaking contexts (Gunsekera, 2005). In fact, it is also common to use the unanglicized *ne* (the Sinhala term itself) for emphasis in SSLE instead of *no*. This is similar to using *no* and *nah* interchangeably in Indian English and Pakistani English (Shakir & Deuber, 2023). There is no evidence of the use of *ne* instead of *no* in the current data. I have not seen this use in another study or any other data although as a Sri Lankan I am aware of this use.

5.6 The functions of Co-occurring PMs

There are 11 different co-occurring combinations in the current data. The combinations are *I mean like*, *I mean you know*, *like I mean*, *like you know*, *no I mean*, *you know I mean*, *you know like*, *you know sort of*, *well I mean*, *well yeah you know*, *well you know*. *You know* and *I mean* both seem to have a higher tendency to combine with other PMs. For example, *you know* is combined with another PM in seven of the 11 co-occurrences while *I mean* contributes to six of the combinations. Lohmann and Koops (2022) suggest that a motivation for two PMs to co-occur is the overlap of their functions. This is true for some of the combinations such as *I mean like*. Both *I mean* and *like* have overlapping functions such as for hesitation, hedging, and exemplification. Example 5.64 is an instance where this co-occurring PM is used.

Example 5.64

<10EF> and erm people can start showcasing what they can do erm er apart from that I **mean like** it's it's it's a struggle it's a gamble and I think i=it's like one step at a time you just once you put your foot in er I would say not to lo=look back just keep taking all those steps er facing all the hurdles and the stumbles that you know erm you come across and just move forward not to look back

In this example, the joint meaning can be categorised as exemplifying and an elaboration. Both *I mean* and *like* are used for exemplification although elaboration is a function of *I mean* but not of *like* in this corpus. Example 5.60 is from an interview between an entrepreneur and an interviewer. The interviewer asks the entrepreneur <10EF> what kind of advice she would like to give someone who wants to start a new business. The co-occurring PM introduces two negative examples of what a new business can be: a struggle and a gamble. This exemplification is also an elaboration of the entrepreneur's ideas.

In the corpus, the 11 different co-occurring PMs perform a total of 10 functions across their 83 occurrences. Table 5.4 presents the classification of functions for co-occurring PMs, showing that the co-occurring PMs generally perform a combination of functions.

Function	Co-occurring PM
Clarification and appeal to common knowledge	You know I mean, you know like
Elaboration	I mean like, like I mean
Exemplifying and appeal to common knowledge	You know like, like you know
Focusing and appeal to common knowledge	Like you know
Focusing and introducing a quote	Like you know
Focuser and launching a new piece of information	Like you know
Hedging and appeal to common knowledge	You know sort of
Nuancing and appeal to common knowledge	I mean you know

Turn taking and appeal to common knowledge	Well yeah you know, well you know, well I mean, no I mean
Hesitation and word search	Like you know

Table 5.4: The functions of the co-occurring PMs

The functions of the co-occurring PMs are not analysed quantitatively as the limited data sample would not guarantee the reliability of the findings. However, the most common co-occurring PM in this data set is analysed to get an idea of the role that these PMs play. There are 54 instances of *like you know* among the 83 co-occurring PM instances, and they fulfil five different functions. The different functions are focusing and appeal to common knowledge, exemplifying and appeal to common knowledge, focusing and introducing a quote, hesitation and word search and focuser and launching a new piece of information. Table 5.5 presents the functions of the co-occurring PM *like you know* in CSSLE.

Function of <i>like you know</i>	Frequency	Percentage
Focusing and appeal to common knowledge	34	63%
Focusing and introducing a quote	07	13%
Exemplifying and appeal to common knowledge	07	13%
Hesitation and word search	01	2%
Focuser and launching a new piece of information	05	9%

Table 5.5: the functions of the co-occurring PM *like you know* in CSSLE.

Two of the most common functions are illustrated in examples 5.65, 5.66 and 5.67 to show the functional variation according to the PM combination in co-occurring PMs. Example 5.61 is an example of the function of focusing and appeal to common knowledge. 63% of *like you know* is used for this function. This high use is not surprising, since focuser is the most prominent function of *like*.

Example 5.65

<03AF> I like being

<01HM> people who say not nice things

<03AF> I like being criticized because then I'll **like you know** change for the better

<01HM> the nasty comments

Like, as a single unit, functions as a focuser as presented in Section 5.4.3 but it does not appeal to common knowledge when used on its own. As a single unit, *you know* fulfils a combination of functions such as clarification and appeal to common knowledge and hesitation and appeal to common knowledge as illustrated in Section 5.4.1. In co-occurring PMs, the functions of the individual PMs are combined and presented as a joint function. In example 5.61, <03AF> points out that negative criticism will make her change for the better. She draws the attention of the interlocutor to the idea of this kind of change by using *like*, and simultaneously she appeals to the common knowledge of the interlocutor with *you know* because changing for the better is a common idea in society. Aijmer (2002) and Lohmann & Koops (2016, 2022) state that PMs that share similar functions can combine with each other. However, *like you know* is an exception as either *like* or *you know* do not share the combined functions. Two very different functions have come together to give a pragmatic interpersonal meaning in this context.

Example 5.66 is an example of *like you know* playing the role of focusing and introducing a quote.

Example 5.66

<8EM> during President Clinton's time in the US er US government signed an ag= a duty-free agreement with Africa called AGOA African Growth Opportunity Act er then I thought **like you know** this is a huge opportunity for me to go and explore er that I I spoke to er my good old friend my chairman Gloria Vanderbilt

In example 5.67, *like you know* introduces a quote in the form of a loud thought. *Like you know* as one unit focuses and introduces the thought. Both *like* and *you know* introduce quotations as seen in the current data in sections 5.5.1 and 5.5.3. In this co-occurrence, *like* performs as a focuser and not as a quotative while *you know* fulfils the role of introducing the thought in the form of a quotation.

This can be taken as a combination that confirms Aijmer (2002) and Lohmann & Koops' (2016, 2022) claims.

Example 5.67

<11AM> er see that's the stuff I do right III do things **like you know** murder mysteries horror

<01HM> for sure

<11AM> ye=er

Example 5.63 shows *like you know* functioning to exemplify and appeal to common knowledge. This is a conversation between a dramatist and an interviewer. <11AM>, the dramatist is explaining the kind of drama he produces. The examples of the genres are listed and simultaneously there is a sense of appeal to the interviewer to know the genres of drama. *Like* is used as an exemplifier on its own and *you know* is used as an appeal to common knowledge. However, these are not similar functions.

What is evident in this limited analysis is that diverse functions of individual PMs co-occur to achieve a combination of functions that go hand in hand. The similarity or difference of the functions of two PMs do not decide the probability of combining. It is suggested that the speaker intention plays a more important role than PM similarity or difference.

5.7 Conclusions

The analysis in this chapter has explored the presence, the meanings and patterns of pragmatic functions in Standard Sri Lankan English. The SSLE speakers in this corpus use 11 PMs and 11 co-occurring PMs. There are two PMs that can be regarded as nativized PMs: *no* and *aah*. The PMs have a clear function in SSLE. The 11 PMs perform 23 functions. Some of the PMs are highly multifunctional and they all cover structuring and interpersonal functions. The functions for most of the PMs could be determined by following the classifications used in previous research despite shortcomings such as overlaps, differences and using different terminology for similar functions. Some of the PMs share similar functions, but they are not all interchangeable. For example, one common function among five pragmatic markers is emphasis. Although the PMs *aah*, *no*, *right*, *well*

and *you see* fulfil this function, only some of them are used interchangeably. Each speech context or discourse situation demands different PMs which restrict PMs being used interchangeably.

The primary purpose of the PMs used by SSLE speakers in this data is to gain cooperation and maintain a cordial relationship with the speaker. This is visible from the most frequently used PM which is *you know*. It accounts for 49% of the total PMs in the data. *You know* is mainly used for clarification and an appeal to common knowledge. At an interpersonal level this function makes the addressee feel included and at a discourse structural level, the effort to clarify ensures that there is no breakdown in conveying the message of an utterance. This slant towards ensuring clarity is further confirmed with the results yielded for the behaviour of *I mean* and *right* as well. *I mean* is the second most frequent PM and it is mainly used for clarification, exemplification, elaboration, and reformulation. Although *right* is not in the top three highly used PMs, it is also used mostly for explanations, justifications and conclusions confirming the speakers' disposition to ensure that the ideas are communicated clearly. It should be noted that this high use of *you know* and *I mean* is reflected in studies that investigated PMs in AmE and BrE as well. By contrast, some functions behave very differently to AmE and BrE. One example is that compared to these two varieties, SSLE speakers use double the percentage of *I mean* for self-repair. The reason for this discrepancy could be the type of data that is observed in the studies of comparison.

Some conclusions about the functions of other PMs in the corpus can be drawn. The third most frequent PM used by speakers of SSLE is *like*. *Like* is used 1.7 times per one thousand words (ptw). This rate is more similar to Indian English (IE), which reports 1.51 ptw, than to British English (BrE) at 0.49 ptw. This pattern shows that the frequency of use for *like* is more similar to IE than AmE or BrE. In terms of functional use, SSLE speakers use *like* as a focuser the most. *Well* is used mostly for taking a turn that imply a polite interruption. It is reported that *well* is more commonly used than *you know* and *I mean* in both AmE and BrE, but in SSLE, this is reversed. *Yeah* is more often used as a turn taker than *well* in AmE, but SSLE shows *well* as a more popular choice than *yeah*.

Similar to nativised PMs, indications of nativised functions too were discovered. One example is that SSLE speakers use *yeah* mostly to mark ending a turn. This could be a nativised function of *yeah* as there is no record of this use evidenced in other varieties of English. Another example is that the functions of *right* as a progression check question tag and emphatic lexis show the influence of the native languages of Sri Lanka. For instance, both Sinhala and Tamil use the equivalent of *right* as progression check question tag and for emphasis. However, the same functions are recorded for other varieties of Englishes as well. Whether Standard Sri Lankan English speakers use *right* because of influence from Sinhala and Tamil or because of the influence from other varieties of English such as British and Canadian Englishes is to be investigated in further research.

Kind of and *sort of* are used for metacommenting, hedging, and qualifying in CSSLE. It can be seen that *kind of* is more frequent than *sort of* which is a contrast to the distribution pattern in BrE. The least used PMs are the two nativised PMs. They are mainly used as emphatic lexis and mark speaker turn. A conclusion that is drawn is that the nativised PMs show strong etymological connection to the two native languages of Sri Lanka, Sinhala and Tamil.

The data shows that SSLE speakers do not use PMs as fillers, which is often regarded as a characteristic of disfluency. For example, the PMs *you know*, *like*, *I mean* and *well* which indicate hesitation or used as fillers in other varieties show a very low number of occurrences in this corpus. Therefore, we can arguably omit PMs used as fillers as a feature of SSLE in semi-formal to formal settings. An investigation into informal conversations in SSLE might allow for a stronger claim to be made in future.

Finally, PMs in SSLE play a prominent role in maintaining the relationship between the speaker and addressee. A noticeable number of PMs are used despite the high formality level of the data implying that PMs are not perceived as a feature of weak language by SSLE speakers. SSLE use of PMs shows a broad similarity to BrE and AmE in their functional distribution but there are some patterns that are unique to SSLE. SSLE shows influence from BrE, AmE and IE in varying degrees in the use of PMs. An assumption is that these similarities and differences in terms of function heavily

depend on the type of data as well. The features of data such as speaker demography, medium and mode of communication and degree of formality seem to affect the results, perhaps more than variety of English. I would also like to note in the conclusion that it is difficult to find studies that are comparable to this one in terms of variety of English, speaker demography, data type and research design, and these are needed so that a successful comparison can be made between varieties.

Chapter 6: Gender and Pragmatic Marker Variation

6.1 Introduction

Gender is an important social demographic category that is explored with regard to language variation. It is also one of the most politicised and problematised social demographic categories (Cheshire, 2002). In this chapter, PMs are employed to give a genderlectal analysis of SSLE speakers. As Östman (1995) mentions, PMs are “windows” (p.100) through which we get a glimpse of a speaker’s attitudes and opinions. They mark a speaker’s subjective attitude (Schoning et al., 2023), and help to establish and sustain common ground between two speakers. For example, when a speaker uses the PM *you know* it immediately invites the addressee to pay attention to the idea that is shared. PMs are also often claimed to mark a person’s gender. For instance, they signal what is regarded as appropriate behaviour for women and men in diverse discourse contexts (Eckert & McConnell-Ginet, 2013; Vine & Holmes, 2023). Schweinberger (2018b) reports that *you know* and *eh* have been identified as markers of masculine communication. For instance, in his study based on the New Zealand component of the International Corpus of English (ICE-NZ), he claims that *eh* indirectly marks male identity because it indexes stances that men value such as covert prestige. *Eh* is regarded as a stigmatised type of a PM and men achieve covert prestige by using *eh* frequently. Therefore, this chapter intends to understand how male and female speakers of SSLE convey their perspectives and attitudes through the use of PMs. As reported in Section 6.6, differences in PM usage between male and female speakers can be observed in the CCSLE. This chapter discusses the implications drawn from the data to show how women and men use PMs in Standard Sri Lankan English.

The current chapter examines the question whether there is genderlectal variation in the distribution, use and functions fulfilled by the PMs in Sri Lankan English (SLE). In existing research genderlectal variation has been evident in discourse marker studies that explored hedges, politeness markers, adjectives, tag questions and hypercorrect grammar to name a few (Lakoff, 1975; Priesler, 1986; Holmes, 1995). Since hedging and politeness are two key functions of PMs, it is only natural to assume that gender may be reflected in the use of PMs even though some research claims it to be

non-significant. This point will be further discussed under Section 6.3 where gender neutral research outcomes will be discussed, alongside evidence of the impact of gender on the use of PMs. This chapter intends to present a brief overview of previous literature on the impact of gender on different aspects of language and then discuss previous genderlectal research specifically on PMs in Kachru's (1985) inner circle varieties and outer circle varieties. This will be followed by a discussion of the data from the present research in Sections 6.4 and 6.5 to explore whether gender as a social variable influences the frequency and the functions of pragmatic markers. The findings will be consolidated to facilitate a comparison with similar previous research. Finally, the conclusions are drawn on the impact of gender on the use of PMs in spoken Standard Sri Lankan English (SSLE) in Section 6.6.

A caveat is necessary in the use of terminology since there is much discussion and debate on the use of the words *sex* and *gender* in the literature. There is research to warn us of the problematic nature of the use of binaries in gender-based language analysis (Coates, 2015). The terms are more complex than the labels suggest (Cheshire, 2002). The two terms *sex* and *gender* fundamentally recognise biological and sociocultural differences respectively: while *sex* is a biological construct, *gender* is a social construct. Sex difference is more visible while gender is not. Gender differences may depend on sex differences. Sex differences attribute many tasks to men and women which lead to gender constructs. For example, work that requires great physical strength requires men. Equally, there are many social roles that are not determined by biological sex. Most of the variations in language are a result of gender rather than biological sex if analysed carefully. It is the gender construct that prompts people to speak in the way they do. Only a prosodic aspect such as vocal pitch may depend on sex. Yet, it is also difficult to clearly see the boundaries of sex and those of gender, and the distinction between gender and sex is sometimes used in indiscriminate way in variationist studies. In this background, since gender is a social construct similar to language, the term *gender* is used to in the present study.

6.2 Linguistic Differences Based on Gender: A Broad Survey of Previous Research

There has been much research on the impact of gender on language. A vast majority of the research has been conducted on BrE and AmE. This section attempts to briefly look at the impact of gender on some key areas of language, specifically vocabulary, grammar and conversational practice. In terms of vocabulary, Jespersen (1922, p. 247) asserts that men are “the chief renovators of language”. Coates (2014) states that the time when Jespersen was writing women were held responsible for introducing ephemeral words. As per Jespersen, women use more exaggerated adverbs than men. Adding to Jespersen, Lakoff (1975, p. 54) claims that women use the intensifier *so* more than men.

Another area where women and men differ in terms of vocabulary is the use of vulgar or taboo words. It is generally believed that women use fewer vulgar or taboo words. Again, noticing this difference can be traced back to Jespersen who claims that swearing “is found much more extensively among men than among women”. Lakoff (1975) too claims through anecdotal evidence that men use more expletives than women. This claim has subsequently been validated by various empirical studies (e.g. Jay, 1992; Berger, 2003). Schweinberger (2018a) investigates swearing in the Irish component of the International Corpus of English (ICE-Ireland). He looks at 255 speakers using 249 instances of swearing in a corpus of 153,260 words. He concludes that men seem to use substantially more swear words in same-gender conversations compared with their number in mixed-gender conversations and compared with women in both same-gender and mixed-gender conversations.

Some studies have suggested that there are conversational strategies that are gender specific as well. The present study examines the impact of gender on the use of PMs in conversations. PMs are used to strategise discourse. Therefore, it is relevant to understand the way in which men and women strategize their conversational practice. Coates (2004) claims that back channelling such as the use of *yeah*, *right* or *mhm* is used by women more than men mainly to show the listener’s

support for the speaker. Back channelling is used at moments when the listener support for the speaker is required. She further claims through her extensive survey of existing literature that men do also use back channelling or minimal responses in mixed gender interactions, yet they are often delayed. She concludes that men use delayed back channelling as a tactic to reinforce male dominance. This points that there is gender difference in backchannelling as well.

Hedging as a conversation strategy is another area where gender difference is often claimed (Lakoff, 1975; Preisler, 1986). Lakoff (1975) asserts that women use more hedges than men, and that hedges signify female vulnerability. However, subsequent scholars have discussed this hedging seen not as a mark of vulnerability but as a show of politeness. Additionally, a close examination of the purpose of the hedging would give us greater insights into femininity and masculinity than a quantitative analysis of the hedges. Holmes (1984, 1987) for example, looked at the differences of functions (purpose) of hedges in a study that examined a corpus which contained equal numbers of men's and women's speech. She concludes that women use *you know* more than men when its function is to express confidence and less when its function is to show uncertainty. Coates (2004) attributes the low use of hedges by men to their possible hesitance to speak about sensitive topics. Therefore, Coates assumes that if sensitive topics are discussed, both male and female speakers will use hedges as a valuable resource (p. 90). The functional categorisation of PMs that is part of my study allows us to observe whether PMs such as *I mean, sort of* and *kind of* which act as hedges showing certainty, uncertainty or the softening the effect of a comment are more prevalent in women's speech than men's.

The tag question is another linguistic form that shows genderlectal variation in speech. In early research pertaining to tag questions, Lakoff (1975) points out that tag questions diminish assertions and claims that women use more tag questions than men. In a more empirically tested study, Holmes (1984) performs a function-based analysis of tag questions. She analyses 90 tag forms found in a 43,000-word corpus consisting of equal amounts of female and male speech. She

categorises the functions of the tags as expressing modal meaning (i.e. uncertainty) and affective meaning (i.e. solidarity, politeness). Women and men do not differ greatly quantitatively. However, 59% of the tags used by women in their facilitative, solidarity-stressing meaning whereas only 25% of men's tags were used for the same function. Hepburn and Potter (2011, p. 145) analyse tag questions in 158 British everyday telephone calls and reports that gender is not relevant to their role. They further claim that tag questions do not index a psychological state of weak agreement or strong agreement. They advise that making a general correlation between the grammatical form and the social category of gender misses this type of granularity in data.

Looking at research focused on speech acts rather than individual forms, many studies have been carried out on compliments. Much of the foundational work in a variety of Englishes such as New Zealand English and American English on the use of compliments claims that women give and receive more compliments than men (Holmes, 1988, 1998; Herbert, 1990; Eckert & McConnell-Ginet, 2013). Additionally, women prefer more personalised forms while men prefer impersonal forms (Herbert, 1990). Rees-Miller (2011) looks at two corpora collected in 2008 and 2010 to investigate the nature of compliments in US college students. She reveals that it is not the quantity of compliments that signify gender difference but the type of compliment. Her study finds that both men and women offer comparable numbers of compliments but, the situations in which the compliments are given show a gender demarcation. For example, women freely give compliments in both unstructured settings and goal-oriented settings, while men are more likely to compliment only each other in goal-oriented activities.

This section shows that there is research suggesting that men and women follow different communication styles in the way they use vocabulary, grammar, and communication strategies such as hedges, question tags and compliments. Women seem to use more hedges and question tags while men use more expletives and fewer compliments. Women use more 'standard' language and men consistently used more non-standard speech. Similar to various linguistic phenomena such as

tag questions, expletives, and compliments which are sensitive to gender, the current research investigates whether these patterns can be observed in the use of PMs as well. The next section reviews previous research on the impact of gender on PMs so that prevailing trends are understood.

6.2.1 Gender in World Englishes research

The research detailed in the previous section on the gender-based analysis of the various aspects of language focuses on Kachru's (1985) Inner Circle varieties. Studies on world Englishes also show that gender plays a significant role in language innovation and use. For instance, in Indian English (IE), there is evidence that women initiated the use of the PM *yaar* meaning 'pal, mate, guy' (Lange, 2009). In her analysis of texts from the Indian component of the International Corpus of English (ICE-IND), Lange finds that two thirds of all *yaar*-tokens come from same-sex group conversations. This indicates that *yaar* is freely used by and among women. Similarly, women ushered in the use of existential and topicalized structures (Lange, 2012) and the use of quotative *like* (Davydova, 2015) as mentioned by Suárez-Gómez & Seoane (2021) in their study on the role of age and gender in grammatical variation in world Englishes. Parviainen and Fuchs (2019) explore the particles *also* and *only* in Indian, Hong Kong and Philippine English in terms of gender and other social factors. The study yielded mixed results. Clause-final *also* is more frequent in women's speech in Hong Kong English. Indian English and Philippine English did not show any clear trend related to gender. However, the overall results showed that female speakers show more use of the particles *also* and *only* in the three varieties combined than male speakers.

Tag questions are another feature that has been examined in world Englishes in relation to gender. Lange and Leuckert (2021) performed a gender analysis of the tag questions in ICE-IND. The corpus contained 218,531 words spoken by 242 different speakers. Results show that the indigenous *no* and *na* are used more by female than male speakers. In contrast, canonical tag questions such as *isn't it?* and *innit?* are used more by men than women. The pragmatic functions of these tag questions too have been analysed through a gender lens using a categorisation adopted from Tottie and Hoffmann (2006). The tags are classified according to the meanings they carry such as

informational, confirmatory and attitude. Attitudinal tags which emphasise what the speaker says and do not expect involvement or reply dominate for both female and male speakers. On the other hand, facilitating tags in which the speaker is sure of the truth of what s/he says but wants to involve the listener are used more often by male speakers, whereas confirmatory tags, where the speaker is not sure of what s/he says but wants confirmation, are used slightly more often by female speakers. Even though the differences are relatively small, gender seems to influence the selection of specific tag types in Indian English. Fung et al. (2023) observe tag questions in English and in Malay, used by 227 Malay millennials on Facebook commentaries, and note that males use tag questions three times more than females. They claim that although tag questions are traditionally associated with female speech, they appear to be a characteristic of Malay male Facebook users.

Hansen (2021) studies the quotative system in Ghanaian English, investigating whether women use quotative *be like* more than men. Her study, which investigates the Ghanaian component of the International Corpus of English (ICE-GHA), concludes that women use the quotative *be like* less frequently the older they get compared to their male counterparts in her sample.

However, not all features appear to show differences in use across genders. Fuchs & Gut (2015) investigate the use of the progressive in Nigerian English in apparent time and the impact of the variables age, gender, ethnic group and text category. They use the ICE-Nigerian corpus. They find that although male speakers produced more progressives than female speakers, gender does not influence the frequency of the progressive in a statistically significant way.

6.2.2 Gender in Sri Lanka and in Sri Lankan English

Research into the influence of gender in Sri Lankan English needs to be considered in relation to gender roles in Sri Lanka, and this is something that is explored in perceptual studies. In this section I try to create a picture of the role gender plays in language through such perceptual studies done on gender in Sri Lankan society more broadly. A name at the forefront on feminist studies in Sri Lanka is Dr. Kumari Jayawardena. She states in her publication on feminism and nationalism in the Global South, published almost three decades ago, that “in spite of conditions that appear

favourable to them, women have existed and continue to exist in a situation of subordination” (Jayawardena, 1986, p. 109). The following section looks at how women were perceived in important societal segments such as education and employment historically and in contemporary times.

Compared to other countries in the region, women in Sri Lanka were ahead in accessing education (Jayawardena, 1986). The Dutch who colonized Sri Lanka in 1656 established co-educational parish schools. This policy was continued and further developed by the British who displaced the Dutch in 1796. Influence by the ethos of England and India during the same time which emphasized the state’s responsibility in providing education to the masses, the Morgan Report (1869), among many recommendations on democratizing education, had recommended that girls-only schools should be opened wherever possible (Jayawardena, 1986, p. 119; Wenzlhuemer, 2008, p. 207). Although there was agreement on access to education for both genders, there were debates about which subjects should be learnt at school by girls and boys respectively. Jayawardena (1986) mentions comments by women of the bourgeoisie suggesting that girls should study drawing, needlework and dressmaking instead of music theory or trigonometry, implying that the education system was supporting the ideology that women should be good and useful housewives (p. 121).

Subsequently, in the early twentieth century, women have had much better recognition and leadership roles in the left movement which supported national independence for Sri Lanka. The leftist political movement which started in early 1930s had attracted many radical women, and the women’s movement in Sri Lanka was a result of this political movement for national independence (Jayawardena, 1986). The processes of education together with leftist political movement in Sri Lanka contributed to directing women into non-traditional roles, though Jayawardena (1986) observes that these were only superficially different from the traditional roles. However, Sri Lanka has a society which does not subject women to harsh oppression. Sri Lanka produced the world’s first female prime minister (through the process of a democratic election) in 1960 while maintaining the general patterns of subordination for women in society at large (ibid).

It is in this social context that we should view and understand the little research available on the role gender plays in Sri Lankan English. One of the first remarks on gendered syntax is simply one paragraph by Gunesequera (2005, p. 137) in her seminal work titled *Postcolonial Identity of Sri Lankan English*. She speaks about the gendered use of metaphors, expressions and idioms. Her description is based mainly on lexical differences. As for the lexical items mentioned such as “bloody” and “bugger”, she concludes that the gender differences are beginning to fade (p. 137). One recent study by Gries et al. (2020) examines the role of gender in genitive alternation in SSLE and BrE in the relevant corpora in ICE. They use a variationist approach to understand the choices speakers make when selecting which genitive form to use. They claim that these choices are affected by modality, variety and gender. Modality here refers to whether the language is written or spoken. According to their research, more s-genitives are used in spoken language than in written language, and there are also notable differences in terms of the variety of English, with BrE using more s-genitive than SSLE. Most significantly for this chapter, Gries et al. (2020) conclude that SSLE female speakers use more s-genitives than their male counterparts and BrE male and female speakers. The researchers claim that this indicates Sri Lankan women reacting more to the grammatical cue of definiteness than BrE women. Female SSLE speakers show the difference between definite and indefinite “bigger” (p. 135) than male SLE speakers. Their overall results provide two findings: 1) a stronger gender difference among men and women who use SSLE than BrE speakers, 2) a cross-varietal difference emerging between SSLE and BrE female speech. The researchers attribute this inclination of (female) SLE speakers to use the s-genitive to a transfer of structures from Sinhala, Sri Lanka’s majority language. In Sinhala, the possessum is always preceded by the possessor, which corresponds to the s-genitive. Similar first language influences had been observed in Singapore and Kenyan Englishes as well (Brunner, 2014). They ascribe this divergence of the pattern from that of BrE in female speakers to the female employment rate. There are fewer Sri Lankan females participating in the labour market than British women, and the researchers believe that this makes females associated less with the use

of English since they move in narrower social circles. In other words, the females are less exposed to more BrE-like patterns.

Funke (2022) observed the pragmatic nativisation of thanking in Indian English (IE) and SSLE with a comparison to BrE in ICE corpora. She finds that SSLE shows distinct patterns with regard to age and gender. For example, young females use *thank you* as opposed to *thanks* more frequently than older male speakers. She further concludes that SLE has developed nativised norms for thanking. For example, SLE speakers use *thank you* more often if it is used turn-internally or if a reason is included.

Degenhardt and Bernaisch (2022) investigate apologies in IE and SLE compared to BrE in the relevant components of the ICE corpora. Their study aims to find possible variety-specific apology patterns and the influence of gender and age in the choice of apologies. They focus on *sorry* in comparison to other forms of apologies such as *apologise*, *forgive*, *excuse me* or *pardon*. They find that apologies are more common in BrE than in IE and SLE. They conclude that male speakers do not apologise as often as female speakers in all three varieties of English. However, older male IE speakers and their female counterparts used apologies in a similar manner in their data.

Another feature that shows gender sensitivity in its use in world Englishes is backchannelling. Kraaz and Bernaisch (2022) analyses 3212 instances of backchannels in ICE-GB, ICE-IND and ICE-SL according to age and gender. The study concludes that both IE and SLE speakers use backchannelling more frequently than BrE. Female speakers in ICE-GB and ICE-SL use and receive more *mhm* backchannel than males and the reverse is noted for Indian speakers.

Degenhardt (2023) compares requests in IE and SLE with BrE in ICE corpora. She identifies request choices as a gender sensitive language item. Her study reveals that both SLE females and male speakers and IE female speakers seem to realise requests directly. Indian men differ from this pattern as they use the interaction predictor *you see/ look* more than the other groups. There is no gender difference between Sri Lankan men and women with regard to making requests. A study that

confirms the influence of gender and the social factor age is Degenhardt (2024)'s study on parentheticals. Degenhardt (2024) investigates parentheticals in spoken IE and SLE in the data from relevant ICE corpora. Parenthetical expressions are linguistic entities that are linearly integrated in another linguistic structure but are unrelated to the surrounding linguistic material (Dehé, 2014). She observes the parentheticals *I assume/ believe/ feel/ guess/ suppose/ think* in ICE-IND and ICE-SL. She finds that *I think* is the most frequently used lexical type comment clause in SSLE, and concludes that education is a stronger predictor than gender in the choice of parentheticals. Gender-based research on pragmatic aspects of SSLE seems to suggest mixed results. While gender is a prominent factor that affects pragmatic features such as apologies, it does not influence the use of requests. The results from previous research further reveals that the comparisons between SSLE and other varieties such as BrE and IE too yield mixed results. In sum, it is difficult to conclude if SSLE resembles BrE or IE more closely.

6.3 Research on the Impact of Gender on PMs

Research on the impact of gender on PMs offers mixed results. A possible reason for these mixed or contradictory results could be differences in the context of the speech samples, the size of the samples, how they have been derived and different methods of analysis. Another relevant factor is the lack of agreement on terminology and definitions. As Newman et al. (2008) aptly points out "one researcher's uncertainty verb phrase is another's hedge" (p. 215). These limitations prevent us from getting a coherent picture on the impact of gender on language or the focus of the current research, PMs. However, research also suggests that an analysis of function words can help us understand the underlying intentions of men and women (Pennebaker and King, 1999; Shapiro 1989). Function words are different to content words. Since PMs are a functional class of linguistic devices, the results of the current research can provide empirical evidence about gendered use of language.

Rayson, Leech and Hodges (1997) examine gendered speech styles in the spoken component of the British National Corpus (BNC). They report that female speakers take more turns and talk for longer than males, a finding that supports traditional beliefs about language use by women. They also report that informal interjections such as *yeah, okay, right, ah, aye* are more used by males than females. PMs mark turn taking and such multi-functional interjections act as PMs as well and therefore, this finding is significant to the current study as a point of comparison. In discussing sociolinguistic and stylistic features, Brinton (1996, p. 35) mentions that PMs are gender specific and more common in women's speech. She admits this is a controversial statement but elaborates that PMs express tentativeness and powerlessness, qualities that are usually associated with women's speech.

There is extensive research to show that *you know* is a feature of female speech. Fishman (1978) studied 52 hours of tape-recorded conversations between American intimates in their homes, a very different context to the current study. Alongside other language features such as asking questions, asking *d'you know* and minimal responses, she also observed *you know*. She found that women used *you know* 34 times and men only 3 times. In her analysis she concluded that women strive to achieve a smooth interaction, and men control what will be produced as reality by the interaction. Women use more *you know* than men in Östman (1981)'s study of AmE as well. She investigates 107 instances of *you know* in five conversations and finds that *you know* is indeed sensitive to sex-difference and in particular more used by women than men. Similarly, Macaulay (2002) finds that females use more *you know* than males in Scottish English. His study was based on two corpora that had a word count of about 254,000 which is similar to the current study. His data regarding *you know* showed a total of 846 instances of this PM in the corpus, divided into 57% instances by women and 43% instances by men. Laserna et al. (2014) report that the PMs (*I mean, you know, like*) were more common among women, young participants and more conscientious people. The study observed the speech of 263 participants, with male and female participants roughly in equal numbers. Filled pauses and discourse markers were correlated with age and

gender. They use the Big Five Inventory to carry out a personality analysis of the participants which gauges their conscientiousness. Their explanation for this association is that conscientious people are generally more thoughtful and aware of themselves and their surroundings. Therefore, when having conversations with listeners, conscientious people use PMs such as *I mean* and *you know* to show they like to share or rephrase opinions to recipients. The study concluded that females used the PMs *I mean*, *you know*, and *like* more frequently than males. This implies that women are more conscientious than males. In more recent work on BrE, Beeching (2016) also reveals that *you know* is used by females than males.

Frequency of use might favour one gender, however, research points out that functions too could be gender-biased. In addition to frequency, Koczogh and Furkó (2011) find gender-based differences in the functions of PMs in AmE. They say that females use *you know* less frequently than men for the function of seeking agreement. They also look at *I mean* and find that only women use *I mean* for the clarification of a misunderstood or misinterpreted utterance. In terms of function, Koczogh and Furkó (2011) claim that male speakers use *you know* as a device for topic change less frequently than women. Male speakers use *you know* more often for hesitation than female speakers do. Only men use *I mean* for summarising and providing a conclusion. A similar conclusion is drawn by Zheng (2015) in his study of BrE with regard to functions. His study shows that males used *you know* nearly as often as females. There were functional differences between the use of females and males. For instance, *you know* was mainly used for referential purposes by males and in connection to emotional comments by females.

By contrast, the prevalence of PMs in male speech is reported in research into some of the varieties of world Englishes. In New Zealand English, Stubbe and Holmes (1995) observe that there is no gender difference in the use of *you know* in the middle classes but in the working class, young males use *you know* more than their female counterparts. This is an example of research that shows that more fine-grained analysis might uncover more complex differences. Leimgruber et al. (2021)

investigate social variation in the use of discourse particles that are of non-English origin, and which fulfil the clause-final requirement in Colloquial Singapore English (CSE) in terms of ethnicity and gender. They find that both Indian and Malay male speakers use more non-English discourse particles than females in each ethnicity. For example, the PM *ah* is used by Chinese females less than males while Malay females use *ah* more than males. Indian females use dramatically less *ah* than Indian males. There is ethnicity-based variation in this use. Overall, *ah* is used less by females than males.

There is also research which claims that there are no significant or meaningful differences between female and male speech. Ferrara & Bell (1995) look at three AmE corpora compiled in 1990, 1992 and 1994 respectively which contain tape-recorded elicited personal-experience narratives from 405 different speakers in Texas. The study considers many social factors including gender. *Be like* is the second most frequent among other types of dialogue introducers such as *say* and *go*. Also, over time from 1990 to 1994, the female bias towards using *be like* neutralized and the use is not confined to females. In the early corpus, women used *be like* more than men, but in the later corpus, there was no significant difference. Buchstaller (2008) uses data from USA and UK to understand the use of *be like*. She selects 136 speakers from major dialect areas in the Switchboard corpus for the US data and 64 speakers from Derby and Newcastle for the UK corpus. The data shows that *be like* does not pattern by gender but by class and age for the US corpus. Similarly, in the UK corpus data it is patterned by age but not by gender or class. She claims that global trends of *be like* observed in other studies are not reflected in these locally specific corpora.

In sum, research yields mixed results about how gender affect the PM use. Yet, there are patterns emerging in terms of gender. The patterns are more affected by data type than the variety of English the participants speak.

6.4 Gender-based Results of the Present Study

6.4.1 General Results

The corpus used in this study has a total of 202,557 words, as mentioned in the Section 4.3.5, and is gender balanced in terms of the number of speakers. The female corpus (CSSLF) has 100,837 words and the male corpus (CSSLM) has 101,720 words. Since there is an 883-word difference between the size of the two corpora, the data was normalised to 100,000 to get comparable results. In the raw data a total of 2949 PMs were used by both men and women. Females used 1562 PMs and males used 1387 PMs. The normalised data translates this as 1549 PMs for females and 1363 PMs for males. As a percentage, women use 53% and men use 47% of the total PMs in the current data. In the overall results, females have produced more PMs than males.

There were 68,297 words, 67,546 words and 66,714 words respectively in the artists, academics and entrepreneurs sub corpora respectively. Table 6.1 show the raw distribution and normalised distribution of PMs in total according to profession and gender. The normalised frequencies for 100,000 words are given in bold in the second column. The PM frequency as a percentage is provided in the brackets next to normalised frequencies.

	Distribution											
	Artists				Academics				Entrepreneurs			
	Female		Male		Female		Male		Female		Male	
Total	440	644	554	811	623	922	268	397	498	746	565	847
		(44%)		(56%)		(70%)		(30%)		(47%)		(53%)

Table 6.1: Tokens of total PMs including co-occurring PMs used in the CSSLE corpus

As mentioned in Chapter 4, the following PMs were observed in the corpus (arranged in alphabetical order):

Aah, I mean, kind of, like, no, right, sort of, well, yes/ yeah, you know, you see.

In addition, the following co-occurring PMs were observed accounting for 2.81% of the total production of PMs:

I mean like, I mean you know, like I mean, like you know, no I mean, you know I mean, you know like, you know sort of, well I mean, well yeah you know, well you know.

Out of the total of 2949 PMs, females and males produced 24 and 59 instances of co-occurring PM respectively. This is 24 and 58 instances respectively when normalised. Females produced 29% of the co-occurring PMs while males produced 76%. These co-occurring PMs are not separately discussed with a gender-based analysis due to the limited number of occurrences. Thus, the PMs referred to in the remaining sections exclude co-occurring PMs.

Without the co-occurring PMs, females and males used 1537 and 1328 PMs respectively. The figures are 1525 and 1305 respectively for normalised data. Overall, females and males use 53% and 47% of PMs respectively. Of these, the most preferred PM for both females and males is *you know*, accounting for a total of 47% (656 tokens) by females, while males produced 53% (749 tokens). This is a contrasting result to previous research as traditional results show that *you know* is used more by females than males (Macaulay, 2002; Beeching, 2016).

According to table 6.1, female academics use the greatest number of PMs whereas male academics use the least. In the cohort of academics, female academics use 70% of the PMs while male academics used even less than half of that amount at 30%. The academics cohort shows the highest gender-based difference among all three occupation groups. In both other occupation groups, males have used more PMs than females.

Figure 6.1 shows the overall distribution of PMs in CSSLE for all pragmatic forms. The percentages show as a part of all PMs for each gender.

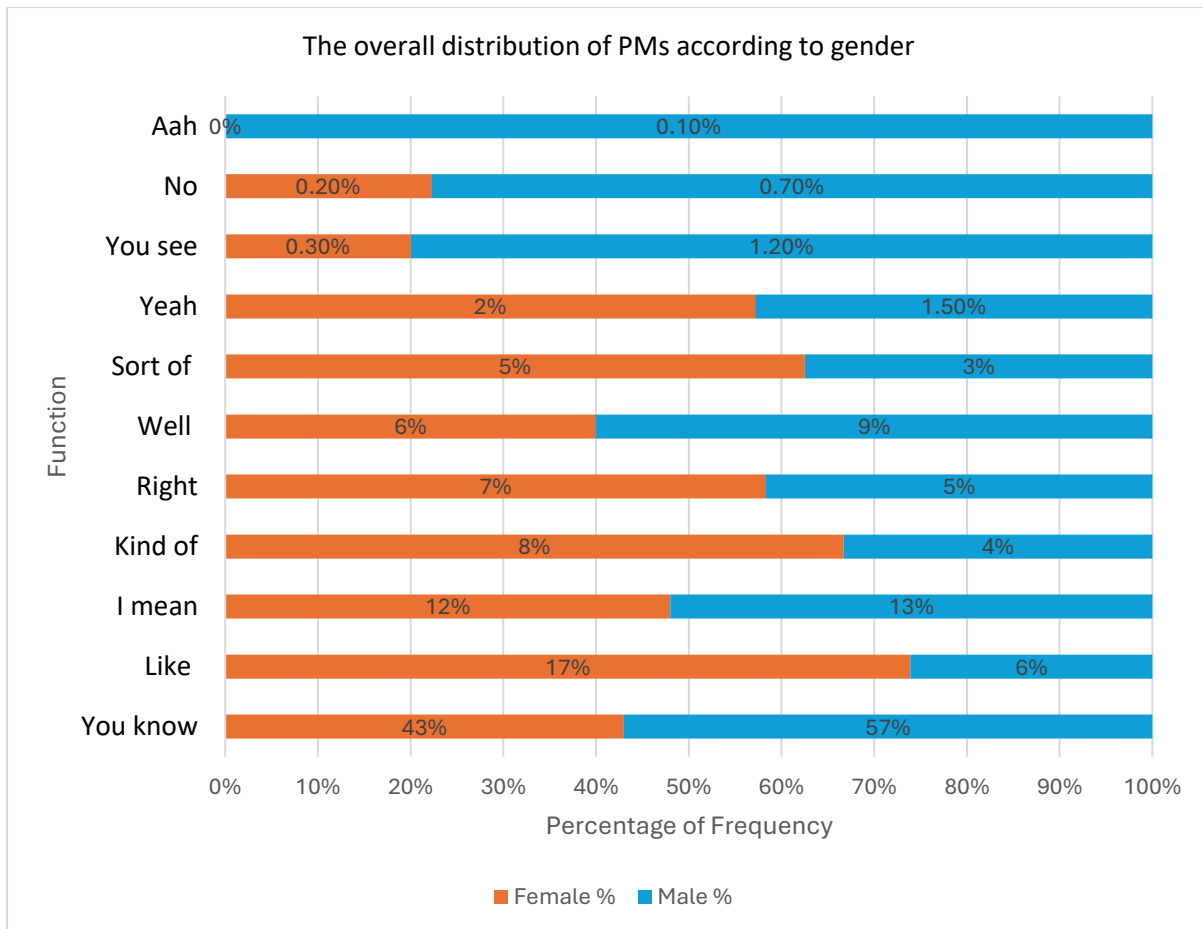


Figure 6.1: The overall distribution of PMs according to gender

As figure 6.1 shows, there are differences of preference according to gender. Both genders use *you know* the most. The second and third most preferred PM for women is *like* and *I mean* while for males, they are *I mean* and *well*. Females use *like*, *kind of*, *sort of*, and *right* more than males; moreover, females use more *yeah* than males while males use more *you see* than females. The usage pattern for *like* is expected as *like* is considered a feature of female speech as seen in AmE (D'Arcy, 2017). There are 2 nativized PMs present in this corpus, *no* and *aah*. There are 14 instances of these nativized PMs used in the whole corpus, and males use them more than females. These differences will be discussed in detail in from Section 6.5 onwards. A clearer and more comprehensive understanding of the PMs can be gained when they are analysed according to the functions. It will give better suggestions for the difference of use according to gender.

Although profession is not a variable I considered in depth, I calculated the frequencies according to profession and gender to gain a more complete and detailed picture of PM use in CSSLE (Corpus of Standard Sri Lankan English). The table below shows the overall distribution of PMs in CSSLE as a percentage according to gender and profession, and this shows some noticeable differences between the different groups of speakers.

PM	Frequency					
	Artists		Academics		Entrepreneurs	
	Female	Male	Female	Male	Female	Male
Arh	0	50%	0	50%	0	0
I mean	22%	24%	19%	11%	11%	13%
Kind of	17%	12%	29%	12%	24%	7%
Like	33%	15%	14%	1%	28%	8%
No	0	58%	8.5%	25%	8.5%	0
Right	5%	16%	48%	12%	9%	10%
Sort of	18%	6%	44%	17%	7%	8%
Well	18%	29%	14%	16%	12%	11%
Yeah/ so yeah/yes	22%	28%	7%	7%	28%	9%
You know	9%	17%	20%	8%	18%	27%
You see/see	9%	43%	9%	29	5%	5%
Total PMs used	15%	18%	22%	9%	17%	19%

Table 6.2: Percentages of the tokens of the PMs used in the SSLE corpus according to gender and profession

Table 6.2 presents the percentage of the total use of each PM by each gender. It is a percentage of the total production of the PMs in CSSLE. By looking at the percentages, an idea about

the share of production per gender and profession can be gained. Males in both the artist and entrepreneur categories use more PMs than females, though the differences between males and females in these groups is fairly small. However, a different and more marked distribution can be seen in the remaining group: female academics use the highest number of PMs, with twice as many instances as male academics. The percentages stand at 22% for female academics as opposed to 9% for male academics as a percentage of the total PMs in CSSLE. Female academics have produced the highest percentage of PMs out of all groups while male academics have produced the least percentage. As expected, *like* is more frequently used among females than males. *Like* is mostly produced by female artists while the male academics cohort has produced the least number. Additionally, the hedging or downtoning devices *kind of* and *sort of* are more frequently used among the female groups, and female academics lead over the other both groups. Overall, males have produced lower numbers of *kind of* and *sort of* than females. This result pertaining to *sort of* and *kind of* is not surprising as much research has attributed these two PMs as characteristic of female speech (Reichelt, 2021).

There is a striking difference in the data from the 24 academics in the corpus. In fact, these 24 speakers have affected the results considerably. Without these 24 speakers, the results would show the opposite, i.e. men producing more PMs than women. Females in the group of 24 artists have produced 44% of the total and males have produced 56%. The females and males in the group of 24 entrepreneurs have produced 47% and 53% respectively. This is in contrast to the academics group which shows that females produce PMs almost twice as frequently as males. The females produced 70% and males, 30% of the total PMs respectively.

The striking difference in the use of PMs among the 24 academics begs the question if profession has an effect on language use. In this particular data set, females have used almost double the quantity of PMs than males. In both the artist and entrepreneur data men have used more PMs than women. With this discrepancy in mind, first I analyse the results for individual PMs in the total corpus in terms of gender in Section 6.5. Following this overall analysis, the sub corpus of the

academics will be discussed in Section 6.6.7. Finally, a summary of the gender-based comparison of the results of the data will be presented in Section 6.6. The next section briefly gives an overview of the distribution of co-occurring PMs.

6.4.2 The Gender-based Distribution of Co-occurring PMs

The figure 5.2 below presents the overall distribution of co-occurring PMs in CSSLE. Overall, the results show that males use more co-occurring PMs than females. Females compared to males use co-occurring PMs sparsely. Only *I mean like*, *like you know*, *you know I mean*, and *you know like* are used by both genders. Males also use *I mean you know*, *like I mean*, *no I mean*, *you know sort of*, *well yeah you know*, and *well you know*, and there are no co-occurring PMs used only by females. A further analysis of co-occurring markers with regard to gender will not be conducted due to their limited occurrences in CSSLE.

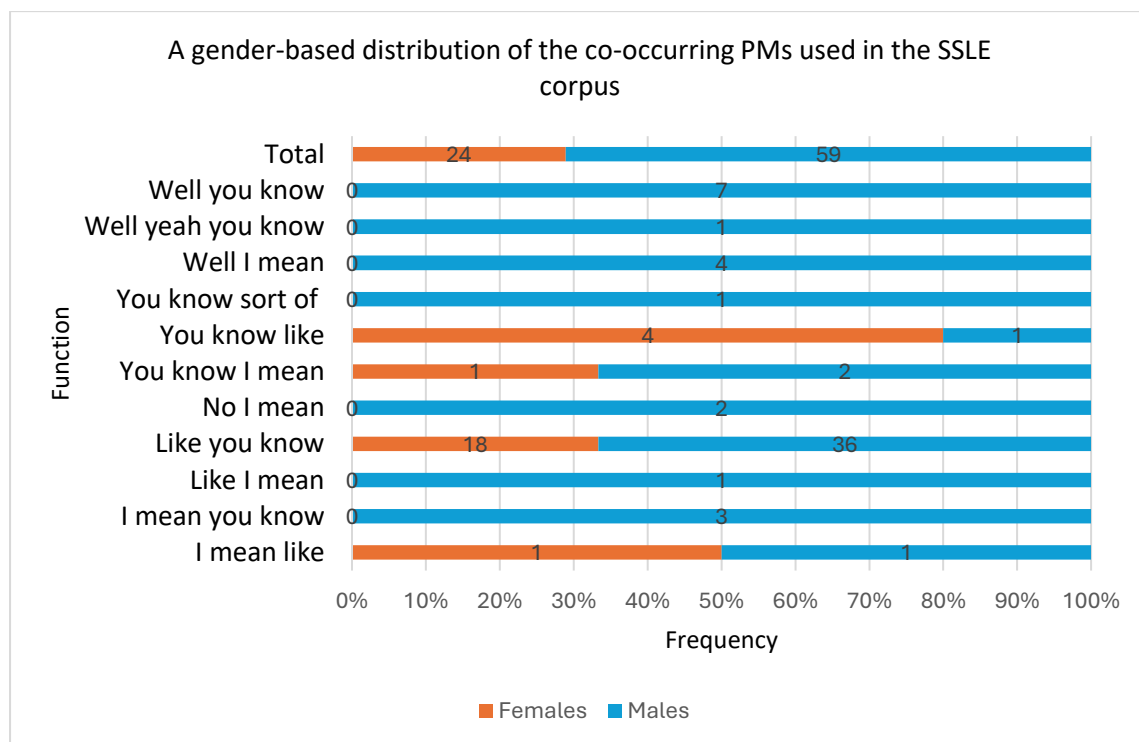


Figure 6.2: A gender-based distribution of the co-occurring PMs used in the SSLE corpus

6.5 Gender-based Data Analysis of Individual PMs

A second caveat should be provided with regard to the raw and normalised data. It was noticed that the difference between raw data and the normalised data is minimal. Therefore, the graphs and the tables from this section onwards are based on the raw data values and percentages.

6.5.1 A Gender-based Analysis of *You Know* in the CSSLE

	No of words (% of the corpus)	<i>You know</i> (%)
Female (artists, entrepreneurs & academics)	100,837 (49.7%)	656 (47%)
Male (artists, entrepreneurs & academics)	101,720 (50.3%)	749 (53%)
Total	202,557	1405

Table 6.3: Distribution of *you know* by gender of speaker in CSSLE

Overall, the results for the distribution of *you know* in the entire corpus show that men use marginally more *you know* than women. At a glance, my result contrasts with much of the previous research on Scottish English, AmE and BrE, in which women use more *you know* than men (Macaulay, 2002; Beeching, 2016). Previous research claims that women use *you know* for functions that indicate uncertainty or to seek cooperation which is a sign of weak or dominated speech (Fishman, 1978; Holmes, 1995). This overall result of the total corpus challenges the traditional bias (as regarded in the dominance approach) towards viewing women's speech as the deficit or weak variant (Lakoff, 1975).

This result differs from Macaulay (2002) for Scottish females and males, who claims that women are likely to use *you know* more often than men. His study reports 57% and 43% of *you know* occurrences for females and males respectively. This is almost the exact opposite result of the current research. The current study reports 47% and 53% instances for females and males

respectively. This difference could be attributed to a few reasons. One reason could be the fact that part of Macaulay's data came from teenagers. Research shows that teenagers' use of PM is different to adults (Stenström, 2014). It can even be attributed to possible differences between the two varieties of English. Further, my study contrasts with Fishman (1978) and Östman (1981) as well. In Fishman's (1978) analysis she concluded that women strive to achieve a smooth interaction, and men control what will be produced as reality by the interaction. Women use more *you know* than men in Östman's (1981) study of AmE. It could be interpreted that *you know* is regarded as necessary by women for interaction to occur smoothly in the data samples in the previous studies.

Although the results from Scottish English, BrE and AmE, show a higher use of *you know* by women, contrasting with my study, it cannot be hypothesised that this is a feature of these particular varieties or inner circle varieties. There are similar findings to mine in other studies with data from the same varieties. Similar results are drawn in Koczogh & Furkó (2011)'s study which reports that the male use of *you know* is slightly higher than female use in data from talk shows in AmE. The calculated D-value for their study is 91.17% and 92.77% for females and males respectively. The D-value is the number of *you know* or *I mean* functioning as a PM as a percentage of the total number of occurrences. The data type is similar to my study and perhaps this could be a reason to similar results. Further, my study shows similar results to Erman (1992) who found that men used *you know* more frequently than women. She studied *you know*, *you see* and *I mean* in female and male British English speakers. She was also investigating whether the production of pragmatic expressions depended on same-sex or in a mixed-sex environment which the results confirmed. The pragmatic expressions are more liberally used in mixed-sex in comparison to same-sex interaction. However, she does conclude that there are gender-specific differences in the use of the PMs in both types of interaction. One difference is that women tended to use pragmatic markers between complete propositions to connect consecutive arguments, whereas men preferred to use them either as attention-drawing devices or to indicate repair work. To see if gender impacts the functional use in

the present study, the functions of *you know* are analysed. Figure 6.3 shows the different categories of *you know* that were determined for this analysis as mentioned in Section 5.4.1.

The distribution of the functions of *you know* as a percentage of the total occurrences for each gender provides more insight to understand the impact of gender on functions.

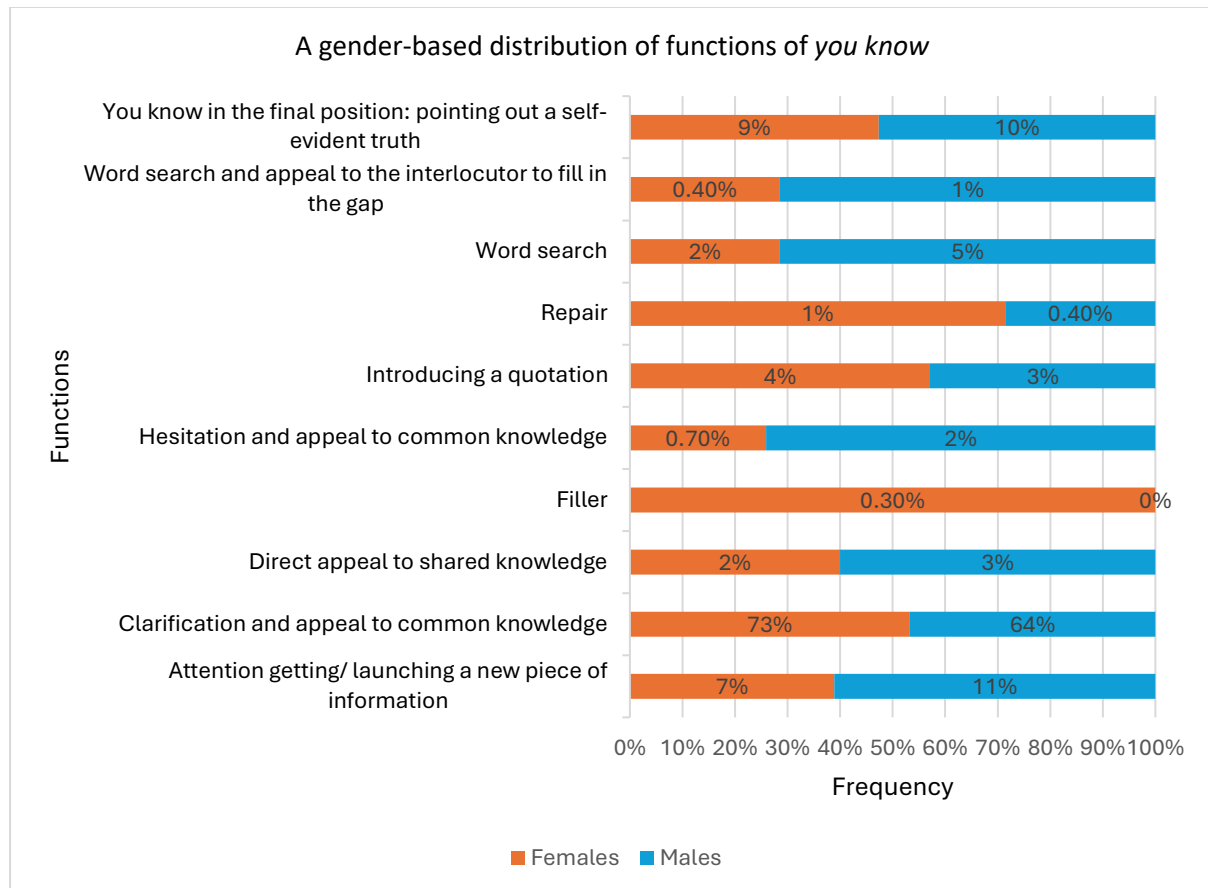


Figure 6.3: A gender-based distribution of functions of *you know*

The analysis of functions shows that both genders use *you know* mostly for clarification and appeal to common knowledge. The percentages of the functions are given as part of the total production of *you know*. For example, females used 9% of their total production of *you know* in the final position pointing out a self-evident truth while males used 10%. Females use 72.5% of total *you know* for this function, and males follow closely with 64.2% of total occurrences of this PM for the same function. It seems that clarifying ideas and establishing a sense of common ground with the interlocutor is important for both the genders. When one says *you know*, one expects that the

listener understands what was meant in the utterance accompanied by the PM. Whether the listener grasps the speaker's message in its entirety is not guaranteed for certain. *You know* used as a PM does not demand a verbal acknowledgement, so does not prompt any overt reassurance from an interlocutor. Yet, by using *you know* in this function, the speaker is satisfied into believing that an idea was clarified, and a common ground was reached with the interlocutor. All speakers seem to give considerable weight to establishing a bond with the interlocutor by using *you know* in this way. This dispels the traditional belief that seeking cooperation is a sign of interpersonal powerlessness especially seen in women (Fishman, 1978) as both genders use *you know* frequently for the same function.

Both genders use *you know* fairly equally to introduce a quotation, with 3.6% and 2.8% for females and males respectively. The quotation can be a verbal quote or a thought that occurred. In example 6.1 a female academic is assuming what other could be thinking of her driving skills and she introduced this thought with *you know*.

Example 6.1

<02AcF> yes I know and and it stunts them because when people you know when they see sometimes women driving and I might be a much better driver they tend to think erm **you know** you are not too good so once I got someone to put a window down and gave him quite a feminist lecture poor guy in the middle of the road

In the example 6.2 given below, a male artist is talking about his last-minute advice as a director of a play to his actors. He introduces his direct quote with *you know*.

Example 6.2

<01HM> okay

<01AM> er talking about stuff and then maybe try and tell them **you know** we should do it like this or why don't we do it like this

We can conclude that *you know* is useful to both genders to introduce a thought or a quote and there isn't a remarkable difference between the genders in this function.

Both genders use *you know* in the final position to point out a self-evident truth roughly equally. As percentages of total use of *you know* this represents 9% of use for females and 10% for males. It seems that there is no difference between the genders for this function. On the other hand, males use *you know* more than women to get attention by launching a new piece of information, with 7.4% and 11% of instances respectively for females and males. Men drawing attention more than women with *you know* is noticed by Erman (1992) as well in her observation of data from BrE. It seems that males seek attention marginally more than women in their interactions. Moreover, men get their attention from the interlocutor by launching a new piece of information. As evident by the percentages of 37% for females and 63% for males, men tend to introduce new information in a conversation more often than women. New information can direct the conversation which can also be interpreted as dominating a conversation. Therefore, in sum, compared to women, men seek more attention and dominate a conversation by presenting new information preceded by *you know*. Examples are given below. In both 6.3, *you know* precedes new information and the interlocutor does not interrupt the speaker for a considerable amount of time in the transcript.

Example 6.3

<01AcM> eight in eighteen eighty Sri Lanka installed its first telephonic er cables so very much on par with what was happening with the rest of the world and just as a sort of side **you know** in Sri Lanka they li=electricity the first light bulb was lit only in eighteen ninety at the Bristol hotel in Fort [...]

You know is used by men more than women for word search. As a percentage of all *you know* use, women and men use 2% and 5% respectively for this function. Speakers try to find the right word by stalling the conversation with *you know*. It could be interpreted that when you mark time with *you know*, the addressee too feels responsible to think about a suitable word until the speaker comes up with it. Men seem to transfer this responsibility to their interlocutor more than women. This could also be a way of engaging the listener and strategically save face of the speaker. Searching for the right word temporarily becomes a joint project until the speaker produces it. Word search as

appeal to the interlocutor to fill in the gap is listed as a separate function as explained in Section 5.4.1. This function too is marginally used more by men than women (F=0.4% Vs M=1%). It can be interpreted that male SSLE speakers prefer more than women to hold the conversation and appeal to the listener to help out when they are searching for a word.

All in all, on the basis of the data of the current study we can conclude that *you know* does index gender, yet not all functions show a gender-based difference. Gender is indexed not just by frequency but by certain functions used more by females than males. The functions and why these functions are important to a certain gender point to characteristics associated with a gender.

6.5.2 A Gender-based Analysis I Mean in the CSSLE

I mean is the second most frequent PM among all speakers. There are 362 instances of *I mean* in the total corpus, accounting for 12% of all PMs. Of these, 52% are produced by women (187 instances) and 48% by men (175 instances).

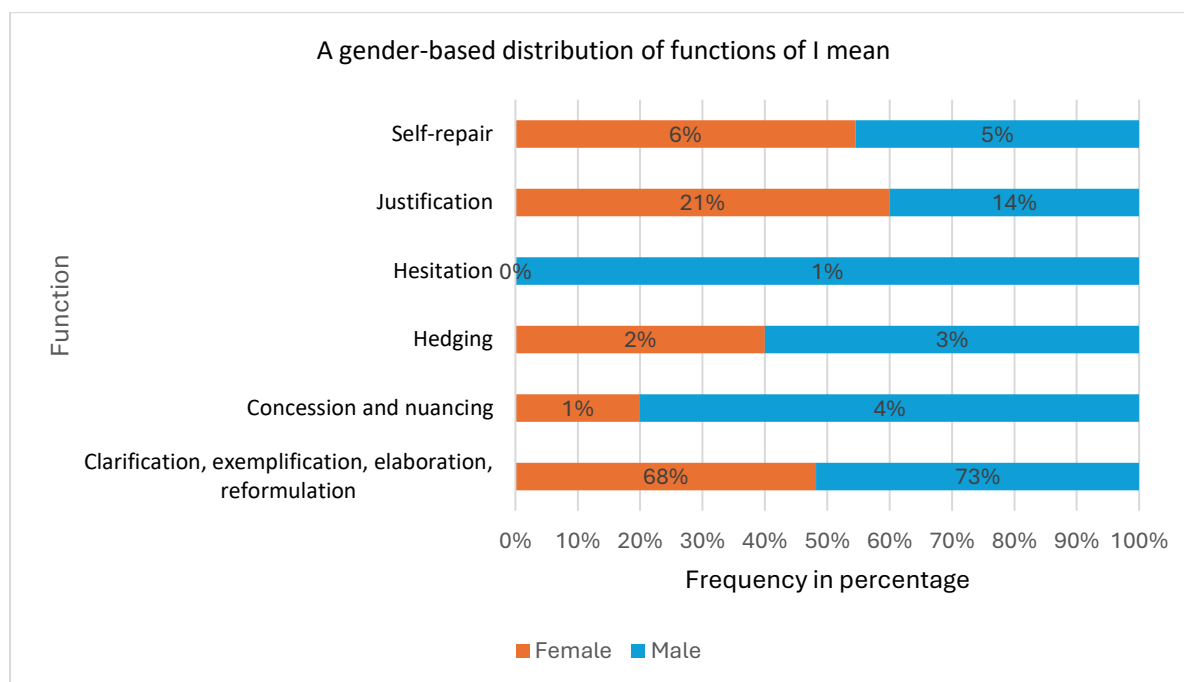


Figure 6.4: A gender-based distribution of functions of I mean

For both sets of speakers, the most frequent function of this PM is clarification, exemplification, elaboration, and reformulation, with 73% of instances produced by males and 68%

by females. It can be concluded that both males and females only marginally differ from each other with regard to clarifications and elaborations. The second more frequent use of *I mean* for both genders is for justification: this accounts for 21% of its total production for females and 14% for males. This indicates that women use *I mean* more than men to justify a declaration. An instance of justification is provided in example 6.4. Moreover, both genders behave similarly in the use of *I mean* for the remaining functions. They tally quantitatively with each other for self-repair, hedging and concession and nuancing in small quantities.

Example 6.4

<01AF> more stuff in the world through FB through internet through so many things so er I er it was not be= no big deal **I mean** nobody nobody said anything no not there sc= **I mean** were teachers who had watched this movie of their schools and er b **I mean** nobody commented on they always what they said was that my I acted well so I mean so

In example 6.4, the speaker <01AF> is an actor. The interviewer inquires about a suggestive and controversial role she played in a film. The interviewer wants to know how this role affected her personal life. The actor defends the role and says that the people that mattered to her personal life including the teachers in her children's school did not react negatively. She uses *I mean* to introduce every justification. There are three instances of *I mean* in example 6.5 and all of them are used for justification.

6.5.3 A Gender-based Analysis Like in the CSSLE

Lorenz (2022) concludes in her comprehensive literature survey on the use of *like* that the effect of gender is inconclusive. Different factors such as proficiency in English, exposure to (naturalistic) English Language use, how long one has stayed in an English-speaking country, the status of English (native speaker versus ESL versus EFL), age and register play a more significant role than gender. However, in the current data set, variables such as the status of English and dialect are

stabilised. Therefore, it is reasonable to assume that the noticeable variation is a result of gender influence on the use of *like*.

Like is the third most frequent PM in the CSSLE, with a total of 337 instances. It is 12% of the total production of PMs. Both males and females use *like* as a focuser the most and both genders use *like* as a focuser in equal measure. Women report 51% while men report 43%. Given below are 6.5 and 6.6 which are examples of focuser function from the respective corpora.

Example 6.5

<04AF> I'm at gigs and stuff and I'm always surrounded by **like** amazing food

<03HM> yeah

<04AF> and drinks and stuff so it's really difficult **like** temptations all

<03HM> cake

Example 6.6

<10EM> er it's one off because he's got a Swiss watch designer to replicate his horse on the watch er then two thousand and fourteen er I call it home town I went back to Galle which Galle Fort which is our hometown and which we still treasure because this campaign was shot er in the Fort I mean that's the Galle Fort lighthouse and that where you go and jump into the sea or whatever this was in the old Dutch fort where there is **like** a jail and stuff like that

In example 6.5, which is from a female artist, *like* focuses on the idea that she is surrounded by “amazing food” at gigs. In this context, she is talking about how difficult it is to stay vegan as in her line of work she is surrounded by tempting food. *Like* also functions as a focuser, specifically for the type of food that tempts her. In example 6.6, the focuser use is also apparent. The speaker directs the focus of the listener to the word “jail” using *like*. In both of these examples, *like* could have been omitted, but by focusing on a word, which is usually a word that leads to a mental picture, the user aims for clearer communication with the interlocutor. Figure 6.5 is a gender-based distribution of the functions of *like*.

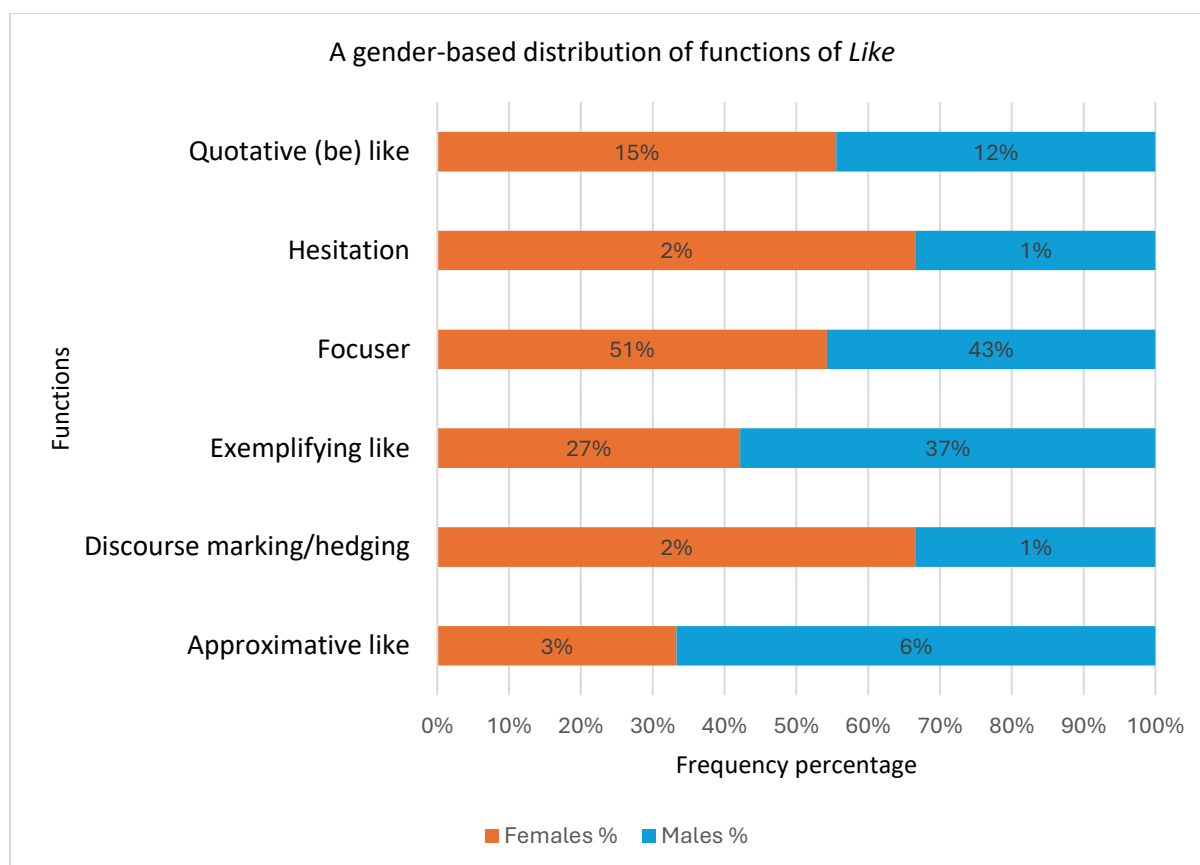


Figure 6.5: A gender-based distribution of functions of *like*

A proportionate interpretation of data as given in the Figure 6.5 above shows that *like* is used more by women than men in most of the functions. The exceptions are exemplifying *like* and approximate *like*. Men use *like* proportionately more than women for the purpose of exemplifying and approximation. Men use 37% of their total use of *like* for exemplifying while women only use 27%. Men use 6% of *like* for approximation which is twice more than the percentage that women use. Both women and men use the function of exemplifying *like* second most frequently. The current data seems to suggest that females and males both generally use *like* to gently draw the attention of the interlocutor to an idea or a claim and provide an example to validate the point.

The third most frequent function of *like* in the corpus for both genders is quotative (*be*) *like*. Women and men show 15% and 12% instances respectively from the total use of *like* for each gender. As found in previous research on other varieties such as British or American English, *like* is

used more often by females to introduce a quote in this data set as well. In general, although men use *like* less than women, the functions of use for both groups fall into a similar order proportionately. Both groups use *like* most prominently as a focuser, to introduce an example and to introduce a quotation.

6.5.4 A Gender-based Analysis of *Well* in the CSSLE

There are a total of 205 instances of *well* in the corpus, and females and males used 44% and 56% of the instances. Figure 6.6 outlines the percentage of frequency for each function of *well*.

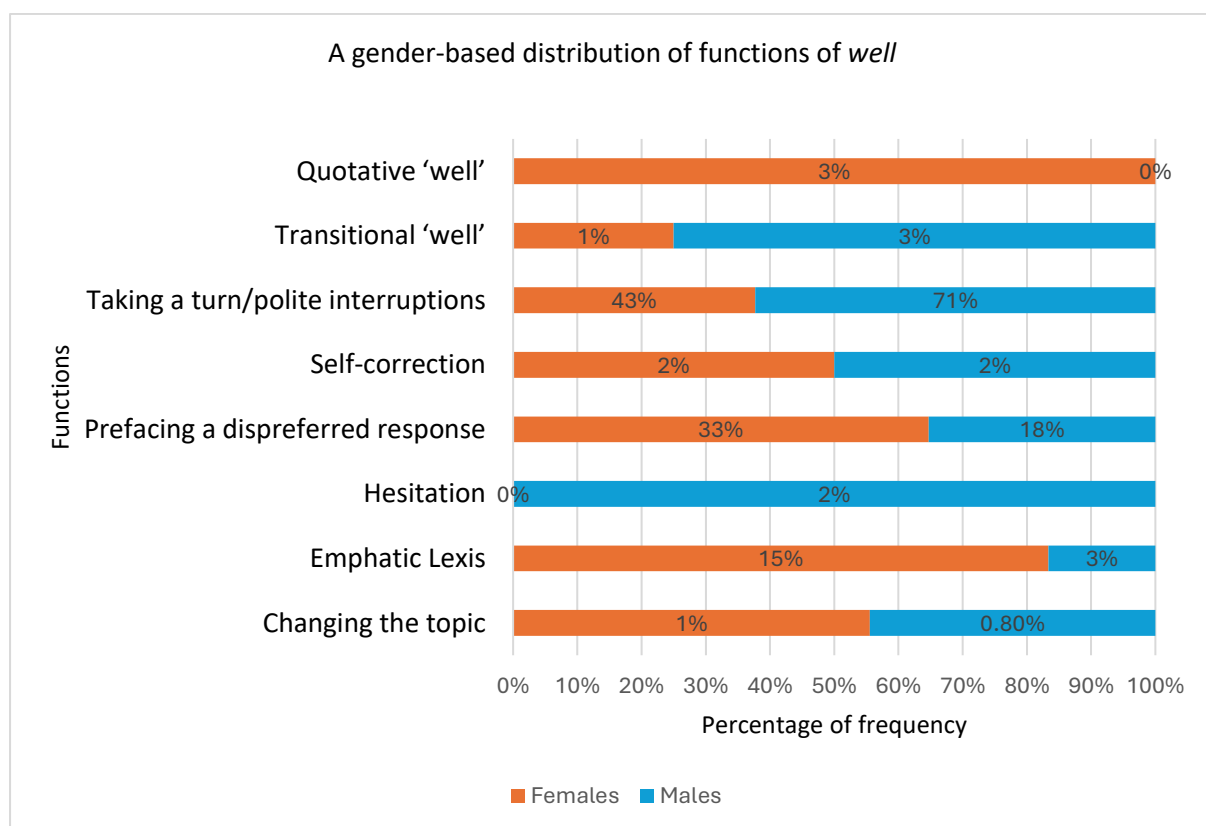


Figure 6.6: The functional distribution of *well*

Clearly men take turns and politely interrupt conversations using *well* more than women. Men use 71% of the total use of *well* for taking turns compared to 43% by women. This is consistent with previous studies considering interruption in general, which conclude that men interrupt more than women (West & Zimmerman, 1983). Tannen (1984) has argued that this trait is not a show of power but a way of collaborative speaking. Men also use *well* to change a topic more often than

women. Women use *well* 33% and men 18% to signal to the interlocutor that the coming utterance is a disagreement, an unexpected or surprising answer. When one prefaces a dispreferred utterance with *well*, it is considered polite. Therefore, this finding suggests that women use *well* more than men to indicate politeness. This is understandable as previous research generally concludes that women strive to maintain politeness in a conversation. Females use *well* for emphasis more than men. Women use 15% of instances of *well* for emphasis and men use only 3%. Holmes (2001) concludes that emphatic lexis is characteristic of female speech. While only women use *well* to introduce a quote, only men use *well* for hesitation, though this use only accounts for a very small number of instances.

6.5.5 A Gender-based Analysis of *Right* in the CSSLE

The fifth most frequent PM in the corpus is *right*. A total of 184 instances are recorded in the data. Figure Females use 62% (114 tokens) and males use 38% (70 tokens) of the instances respectively.

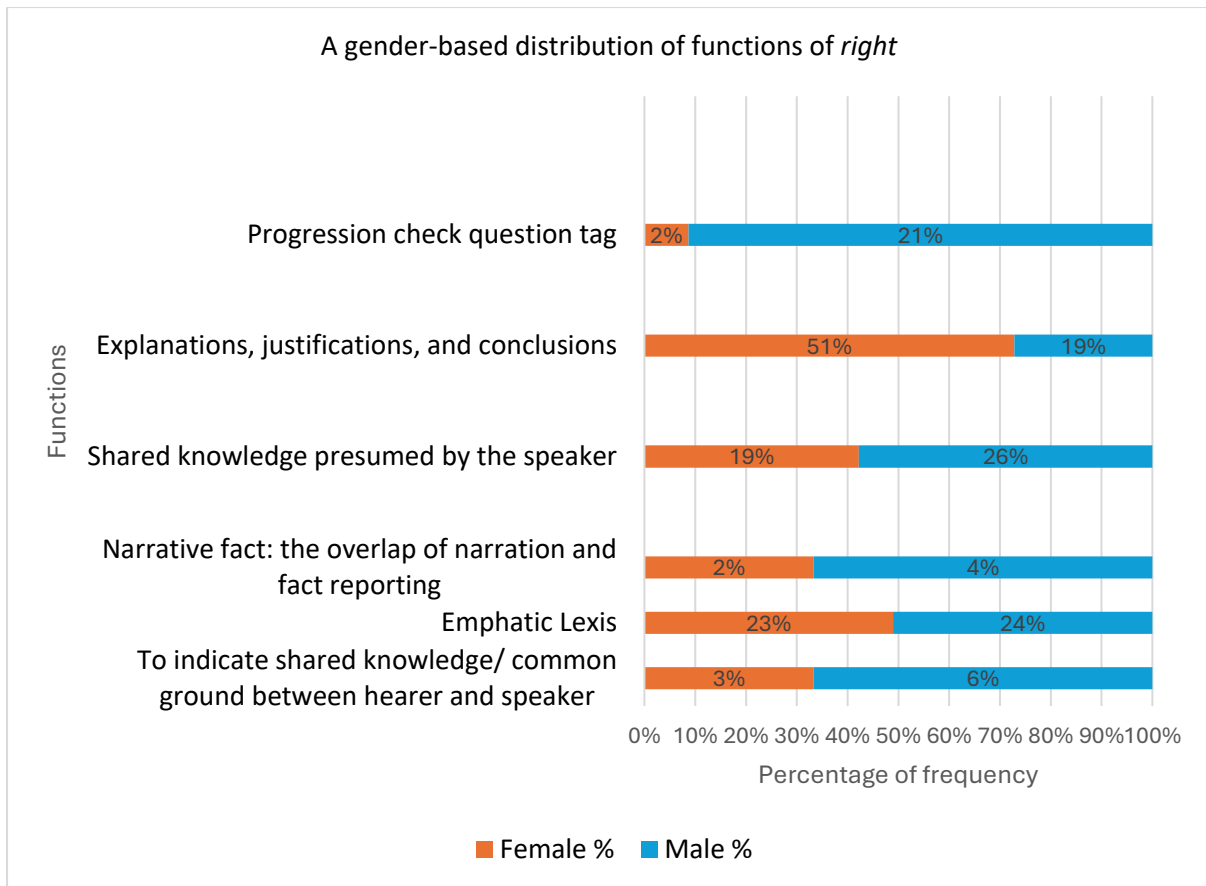


Figure 6.7: A gender-based frequency of *right* in the SSLE corpus

Right can be interpreted as a feature of men's language as men use tokens of *right* more than women for every function except one. The exception is the function of explanations, justifications and conclusions where females use *right* more frequently than males. Women use 51% and men use 19% of the total count for *right* for explanations, justifications and conclusions. Males use *right* mostly to refer to shared knowledge presumed by the speaker at 26% of all *right* usage for men. Men also use *right* as progression check question tag more than women (F=2% Vs M= 21%). This very high use compared to women may indicate that men like to ensure that the listener is closely following the narration. The last function listed in this graph is to indicate shared knowledge/common ground between hearer and speaker. Men use double the amount of *right* which is 6% to indicate shared knowledge or common ground than women who use only 3%. The indication of shared knowledge usually highlight that the speaker includes the interlocutor in the speech context which generally leads to solidarity. It gives the impression that the speaker and the

listener both have had some similar references or experience. This is a sign of inclusion. Therefore, it is suggested that men use the PM *right* as a strategy for fostering solidarity more than women.

6.5.6 The Nativized PMs in CSSLE

The nativized PM *no* is used by either gender only in a few instances. Despite this low frequency, it is worthwhile to analyse the use of *no* as it is an important characteristic of Sri Lankan English. As mentioned in Section 5.5.2, it is surprising that there is evidence of nativised PMs in formal and public speech. It is notable that there are 13 instances of the nativized *no* in CSSLE. This section analyses these instances in detail.

6.5.6.1 No. In this corpus there are two functions of *no*. First, it is used to emphasise the word preceding *no* or the idea given in the utterance.

Example 6.7

<07AcF> I mean we have a victory commemoration right but where is the commemoration of the loss and the death right I mean I'm not necessarily only and this doesn't have to be necessarily only the death of the of the North **no** I mean you know there were a huge large number of you know Sinhalese soldiers who died as well I mean we need to mourn that right where is the mourning

In example 6.7, *no* could be replaced by *you see* which is another PM used for emphasis. This is a quote by a female academic who is talking about the displaced people in the war in the North of Sri Lanka. She is referring to the victory commemoration to celebrate winning the war against the LTTE (known as terrorists or separatist group) in 2009. She says that as much as we are celebrating the victory, we should commemorate the loss of lives in the war and she emphasises that it is not the loss that occurred only in North that has to be commemorated. She uses the nativized PM *no* to emphasis “not only in the North”. In this instance, it is neither a question tag nor an assertion function, but acts as emphatic lexis.

The second function is that *no* is used at turn initial position marking a turn. There are 7 such instances in this corpus. The examples below illustrate this use. In the following instance, *no* is not a direct answer to a question raised by <01HM>, but is used to mark taking a turn. As explained in Section 5.5.2, this use is influenced by the Sinhala language where taking a turn with *no* is a norm.

Example 6.8

<01HM> we talk about er his work with fabric and you work a lot with beauty and and choreography er tell me Senaka <u=?> how many more Miss Sri Lanka's do we need to have to complete the circle

<03AM> **no** I think every year it's lovely to have a fresh set of girls and it's not who wins or anything like that it's seeing these young girls evolve and become somebody else become a celebrity or a actress or something like that

10 of the 13 instances of nativized PM *no* are used by males while only 3 are used by females. This is similar to the finding in Singaporean English where men use more non-English discourse particles (nativized PMs) than females (Leimgruber et al., 2020). Generally, in language change norms men are known to use less prestigious forms than females (Labov, 1966; Milroy & Gordon, 2003). The use of *no* in public speech in SSLE could indicate this phenomenon though the numbers here are very small.

6.5.6.2 Aah. There are two instances of *aah* in the corpus and both are used by men. In both instances, *aah* could be replaced by *you see*. *Aah* is used as emphatic lexis. This is an interjection in both Sinhala and Tamil languages to emphasise the seriousness of a situation. It could be assumed that the vernacular influence is seen in SSLE.

Example 6.13

<01HM> half a sugar I hope

<09AM> yeah taste it Ceylon tea **aah** okay how is it Kumar

Example 6.14

<14HM> mm

<12AcM> who go on and do whatever they want particularly with regard to violating

people's rights **aah** so we don't really have a government in the sense of a cohesive coherent

government with a cohesive coherent plan to govern

The limited nativized PMs are mostly used by men. In a total of 15 instance of *no* and *aah*, only three are produced by women. However, the presence of the nativized PMs in this data suggests promising avenues for further research. A different type of data such as informal conversations between more intimate speakers could reveal the use of these nativised forms in higher numbers as these PMs are generally indicators of familiarity.

6.6 The Sub-corpus of Academics

The corpus of academics contains SSLE speech samples exclusively from speakers who are employed as academics in state and private universities in Sri Lanka or who work in broader fields of academia such as research institutes or data analysis institutes in Sri Lanka. The corpus includes 12 male academics and 12 female academics. The two corpora are almost identical in terms of word count: the male academic sub-corpus has 33915 words and the female one has 33626 words. The academics use the least number of PMs overall compared to the other groups although it is a marginal difference between groups. The percentages of overall PM production for the three occupation groups are 33% by artists, 30% by academics and 36% by entrepreneurs. The striking point is the gender-based difference of PM production between the women and men in the corpus of academics. The difference between the PMs produced by each gender is greatest among academics in comparison to the other two professions. The use of PMs translates as 70% versus 30% for women and men. This data set is from two groups who belong to relatively the same sociolinguistic subculture with similar social status and education, so the imbalance found in the use of the PMs is unexpected and striking.

This difference is prominent in some of the PMs but not in all PMs as illustrated in figure 6.7. The possible reasons for this difference of use are given in the discussion under each PM. Women have produced more tokens than men of the PMs *you know*, *I mean*, *like*, *right*, *kind of* and *sort of*. For example, women lead with a very great margin in most of these PMs. Women and men have produced 71% and 29% *you know* tokens respectively. Women use 64% of *I mean* and men use 36%. Women use 92% of all *like* tokens and men only use 8%. A similar great gap is seen in the production of *right* as well: women use 80% of *right* and men only use 20%. *Kind of* and *sort of* are used with 71% versus 29% and 72% versus 28% by women and men respectively. The reasons for the most prominently contrasting PMs well as the other PMs are investigated in the next sections.

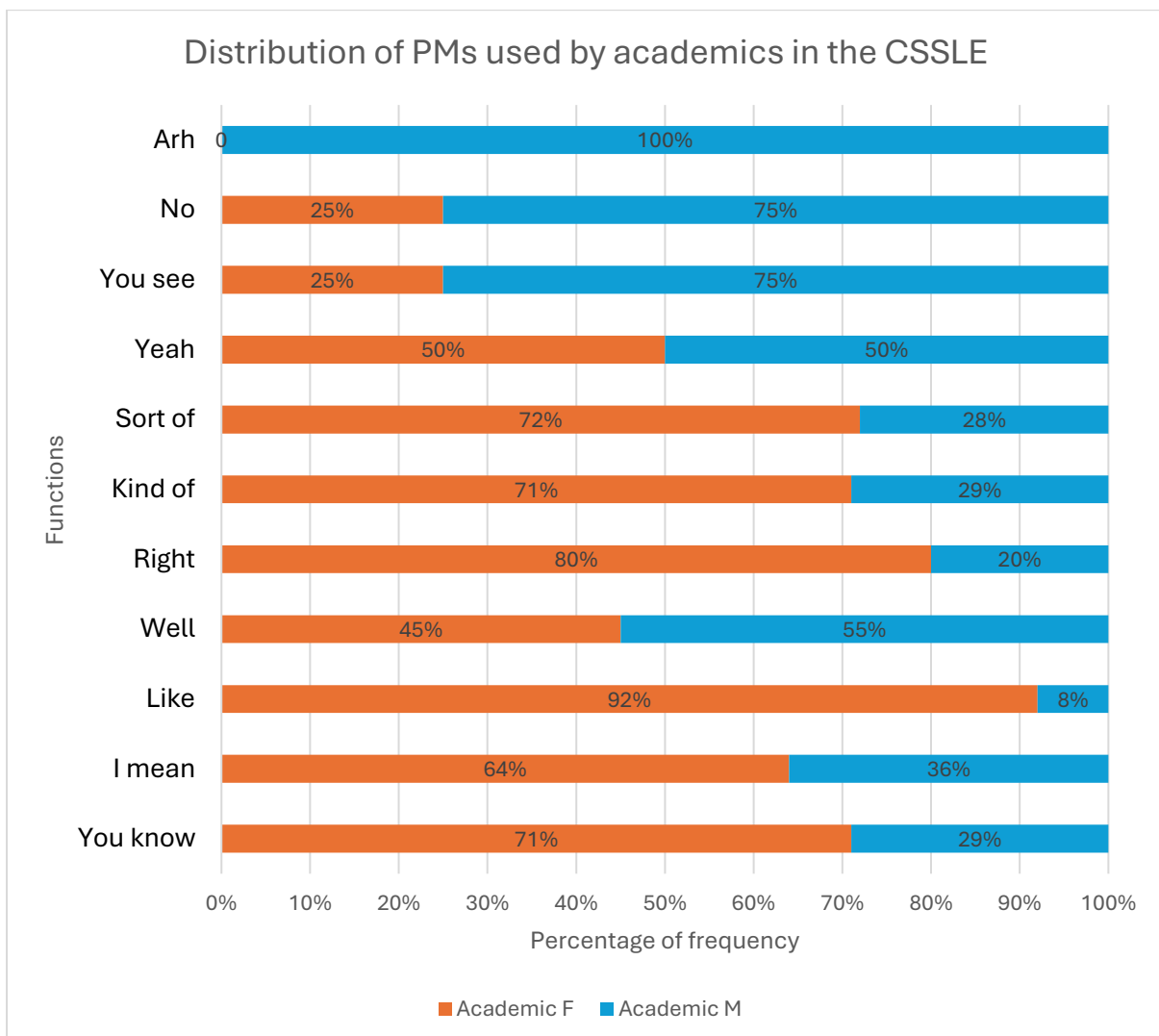


Figure 6.8: Distribution of PMs among academics

6.6.1 *You know in the CSSLEAc corpus*

There is a remarkable difference in the use of *you know* in this group compared to the other two occupation groups: females use *you know* almost twice as often as men. As a percentage it is 71% for women and 29% for men. This is a finding that is consistent with some previous research as well (Macaulay, 2002; Beeching, 2016). It is important to refer to Östman (1981)'s study although it was done four decades ago. Östman observed a result similar to the current study. Östman (1981) studied 5 dinner table conversations in AmE recorded in Dallas, Texas and Berkeley between 1976-77 and 1979-80 respectively. He noticed a difference in one of his data sets compared to the others in his study of the use of *you know* by men and women, where women used less *you know* than men; that data set exclusively included university teachers. Östman argues that the academic world doesn't show "normal" manifestations of women's language" (p. 72). He believes that the female academics consciously avoided overusing *you know*. Four decades later, female academics in another part of the world speaking a different variety of English still do not manifest 'normal' women's language, but in this data they do not mirror the speech patterns of the female academics 40 years ago either. The current study reveals a contrasting result where female academics use more *you know* than male academics. In addition to the general quantitative distribution of *you know*, a functional analysis of *you know* will give us a more precise idea of how it is being used by this group.

You know is one of the PMs that is considered by some linguists as a hedging or corrective device and marks the inferior or powerless speech of women (Erman, 1992; Johnson, 1993). These terms are associated with the lack of authority or status and also with expressing doubt and imprecision. Using more *you knows* than men is a characteristic of female speech in previous research as well (Macaulay, 2002). However, Beeching (2016), Holmes (1986, 1990), Coates (1996) and Fishman (1990) point out that *you know* and comparable expressions can have different functions and do not necessarily show deficit or powerless language for women. It should also be noted that hedging is an essential part of academic writing because it allows the writer to comment on phenomena in a way that do not offer predictions or solutions which are absolute in all

circumstances (Gherdan, 2019; Markkanen & Schröder, 2010). Therefore, we can predict that the academics in the data are prone to use PMs for softening or downtoning ideas. In this section, the aim of my research is to find out if the high use of *you know* by female academics might be interpreted as showing imprecision, weakness, repair, lack of authority, or might instead show different functions that have no connection to powerless language.

In this data with regard to academics, there is a marked difference in the distribution of *you know* among the two genders. Females have a broader spectrum of use for *you know* than males. Only females use *you know* for repair and word search and appeal to the interlocutor to fill in the gap. Academics in general do not use *you know* as a filler or for hesitation and appeal to common knowledge. Although females have produced more PMs, as a percentage of the total PMs for each gender, a different pattern emerges. For instance, although there were 24 instances of attention getting/ launching a new piece of information function for females and 20 instances for the males, as a percentage of the total production of PMs for each gender, males have technically used *you know* more for this function than females.

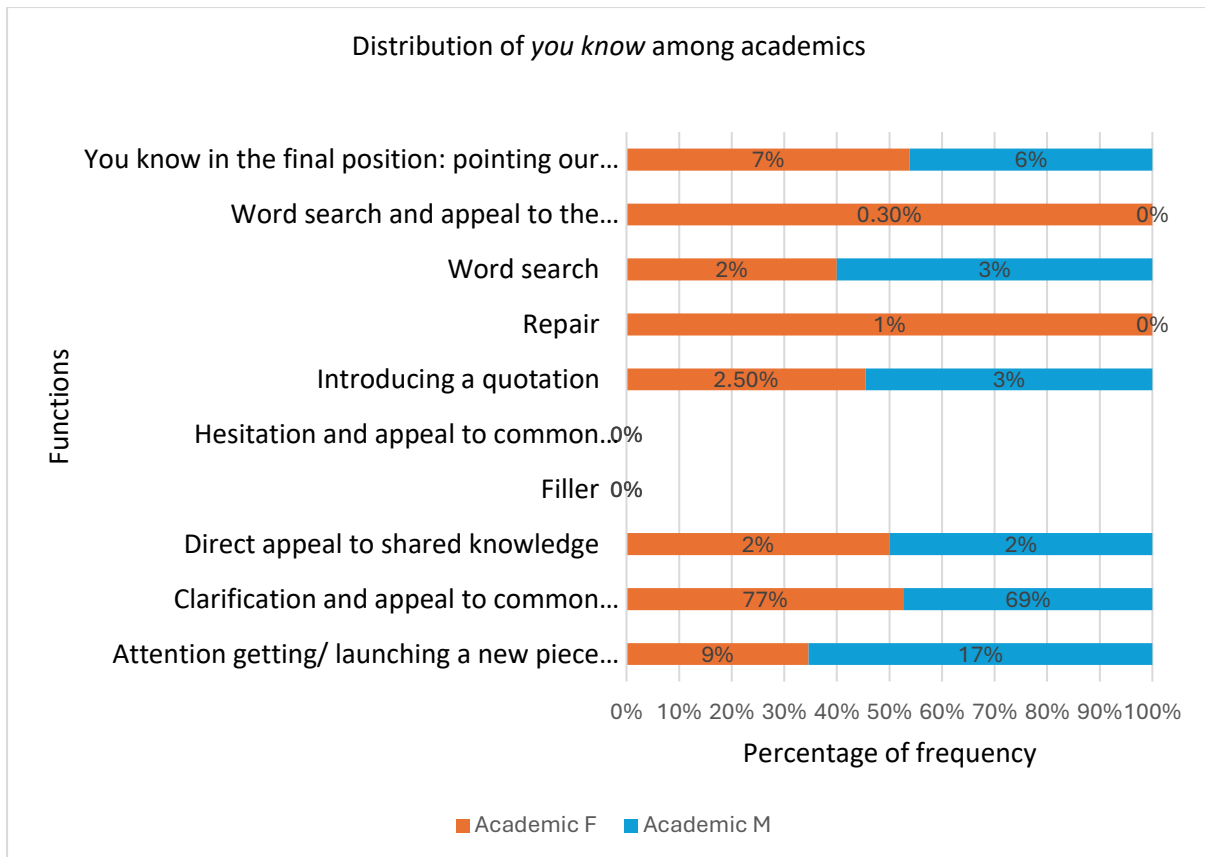


Figure 6.9: The distribution of functions of *you know* according to gender in the SSLEAc sub-corpus

As percentages of the total use of *you know* for each gender, there is no great difference between how both the genders behave for most of the functions. For instance, both groups use *you know* mostly for the function of clarification and appeal to common knowledge. As a percentage of their individual production, they account for approximately 77% and 69%. By clarifying and appealing to the interlocutor for common knowledge, the speaker establishes a sense of cooperation from the interlocutor. Appealing to common knowledge is also a strategy to build solidarity with your interlocutor. Given below is an example of establishing cooperation. The interviews with the academics were mainly on academic subjects. In the example below, the academic is an expert on museums. *You know* prefaces an explanation about museums and the speaker makes an assumption that the interlocutor perhaps already knows the etymology of museum.

Example 6. 16

<12AcF> **you know** the word museum comes from the word muse which is about inspiration so for me museums are places that we go to to to get inspired to think about our future

This assumption of shared knowledge is a pretence on the part of the speaker, perhaps to save the positive face of the interlocuter (Östman, 1981). Based on her study of BrE, Beeching (2016) makes the point that women may use *you know* to express solidarity and maintain positive face. Such direct positive assumptions from the part of the speaker makes the interlocuter feel included in the realm of knowledge of the academic. This inclusion also reduces the status difference between the interviewer and the interviewee. The reduction helps the academics to politely establish cooperation with the interlocuter. This use is well defined by Östman (1981): “The speaker strives towards getting the addressee to cooperate and/or accept the propositional content of his utterance as mutual background knowledge.” It could be assumed that since the nature of their professional work requires academics to share knowledge, they use *you know* to do it in a non-threatening manner. Mauranen (2006) investigates first language (L1) and second language (L2) conversations in the early stages of English as Lingua Franca in Academic writing (ELFA) corpus. She looks at L1-L1 discourse and L2-L1 discourse to find how speakers mitigate genuine misunderstandings. In one of her examples, the L1 user speaker uses *you know* to facilitate understanding. It could be interpreted that *you know* can be used by experts to check understanding when explaining some new information or concepts to a lay person. The conversations between the interviewer and academics in the CSSLE are generally explanations about concepts or issues. Therefore, we can interpret that *you know* is not a mark of weak or powerless language but a strategy to mitigate misunderstanding.

Holmes (1995) reports that “Women tend to use questions, and phrases such as *you know* to encourage others to talk” (Holmes, 1995, p. 2). In the current data set, the females know that they are the interviewee, and they do not need to encourage the interviewer to talk. Instead, it can be assumed that the female academics are ensuring that the interviewer’s cooperation is gained. One could interpret this effort to gain cooperation as an attempt to be polite as well. According to Brown and Levinson:

“politeness, like formal diplomatic protocol (for which it must surely be the model), presupposes that potential for aggression as it seeks to disarm it, and makes possible communication between potentially aggressive parties.” (Brown and Levinson 1987, p. 1)

The idea of politeness in this sense implies that an effort is made to minimise any confrontation in discourse. Therefore, we can say that female academic SSLE speakers try to be polite by using *you know* for clarifications and appeal to common knowledge. Through the appeal, they gain cooperation, and gaining cooperation is a mark of a successful communicator. *You know* also marks politeness. Example 6.17 illustrates

Example 6.17

<06AcF> right erm I mean in some ways **you know** you can't really fault people because this is for most erm students for young kids at school you know the Shakespeare drama competition is maybe the only theater experience that they have and as a result obviously they want to make it very memorable

The first *you know* in example 6.18 is used for clarification and appeal to common knowledge. This is a conversation between an interviewer and a female academic. Sri Lanka has a popular and long-standing Shakespeare drama competition for school children. The academic featured in this conversation is a member of the panel of judges. The interviewer asks her what she would say a judge should really be evaluating at the Shakespeare drama competition. In the response she says that many directors rely heavily on theatrical effects, and she does not endorse it. Yet, in example 6.18 she softens her approach by using *you know* and says that it is no fault of anyone given the novice experience students have about the competition. This is also a way of showing politeness. In this sense, *you know* marks politeness in certain contexts.

The function that both groups use third in frequency is *you know* in final position: pointing out a self-evident truth/ impositional. This is a use that directs speech towards emphasis. Females use 19 instances and males use 6 instances, 5% and 7% of their total use of *you know*. Therefore, we can conclude that men and women behave similarly in emphasising ideas by using *you know*. None

of the academics use *you know* for hesitation or as a filler. It shows that *you know* does not indicate insecurities or hesitancy. All in all, it can be concluded that SSLE speaking academics use *you know* mostly as an interactional and facilitative tool than to show hesitation or insecurity.

6.6.2 Analysis of the other PMs in CSSLEAc

The nature of the academics' use of other PMs also should be understood separately as the results stand in contrast to the other two professional groups. For instance, female academics use more instances of *right* than men as evident in the raw data. Females use 88 instances compared to 22, which is quarter of the instances as females. A close examination shows that women's use of *right* is dominated by one individual, who uses 63 instances out of the 88. However, even if the highest two users of *right* are excluded from both the corpora, the remaining females still use 50% more instances of *right* than males. Figure 6.9 provides the percentages of the frequencies for both genders so that a more nuanced quantitative analysis could be made.

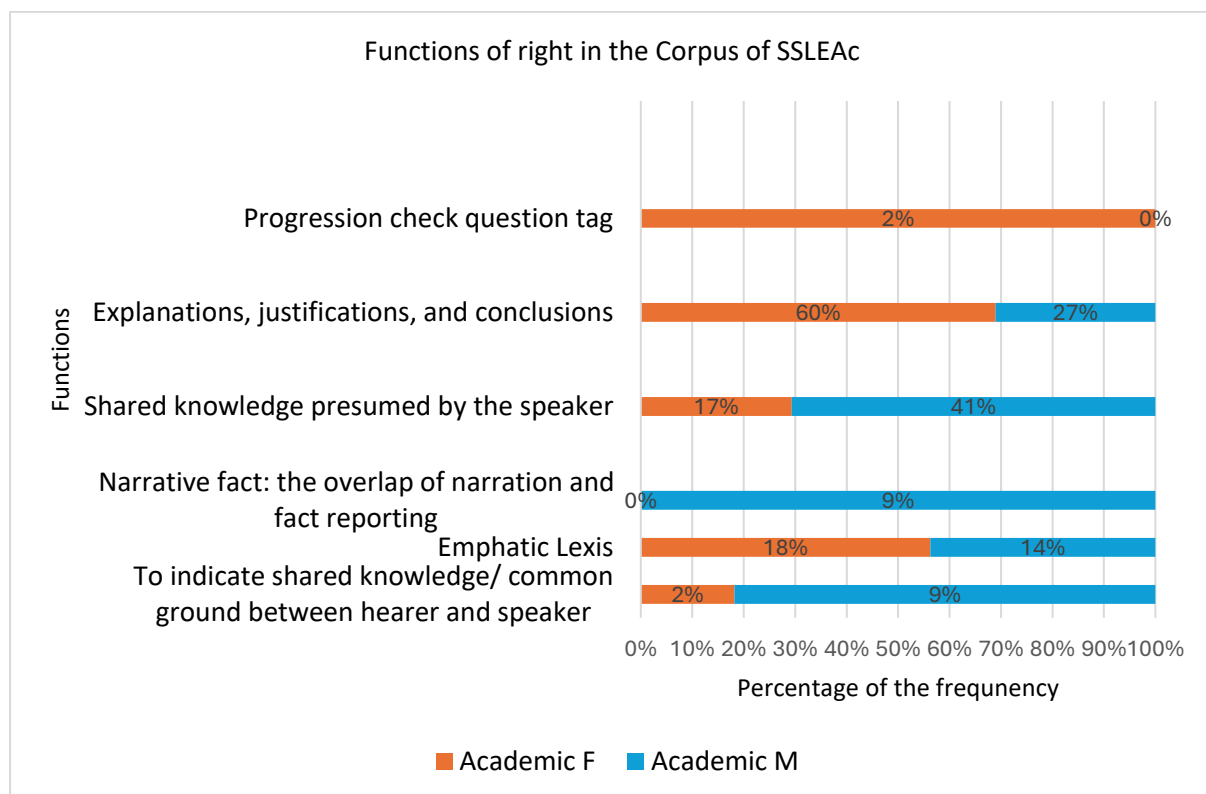


Figure 6.10: Functions of *right* in SSLEAc sub-corpus

In this data, academic females lead males by far in the function of explanations, justifications, and conclusions only. By sharing explanations, justifications, and conclusions with *right* frequently, female academics seem to try to engage the interlocutor more actively in the conversation. Similar to *you know*, it seems that female academics are much more concerned than males to ensure that the listener or the interlocutor is following the ideas expressed by the academic closely.

The remaining results show that the female academics use more hedging devices than male academics. For instance, female academics use both *I mean* and *kind of* more than male academics. Hedging devices are generally associated with women and are considered a mark of 'powerlessness' (Lakoff, 1975). Another finding that is aligned with traditional female PM use is *like*. Female academics use more *like* than males.

The five occurrences of the nativized *no* and *aah* are also important as generally nativized SLE terms are avoided and considered part of informal language. Four instances of the five have been used by males. The academics consciously or unconsciously using these terms imply how deeply and naturally SSLE has become part of the lives of its speakers.

It could be concluded that in comparison to the other two professional groups, female academics make a considerable effort to engage the interlocutor in the conversation and they achieve this by using *you know* and *right* frequently. This is also a show of politeness. Also, the female academics show two other characteristics that are generally found in feminine speech; i.e. hedging and the general use of *like*. All these characteristics that are evident in the female academics' use of PMs indicate that they are behaving in typically feminine ways with regard to language. Making effort to conjoin the interlocutor and softening propositions by hedging are typical features of feminine speech.

6.7 A Summary of Findings

The aim of this chapter was to find possible gender differences in the use of PMs in SSLE. With regard to CSSLE, it can be concluded that there are some gender-based differences in the PM use of men and women. The quantitative analysis revealed that the most preferred PM for both

genders is *you know*. *You know* is mostly used for clarification and appeal to common knowledge. *You know* tends to be used more by men in SSLE than women and this is a contrasting result to the findings based on some of the research on inner circle varieties of English. The results are based on corpus data from three professional groups who speak SSLE and out of them, the academics show an opposite behaviour in terms of gender impact on PM use than the others in the corpus. For instance, female academics use more PMs than males in the academics' sub corpus. Female academics use *you know* in a higher percentage of use than males to clarify points in a conversation in a non-threatening manner. This use indicates that female academics are more polite than male academics. This female use of PMs closely replicates traditional women's use of language as noted in previous research. *I mean* is the second most frequently used PM in the corpus and there is no great difference in its use according to gender. Both genders behave similarly in all functions of *I mean*. The third most used PM is *like* and it is used more by females than males. Men and women use *like* differently in certain functions, yet similar to other varieties such as British or American English, females use *like* more often to introduce a quote than men. *Well* is used quantitatively more by males than females. It is used as a polite turn taker or polite interruption by males more than females which implies a negatively polite domination in conversations by males. Females on the other hand use *well* to politely point out a dispreferred response. In conclusion, *well* works as a negative politeness marker for both genders.

Females use the PMs *kind of* and *sort of*, which act as hedging devices, more than males. This is a result that is expected in traditional views of language behaviour. It can be assumed that these hedging devices are used mainly to soften a remark or to avoid being seen as overconfident. Nativised PMs are used more by males than females which is again a traditional finding in gendered language behaviour. The fact that nativised PMs are present even marginally in semi-formal to formal discussions also indicate that speakers have accepted the nativised norms of SSLE confidently. A nativised function is also recognised in the use of the nativised PM *no* for turn taking. The nativised

PM *aah* is used as emphatic lexis in place of *you see* in SSLE. The analysis shows that both *no* and *aah* work as PMs due to similar PMs in the L1 of the speakers.

Overall, SSLE speakers show marginally different patterns of PM use to traditional research biases with regard to BrE and AmE. For instance, as SSLE females speakers use more *you know* than men overall, it could be stated that SSLE speakers do not follow the same gender norms as other varieties of English. This result is interesting given the arguably different cultural contexts.

Chapter 7: Age and Pragmatic Marker Variation

7.1 Introduction

A synchronic analysis of one language aspect among speakers in different age groups informs us about how that particular language feature is used by the individual age group as well as how it has progressed over time. Baily (2002, p. 312) states that the “synchronic approach” to the study of language change and the study of change in progress, forms one of the cornerstones of research in language variation. Research has shown a clear variation in the use of pragmatic markers with respect to age stratification (Andersen, 2001; Reichelt, 2021). In fact, compared to gender, age stratification has an even higher correlation with language variation as age stratification can occur without being accompanied by systematic gender differences (Labov 2001, p. 449). Certain PMs are associated with more frequent use by certain age groups (D’Arcy, 2017). Palacios Martinez (2015)’s study of Bergen Corpus of London Teenager Language (COLT) and Linguistic Innovators Corpus (LIC) finds that *innit* is a classic linguistic feature of teenagers from middle and working classes speaking London English. In contrast, the counterpart adult corpus showed a marginal use of *innit* tokens in working class adults. Although the impact of age on language behaviour is well established in research pertaining to the inner circle varieties of English, there is little research with regard to outer circle varieties (Suárez-Gómez & Seoane, 2021).

One of the objectives of this study is to observe how different age groups use PMs and notice how PMs evolve with each generation as well. As seen in Chapter 6, gender has an effect on the use of PMs. Age adds an additional dimension to this finding. One of the motivations for the analysis presented in this chapter is to find out if such age-specific features are evident in the data. If there are age-based features, we are able to understand generation-based characteristics of PM use for Standard Sri Lankan English. Such information is useful to understand the concept of appropriacy in language use. One can discern the current language use as opposed to language use from an older generation by observing the frequency patterns of PMs. Another objective is to describe how age-related differences can indicate change in progress. The chapter broadly answers the research

question whether there is any observable variation in PM use according to different age groups, and how this can be explained. The patterns also will reveal how PMs use has evolved over the generations. The study investigates which PMs have remained stable in terms of their frequency and function over time.

The chapter opens with a brief literature-based discussion of age as a sociolinguistic variable in Section 7.2. Then, a broad literature review of research on the relationship between age and PMs are presented in Section 7.3. This is followed by the theoretical framework for analysing age and language in the current study in Section 7.4. Section 7.5 is devoted to the discussion of the age-based results of the present study. This is followed by a data analysis on how age reacts with each PM in Section 7.6 and the chapter will conclude with a summary of findings.

7.2 Age as a Sociolinguistic Variable

Research on age as a linguistic variable began in the 1970s and 1980s (Eckert, 1984). It is assumed that each generation reflects the language as it existed when that generation acquired the language (Murphy, 2010). Therefore, age is a reliable sociolinguistic variable to measure, observe or analyse language variation. When you look at aging as a movement through time, age is also a person's place at a given time in the social order, the system of social structures (Eckert, 2017). Age is a personal experience as well as a group experience with people in the same age when you share a similar history or events of time. Previous research assigns speakers ages from many perspectives such as physiological, psychological, chronological, functional and social. Each perspective has its own strengths and weaknesses. Much of language related research uses the chronological age, the number of years a person has lived, as the age marker. However, Eckert (2017) cautions the researcher with her comment that chronological age can only provide "an approximate measure of the speaker's age-related place in society" (p.155). For example, the chronological age may not always mirror social and biological development of a speaker. She urges the researchers to look at life experiences of the participants rather than the chronological age. This is possible in anthropological research where the researcher has reliable evidence of such details. However, she also says that

since individual differences in one age group are small in relation to life span, chronological age is an acceptable measure. The current research too uses the chronological age of the SSLE speakers to determine age-related variation with regard to PM use. As reiterated in the Section 4.5.1., the speakers are selected from online videos with limited meta data. Therefore, the only measurement of age the current study could rely on was chronological age.

7.3 Research on the Impact of Age on PMs

Previous research shows that there is a clear association between age and the PM use. Adolescents are mostly highlighted as the innovators of PM use (Stenström, 2014; Tagliamonte, 2005; Andersen, 2001, Miller and Weinert, 1995). Östman (1981) mentions *you know* as associated with informal speech and slang, and because of its potential to be socially stigmatised, it is likely to be used more frequently during adolescence by speakers in American English. PMs perform certain functions in a conversation. Some of these functions, such as focusing or approximation, are performed by speakers after a certain age or after maturity. For example, in a study on Scottish English Miller and Weinert (1995) report that children in the age groups of 8 and 10 produced far fewer instances of *like* than those in early 20s. Thus, Miller and Weinert (1995, 1998) propose that the PM *like* is acquired after the age of 10. Some of the PMs are acquired after a certain age as the functions associated with the PM are acquired later. Similarly, much research attests that *like* is a PM that has a high frequency of use among adolescents and pre-adolescents (Dailey-O’Cain, 2000; Müller, 2005; Levey, 2006). Andersen (2001) compares the COLT corpus with BNC, a teenage corpus and an adult corpus of British English respectively and says that *like* is most common among adolescents but that it has been adopted to some extent by adult speakers aged 41 and lower. It is used to a very little extent by people over 45. Andersen (ibid.) mentions that in his study of COLT data that it is primarily a feature of adolescent girls. In his analysis of *like* with regard to ethnicity, he discovers that it is a feature of white adolescent speech. He quotes Ferrara & Bell (1995) to support his claim. Ferrara and & Bell (ibid.) have found that in the US, white speakers help to spread *like* more than black and Hispanic speakers. Andersen further analyses his data according to the location of

speakers and finds that it is a suburban feature. He narrows down the use of *like* to the prototypical user who is a white 17-year-old girl from the highest social class who attends the boarding school in Hertfordshire (2001, p. 294). With regard to Irish English, Schweinberger (2012) claims that as speakers exceed their mid-thirties, *like* use decreases significantly and all variations of *like* are confined to younger speakers. Denis and Tagliamonte (2016) reveal that similar to *like*, *right* is also associated with a younger cohort of speakers. *Right* as an utterance-final tag in Canadian English is most frequently used among speakers born after 1970 (under 30 years old at the time of the study).

D'Arcy (2017), who has done extensive research on *like* in American English, presents findings that reach beyond what is found in British and Irish English. She claims that *like* is not confined to the use of adolescents alone and that it is a feature shared across speech communities. There are individuals who overuse it and others who rarely use it. Rather than age, she mentions that social factors such as whiteness and suburbanity play a role. According to her study of 800 years of the development of *like* in American English, D'Arcy (ibid.) further states that *like* began to expedite its use among speakers born in the 1970s.

While most research revolves round the frequency of PM use and its correlation with age, some research reveal that the functions of PMs also correspond with age. For example, Erman (2001) finds a very clear difference between use of *you know* in adults and adolescents. Adults use *you know* as a text monitor. Moreover, adults use *you know* for thematic organisation of the text and as a cohesive device to bracket utterances. In contrast, adolescents use it more to communicate, in the sense of maintaining an open channel between the speaker and the hearer. The speaker wants to ensure that the hearer can properly understand her/him. Therefore, *you know* is a social and metalinguistic monitor. Similarly, Andersen (2001)'s comparison of BNC and COLT to investigate *like* shows that there is a narrower range of functions in the adult corpus than the teenage corpus. For example, there were no instances of *like* as a quotative or any metalinguistic functions. D'Arcy (2017) notes that although this use of *like* as a quotative might be common to the whole of the speech community, it is difficult to define a specific context of use.

There are similar patterns emerging in the research based on world Englishes. However, the findings in outer circle varieties and expanding circle varieties do not always mirror the patterns observed in inner circle varieties. In terms of the distribution of nativized PMs, a study based on colloquial Singaporean English reveals that some nativized PMs have increased over time while some others do not show any noticeable change. Li, Lorenz & Siemund (2023) investigated oral interviews conducted between 1979-2009 with people born between 1899 and 1983. They divided speakers producing nativised PMs into two age groups. The older PM group included *ah* and *lah*. The younger PMs included *leh*, *lor*, and *meh*. Even speakers born after 1920 use the PMs *ah* and *la* and use of these PMs have increased over the years, whereas *leh*, *lor* and *meh* are used rather less frequently among all speakers. This is surprising as generally nativized features are evident in later generations as earlier generations try to closely replicate the colonial tongue. The study concludes that individual preferences for the PMs seem to rank higher than all other social factors. Lange (2009) shows a pattern where the nativized PMs *yaar* and *na* in Indian English are heavily used by 18-25 age group whereas a sharp decrease in overall use is evident for the age group from 42 onwards. The same is observed for the nativized *no* as well. The youngest age group favours the particle *na* but the older generation shows a different frequency of use, although it is found in all types of their conversations. In a closer examination of these features in Indian English in ICE-India, the same corpus she examined in 2009, Lange in 2012 confirms that the youngest age group (18-25) prefers the invariant tags *isn't it/ is it* compared to the four older groups (26-33/ 34-41/ 42-49/ 50+). This group uses this invariant tag more frequently than the nativized *no/na* tags. Although there are variational studies on South Asian Englishes, there aren't many on age variation. In the literature I surveyed on pragmatic marker variation according to age, there are no studies on Pakistani English or any other South Asian English, with the exception of Indian English.

7.4 Theoretical Framework for Analysing Age and Language in the Current Study

My study is influenced by the apparent-time hypothesis. Apparent-time hypothesis proposes that age-based variation indicates linguistic change in progress. It rests on the idea that speakers do

not change their way of speaking significantly with time and life experiences. Therefore, we can compare the speech of older and younger speakers and make assumptions about language forms at a single time point (Tagliamonte, 2012). It assumes that if language change is taking place the older generation will use earlier forms of language, and the younger generations will use later forms of language. The patterns of frequency in the use of a linguistic feature according to speaker age can manifest change in progress (Sankoff, 2006). However, this hypothesis clashes with the idea of age-grading. In age-grading, speakers change the way they speak with age and therefore, the language variation between generations is a result of aging rather than change in the language. This conflict has been recognised by the sociolinguists who accept that some features reflect ongoing diachronic change while others relate to the aging of individual speakers (Boberg, 2004). According to Bailey (1991; 2002) and others (Boberg, 2004; Eckert, 2017), apparent time data is not the best in that it is only a substitute for real time evidence and cannot be accepted to represent diachronic linguistic variations. They propose that in order to get an accurate view of language change both apparent time and real time data should be considered. This suggestion too is not without issues. It is difficult to find studies that have similar settings and similar speech communities done both in real time and apparent time. In addition, matching the methodology of the two kinds of study is a challenge. The current study does not have a comparable study which uses real time data from speakers of the same speech community and therefore it can only observe apparent time data to draw conclusions.

As mentioned before, it should also be noted that age was not a criterion for the selection of participants for this study in its early stage. The only criterion was that the speakers use the standard variety of Sri Lankan English (SSLE) and that there be equal number of females and males. An initial survey of the data suggested that age appears to influence the PMs in SSLE, and thus, an age-based analysis was conducted. Data that match the dialect of English, gender, occupation and age was also unavailable at the point of data collection. Therefore, there are no equal number of participants for each age group in the analysis. This study follows an etic approach to grouping the speakers. An etic approach groups the speakers arbitrarily according to equal spans of time such as decades as

opposed to an emic approach where speakers are grouped based on shared experience of time such as childhood, adolescence, and young adulthood (Trudgill, 1974; Wolfram, 1969; Eckert, 2017). Some studies (Lange, 2012) group participants into 5-year age ranges as in 18-25, 26-33 and so on, and others group speakers into decades. I divided the speakers in my study into three broad social generations following Strauss & Howe (1991) and Colburn (2017)'s generation classifications. Strauss and Howe (1991) hypothesised that individual generations share common characteristics and attributes. Their generational theory hypothesised that each generation has original characteristics with very little opposing or diversity between other generations. There are 5 living social generations within the 20th and 21st centuries. As stated in Colburn (2017), they are Silent Generation (1928-1945), Baby Boomers (1946- 1964), Generation X (1965 – 1980), Generation Y (1981-1997), Generation Z (1998-2010) and Generation Alpha (2011-2025). I employed this classification as the 72 speakers in the study who are born between 1943 and 1997 fitted into three of the generations: Baby Boomers (1946-1964), Generation X (1965-1980) and Generation Y (1981-1997). Having three groups rather than many more groups ensured a reasonable number of speakers in each group so the data can be compared.

The age of the speakers was located through Google searches. As most of the participants were public figures this demographic detail was easily available for the majority of speakers. The exact age of 58 of the speakers were available in online data. For those speakers whose ages were not publicised and easily retrievable, I also looked at LinkedIn profiles to calculate age by looking at educational milestones. Some of them mentioned their primary and secondary schooling, and since Sri Lankans start formal primary at the age of 6, it was easy to find the precise birth year. There were 14 participants for whom the birth year was not precisely mentioned. I guessed their age by what is mentioned as life experiences in their interviews and by their various profiles online. The age of each individual speaker is provided in Appendix H.

As stated above, the age groups in my data fits into three of 6 living generations, excluding Silent Generation, Generation Z and Generation Alpha. There were 2 speakers who were in the

borderline for the Silent Generation, born in 1943 and 1945. I added them to the Baby Boomers as it is meaningless to analyse data for just two speakers from the Silent Generation. Table 7.1 gives the distribution of the participants of this study according to the above-mentioned three social generations. It also includes information on the number of words in the corpus for each generation.

Social Generation	Date of birth	Gender (female)	Gender (male)	Total	Number of words
Baby Boomers	1946- 1964	11	17	28	79,293
Generation X	1965-1980	13	13	26	73,024
Generation Y	1981-1997	12	06	18	50,240

Table 7.1: Social generations and the distribution of the speakers in the SSLE corpus

7.5 Age-based Data Presentation and General Results

7.5.1 General Results

In this section, the general results for the PMs divided into the three social generations are presented. First, the number of words needed to be normalised as there were not equal numbers of speakers or equal numbers of words in each generation. Therefore, the tokens have been normalised per 100,000 words. The number of tokens was divided by the number of words, and it was then multiplied by 100,000. The raw data is normalised to get the figures per 100,000 words for all age categories as then, the groups can be compared.

Social Generation	Year of birth	Speakers	N words	Tokens	Percentages of overall PM use	Normalised tokens per 100,000 words
Baby Boomers	1946-1964	28	79,293	966	28%	1218
Generation X	1965-1980	26	73,024	1177	38%	1609
Generation Y	1981-1997	18	50,240	723	34%	1439

Table 7.2: The normalised tokens per 100,000 words

Table 7.3 gives the raw data, percentages and the normalised number of tokens for the overall use of all PMs by all three social generations. There is a total of 2866 tokens of PM. Baby Boomers (BB), Generation X (Gen X), Generation Y (Gen Y) use 966, 1177 and 723 tokens respectively in their relevant corpora. These tokens are have been normalised to 1218 (BB), 1609 (Gen X), and 1439 (Gen Y). The data is presented in the order of raw numbers, percentages and normalised data (within brackets). The percentage shows each PM frequency as a portion of the total PM production of each social generation.

PM	Total use Baby Boomers (BB) (normalised frequency)	Total use Generation X (Gen X) (normalised frequency)	Total use Generation Y (Gen Y) (normalised frequency)	Total
You know	541 – 55.9% (682)	599 – 50.9% (820)	265 – 36.6% (527)	1405 – 47.5% (2029)
I mean	143 – 14.7% (180)	131 – 11.1% (179)	88 – 12.1% (175)	362 – 12.5% (534)
Like	60 – 6.2% (76)	73 – 6.2% (100)	204 – 28.2% (406)	337 – 13.6% (582)
Well	94 – 9.6% (118)	73 – 6.2% (100)	38 – 5.2% (76)	205 – 6.8% (294)
Right	24 – 2.4% (30)	132 – 11.2% (181)	28 – 3.8% (56)	184 – 6.2% (267)
Kind of	44 – 4.5% (55)	79 – 6.7% (108)	47 – 6.4% (93)	170 – 6% (256)

Sort of	21 – 2.1% (26)	69 – 5.8% (94)	30 – 4.1% (60)	120 – 4.2% (180)
Yeah	15 – 1.5% (19)	15 – 1.2% (20)	16 – 2.2% (32)	46 – 1.6% (71)
You see	13 – 1.2% (15)	03 – 0.2% (04)	06 – 0.8% (12)	22 – 0.7% (31)
No	09 – 0.9% (11)	03 – 0.1% (04)	01 – 0.1% (02)	12 – 0.4% (17)
Aah	02 – 0.1% (02)	0 – 0% (0)	0 – 0% (0)	02 – 0.04% (02)
Total	966 – 100% (1218)	1177 – 100% (1609)	723 – 100% (1439)	2866 (4266)

Table 7.3: Raw data for PM use for the three social generations

Figure 7.1 shows the trends more clearly.

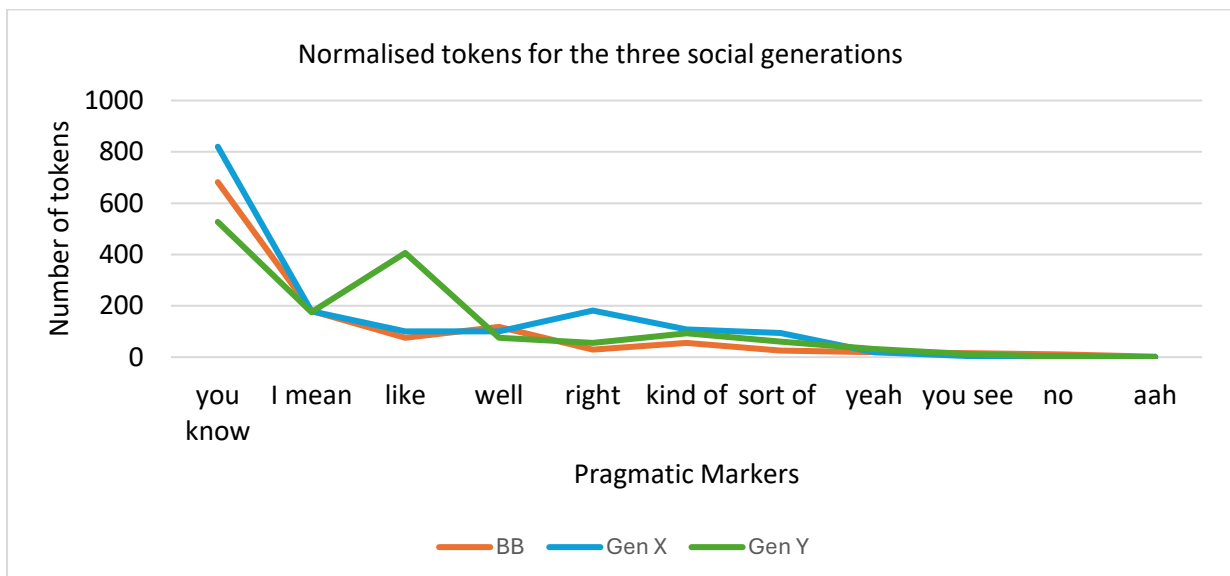


Figure 7.1: Normalised tokens for the three social generations

Gen X uses the highest total number of PMs followed by Gen Y and BB respectively; as percentages, 38%, 34% and 28% respectively. The difference between the three generations is small yet the difference between BB and Gen X is a 10% decrease. The PM with the highest frequency of use for all three social generations is *you know*. Gen X uses the highest number of *you know* among the three generations. Gen X uses 40% of total occurrences of *you know* while BB and Gen Y use 34% and 26% respectively.

Like is used most by Gen Y. *Like* has increased over the generations and Gen Y uses *like* four times more than Gen X. As a percentage of the total use of *like*, BB, Gen X and Gen Y use 13%, 17%, and 69% respectively. *Right* is particularly highly used by Gen X. The percentages of use for *right* for the three generations are 11%, 68% and 21% respectively. Gen X uses six times more than Gen X. Gen Y uses one third of the frequency of Gen X.

I mean is used almost equally among all three generations. *Kind of* shows the least use with BB and then doubles with Gen X. Gen Y uses *kind of* marginally less than Gen X. The frequency of *sort of* has increased threefold from BB to Gen X and the frequency shows a marginal drop with Gen Y. *Yeah* as a PM has steadily increased through the generations. As percentages of the total production of *yeah* the three generations report 27%, 28% and 45% respectively. BB uses 48% (15 instances) of occurrences for *you see*, but use decreases to just 13% (four occurrences) for Gen X. However, use among Gen Y speakers rises considerably to 39% (12 occurrences). The nativized PMs *no* and *aah* seem to be features that are dwindling in their use among Gen X and Y compared to BB at least in these semi-formal interview contexts of SSLE speakers, though the very low frequency in any group makes it difficult to draw clear conclusions here. *Aah* is used just in two instances by the BB cohort and does not occur among Gen X and Gen Y speakers. These trends will be analysed and discussed in detail in the data analysis section.

The chart below offers a comparison of the PMs with the three highest frequencies of overall PM use. The most frequently used PM among all three generations is *you know*. Each generation has used the other PMs at varying frequencies.

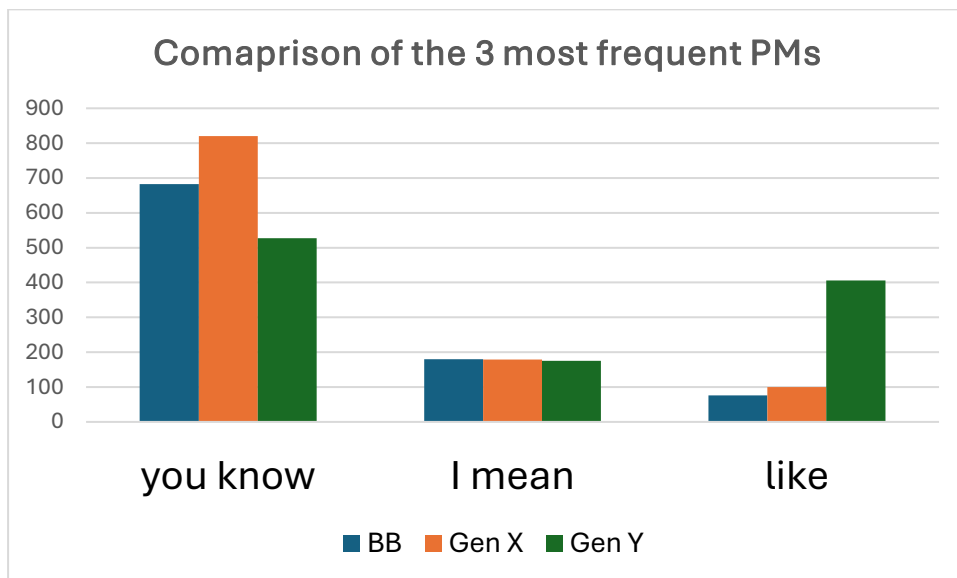


Figure 7.2: A normalised token comparison of the most popular PMs

This chart gives a close-up view of the three most popular PMs in the data. *You know* reports the highest frequency for all three generations while *I mean* shows similar frequency for all speech communities. *Like* is most frequently used by Gen Y. In fact, the use of *like* has gradually increased from BB to Gen X and increased exponentially in Gen Y.

7.6 Data analysis of Individual PMs

This section looks at the PMs individually across the three social generations to understand the trajectory of each PM closely. The analysis of PMs is presented according to frequency across the whole dataset.

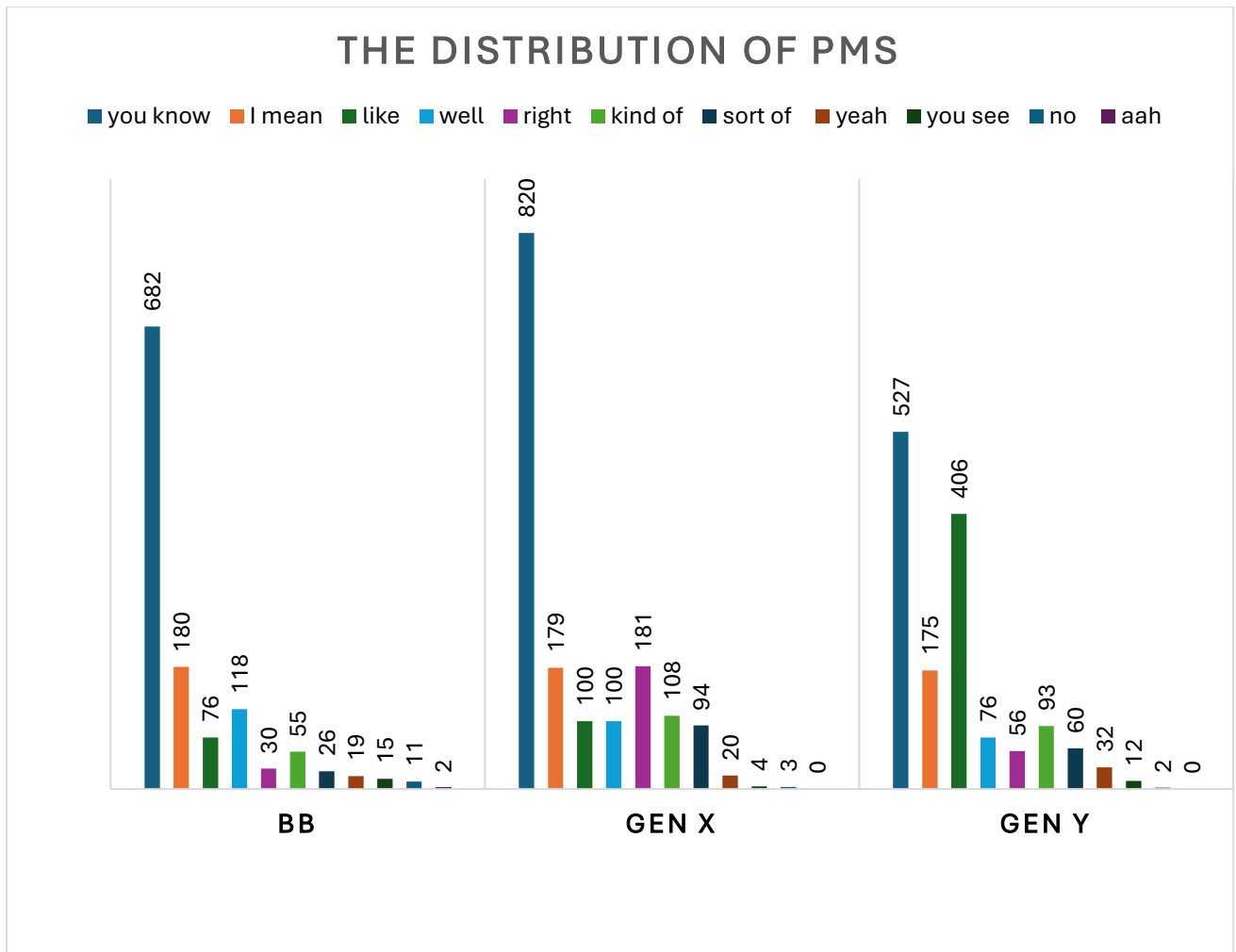


Figure 7.3: The distribution of pragmatic markers (normalised frequencies)

7.6.1 You know

You know has the highest frequency for all three speaker groups. It is widely used across age categories. This is no surprise given that *you know* is instrumental in maintaining cooperation between two speakers and the data is from spoken discourse in a public platform where cooperation and face-saving are particularly important. The social generations, i.e. BB, Gen X, Gen Y proportionally use *you know* in 34%, 40% and 26% of cases respectively. Gen X uses *you know* the most. For this group there are a total of 820 instances while BB and Gen Y use 682 and 527 instances respectively (here and in the following discussion all frequencies given are normalised). However, the overall trend suggests that *you know* might decrease in use with the generations that follow. Another PM could be replacing some of the functions of *you know*. Yet, it still leads in terms of overall use:

there is a noticeable gap between *you know* and *like*, which is the next most frequent PM. There are 2029 normalised tokens of overall *you know* use among all three generations, while there are only 582 instances of *like* overall. Nevertheless, it should be noted that *like* is the strongest competitor for *you know* especially because there is a high use of *like* in Gen Y data. In the analysis of *like*, the study will take special note of any instances where the functions of *like* may overlap the functions of *you know*.

An analysis of the functions of *you know* according to the three social generations will provide us with a better understanding of how each generation uses this PM. There are differences between groups in terms of function as evident in this data. The table below gives the distribution of *you know* according to functions in the SSLE corpus. The frequencies have been normalised to 100,000 words, and normalised tokens are in bold for clarity. The percentages show functions as a part of the total production of *you know* by each generation. For example, BB uses 13.3% of all *you know* with the function of attention getting. Similarly, Gen X and Gen Y use 5.1% and 11% of their total production of *you know* for attention getting.

Function	Baby Boomers (BB) -% of the total production of <i>you know</i> by BB	Normalised tokens	Gen X - % of the total production of <i>you know</i> by Gen X	Normalised tokens	Gen Y - % of the total production of <i>you know</i> by Gen Y	Normalised tokens
Attention getting/ launching a new	13.3%	91	5.1%	42	11%	58

piece of information						
Clarification and appeal to common knowledge	66.5%	454	67.9%	557	70.9%	374
Direct appeal to shared knowledge	1.6%	11	3%	25	3%	16
Filler	-	-	0.3%	03	-	-
Hesitation and appeal to common knowledge	0.8%	06	2.4%	20	1.8%	10
Introducing a quotation	2.7%	19	3.5%	29	3%	16
Repair	0.8%	06	0.4%	04	1.1%	06
Word search	1.9%	13	5%	41	4.5%	24
Word search and appeal to the interlocutor to fill in the gap	0.5%	04	1.2%	10	0.3%	02
You know in the final position: pointing out a	11.4%	78	10.8%	89	4.1%	22

self-evident truth						
-----------------------	--	--	--	--	--	--

Table 7.4: The functional distribution of *you know* according to age

All three social generations use *you know* mainly for clarification and appeal to common knowledge. As illustrated in table 7.4, all three generations use more than 65% of the total production of *you know* for clarification and appeal to common knowledge. This is in line with the overall findings of this study as well. Clarification and appeal to common knowledge, which helps to establish solidarity among speakers, seem to be an important function among all three generations. This inclusion of the listener so that s/he feels closer in the relationship with the speaker is illustrated in example 7.1.

Example 7.1

<01HM> do you play with them

<07AF> er yeah the first thing I do is take them **you know** in into the house and er like and my

In this example, <01HM> is an interviewer and <07AF> is the principal dancer in a pioneering traditional dance academy in Sri Lanka. The interviewer asks from <07AF> about her dogs. <07AF> explains about how she spends time with the dogs, and she includes the speaker by using *you know* and appeals to the common understanding of the listener. She seems to suggest that taking the dog into the house is a common action even to the listener. Thereby, she establishes a common ground with her interlocutor.

Gen X shows the highest frequency in the use of *you know* for six of the function compared to the other two generations. For example, Gen X has the highest frequencies for clarification and appeal to common knowledge, hesitation and appeal to common knowledge, introducing a quotation, word search, and word search and appeal to the interlocutor to fill in the gap. It is also the

only generation that uses *you know* as a filler, although as Chapter 5 pointed out, PM use for hesitation is not a typical feature of SSLE speakers.

Although the difference is very small, Gen X uses *you know* twice as often as BB for direct appeal to shared knowledge. BB generation devotes only 1.6 % of *you know* use for direct appeal to shared knowledge whereas both Gen X and Gen Y each use 3% of *you know* for this function. The age gap between two speakers appears to be a reason for the less frequent use of direct appeal to shared knowledge. The BB speakers in the interviews are the older of the two interlocutors in 22 of the 28 interviews in the BB cohort and therefore, BB could imagine that there is not much of shared knowledge with the younger interviewer. Therefore, the BB speaker does not directly appeal to shared knowledge. For example, in many of the transcripts there is evidence of the BB speaker referring to a historical past that s/he assumes the interviewer is not familiar with. There the context does not give rise to use *you know* for shared knowledge. Here are a two of such references to past events from the corpus of BB cohort.

Example 7.2

<02EM> then I er suddenly somewhere way back in er seventy seven when I was doing Queen Elizabeth construction the container terminal first Sri Lankan container terminal not in Singapore not any where in it was where planned and was starting going on in Colombo port and the Queen Elizabeth quay was to be made

The speaker in 7.2, <02EM> was born in 1949 and the interviewer was born in 1995. The YouTube video was made in 2018. Therefore, <02EM> was 69 years old and the interviewer was 23 years old at the point of the recording. The phrase *way back in er seventy seven* indicates that <02EM> is conscious of the age difference between the two speakers. Therefore, the BB, in this case <02EM> seems less likely to rely on any appeal to shared knowledge.

Example 7.3

<16HM> I guess back then you don't you don't have pen drives

<05EM> no [laughter]

<16HM> <u=?>

<05EM> no [laughter] it was very vinyl and it was it was just instinctive how you play music or how you read people and er it was a great thing it was a great time I had there down south and er from there I played music er on the weekends and I in the weekdays a few of the weekdays but I used to always shuttle back after work after work in the nights at about five six in the morning get back home have a shower get back to hotel school so it was kind of a long hours of work but you see what drove me was the passion for the music and what I was doing so the the time really didn't matter never looked at my watch **you know** so I just moved on and I and I was happy doing that and then I thought okay it's time to come to Colombo and make some noise in Colombo

In example 7.3, <05EM> is 56 years old and the interviewer is 22 years old. <05EM> is describing the time when he was very young and was working as a DJ in a place far from his home. <05EM> is describing how he carried the music with him. The young interviewer jokingly refers to pen drives to carry music and the elderly entrepreneur laughs and delves into history. He talks about vinyl records and how he worked as a DJ. He gives a sequential account of a typical day in his life in that era but he does not use *you know* even once for shared experience as the age gap between the two speakers is prominent. He uses *you know* once in final position, pointing out a self-evident truth. The self-evident truth is that because was so passionate about his work, he never cared how much time had passed so he never wanted to look at his watch. He first mentions “what drove me was the passion” and then this is backed by the phrase “never looked at my watch” which is the self-evident truth. When you are passionate about something, you do not feel the time pass. He sums up this self-evident truth and presents it preceded by *you know*.

Example 7.4

<06AcM> [laughter] so but er we grew up in this area where there's a lot of land I mean there would have been not more than thirty forty houses in that whole village er ran around the paddy fields the tank we we didn't buy anything from the stores the vegetables fruits the milk eggs all came from the land ma=my mother was a great lady

In example 7.4, the speaker <06AcM> was 55 years old at the time of the interview and the interviewer was 37 years old. <06AcM> is explaining what rural village life looks like in detail because rural life has changed so much in the modern world that is familiar to his interlocutor. The speakers in these instances cannot make a direct appeal to shared knowledge to the interlocutors because the context they are referring to is very alien to the interlocutors. Therefore, I believe a broad age gap between the two speakers may be part of the reason for fewer instances of direct appeal to shared knowledge.

Out of the total use of *you know*, BB and Gen X use *you know* in the final position pointing out a self-evident truth in 11.4% and 10.8 % of cases respectively. It is approximately twice more frequent than for Gen Y speakers, at 4.1%. When *you know* is used to embed a self-evident truth, the supposedly truthful idea is presented as a strong opinion. At the end of the utterance, the speaker sums up a point and presents it as a strong idea. In the example 7.5 below *you know* is in final position of the utterance.

Example 7.5

<4AM> I don't want to underestimate the problem we are facing today it is a global crisis so that's the next point **you know** and er the third thing is

This conversation took place during the Covid pandemic in 2020, when Sri Lanka was under a lockdown. In this example, <4AM> who is a male artist is talking about the difficulties people face during the lockdown and how to circumnavigate them. He is summarising points he has already made and moving onto his second point. There is no hedging used in putting forward this idea. He

directly introduces his second point indicating a certainty about his idea. He emphasises the self-evident truth that the pandemic is a global crisis. Therefore, when *you know* is used in the final position to introduce a self-evident truth, there is a sense of certainty about this idea.

7.6.2 Like

The next most frequent PM among the social generations is *like*. *Like* is especially frequent among Gen Y. As a percentage, the three social generations have used it in 13%, 17% and 70% of cases respectively. There is therefore a very steep increase in the use of *like* among the Gen Y cohort. Yet, *like* is not an invention of Gen Y as it is present in BB and Gen X as well. The youngest member of the Gen Y cohort is 27 years old and the oldest is 43 years old as of 2024. The speakers belonging to Gen Y are born between 1981-1997. Since there is an increase of use from BB to Gen X, we can say that by the 1980s *like* had become part of the PM repertoire. D’Arcy (2017) attributes the increase of *like* to those who were associated with Jazz, Cool and Beat groups of New York and notes that it increased in use among speakers born in the 1970s. This is reflected in the SSLE data shown in this corpus. *Like* use only marginally increases from BB to Gen X. The noticeable huge increase is from Gen X to Gen Y with people who are born in the 1980s and 1990s. It can be argued that while stylistic and social factors such as the dialect of the speakers are kept constant, the increase in the use of *like* in apparent time reflects a diachronic increase in the use of *like* in SSLE. This is a similar hypothesis to that presented by Labov (1963) when he analysed the centralised onset of (ay) and (aw) on Martha’s Vineyard. He noticed in his comparison of 5 age groups that the centralisation of (aw) and (ay) had increased with the younger age groups. He interpreted this apparent time development as a diachronic increase as well. His hypothesis might not be an absolute validation, but it provides a possible explanation in the absence of real time data.

D’Arcy (2017) looks into to the origin of the spread of *like* and says that she does not attribute the rising rates of *like* across Englishes to the New York social groups which originally propagated it in 1970s. This trend was seen in SSLE in the 1980s, but Andersen’s (2001) study provides a better explanation for the increased *like* in SSLE more recently. As mentioned in section

7.3., Andersen (2001) attributes *like* to the speakers in their teens and twenties, a cohort born in 1970s and 1980s. His comments are based on data from the Bergen Corpus of London Teenage Language (COLT) which was collected in 1993. He also mentions that *like* is seen to some extent among people below 41 and it is used very little by people over 45. These trends are reflected in SSLE as well. By the 2000s, *like* is clearly a well-established, versatile PM that expresses various mental processes including reporting thoughts and speech. Going beyond a mere frequency analysis of its occurrence, a functional analysis of *like* according to social generations will reveal the changes of use of *like*.

Function	BB tokens as a %	Normalised tokens	Gen X tokens as a %	Normalised tokens	Gen Y tokens as a %	Normalised tokens
Approximative like	5%	04	3%	03	3.5%	14
Discourse marking/hedging	-		-		3%	12
Exemplifying like	37%	28	45%	45	22%	90
Focuser	52%	39	31%	31	54%	221
Hesitation	3%	02	1%	01	2%	08
Quotative (be) like	3%	02	14%	19	15%	62

Table 7.5: The functional distribution of *like* according to age

As shown in table 7.5, BB generation and Gen Y use *like* as a focuser most often. In half of the total occurrences of *like*, it is used as a focuser by both these generations. Gen X uses *like* mostly to introduce and exemplifier. The second more frequent use of *like* is for exemplifying purpose in Gen BB and Gen Y. *Like* is least used across all generations for hesitation. The use of *like* as a quotative has tripled with Gen Y compared to the other two generations. There is a similarly marginal use of *like* for approximation among all three generations.

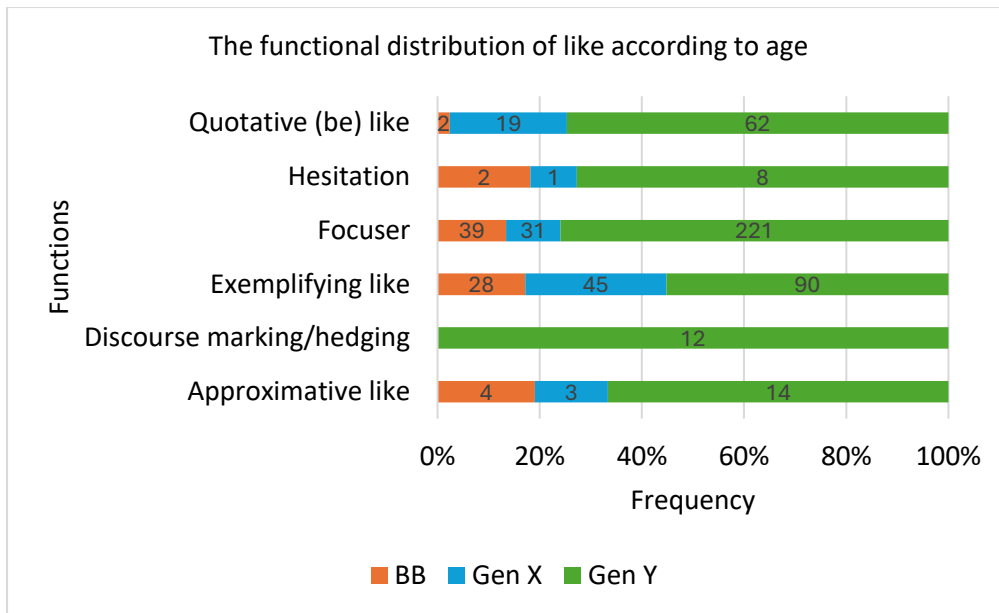


Figure 7.4: The distribution of functions of like according to age

Underhill in 1988 (p. 234) notes that in trying to collect data for his study of *like*, he does not hear *like* as much as he used to and suggests that *like* may even be becoming archaic. Ironically in the present data the Gen Y speakers who are born between 1981 and 1997 use *like* with the greatest frequency among all three generations. This increased use is reflected in the increased values for all functions of *like* as well from BB generation to Gen Y. Gen Y has an added function that is not found in the other two social generations; that is the function of discourse marking/hedging. An example of this function from one of the Gen Y speakers is provided in 7.6. In this instance, the speaker is describing her recent decision to eat healthier food and she says something to the effect that your whole self is completely described by what you eat. This is a strong statement, and she uses *like* in the final position of the utterance soon after this strong statement to downtone what she just said. The *like* here allows the speaker to reorient her statement and soften the message she is giving, and it is similar to the use of *sort of*. The strong statement that you are equal to what you eat is hedged by *like*.

Example 7.6

<04AF> I'm making better decisions it does affect you because you are what you eat right **like** and I come from a Burgher family and they love their meat like I couldn't have any meal

Example 7.6 shows an instance of *like* in sentence-final position which is considered a typical feature of northern BrE. *Like* occurs in sentence initial, medial and final positions. Hancil (2021) states that final *like* is more established in varieties in the British Isles. Hancil (ibid.) states that final *like* is mainly used for hedging. It can be noted that SSLE includes this northern English feature although the reason for the influence is not clear.

All the functions of *like* have increased by five-fold or more from BB to Gen Y. This shows that *like* has attained a prominent place in PM use as generations progress. It seems to be the one PM that fits many purposes for Gen Y, and therefore it is liberally used in spontaneous conversations. For instance, *like* as a focuser has increased exponentially from Gen X to Gen Y. This drastic increase shows *like* to be a serious competitor to the other PMs. For example, *like* has potentially replaced *I mean* and *you know* in instances where the PM deals with attention. This could be a reason why *I mean* and *you know* do not show the same growth in frequency as *like*. Given below are two examples from Gen Y cohort where *like* is used as a focuser instead of *I mean* or *you know*. *Like* in both these examples could be replaced by *I mean* or *you know* had the speakers belonged to different social generations.

Example 7.7

<01EM> I've just actually discussed this even with some of my clients saying what do you think do you think we need to make it a little more **like** general population specific

Example 7.8

<11EF> er well with the main production of the er food yes I'm doing it but my family is very supportive er my sister has **like** the creative eye so she helps me with the packaging

There is also a drastic increase in the quotative (*be*) *like* function. This is a trend that is seen all around the world with regard to *like* in inner circle varieties of English (D'Arcy, 2017). D'Arcy (2017)

mentions that of all functions this function is the most recent innovation for *like* in American English. She states that it was first noticed in American English in 1980s in Butters (1982). The data in this corpus only records two instances of quotative *(be) like* for BB generation, given below.

Example 7.9

<10AF> and then Sam oh he'll he'll let me

<01HM> sage

<10AF> yeah he he's like okay now time to get back into the stable [laugh]

In example 7.9, <10AF> who is born in 1964 is commenting on an imaginary conversation with her husband. The fact that she is born in 1964 classifies her as a borderline BB. Therefore, she is perhaps more likely to be influenced by Gen X through life experiences and thus be familiar with quotative *like*.

Example 7.10

<2EM> I was **like** er er I said can I get you employed I have a job for you I said I would

The speaker in example 7.10 is born in 1949. He first uses the *(be) like* construction and then corrects it. Perhaps, he could be monitoring his speech for formal purposes and therefore switching to direct reported speech with *say* which is more prevalent in his generation. It is Gen X (1965-1980) who uses *(be) like* more frequently in this data and this can be explained because *(be) like* first appeared in SSLE in 1980s. Butters (1982) also mentions that *(be) like* is used to introduce an unuttered thought. It is now used to present direct speech or uttered speech in addition to unuttered thoughts. The quotation below from a Gen Y female speaker is a good illustration of how much *(be) like* has expanded. The speaker is born in 1997 and she was 24 years old at the point of this recording done in 2021.

Example 7.11

<07EF> where I was **like** you know you know what I think I want to start my own thing and I had a few friends who are **like** okay we'll support you we'll make the website and then I met like my dietician my partner we were **like** we'll do it together

All three instances of *(be) like* given here are examples of reported speech. The first instance of **like** expresses what the speaker told her friends. The second **like** introduces what her friends told her. The third **like** introduces what the speaker and her partner spoke. However, there is an idea that this is not exactly what was said but something along the lines of these ideas was expressed. This is why *(be) like* is different to other quotatives such as *say* or *go* as it doesn't necessarily report the exact words. *(Be) like* is no longer confined to unuttered thoughts. All in all, *like* is a thriving PM in SSLE because it has expanded in its functions and has taken over some of the functions of other PMs. This aggressive development suggests the characteristics of surviving PMs in the future.

7.6.3 *I mean*

As shown in Section 5.4.2 in Chapter 5, *I mean* in my data is used mostly for clarification, exemplification, elaboration, reformulation. All three social generations use *I mean* equally in this data. The normalised values for total use are 180, 179 and 175 for 100,000 words. As a percentage, the frequencies are 34%, 33% and 33% respectively. The overall use of *I mean* is stable across all three generations. The Table 7.6 displays the frequencies of functions among the three social generations.

Function	BB as a % of all tokens	Normalised tokens	Gen X % of all tokens	Normalised tokens	Gen Y % of all tokens	Normalised tokens
Clarification, exemplification, elaboration, reformulation	69%	124	73%	131	68%	119
Concession and nuancing	3%	06	2%	03	2%	04
Hedging	1%	02	4%	08	4%	06

Justification	16%	29	17%	30	24%	42
Self-repair	11%	19	4%	07	2%	04

Table 7.6: The functional distribution of *I mean* according to age

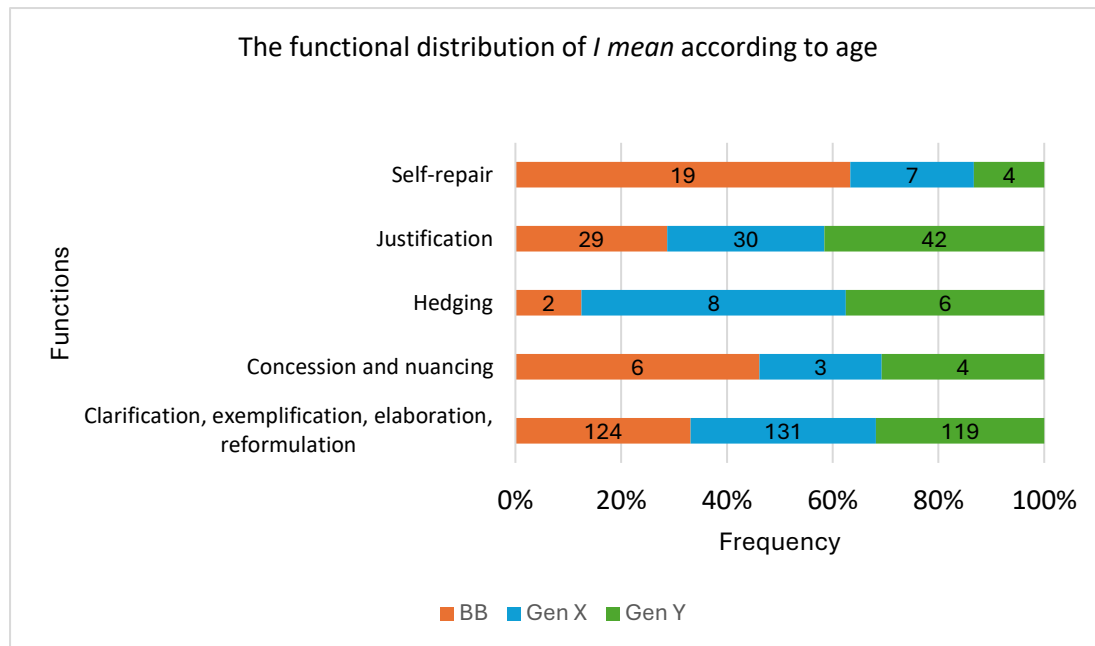


Figure 7.5: The functional distribution of *I mean* according to age

All three generations use *I mean* mainly for clarification, exemplification, elaboration and reformulation. The use has slightly increased among Gen X but has gone back to a similar use to BB for Gen Y. There seem to be a slight anomaly in Gen X in its use of *I mean*. It is difficult to interpret the reason for this very slight but gradual increase of *I mean* among Gen X. As for the slight decrease among Gen Y, it could be assumed that perhaps another PM with similar functions is replacing *I mean*. *Like* is a rising PM and exemplification is a function that overlaps with functions of *I mean* and may be replacing *I mean* in this function. All three social generations use *I mean* for justification and this function is moderately increasing with each passing generation. It indicates that *I mean* has established itself for the use of justification among all three social generations. Gen Y uses *I mean* more often than the other generations to signal justification in the utterance following. There is an upward trend to use *I mean* for hedging. This use is at a minimum with BB but Gen X shows more frequent use and it has marginally decreased in Gen Y. *I mean* as a self-repair strategy is frequent

among BB but it seems to lose this function in Gen X and Gen Y. The same pattern is observed in *you know* for self-repair as well. BB and Gen Y use the same number of instances of *you know* for self-repair while there is a slight dip in Gen X. *I mean* has decreased in use for self-repair as the generations progress.

7.6.4 Well

Well is a PM used frequently among BB and Gen Y compared to Gen X. BB and Gen Y almost match each other in their use of *well*. Figure 7.6 outlines the frequencies of the functions of *well*.

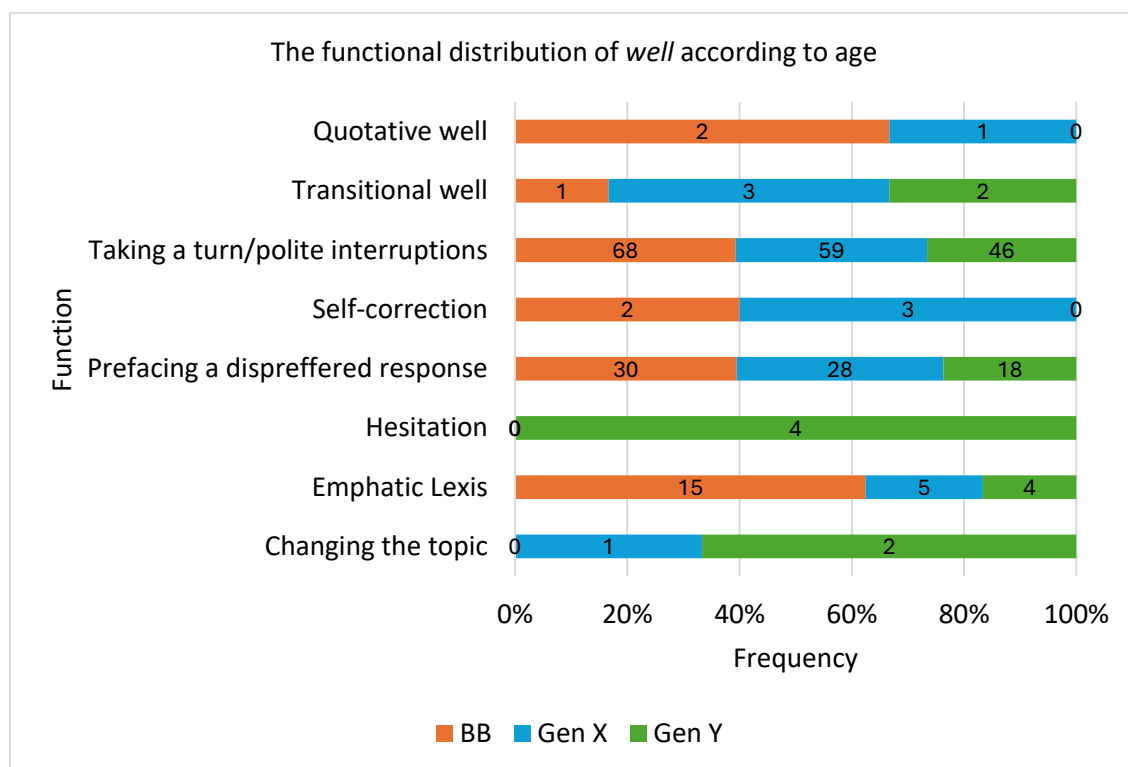


Figure 7.6: The functional distribution of *well* according to age

The use of *well* has decreased over the generations as the total use for this marker is 118, 100 and 76 respectively. This indicates that *well* is a feature of an older generation in this data set. The data shows eight functions for *well*. Taking a turn and prompting polite interruptions is the most prominent function for all three social generations. This use, too, is decreasing, which is a reflection of the decrease in overall use of the PM. However, *well* for taking a turn/polite interruption and *well* prefacing a dispreferred response are in frequent use indicating that they are important functions to

all generations. *Well* prefacing a dispreferred response expresses the speaker's disagreement and politely corrects a misunderstanding.

Example 7.12

<01HM> so you get annoyed when you get described as Professor JB's daughter?

<02AcF> not annoyed

<01HM> <u=?> it irks you

<02AcF> **well** no it's just that women have been defined by normally whose father they are whose wife they are and that kind of as a woman not irks me but I wish the system was different

In 7.12, <02AcF> is an academic and her father too is a renowned academic in Sri Lanka. Therefore, the interviewer is asking whether <02AcF> gets annoyed when she is described as this professor's daughter. The second comment "it irks you" implies that the interviewer is expecting a positive response to this question. <02AcF> avoids being impolite by using *well* to introduce the dispreferred response and the explanation for her answer.

BB and Gen X use *well* at a similar frequency for this function while Gen Y shows a lower frequency. One notable difference is that *well* is used often by BB for emphasis, but much less by the next two generations. There are certain functions that correlate even more prominently with age. For example, *well* used for changing the topic is confined to Gen X and Y, although this observation is based on very few tokens. One possibility is that this function was innovated by Gen X in SSLE. Another interpretation is that its use in BB is not recorded in the current data; perhaps the context of the current data does not capture this use. However, according to this limited data we can assume that *well* is adaptable as it has found a new function or a new meaning with Gen Y, and this is *well* for hesitation. This is also a very rare function as only two instances are recorded in the present data. Both instances are from one speaker.

Example 7.13

<01HM> er any place that you haven't been to which you are dying to visit you want to go

<07AM> yeah er well er

<01HM> mine <u=?> one day

Example 7.14

<07AM> you know sort of cliché but that's just the one thing we've been well erm yeah

mostly I think it's just the experience being a in a different place I think we are we are big on

food different types of food and er try to recreate it when we get home basically

In example 7.13, the speaker is buying time to give a well thought-out answer to the question. In this instance he is clearly using *well* for hesitation. The filled pause *er* precedes *well* to create more time for the speaker to continue. In example 7.14, the same speaker uses *well* again with the filled pause *erm* to show hesitation in forming his ideas. However, it should be noted that with the very few tokens that are in the data, it is difficult to draw any conclusions with regard to patterns that reflect age. While *well* for hesitation is only found with one speaker in Gen Y, *well* for self-correction and as a quotative are only used by BB and Gen X. Based on the very limited data, it can be assumed that these uses have become redundant among Gen Y. Nevertheless, when one function discontinues with a generation, it finds its way to a new generation with a new purpose. These generation specific developments show that PMs that have a risk of redundancy can revive through innovative use.

7.6.5 *Kind of and sort of*

Kind of and *sort of* are discussed together here as they are almost synonymous. They are both used only for metacommenting, hedging and qualifying in the entire corpus. Metacommenting refers to linguistic distancing 'from the responsibility for using words which are inappropriate because they are technical, trite, too informal, too formal etc.' (Aijmer, 2002: 209). Hedging and qualifying refer to diminishing the impact of certain words on the listener and down toning the speaker's attitude. There are 170 raw instances of *kind of*, meaning a normalised frequency of 256.

The percentages for BB, Gen X and Gen Y are 21.4%, 42.1% and 36.3% of the total respectively. There are a total of 120 raw tokens of *sort of* (normalised frequency 180). The percentages for the three generations are 14.4%, 52.2% and 33.3% respectively. Both *kind of* and *sort of* are used by Gen X more than the other generations. The comparison of the total use for these two PMs suggests that SSLE speakers use *kind of* more often than *sort of* although Gen X seems to prefer *sort of* over *kind of*. The overall result however, is a contrast to BrE. A 2009 study observed that the preferred variant in BrE varieties is *sort of* (Miskovic-Lukovic, 2009, p. 619). Both PMs show a similar trajectory over the three generations.

Function	BB %	Normalised tokens	Gen X	Normalised tokens	Gen Y	Normalised tokens
Metacommenting, hedging and qualifying - kind of	21.4%	55	42.1%	108	36.3%	93
Metacommenting, hedging and qualifying – sort of	14.4%	26	52.2%	94	33.3%	60

Table 7.7: The functional distribution of *kind of* and *sort of* according to age

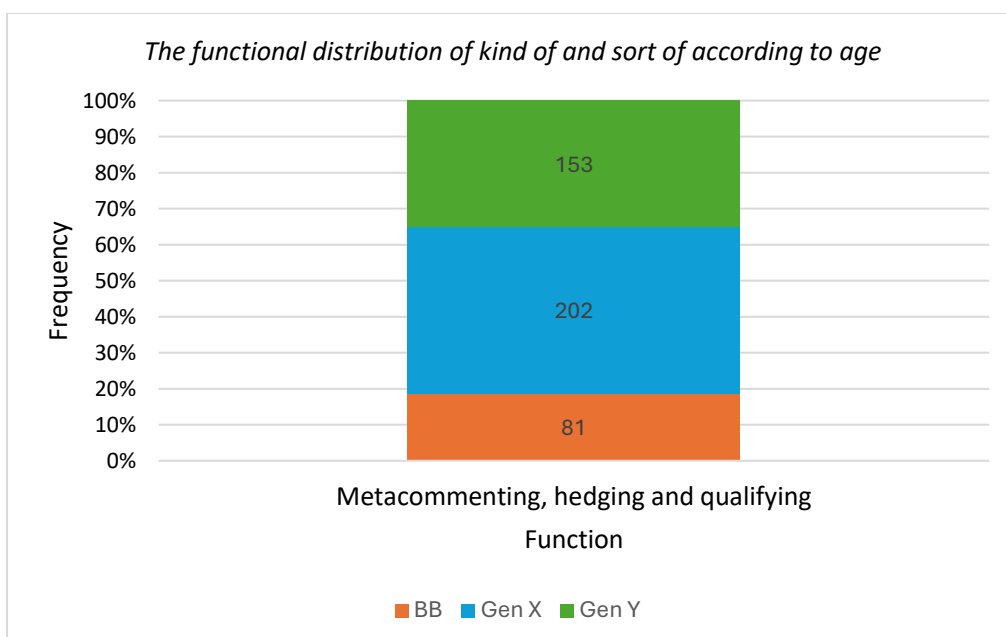


Figure 7.7: The functional distribution of *kind of* and *sort of* according to age

Kind of and *sort of* have more than doubled their use among Gen X compared to BB. They have declined marginally among Gen Y speakers, which might suggest that Gen X is more cautious in their language than BB. A high use of metacommenting, hedging and qualifying is used when you express measured ideas in a way that is aware of the context and the interlocutor. This is natural in a conversation that happens in a semi-formal setting on a public platform such as YouTube. However, the fact that Gen X speakers use this feature more often than the other generations suggests that they are particularly concerned about expressing their ideas and opinions cautiously. The only other PM used for hedging in this corpus is *like* and only Gen Y uses *like* as a hedging device. Perhaps the strategies of hedging is evolving between Gen X and Y and *like* is taking over as the new strategizer. This may be the reason for the decrease of *kind of* and *sort of* in Gen Y. Gen X uses 47.15% of *sort of* and *kind of* while Gen Y uses 34.8%. Despite the marginal drop from Gen X to Gen Y, it indicates that *kind of* and *sort of* are characteristic of Gen X speech and continue in Gen Y speech as well. According to Labov's (1994) definitions about language variation, the high use of these two PMs in Gen X could be a marker of language behaviour for that generation. As Reichelt (2021) reiterates, this type of a conclusion requires attitudinal data²⁰ which the current study lacks. The marginal drop in the use of *sort of* and *kind of* among Gen Y may indicate that Gen Y may have found an alternative PM for metacommenting, hedging and qualifying. The alternative could well be *like* as there is a drastic increase of *like* among Gen Y. Further, hedging as a function of *like* is seen only in Gen Y. Given below is an instance where a Gen Y speaker uses *like* for hedging.

Example 7.15

<04AF> it's it's amazing

²⁰ Attitudinal data in this context refers to speaker intentions, motivations and opinions about the discourse situation

<03HM> my

<04AF> it's one of my **like** favourite things

<03HM> yeah

<04AF> in one platter

Like in this example can be easily replaced by *kind of* or *sort of*. In example 7.16 of a Gen Y speaker, all three hedges can be seen.

Example 7.16

<03AcF> and then I like I kind of did I had a sort of geographical shift in my interest because in my undergrad I was doing more sort of the European and American histories and which was sort of standard erm

The speaker seems to be indecisive about which PM to choose for the hedging purpose, and ultimately uses all of them to hedge more. Clearly the natural first choice is *like*, which is then quickly changed to *kind of* and then to *sort of*. This can be taken as an excellent example of future speaker choice according to preference in SSLE. Being a Gen Y speaker, she shows that her first choice is *like*, and *like* is most frequently used by Gen Y speaker. Generally, SSLE speakers prefer *kind of* over *sort of* and therefore, her second choice was *kind of* and the third was *sort of*. She reflects the choice of PMs for a Gen Y Standard Sri Lankan English speaker.

7.6.6 *Right*

Right as a pragmatic marker is definitely a feature of Gen X. As a percentage of total use of *right*, the generations use 11%, 68% and 21% of instances respectively. In some of its functions, it works as a progression check question tag. A speaker will use *right* in his or her speech to see if the listener is following the conversation. Yet, the speaker does not expect a verbal answer. It is used as a way of maintaining the conversational engagement with the listener. In other functions, it works as a

key to tapping assumed common knowledge or sharing knowledge as mentioned in the Section

5.4.5.

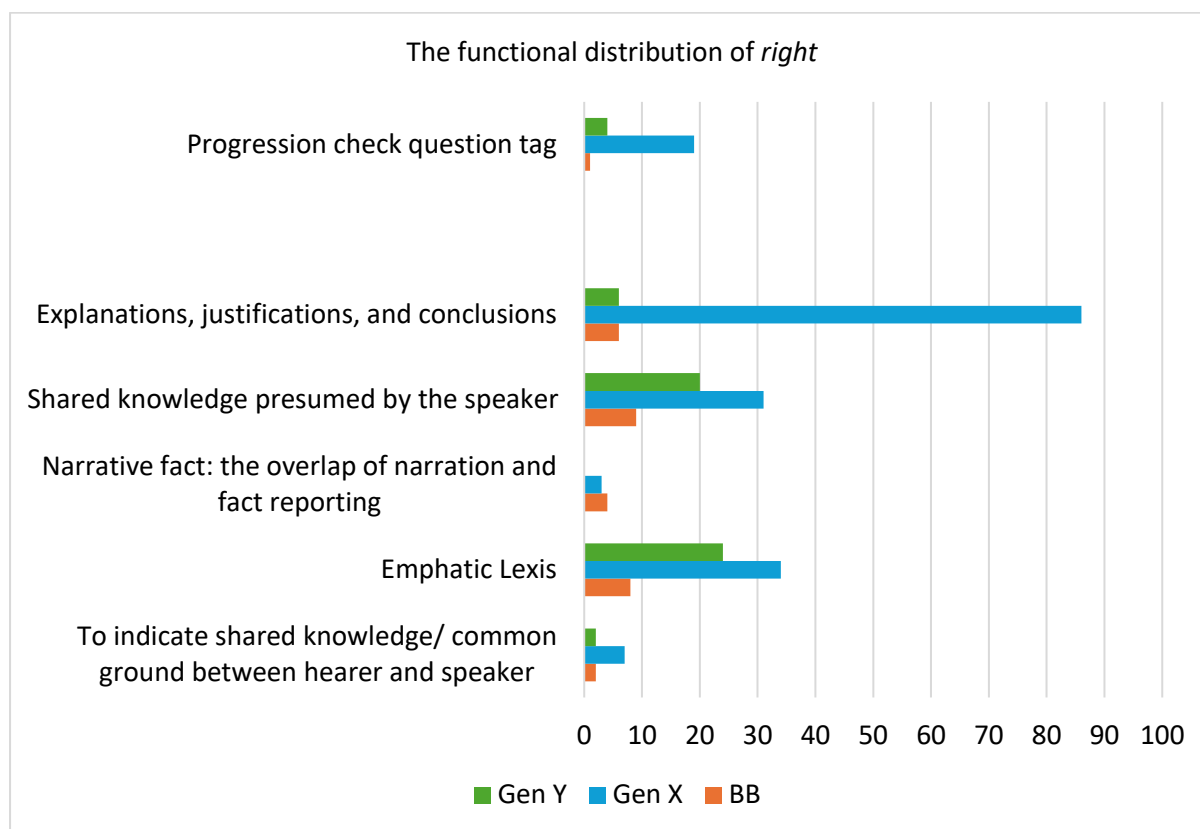


Figure 7.8: *The functional distribution of right*

Right is used most frequently by Gen X in all the functions. Gen X uses it mostly for explanations, justifications, and conclusions. BB and GenY show very low frequency in this use, though there are instances for both groups. Gen X uses *right* in a similar way to *you see* in these instances, as illustrated in the following example 7.17.

Example 7.17

<04EF> you know I wor=I worked in different types of organization and my last appointment was Apollo Hospitals

<17HF> okay

<04EF> **right** so I had to give it up because of my son

<17HF> okay

<04EF> **right** he was little at that time but now he is a grown up boy you know

In this instance, <04EF> a 59-year-old female entrepreneur is explaining some of her personal experiences to the listener to give her context. The explanations are marked by *right*. These could be replaced by *you see* if needed. The listener is acknowledging the explanations by using the backchannel *okay*. This use of *right* is a characteristic of Gen X as evident in this data set. The reason for this sudden spike in Gen X is unclear. It is not an over production of one speaker. As explained in Section 5.4.5, *right* is influenced by the Sinhalese and Tamil use of *hari* [ha'ri] and *sari* [sa'ri], which are confirmation markers. Perhaps there is an increase in *hari* among Gen X and therefore it transfers to SSLE. Without comparative data from Sinhala, this can only be an assumption. However, we can conclude that *right* is innovative in being popular among those born between 1965 – 1980.

7.6.7 Yeah

Yeah has considerably increased with each social generation. For example, BB uses a 26.3%, Gen X uses 29.1% and Gen Y uses 44.4% of total use. It is a PM that is becoming frequent among more recent generations. Its functions seem to be generation specific. For example, it functions to mark explanations, justifications, and conclusions only among the BB cohort, and topic shift only among Gen X according to this data, though there are only a small number of tokens for each use.

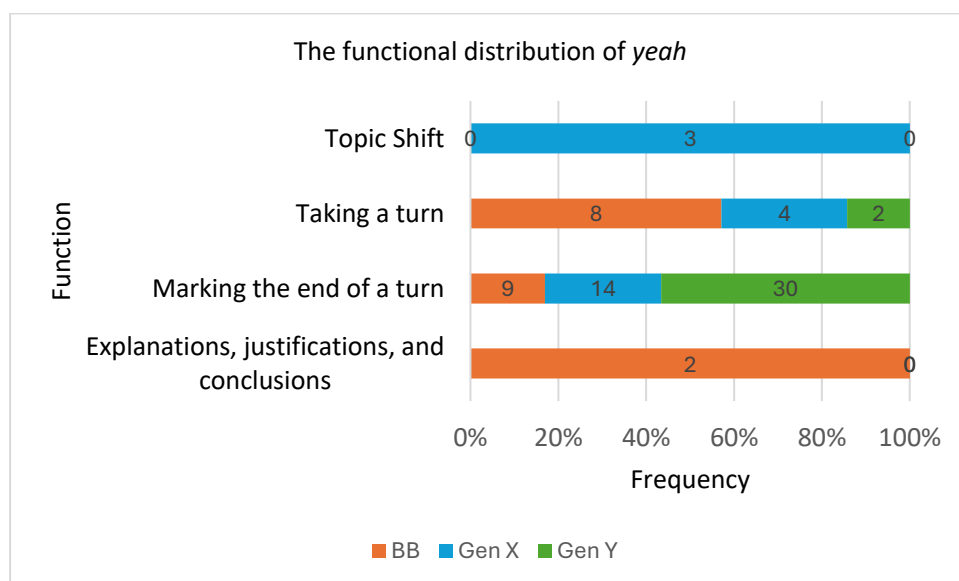


Figure 7.9: The functional distribution of *yeah*

Use of *yeah* to mark the end of a turn has almost doubled in use with each generation, and Gen Y mainly uses this PM to mark the end of a turn. This function is a therefore a noticeable characteristic of Gen Y. The interview data show that *yeah* draws a close to what the interviewer has been discussing as illustrated in chapter 5. This is a very common use among younger speakers as evident in this data. Use of this PM in turn-initial position has diminished.

7.6.8 You see

The normalised frequency of *you see* in the current data is 31 across all three groups, which is split as shown in figure 7.10.

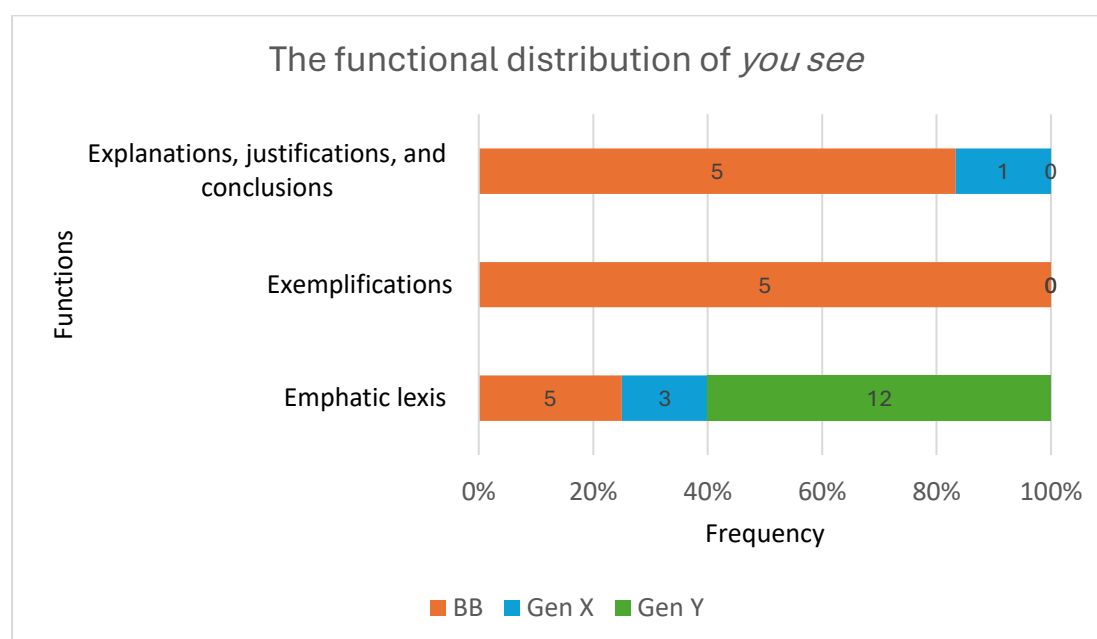


Figure 7.10: The functional distribution of *you see*

You see which has a very few tokens in this data is mostly used by the BB cohort. All three social generations use *you see* for emphasis, but BB is the only generation that uses *you see* for a range of functions. Gen Y uses it for emphasis the most frequently and the data demonstrate that this is its only function among this speaker group. All three generations use a wide range of PMs for emphasis. They have the choice of *well*, *right*, *you see*, *no* and *aah*. It would be useful to know which contexts give rise to each type of PM for emphasis and whether they are interchangeable. The data

shows that BB uses *well* mostly for emphasis, Gen X mostly uses *right* and Gen Y mostly uses *you see* for emphasis. As this data analysis shows in the sections 7.6.9 and 7.6.10, *no* and *aah* are used for emphasis by BB generation. Naturally, the more prevalent PMs such as *well* and *right*, with a wider range of functions than *you see*, are more easily adopted by speakers. Another observation is that only the BB generation uses *you see* for exemplification. Gen X and Gen Y are more likely to use *like*, a more common and trending PM for exemplification, which could be a reason for this diminished use of *you see*.

7.6.9 No

Gunsekera (2005) opens her chapter on Sri Lankan English Syntax with the utterance “Raining no, how to come?”. She explains that this utterance is typical of Sri Lankan English syntax. The normalised frequency of *no* is 17. The BB generation has 12, Gen X has three and Gen Y has two instances respectively. Gunsekera (2005) analyses the functions of *no* as emphasis and as a question tag. The current study notices an additional function, where it is used to mark turn taking. As for the origin of this feature, as mentioned in Chapter 4 and as supported by Gunsekera (2005, p. 129), it is an example of language transfer from Sinhalese to English. *No* is used most frequently by BBs and then there is a noticeable drop in the number of tokens with Gen X and a further drop with Gen Y. Perhaps, since *no* is regarded as informal, speakers have avoided it in this particular speech context.

As mentioned in Section 5.2, *no* as a question tag was omitted from the search for PMs. Two functions are reported: emphatic lexis and taking a turn. BB generation uses *no* for both these functions. Gen X uses *no* for emphatic lexis only while Gen Y uses it to mark taking a turn. This pattern suggests that *no* is feature of BB and perhaps, Gen X and Y try to avoid it. This result is similar to Lange’s (2012) study of IE where she found that the use of the nativized pragmatic marker *na* and *no* had reduced with the youngest age group in her study.

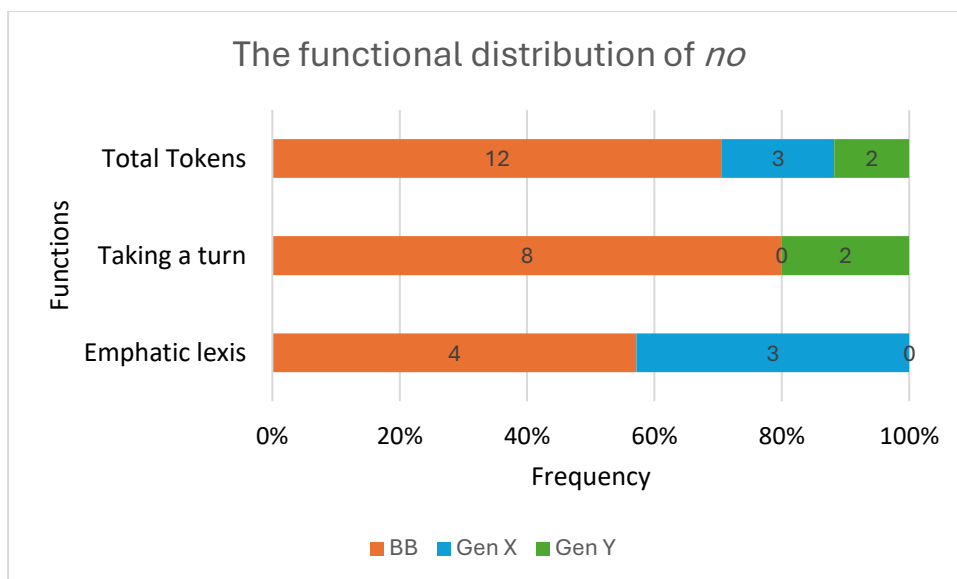


Figure 7.11: The functional distribution of *no*

7.6.10 Aah

The current data only records two instances of *aah* as a nativized PM. Two speakers in the BB cohort use one token each, each for separate functions. These functions are emphatic lexis and explanations, justifications, and conclusions. This indicates that *aah* is a PM marker that has the potential to be a versatile marker in SSLE. Meyler (2011) in his *Dictionary of Sri Lankan English* gives three examples of *aah*. His examples are from Sri Lankan English literature. One of the examples is “Why you want to get all high and mighty, ah?” (p.3). What was noted is that the example that Meyler gives of *ah* is from a novel called *Jam Fruit Tree* and the speaker in this literary text is a person from the BB cohort. Therefore, it can be proposed that there is evidence from literature to suggest that this is a feature of BB in SSLE.

7.7 A Summary of Findings

This chapter explored the correlation between age and pragmatic markers in SSLE in apparent time. The data was divided into age cohorts of social generations for this purpose. The following conclusions were drawn from the data analysis.

There are correlations between age and PM use in SSLE, which are mainly evident in their frequency of use. Generation X uses the highest number of PMs. *You know* is the most frequently

used PM among all social generations. In fact, statistically one third of PM production consists of *you know* for every person in this SSLE corpus. However, there is a marginal but steady declining pattern for the use of *you know* with each progressing social generation. Some of the functions of *you know* have decreased with the younger generations, such as the use of *you know* for clarification and appeal to common knowledge. Although this is one of the most recurring functions, its use has slightly declined. The data suggests that another PM which is gaining popularity with younger generations could have replaced *you know* in this use. In the case of clarification and appeal to common knowledge, the data suggests that *you know* is replaced by *sort of* and *kind of*. Gen X has used *you know* in its entire range of functions. The data suggests that if there is a broad age gap between two speakers there will be less instances of direct appeal to shared knowledge. Further, if an interlocutor is younger than a speaker, there is a tendency for the speaker to use *you know* more often in the final position pointing out a self-evident truth. However, this claim needs to be tested with a larger body of data to be validated.

Another conclusion is that frequency of use indicates which generation owns a certain PM. For instance, *you know* is a popular PM that is very often used by all three generations. This makes it difficult to assign it to one particular generation but allows us to classify that as a firm feature of SSLE. In contrast, PM *like* shows a very clear prevalence among speakers of one generation: it is 'owned' by Gen Y. Similarly, *right* is 'owned' by Gen X. The noticeable spike in the frequency of use brings us to these conclusions. The analysis of frequency also shows that *I mean* is a PM that is used frequently and it has established itself and is stable among all its users.

A functional analysis shows the development of a PM across generations. We can clearly see how some PMs have become stronger by expanding their repertoire of functions and how some have become weaker by becoming more limited to fewer functions. *Like* is an example of a pervasive PM while *right* is becoming weaker comparatively. A comparatively non-traditional PM is *yeah*. The data shows that this PM is gaining popularity with the younger generations and therefore, it can be assumed that it too will be a strong PM among future generations.

One more conclusion is that there may be competition among PMs, though this is an open question. My interpretation of the data suggests that the rise in the use of *like* is linked to the slight decrease in the use of *you know*. There is also a decline in the use of *you know* as a quotative and a high increase of *like* for this purpose. In this light, *like* is the strongest competitor to *you know*. One of the most popular functions of *you know* is for clarification and appeal to common knowledge and even this has declined over the years. The PMs *kind of* and *sort of* too imply clarification and their use has increased with each generation. Therefore, it seems likely that these two PMs also compete with *you know*. However, as mentioned earlier this is one interpretation and there could be other explanations that do not come to light in this data.

PMs display great survival tactics. For instance, when a PM loses its function to a more vibrant contender, it may ensure its own survival by flexibly adapting a new function. For instance, the traditional hedging PMs are *kind of* and *sort of* but this data shows that *like* is used for hedging in Gen Y, the youngest social generation group in the data.

The use of nativized PMs has diminished over the generations. Perhaps this data is not the best choice to explore nativized markers as the speakers are in a semi-formal interview on a public platform. In a context as this, younger generations seem to avoid the use of the nativized PMs *no* and *aah*.

This chapter analysed how age interacts with the use of PMs among SSLE speakers. The findings reveal that there are correlations between age and PM use in SSLE, with some PMs considerably more frequent in certain generations. Moreover, the evidence suggests that this distribution is in some cases similar to other varieties such as BrE and AmE. The trajectory of PMs over the generations also reveal that there are dynamics of use among PMs and some PMs are being replaced by others in particular functions or overall use.

Chapter 8 : Conclusions

8.1 Introduction

The main objective of this study has been to describe the pragmatic markers in Standard Sri Lankan English in terms of their functions and describe the ways in which gender and age impact them. The data clearly demonstrates the complexity and richness of PMs in SSLE, even though they are seemingly little words with no impact on the propositional or the main meaning of an utterance. PMs manifest valuable information about how interlocutors manage their speech with each other through elements outside the propositional stance. At a general level, the PMs contribute to the smooth flow in a conversation by establishing cooperation from the interlocutor, allowing time to plan speech, facilitating turn taking and expressing politeness. At another level, they indicate upcoming disapproval, express the certainty level of an argument, present justifications and clarifications, and bring the focus of the interlocutor to the most important message in an utterance. The PMs that are used to achieve these subtle meanings or functions adjust their role according to each speech situation. It is indeed fascinating to understand the flexibility and dexterity of the PMs in the data of this study.

In this chapter, I summarise and synthesize the major findings from the study. The summary of findings is organized according to each research question of the study. I go on to discuss the implications that generate from this thesis for Standard Sri Lankan English specifically, then, PMs in world Englishes and finally, for variationist studies on PMs. I conclude this chapter with suggestions for future directions in studies connected to the present research.

The research questions are repeated below for ease of reference:

1. What pragmatic markers are used by SSLE speaker in semi-formal to formal conversations?
2. What are the functions of each of these PMs in the corpus?
3. How does gender interact with PMs in terms of frequency and function?
4. How does age interact with PMs in terms of frequency and function?

A custom-made corpus of standard spoken Sri Lankan English (SSLE) was compiled with a word count of 202,557 to explore the presence of PMs. The corpus contained speech samples of SSLE spoken by 72 speakers with equal gender representation. This particular variety was selected because of its perceived prestige; indeed, it is the variety that is promoted in the ESL classroom in Sri Lanka. This custom-made corpus is, alongside the research findings, a further tangible contribution of this thesis. This corpus (CSSLE) is, at the time of writing, only the second existing corpus resource of spoken data in Standard Sri Lankan English (SSLE), alongside ICE-SL. As noted in Section 4.3, the considerations taken into account when compiling the CSSLE offer some potential advantages for researchers interested in specific aspects of interaction in SSLE. For instance, CSSLE has metadata such as occupation, gender and age for all the speakers. CSSLE will be freely available and will be hosted in the Digital Humanities Lab of University of Colombo, Sri Lanka (dhlab.cmb.ac.lk). This will open up new areas of investigation for other researchers working on Sri Lankan English, adding to the growing collection of resources for the study of World Englishes. Since it is a relatively moderate-sized corpus it makes available a body of relevant and reliable data that can be easily manually analysed or can be digitally processed.

The remainder of this chapter is organised as follows. The next section summarises the major findings with regard to research questions 1 and 2, namely the PMs that are used by SSLE speakers in semi-formal to formal conversations and their functions. Section 8.3 outlines the findings about the impact of gender on the use of PMs. Section 8.4 presents a summary of findings on the prominent patterns of PM use in three different age groups. Section 8.5 discusses some limitations of the study. Finally, Section 8.6 proposes directions and suggestions for further research.

8.2 PMs use and function: key findings

The data contained a total of 2949 PMs. This figure shows that even a medium-sized corpus such as the CSSLE can yield a meaningful amount of data for both quantitative and qualitative study. It suggests that it is achievable to carry out exploratory studies in pragmatics, discourse, and other

linguistic aspects of varieties of English that are less represented in corpus resources by developing smaller-size corpora. The first research question of the study was to identify the PMs that would appear in this corpus of 72 formal to semi-formal conversations. The following were identified as PMs in the CSSSLE: *Aah, I mean, kind of, like, no, right, sort of, well, yeah/so yeah/yes, you know, you see/see*. There are a total of 11 PMs excluding co-occurring PMs used by SSLE speakers in Sri Lanka in formal to semi-formal interviews. 11 co-occurring PMs were also identified: *I mean like, I mean you know, like I mean, like you know, no I mean, you know I mean, you know like, you know sort of, well I mean, well yeah you know, well you know*. This repertoire of PMs in the CSSLE shows that they closely follow the same forms of PMs used in inner circle varieties. The tendency is for a speaker to use an average of 50 PMs for every 2750 words in the present study. This indicates that approximately 1.4 % of our speech consists of PMs. This finding is key to understand that despite the bad reputation PMs seem to have in perception studies, fluent speakers in Sri Lanka do use PMs to achieve their communication goals even in highly public formal settings.

The second key finding is that in addition to PMs that are typically found in other varieties of Englishes, two nativized PMs too were present in the data. These are *aah* and *no*. These two features are a distinctive feature of SSLE. The nativized PMs have similar counterparts in Indian English and in Singaporean English. For instance, *no* is a question tag that is used in Indian English and several other South Asian varieties of English as well. In SSLE, *no* is a turn taking and emphasis marker in addition to being a question tag. The current study did not consider the instances of *no* performing the role of a question tag, as then it becomes a part of the propositional meaning. The second nativized PM is *aah*. *Aah* is a feature that is found in Singaporean English as well. Lim (2007) notes that *aah* is used at the end of a statement or a question. The question may be either rhetorical or one which requires a response. *Ah* in SSLE too is used at the end of a statement and performs a similar function. These similarities suggest that PMs are similar across regional varieties as well, although the indigenous languages that influence such PMs are not mutually intelligible. Regional varieties do not even share the same language family as in the case of Malay and Chinese that influence Singaporean English and

Sinhala and Tamil that influence SSLE in Sri Lanka. Malay and Chinese are from the Sino-Tibetan language family, Sinhala is from Indo-European language family and Tamil is from the Dravidian language family. The presence of the nativized PMs in the data also indicates that SSLE speakers use them in formal to semi-formal settings. Despite the negative reputation that PMs broadly have as fumbles or features of deficit language, the data shows that even nativized PMs are an integral part of the speech of SSLE speakers in high profile discourse situations such as recorded data intended for public dissemination.

Another finding with regard to the use of nativized PMs is that the Standard Sri Lankan English speakers use the nativized PMs rather sparsely. For instance, there are only 15 instances of nativized PMs among the total of 2949 instances. This implies that the nativized PMs are not equal to the forms that are shared with the inner circle Englishes. It further implies that the SSLE speakers do not use localization processes of PMs in SLE to highlight their language identity strongly. It can be assumed that the SSLE speakers are reluctant to take the ownership of their use of language. This observation is indicative of the speakers' perception about what is appropriate for different communicative settings. It seems from this research that in this formal/semi-formal register, even when the speakers are relaxed with their interviewer, they are still opting out of nativized PMs. Further comparisons and research are needed in different communicative settings to test whether this absence of nativized PMs in this register is a feature of the language variety or of the setting.

The third key finding is that the 11 PMs in the CSSLE perform 23 functions. The PMs are mainly used to clarify information and find common ground in the domain of the present study at least. The interview setting in CSSLE data is a very interactive one, where a lot of information is given, and where the interviewee might wish to make a good impression, perhaps to seem likeable and friendly and make themselves understood. Therefore, they might be particularly inclined to use many PMs for these functions. This function is achieved mainly by using *you know*, which is the most used PM, accounting for 49% of all PM use. The frequent use of *you know* in SSLE indicates that the SSLE speakers are keen to ensure that their ideas are understood clearly while expressing solidarity with

the interlocutor. This keenness to ensure clarity is supported by the fact that the second most frequently used PM *I mean* is used mainly for clarification, exemplification, elaboration, reformulation. It can be assumed that clarifying a point in a way that directly engages the speaker in an effort to show cordiality is a mark of SSLE.

The fourth key finding is that the present data has evidence of two possibly variety-specific functions. One is using *yeah* as a turn ending PM. Most instances of the PM *yeah* in this data indicate the ending of an utterance. It can be stated that *yeah* is used to partition thoughts so that it is indicated to the listener that the speaker will not talk about a certain topic any further. Additionally, the results of the present study lead to the hypothesis that *right* as a progress check is a nativized function. *Right* is used commonly among other varieties of Englishes as well. However, Sinhala and Tamil, the two native languages of Sri Lanka have an exact equivalent of *right* used for similar functions. Therefore, the comparatively high use of *right* especially as a comprehension check question could be attributed to the native language influence. Thus, it can be considered a nativized function. In sum, the present SSLE data from a formal discourse context reveals nativized PMs and nativized functions PMs, in addition to the presence of a repertoire of PMs closely resembling the PMs commonly used in inner circle varieties such as AmE and BrE.

The present study made some attempt to compare SSLE with other varieties of English as a comparison is helpful to identify the distinctive SSLE features. The use of PMs overall resembles patterns observed in AmE, BrE and IE to varying degrees. Both AmE and BrE report that *you know* is a highly frequent PM in these varieties (Beeching, 2016; Östman, 1981), as in SSLE. There are contrasting frequency reports for some other PMs. Research report that BrE prefers *sort of* over *kind of* but the opposite is true for SSLE. The frequency of functions too might differ from other varieties. For example, a common function of *like* in SSLE is as a focuser. No occurrence of *like* as a focuser is reported in the data in a comprehensive study of BrE (Beeching, 2016). The frequency pattern in terms of the use of *like* per thousand words in SSLE is more similar to Indian English than British, New Zealand, Canadian and Philippines Englishes. As stated in chapter 2, SSLE speaker attitude surveys

show SSLE speakers prefer to align SSLE as a variety akin to its input variety which is BrE. Therefore, this resemblance to IE with regard to the use of *like* might be a surprise to SSLE speakers. It also may indicate that the actual use could be different to the perceptions of the speaker.

8.3 The impact of gender on PMs: key findings

In addition to the functional analysis, the investigation into the interaction of PMs with gender provides some valuable insights into the differences between women's and men's language. The most preferred PM for both the genders is *you know*, in line with the overall corpus finding. For the rest of the PMs in the corpus, there are differences of preference according to gender. The results were considerably affected by the 24 academics in the corpus. If their results are removed, the study would show opposite results which would be that men use more PMs than women. However, since the women in the academics group produced twice as many PMs as males, the results were reversed.

Overall, the current study shows contrasting results to some of the previous research which shows that women use more *you know* than men (Macaulay, 2002; Beeching 2016). Nevertheless, this study agreed with the results of one previous study (Erman, 1992), where men used *you know* more frequently than women in British English. Erman's study had observed 12 face to face conversations which contained 65000 words from equal number of female and male speakers in the London Lund Corpus. The genderlectal analysis of the results showed that both genders use *you know* mostly for clarification and appeal to common knowledge. This result was expected after the analysis of the overall functions. It seems that clarifying ideas and establishing a sense of common knowledge with the interlocutor is important for both genders. This contradicts the traditional belief that seeking cooperation is a sign of interpersonal powerlessness especially seen in women, as both genders use *you know* frequently for the same function. *You know* is also used fairly equally used by both genders to introduce a quotation. The quotation is either a direct verbal quote or sharing a mere thought. *You know* is least used by both groups as a filler. A genderlectal variation is that males

use *you know* more than women to get attention by launching a new piece of information. *Like* is used as a focuser which is slightly similar to getting attention as used in *you know*. However, *like* as a focuser only focuses on either some new information or the most important information in an utterance. Women and men both use *like* as a focuser. However, women have produced 78% and men have produced 22% of the total occurrences of *like* as a focuser. This might indicate that men and women in SSLE utilize different PMs to attract the attention of the listener. Therefore, a gender-based functional use is evident in this data in addition to the frequency association.

The genders behave fairly symmetrically in the use of certain PMs. A case in point is *I mean*. There is only a 4% difference between female and male use of *I mean*. Females have used a marginally higher number of *I mean* than men. Overall, women and men use *I mean* in a similar manner. In contrast, there is a noticeable gender influence on the use of *like*. *Like* is used considerably more by females than by males. Both males and females use *like* as a focuser the most. Men use *like* as an exemplifier more than women. Beeching (2016) interprets that *like* as an exemplifier is a face-saving device. Therefore, we can conclude that men use *like* for face saving more than women by referring to an example preceded by *like* to validate a point mentioned earlier in the conversation. It can be assumed that face saving is more important to Sri Lankan SSLE male speakers than to women although the reasons for such behaviour cannot be predicted by this data. Similar to BE and AmE, *like* is more often used by females to introduce a quote in this data set as well.

Females use *well* less frequently than males in the data. Men take turns and politely interrupt conversations using *well* more than women. This is consistent with previous studies which show that men interrupt a conversation more than women (West & Zimmerman, 1983). *Well* serves as another strategy to facilitate interruptions. Females use *well* more than men to preface a dispreferred response. Similar to using *like* for exemplifications by men, it could be assumed that women save face by using *well* to front a dispreferred response. It can be proposed again that men and women do display genderlectal variation with regard to PM use. While only women use *well* to introduce a quote, only men use *well* for hesitation.

As a total, females use *right* more frequently than males in this data. Yet, when analysing the data in terms of function, this pattern for *right* is true only for the function of explanations, justifications and conclusions. In all the other functions, males use *right* more than the females. For instance, males use *right* more than females to refer to shared knowledge presumed by the speaker and as progression check question tags. Therefore, it could be concluded that *right* is a PM associated with male speech than female speech.

Since an opposite result to the rest of the data was noticed in the sub-corpus of academics, they were analysed separately. Female academics produced 70.1% and male counterparts produced 29.9% PMs in this group. In contrast to the other two professions, female academics led males in the frequency of use for every PM except *well*, *you see/see*, and the nativized PMs. Further, females have a broader spectrum of functions for *you know* than males. Female academics use *you know* as a strategy for conveying politeness. The analysis concluded that SSLE speaking academics use *you know* as an interactional and facilitative tool rather than to show hesitation or insecurity. The results also indicate that *you know* is used by women to mitigate misunderstanding between expert and a lay person. Therefore, the participants in this study suggest that *you know* is not used to display weak or powerless feminine traits as claimed by some early research on *you know*, but rather to convey cordial, thoughtful and polite communication.

The results for the PM *right* revealed that female academics use *right* more frequently than males. This indicates that female academics are much more concerned than males to ensure that the listener or the interlocuter is following the ideas expressed by the listener closely. The female academics use more hedging devices such as *I mean* and *kind of* than male academics. This is not a surprise given that academic language is characterized by hedging according to previous research. Therefore, it can be assumed that professional language has influenced the personal language of this cohort of speakers.

Out of the 11 different types of PMs in this data, 5 PMs are used for the function of emphatic lexis among other functions. These PMs are *aah*, *no*, *right*, *well* and *you see/see*. Males use four of

these five PMs more than females as emphatic lexis which indicate that males use PMs as a language feature for emphasis more than females; the exception is *well*. Overall, the analysis indicate that males use various PMs for emphasis, to draw attention and to check progress than females. Females use PMs to introduce quotations or thoughts, show politeness and ensure that the interlocutor is following the conversation closely.

The genderlectal analysis shows that the nativized PMs are used by males more than the females. Generally, the nativized PMs or any nativized feature is avoided by SSLE speakers in formal to semi-formal settings. Speakers sometimes hesitate to use the nativized features because they may think it is not appropriate in a formal setting as it may be perceived as sub-standard. However, the few instances that are discovered in the data have been used by males mostly. It implies that men promote the use of nativized forms than women. In sum, the genderlectal variation noticed in the use of PMs in the current study confirms that gender has an impact on some PM use but not all. In this corpus, women tend to adopt a more inclusive polite and supportive manner of communication using PMs while men tend to save face, gain more opportunities to talk, emphasise ideas and draw attention to points in the discourse with PMs.

8.4 The impact of age on PMs: key findings

This section outlines the conclusions derived from the analysis on how the age-related variation is reflected in the data. It is assumed that each generation reflects the language as it existed when that generation acquired language (Murphy, 2010). Age-related differences in language can indicate how language evolves with time. As this study is based on limited data from a relatively small number of speakers, it cannot provide a comprehensive analysis of age-based PM development or impact of age on PM use. However, conclusions drawn from the age-based analysis will invariably reveal new information with regard to language change in SSLE, as well as the impact of age on language use. It also will provide a basic blueprint for anyone interested in pursuing larger and further research into the relationship between age and PMs.

The data was divided into three social generations for comparative analysis. They are Baby Boomers (1946-1964), Generation X (1965-1980), and Generation Y (1981-1996). Gen X uses the highest total number of PMs followed by Gen Y and BB respectively. As a percentage their use represents 38% (Gen X), 34% (Gen Y) and 28% (BB) of the total of all PMs. The study drew conclusions about the notable patterns of use for each PM in each generation. A simple frequency analysis showed that the PM with the highest frequency of use for all three social generations is *you know*. Although there is a high use of *you know* for all three generations, there is a marginally downward trend in the data which suggests that *you know* seems to decrease with younger speaker groups. This is the trend that has been observed in other varieties such as BrE as well (Erman, 2001).

The frequency analysis suggests that some of the PMs are characteristic of certain generations. For instance, *right* is used 68% of the total occurrences by those born between 1965-1980 (Gen X). *Right*, therefore, can be considered a generation specific PM in the present data. *Like* is mostly used by Gen Y. *I mean* is used almost equally among all the generations, but it is most frequent for Gen Y. *Yeah* is used 27% by BB, 28% by Gen X and 45% by Gen Y. This steep increase among Gen Y suggests that it is a feature of the youngest generation of the SSLE speakers.

The data also showed the trajectory of the PMs. The use of *you know* shows a decline among youngest generation of the speakers, while *I mean* has a steady trajectory through all generations. *Like* has increased in use with the generations. *Well* clearly shows a declining trend as generations progress. The frequency of *yeah* as a PM has steadily risen through the generations. The nativized *no* and *aah* have dwindled in their use among Gen X and Gen Y compared to BB in these formal to semi-formal interviews, though the very low frequency in any group made it difficult to draw clear conclusions here.

Chapter 7 also analysed the functions of each PM according to age. The most common function unsurprisingly shared by all three generations is *you know* used for clarification and appeal to common knowledge. All three generations also show a high frequency of *like* as a focuser. An analysis of *you know* for all three generations showed that there are functional differences between

groups. For example, Gen X shows the highest share of frequency in the use of *you know* in six functions compared to the other two generations. It is also the only generation that uses *you know* for hesitation, and as Section 5.4.1 pointed out, the PM use for hesitation is not a typical feature of SSLE speakers. A close examination of data seems to indicate that a broad age gap between two speakers may be part of the reason for fewer instances of direct appeal to shared knowledge. The age gap between the speakers acts as a barrier for shared experiences. Therefore, conversations between such conversation partners do not give rise to opportunities to use *you know* for this function.

Like is most frequently used among Gen Y. All the functions of *like* have increased five-fold from BB to Gen Y denoting that *like* is a prolific PM. This high frequency suggests that *like* is a competitor to other PMs. For example, *like* has possibly replaced *I mean* and *you know* in instances where these PMs deal with attention. Both *I mean* and *you know* have reduced in terms of their functions and frequency as generations progress while *like* has continued to grow. The examples from the data also showed that speakers used *like* in instances they can either use *I mean* or *you know*. There was also a drastic increase in the quotative (*be*) *like* function, which is also a feature in BrE and AmE.

All three generations used *I mean* mainly for clarification, exemplification, elaboration and reformulation. *I mean* used for justification moderately increased with each generation. Similar to *you know*, the use of *well* and *you see* decreased over the generations. This indicated that *well* and *you see* are also features of the speech of an older generation in this data set. The analysis of *kind of* and *sort of* showed that SSLE speakers use *kind of* more often than *sort of*. This is in contrast to BrE, which favours *sort of* as opposed to *kind of*. *Yeah/so yeah/yes* considerably increased with each social generation. Its functions seemed to be generation specific. For example, it occurred for explanation, justification, and conclusion only among the BB cohort and for topic shift only among Gen X, though there were only a small number of tokens of each use. Use of *yeah/so yeah/yes* to

mark the end of a turn almost doubled in use with each generation and Gen Y mainly used this PM to mark the end of a turn.

Previous literature showed the nativized PM *no* to have two functions: emphasis and a question tag. The current study reports an additional function, which is marking turn taking. The use of *no* decreased as the generations progress. This pattern in this data set suggested that *no* is a feature Gen X and Gen Y avoid. The nativized PM *aah* is only used by the speakers in the BB cohort implying it is a feature confined to that generation.

The findings from the gender and age-based analysis adds new information to SSLE research about the profile of the SSLE speaker. It has revealed empirical information with regard to characteristics of male and female SSLE speakers as well as language characteristics of older and younger speakers. Further, this study has revealed new information with regard to previously unresearched functions of nativized PMs. The study contributes to the larger body of research on PM variation, demonstrating that SSLE can be recognized as a variety that shows genderlectal variation in the use of PMs similar to other varieties such as IE and Singaporean English in the South Asian region and BrE, AmE and Canadian English in the inner circle varieties. Similar to *eh* in Canadian English, SSLE displays its own nativized PMs. This indicates that SSLE is separating from the inner circle variety norms and this is a trend witnessed in world Englishes broadly.

8.5 Implications with regard to ESL teaching in Sri Lanka

As explained in Section 2.3 and Section 4.3, SSLE is the norm providing dialect for ESL learners in Sri Lanka. Therefore, research that investigates SSLE can be valuable for ESL practitioners. As mentioned in those sections, SSLE is shrouded in many misconceptions and the teachers do not always know what features entail SSLE. The learners as well as the practitioners should be made aware of SSLE usages. In this light my findings could be used in pedagogy. This study is based on a corpus of contemporary SSLE speakers. Thus, it is highly valuable in terms of its relevance to the linguistic situation in Sri Lanka, and specifically to the use of English in current times. Corpus based material is now considered the most appropriate teaching input (Biber & Reppen, 2002; Reppen,

2011; Szudarski, 2022). Therefore, this research can inform ESL lesson material preparation.

Moreover, this study explores an area of SSLE that is least known about: the PMs. Therefore, it can encourage ESL teachers and students to think in new ways about the features of SSLE in natural speech.

8.6 Limitations

The creation of an ad-hoc corpus for this study is both a strength and a possible limitation of the research. Although its size allowed me to look at the data closely and in depth, it was not always possible to find a substantial number of examples or instances of certain PMs to compare and observe trends. Similarly, the decision to restrict the data collection to a particular setting, while advantageous in terms of controlling variables and ensuring ease of access, also brought some limitations. PMs are most prevalent in spontaneous informal speech, while this corpus contains formal to semi-formal speech, where the speakers are generally conscious of the language they use. It was surprising to see that 1.4% of total speech of the participants even in this formal to semi-formal setting consisted of PMs. This finding in this unlikely terrain for PMs further proved how much an integral part of speech PMs are. Yet, an informal context would have provided a richer set of data for observation. It should also be noted that the conversations in the corpus come from recorded data. This may have caused the interlocutors to adjust their speech which may compromise the naturalness and spontaneity of the data.

The current study only observed the use of PMs of individual speakers. It did not take into account the impact of the gender of the interlocutor. Only the interviewee's language was included in the analysis. Further interactions of variables, such as the whether the choice of PM depends on the gender of the interlocutor rather than the speaker's gender, are left to be explored in future studies.

8.7 Future directions

The careful design of the present study has demonstrated that even a bespoke corpus allows for a rich description of the formal and functional range of PMs in SSLE, and their relationship to speaker gender and age. Further research drawing on larger and more varied corpus data can build on this framework and provide a clearer overview of PMs in SSLE across a wider range of interactional contexts.

As noted in Chapter 2, there are many varieties of SLE, including non-standard, Burgher English and learner English. The present research only focused on SSLE: it would be valuable to obtain similar findings for these other varieties of SLE. In particular, one of the findings of this study was that SSLE speakers use PMs least as a filler. SSLE speakers are considered fluent speakers of English in Sri Lanka. Therefore, a study based on speakers considered to be non-fluent may reveal whether PMs help speakers to mask disfluency or help in any other way to establish communication. Expanding the corpus sources used for research would also give greater access to conversations between same gender or mixed gender interlocutors, which in turn would assist with determining whether the gender of the speaker or the gender of the interlocutor is more relevant in selecting PMs in conversations.

We also need more authentic corpora that facilitate SLE research in all directions. There is a wide gap in terms of research corpora in SLE. Since we now have a digital humanities lab at University of Colombo, Sri Lanka which collects and stores researchable corpora on SLE, new corpora can be shared among the SLE researchers. There are thus a great number of important issues to explore with regard to PMs in SLE. Empirical studies in these directions will build on the findings of the present research and increase our knowledge of the pragmatics and discourse of SLE. It is important to note that the field of PMs and the field of world Englishes are continually evolving, and new research always challenges or refines our understanding of these phenomenon. Therefore, the current study plays its part as the stepping stone to propel further research on PMs as used by the speakers of English in Sri Lanka.

References

- Aijmer, K. (1984). 'sort of' and 'kind of' in English conversation. *Studia Linguistica*, 38(2), 118-128.
<https://doi.org/10.1111/j.1467-9582.1984.tb00738.x>
- Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. John Benjamins.
- Aijmer, K. (2013). *Understanding pragmatic markers: A variational pragmatic approach*. Edinburgh University Press.
- Aijmer, K. (2014). Pragmatic markers. In K. Aijmer & C. Rühlemann (Eds.). *Corpus pragmatics: A handbook* (1st ed., pp. 195-218). Cambridge University Press. <https://doi-org.libproxy.ucl.ac.uk/10.1017/CBO9781139057493>
- Aijmer, K. & Simon-Vandenberg, A. (2006). *Pragmatic markers in contrast*. Elsevier.
- Aijmer, K. & Simon-Vandenberg, A. (2009). Pragmatic markers. In J.O. Ostman & J. Verschueren (Eds.), *Handbook of Pragmatics* (1st ed., pp. 1-29). John Benjamins.
- Aijmer, K., & Simon-Vandenberg, A. (2011). Pragmatic markers. In J. Zienkowski, J. O Östman & J. Verschueren (Eds.), *Discursive pragmatics* (1st ed., pp. 223-247). John Benjamins.
- Allwood, J. (1996). Talspraksfrekvenser. Gothenburg Papers in Theoretical Linguistics, 20, University of Gotenborg, Department of Linguistics.
- Al Rousan, R., Al Harahsheh, A., & Huwari, F. (2020). The pragmatic functions of the discourse marker bas in Jordanian spoken Arabic: Evidence from a corpus. *Journal of Educational and Social Research*, 10(1), 130-142. <https://doi.org/10.36941/jesr-2020-0012>
- Amarasuriya, H. (2010). Discrimination and social exclusion of youth in Sri Lanka. In R. Gunatilaka, M. Mayer & M. Vodopivec (Eds.), *The challenges of youth employment in Sri Lanka*. The World Bank: Washington DC.
- Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation*. John Benjamins.
- Anthony, L. (2024). AntConc (Version 4.3.1) [Computer Software]. Tokyo, Japan: Waseda University.
 Available from <https://www.laurenceanthony.net/software>
- Anwar, B., Rasool, S., & Kamran, M. (2020). Indigenization in multilingual setting: Evidence from

- Pakistani written discourse. *Pakistan Social Sciences Review*, 4(1), 66-78.
- Attanapola, C. T., & Lund, R. (2013). Contested identities of indigenous people: Indigenization or integration of the Veddas in Sri Lanka. *Singapore Journal of Tropical Geography*, 34(2), 172–187. <https://doi.org/10.1111/sjtg.12022>
- Bailey, G. (2002). Real and apparent time. In J.K. Chambers, P. Tudgill & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (1st ed., pp.312-332). John Wiley & Sons.
- Baker, P. (2008). 'Eligible' bachelors and 'frustrated' spinsters: Corpus Linguistics, gender and language. In K. Harrington, L. Litosseliti, H. Sauntson & J. Sunderland (Eds.), *Gender and Language Research Methodologies* (pp.73-84). Palgrave Macmillan.
- Barbieri, F. (2009). 'Quotative be like in American English: Ephemeral or here to stay?', *English World-Wide*, 30(1), 68-90.
- Barnes, B. (1995). Discourse particles in French conversation: (eh) ben, bon, and enfin. *The French Review*, 68(5), 813–821.
- Bazzanella, C. (2006). Discourse markers in Italian: Towards a "compositional" meaning. In K. Fischer (Ed.), *Approaches to Discourse Particles* (1st ed., pp. 449-464). Elsevier.
- Beeching (2002). *Gender, politeness & pragmatic particles in French*. John Benjamins.
- Beeching, K. (2012). Semantic change: Evidence from false friends. In P. Lauwers, G. Vanderbauwhede & S. Verleyen (Eds.), *Pragmatic Markers and Pragmaticalization: Lessons from False Friends: Vol. 44*. (1st ed., pp. 11–36). John Benjamins Publishing Company.
- Beeching, K. (2016). *Pragmatic markers in British English: Meaning in social interaction*. Cambridge University Press.
- Beeching, K., Howie, G., Kirjavainen, M., & Piasecki, A. (2022). Discourse-pragmatic markers, fillers and filled pauses: Pragmatic, cognitive, multimodal and sociolinguistic perspectives. *Pragmatics & Cognition*, 29(2), 181–194. <https://doi.org/10.1075/pc.00025.bee>
- Berger, C. (2003). *The myth of gender-specific swearing*. Verlag für Wissenschaft und Forschung.

- Bernaisch, T., Gries, S. T., & Mukherjee, J. (2014). The dative alternation in South Asian English(es): Modelling predictors and predicting prototypes. *English World-Wide*, 35(1), 7–31.
<https://doi.org/10.1075/eww.35.1.02ber>
- Bernaisch, T. (2015). *The lexis and lexicogrammar of Sri Lankan English / Tobias Bernaisch*. (1st ed.). John Benjamins Publishing Company.
- Bernaisch, T. Mendis, D. & Mukherjee, J. (2019). *Manual to the International Corpus of English – Sri Lanka*. Giessen: Justus Liebig University, Department of English.
- Bernaisch, T. (2021). Genderlectal variation in the English-speaking world. In T. Bernaisch (Ed.), *Gender in World Englishes* (1st ed., pp.1-22). Cambridge University Press.
- Bernaisch, T. (2021). Hedges and gender in the inner and expanding circle. In T. Bernaisch (Ed.), *Gender in World Englishes* (1st ed., pp.1-22). Cambridge University Press.
- Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24(2), 199–208. <https://doi.org/10.1017/S0272263102002048>
- Biber, D. & Reppen, R. (2015). *The Cambridge handbook of English Corpus Linguistics*. Cambridge University Press.
- Boberg, C. (2004). Real and apparent time in language change: Late adoption of changes in Montreal English. *American Speech*, 79(3), 250-269. <https://doi.org/10.1215/00031283-79-3-250>
- Bogdanova-Beglarian, N., Blinova, O., Sherstinova, T., Gorbunova, D., Zaides, K., & Popova, T. (2020). Pragmatic Markers in Dialogue and Monologue: Difficulties of Identification and Typical Formation Models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 12335, pp. 68–78). Springer International Publishing AG. https://doi.org/10.1007/978-3-030-60276-5_7
- Bolly, C., Crible, L., Degand, L. & Uygur-Distexhe, D. (2017). Towards a model for discourse marker annotation: from potential to feature-based discourse markers. In C. Fedriani & A. Sansò (Eds.), *Pragmatic markers, discourse markers and modal particles: New perspectives* (1st ed., pp. 71-98). John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.186>

- Botha, W. (2018). A social network approach to particles in Singapore English, *World Englishes*, 37(2), 261-181.
- Brinton, L. (1996). *Pragmatic markers in English: grammaticalization and discourse functions: Vol. 19. Topics in English Linguistics*. Mouton de Gruyter.
- Brown, P. and Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Brunner, T. (2014). Structural nativization, typology and complexity: noun phrase structures in British, Kenyan and Singaporean English. *English Language and Linguistics*, 18(1), 23-48.
- Buchstaller, I. (2002). *He goes and I'm like: The new quotatives re-visited*. University of Edinburgh Postgraduate conference Proceedings. University of Edinburgh.
- Buchstaller, I. (2008). The localization of global linguistic variants. *English World-Wide: A Journal of Varieties of English*, 29(1), 15–44. <https://doi.org/10.1075/eww.29.1.03buc>
- Cameron, D., (1990) Demythologizing Sociolinguistics. Reprinted in: N. Coupland, A. Jaworski (Eds.), 1997. *Sociolinguistics: A Reader and Coursebook*. (pp. 55–67) MacMillan.
- Caprario, M., Taguchi, N., & Reppen, R. (2022). Corpus-informed instruction of the pragmatic marker I mean. *Language Learning Journal*, 50(4), 460–474. <https://doi.org/10.1080/09571736.2022.2088439>
- Caron-Prague, J. & Caron, J. (1991). Psychopragmatics vs. sociopragmatics. The function of pragmatic markers in thinking-aloud protocols. In J. Verschueren (Ed.), *Pragmatics at issue: Selected papers of the International Pragmatics Conference 1987*. (pp. 29-36) John Benjamins.
- Carter, R. & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide to spoken and written grammar and usage*. Cambridge University Press.
- Cencini, M. (2002). Transcribing and encoding spoken data. *OCLC Systems & Services*, 18(4), 205-211.
- Chambers, J.K. (2009). *Sociolinguistic theory*. Wiley-Blackwell.
- Chandralal, D. (2010). *Sinhala* (1st ed.). John Benjamins Pub. Co.
- Chen, Y. & Adolphs, S. (2023). Towards a speech-gesture profile of pragmatic markers: The case of

- “you know”. *Journal of Pragmatics*, 210, 36-51.
- Cheshire J. (2002). Sex and gender in variation research. In J.K. Chambers, P. Tudgill and N. Schilling (Eds.), *The handbook of language variation and change* (1st ed., pp. 423- 443). John Wiley & Sons.
- Coates, J. (1989). Gossip revisited: language in all-female groups. In J. Coates & D. Cameron (Eds.), *Women in their Speech Communities* (1st ed., pp. 94-121). Longman.
- Coates, J. (2014). *Women, men and language*. Pearson Longman.
- Coates, J. (2016). *Women, men and language: A sociolinguistic account of gender differences in language* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315645612>
- Colburn, R. (2017, April 6). *Here’s everything you need to know about the 7 living American generations*. Avclub. <https://www.avclub.com/here-s-everything-you-need-to-know-about-the-7-living-a-1798260040>
- Columbus, G. (2009). A corpus-based analysis of invariant tags in five varieties of English. *Corpus Linguistics*, 69, 401–414. https://doi.org/10.1163/9789042025981_022
- 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.
- Crible, L. (2018). *Discourse markers and (dis)fluency*. John Benjamins.
- Crible, L., & Degand, L. (2021). Co-occurrence and ordering of discourse markers in sequences: A multifactorial study in spoken French. *Journal of Pragmatics*, 177, 18–28. <https://doi.org/10.1016/j.pragma.2021.02.006>
- Crowdy, S. (1995). The BNC spoken corpus. In G. Leech, G. Myers & J. Thomas (Eds.), *Spoken English on Computer: Transcription, Mark-Up and Annotation* (pp. 224–234). Longman.
- D’Arcy, A. (2007). ‘Like’ and language ideology: Disentangling fact from fiction. *American Speech*, 82, 386-419.

- D'Arcy, A. (2012). The diachrony of quotation: evidence from New Zealand English. *Language Variation and Change*, 14(1), 60-88.
- D'Arcy, A. (2017). *Discourse-pragmatic variation in context: Eight hundred years of LIKE*. (1st ed.). John Benjamins Publishing Company.
- Davidse, K., Brems, L., & De Smedt, L. (2008). Type noun uses in the English NP: A case of right to left layering. *International Journal of Corpus Linguistics*, 13(2), 139-168.
- Davydova, J. (2015). Linguistic change in a multicultural setting. A case study of quotatives in Indian English. In P. Collins (Ed.), *Grammatical change in English world-wide* (pp. 297-334). John Benjamins.
- Davydova, J. (2016). Indian English quotatives in a real-time perspective. In E. Seoane & C. Suárez-Gómez (Eds.), *World Englishes: New theoretical and methodological considerations*. (1st ed. pp. 173- 204). John Benjamins Publishing Company.
- Dehé, N. (2014). *Parentheticals in spoken English: The syntax–prosody relation*. Cambridge University Press.
- Degand, L. & Crible, L. (2021). Discourse markers at the peripheries of syntax, intonation and turns: Towards a cognitive-functional unit of segmentation. In D. Van Olmen & J. Šinkuniene (Eds.), *Pragmatic markers and peripheries* (pp. 19- 48). John Benjamins Publishing Company.
- Degand, L. (2000). Causal connectives or causal prepositions? Discursive constraints. *Journal of Pragmatics*, 32(6), 687–707. [https://doi.org/10.1016/S0378-2166\(99\)00066-1](https://doi.org/10.1016/S0378-2166(99)00066-1)
- Degand, L., Cornillie, B., & Pietrandrea, P. (2013). Modal particles and discourse markers: Two sides of the same coin? In L. Degand, B. Cornillie, P. Pietrandrea (Eds.), *Discourse markers and modal particles: categorization and description*. (1st ed. pp. 1-18). John Benjamins Publishing Company.
- Degenhardt, J. (2023). Requests in Indian and Sri Lankan English. *World Englishes*, 42(3), 523–543. <https://doi.org/10.1111/weng.12573>

- Degenhardt, J. (2024). Parentheticals in spoken Indian and Sri Lankan English. *World Englishes*, 43(3), 1-18. <https://doi.org/10.1111/weng.12696>
- Degenhardt, J., & Bernaisch, T. (2022). Apologies in South Asian varieties of English: A corpus-based study on Indian and Sri Lankan English. *Corpus Pragmatics: International Journal of Corpus Linguistics and Pragmatics*, 6(3), 201–223. <https://doi.org/10.1007/s41701-022-00117-8>
- Denis, D. (2015). *The development of pragmatic markers in Canadian English*. (Publication No. 1762556742) [Doctoral dissertation, University of Toronto]. ProQuest Dissertations & Theses Global.
- Denis, D. & Tagliamonte, S. (2016). Innovation, right? Change, you know? Utterance-final tags in Canadian English. In Pichler (Ed.), *Discourse-pragmatic variation and change in English* (1st ed., pp. 86-112). Cambridge University Press.
<https://doi.org/10.1017/CB09781107295476.005>
- Denke, A. (2009) Native-like performance. Pragmatic markers, repair and repetition in native and non-native English speech. VDM Verlag Dr. Müller.
- de Klerk, V. (2005). Procedural meanings of well in a corpus of Xhosa English. *Journal of Pragmatics*, 37(8), 1183–1205. <https://doi.org/10.1016/j.pragma.2004.11.001>
- De Silva, M., Fonseka N. & Mihlar, F., *The forgotten victims of war: A border villages study*. Neelan Thiruchilvam Trust. <https://srilankabrief.org/wp-content/uploads/2019/10/The-Forgotten-Victims-of-War-A-Border-Village-Study-FINAL.pdf>
- De Silva, P., & Punchihewa, A. (2011). *Socio-anthropological research project on Vedda community in Sri Lanka*. Department of Sociology, University of Colombo. (https://www.researchgate.net/publication/235335142Socio_Anthropological_Research_Projecton_Vedda_Community_in_Sri_Lanka).

Department of Census and Statistics, Sri Lanka (2012). Census of population and housing.

http://www.statistics.gov.lk/PopHouSat/CPH2011/Pages/Activities/Reports/CPH_2012

Dér, C. I., Marko, A. (2010). A pilot study of Hungarian discourse markers. *Language and Speech*, 53(2), 135-180.

Diskin, C. (2017). The use of the discourse-pragmatic marker 'like' by native and non-native speakers of English in Ireland. *Journal of Pragmatics*, 120, 144–157.

<https://doi.org/10.1016/j.pragma.2017.08.004>

Diskin-Holdaway, C. (2022). You know in L1 and L2 English. In E. Peterson, T. Hiltunen & J. Kern (Eds.), *Discourse-Pragmatic Variation and Change* (1st ed., pp. 193–211).

<https://doi.org/10.1017/9781108864183.01>

Drummond, K. & Hopper, R. (1993). Some uses of yeah. *Research on Language and Social Interaction*, 26, 203–212.

Edmondson, W. (1981) *Spoken discourse: A model for analysis*. Longman.

Egbert, J., Biber, D., & Gray, B. (2022). *Designing and evaluating language corpora: A practical framework for corpus representativeness*. Cambridge University Press.

Ekanayaka, T. N. (2020). Sri Lankan English. In K. Bolton, W. Botha and A. Kirkpatrick (Eds.), *The handbook of Asian Englishes* (1st ed., pp. 337-353). John Wiley & Sons.

Eckert, P. (2017). Age as a sociolinguistic variable. In F. Coulmas (Ed.), *The handbook of Sociolinguistics* (pp. 151–167). Blackwell Publishing Ltd.

<https://doi.org/10.1002/9781405166256>

Eckert, P. (2018). *Meaning and linguistic variation: The third wave in Sociolinguistics*. Cambridge University Press. <https://doi.org/10.1017/9781316403242.004>

Eckert, P. & McConnell-Ginet, S. (2013). *Language and gender*. Cambridge University Press.

- Erman, B. (1992). Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change*, 4(2), 217-234.
- Erman, B. (2001). Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics*, 33(9), 1337–1359. [https://doi.org/10.1016/S0378-2166\(00\)00066-7](https://doi.org/10.1016/S0378-2166(00)00066-7)
- Farahani, M. V. & Ghane, Z. (2022). Unpacking the function(s) of discourse markers in academic spoken English: a corpus-based study. *The Australian Journal of Language and Literacy*, 45, 49-70.
- Fedriani, C., & Sansò, A. (2017). Pragmatic markers, discourse markers and modal particles: What do we know and where do we go from here? In C. Fedriani & A. Sansò (Eds.), *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*. (1st ed., pp. 2-33). John Benjamins Publishing Company.
- Fernando, C. (1976). English and Sinhala bilingualism in Sri Lanka. *Language in Society*, 6, 341-360.
- Fernando, C. (1996). The post-imperial status of English in Sri Lanka 1940–1990: From first to second language. In J. Fishman, A. Conrad & A. Rubal-Lopez (Ed.), *Post-Imperial English: Status Change in Former British and American Colonies, 1940-1990* (pp. 485-512). De Gruyter Mouton. <https://doi-org.libproxy.ucl.ac.uk/10.1515/9783110872187.485>
- Fernando, S. (2010). Does Sri Lankan English (SLE) have dialects, and can one dialect be identified as Standard SLE? In S. Fernando, M. Guenesekera, A. Parakrama (Eds.), *English in Sri Lanka: Ceylon English, Lankan English, Sri Lankan English*. (1st ed., pp. 296-305). Sri Lanka English Language Teachers' Association
- Fernando, D. (2014, October 3-5). *Do 'hall' and 'holl' still matter?: Perceptions of Sri Lankan English pronunciation* [Conference session]. 8th International Sri Lankan English Language Teachers' Association conference, Colombo, Sri Lanka.
- Fernando, D. (2015) Words from the 'margins': Exploring Sri Lankan English borrowings in the

- classroom. *Kalyani: Journal of the University of Kelaniya* XXX, pp 39-62.
- Fernando, S. (1985). Changes in Sri Lankan English as reflected in phonology. *University of Colombo Review*, 5, 41-53.
- Fernando, D. (2007). 'Good' English or Sri Lankan English: a study of English teachers' awareness of their own variety. [Unpublished MA thesis]. University of Reading.
- Fernando, S. (2008). When is a 'hall' a 'hole'? Issues and guidelines in Sri Lankan English pronunciation. In D. Fernando, & D. Mendis (Eds.), *English for equality, employment and empowerment: Proceedings of the 4th international conference of the Sri Lanka English Language Teachers' Association*, (pp. 71-81). SLELTA.
- Ferrara, K. & Bell, B. (1995). Sociolinguistic Variation and Discourse Function of Constructed Dialogue Introducers: The Case of Be + like. *American Speech*, 70 (3), 265-290.
- Fischer, K. (2006). Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In K. Fischer (Ed.), *Approaches to discourse particles*. (1st ed., pp.1-20). https://doi-org.libproxy.ucl.ac.uk/10.1163/9780080461588_002
- Fishman, P. M. (1978). Interaction: The work women do. *Social Problems*, 25(4), 397-406. <https://doi.org/10.2307/800492>
- Fox Tree, J. E. & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics*, 34, 727-747.
- Fox Tree, J. E. (2007). Folk notions of um and uh, you know, and like. *Text & Talk*, 27(3). 297-314.
- Fraser, B. (2006). On the universality of discourse markers. In K. Aijmer & A.M. Simon-Vendenbergen (Eds.). *Pragmatic Markers in Contrast: Vol 2* (pp. 73-92). Elsevier.
- Fraser, B. (2010). Pragmatic competence: The case of hedging. *Studies in Pragmatics*, 9, 15-34.

- Fraser, B. (2013). Combinations of contrastive discourse markers in English. *International Review of Pragmatics*, 5, 318–340.
- Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics*, 35, 23–45.
- Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics*, 28(3), 410–439.
<https://doi.org/10.1093/applin/amm030>
- Fung, K. T. M., Ting, S. H., & Chuah, K. M. (2023). Detecting female and male language features in Facebook comments by Malaysian millennial users. *Journal of the Southeast Asian Linguistics Society*, 16(1), 37–55.
- Funke, N. (2022). Pragmatic nativisation of thanking in South Asian Englishes. *World Englishes*, 41(2), 136–150. <https://doi.org/10.1111/weng.12517>
- Funke, N., & Bernaisch, T. (2022). Intensifying and downtoning in South Asian Englishes: Empirical perspectives. *English World-Wide*, 43(1), 33–65. <https://doi.org/10.1075/eww.21064.fun>
- Furkó, P. (2017). Manipulative uses of pragmatic markers in political discourse. *Palgrave Communications*, 3(1). <https://doi.org/10.1057/palcomms.2017.54>
- Gargesh, R. (2019). South Asian Englishes. In C. L. Nelson, Z. G. Proshina & D. R. Davis (Eds.), *The handbook of World Englishes* (pp. 105–134). John Wiley & Sons, Inc.
<https://doi.org/10.1002/9781119147282.ch7>
- Gleason, H. A. (1961). *An introduction to descriptive linguistics* (Rev.ed.). Holt, Rinehart and Winston.
- Gonzalez, M. (2004). *Pragmatic Markers in Oral Narrative: The case of English and Catalan* (1st ed.). John Benjamins Publishing Company.

- Goonetilleke, D. C. R. A. (2003). A paradigm for ex-colonies of Britain. In C. Mair (Ed.), *The politics of English as a world language: new horizons in postcolonial cultural studies* (pp. 337-358). Rodopi.
- Granger, S., Dupoint, M., Meunier, F., Naets, H. & Paquot, M. (2020). The International Corpus of Learner English. Version 3. Louvain-la-Neuve: Presses universitaires de Louvain.
<http://hdl.handle.net/2078.1/229877>
- Gries, S., Heller, B. & Funke, N. (2020). The role of gender in postcolonial syntactic choice making: evidence from the genitive alternation in British and Sri Lankan English. In T. Bernaisch (Ed.), *Gender in World Englishes* (pp. 121-146). Cambridge University Press.
- Gunasekera, M. (1989). *Discourse genres in English newspapers of Singapore, South India and Sri Lanka*. [Unpublished doctoral dissertation], University of Michigan, Ann Arbor.
- Gunasekera, M. (2005). *Postcolonial Identity of Sri Lankan English*. Colombo: Katha Publishes.
- Gut, U. & Fuchs, R. (2017). Exploring speaker fluency with phonologically annotated ICE corpora. *World Englishes*, 36(3), 387–403. <https://doi.org/10.1111/weng.12278>
- Haselow, A. (2019). Discourse Marker Sequences: Insights into the Serial Order of Communicative Tasks in Real-Time Turn Production. *Journal of Pragmatics*, 146, 1-18.
- Hale, S. (1999). Interpreters' treatment of discourse markers in courtroom questions. *Forensic Linguistics*, 6, 57–82.
- Han, D. (2011). Utterance production and interpretation: A discourse-pragmatic study on pragmatic markers in English public speeches. *Journal of Pragmatics*, 43(11), 2776–2794.
<https://doi.org/10.1016/j.pragma.2011.04.008>
- Hancil, S. (2021). The final particle like in Northern English. In A. Haselow & S. Hancil (Eds.), *Studies at the grammar-discourse Interface discourse markers and discourse-related grammatical phenomena* (pp.229-244). John Benjamins.

- Hansen, M. M. (1998). *The Function of Discourse Particles: A Study with Special Reference to Spoken Standard French*. John Benjamins.
- Hansen, B. (2021). Localisation, Globalisation and Gender in Discourse- Pragmatic Variation in Ghanaian English. In T. Bernaisch (Ed.), *Gender in World Englishes* (pp. 23-46). Cambridge University Press. [Doi: 10.1017/9781108696739.004](https://doi.org/10.1017/9781108696739.004)
- Haselow, A. (2019). Discourse marker sequences: Insights into the serial order of communicative tasks in real-time turn production. *Journal of Pragmatics*, 146(1), 1–18.
<https://doi.org/10.1016/j.pragma.2019.04.003>
- Heine, B., Kaltenböck, G., Kuteva, T., & Long, H. (2021). English Discourse Markers. In *The Rise of Discourse Markers* (pp. 91–143). Cambridge University Press.
- Hepburn, A., & Potter, J. (2011). Recipients designed: Tag questions and gender. In S.A. Speer & E. Stokoe (eds.) *Conversation and Gender* (pp. 135–152). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511781032.008>
- Herat, M. (2006). Substitute one in Sri Lankan English. *Leeds Working Papers in Linguistics and Phonetics*, 11, 65–77.
- Herbert, R. K. (1990). Sex-based differences in compliment behavior. *Language in Society*, 19(2), 201–224.
- Höhn, N. (2012). And they were all like “What’s going on?”: new quotatives in Jamaican and Irish English’. In M. Hundt and U. Gut (Eds.), *Mapping Unity and Diversity World-Wide: Corpus-Based Studies of New Englishes* (pp. 263-290). John Benjamins.
- Holmes, J. (1984). Hedging your bets and sitting on the fence: some evidence for hedges as support structures, *Te Reo*, 27, 47-62.
- Holmes, J. (1986). Functions of you know in women’s and men’s speech. *Language in Society*, 15(1), 1-21.

- Holmes, J. (1987). Hedging, fencing and other conversational gambits: an analysis of gender differences in New Zealand speech. In A. Pauwels (Ed.), *Women and Language in Australian and New Zealand Society* (pp. 59-79). Australian Professional Publications.
- Holmes, J. (1995). *Women, men and politeness*. Longman.
- Holmes, J. (1997). Women, language and identity. *Journal of Sociolinguistics*, 1(2). 195-223.
- Holmes, J. (1998). Complimenting: a positive politeness strategy. In J. Coates (Ed.), *Language and Gender: A Reader* (pp. 100–120). Blackwell.
- Huang, L. (2019). A Corpus-Based Exploration of the Discourse Marker Well in Spoken Interlanguage. *Language and Speech*, 62(3), 570–593.
<https://doi.org/10.1177/0023830918798863>
- Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text*, 18 (3), 349-382.
- Hyland, K. (2002). Hedges, boosters and lexical invisibility. *Language Awareness*, 9, 179-197.
- Ilangakoon, S. R., Wijeratne, W. M., Senaratne, C. D., (2021). Cultural and core borrowings reclassified: A corpus-based study of Sri Lankan English vocabulary, *International Journal of Multidisciplinary Studies*, 8(III), 151-165.
- Jain, M. (1974). Error analysis: source, cause and significance. In Richards, J. (Ed.). *Error analysis: Perspectives on second language acquisition* (pp.189-215). Longman.
- Jansz, M. (2019). *The problem of being a native: An exploration of Sri Lankan first language English speaker identities*. [Unpublished M. Phil thesis]. University of Kelaniya, Sri Lanka.
- Johnson, E. (1993). Gender Differences in spoken American English. *Revista Signos*, 26(33), 103.
<https://www.proquest.com/scholarly-journals/gender-differences-spoken-american-english/docview/1347839947/se-2>
- Jucker, A. H. (2018). Data in Pragmatic Research. In A. H. Jucker, K. P. Schneider, W. Bublitz (Eds.), *Methods in Pragmatics* (1st ed., pp.3-36). De Gruyter Mouton.

- Jucker, A. H. (2021). Features of orality in the language of fiction: A corpus-based investigation. *Language and Literature (Harlow, England)*, 30(4), 341–360.
<https://doi.org/10.1177/09639470211047751>
- Jucker, & Ziv, Y. (1998). And people just you know like “wow” discourse markers as negotiating strategies. In A. H. Jucker & Z. Yael (Eds.), *Discourse markers: Descriptions and theory* (1st ed., pp. 171-201). John Benjamins Publishing Company.
- Jay, T., (1992). *Cursing in America: A psycholinguistic study of dirty language in the courts, in the movies, in the schoolyards and on the streets*. John Benjamn.
- Jayawardena, K. (1986). *Feminism and Nationalism in the Third World*. Zed Books Ltd: London.
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world* (pp. 11–34). Cambridge University Press.
- Kachru, B. B. (1992). World Englishes: approaches, issues and resources. *Language Teaching*, 25(1), 1–14. <https://doi.org/10.1017/S0261444800006583>
- Kallen, J. L., & Kirk, J. M., (2012). SPICE-Ireland: A User’s Guide.
- Kandiah, T. (1979). Disinherited Englishes, the case of Lankan English. *Navasilu*, 3, 75-89.
- Kandiah, T. (1981). Lankan English Schizoglossia. *English World-wide*, 2(1), 63-81.
- Karunakaran, T. (2011). English borrowings in Jaffna Tamil. *Language in India*, 11(11), 512–528.
- Kerswill, P. & Trudgill, P. (2009). The birth of new dialects. In P. Auer, F. Hinskens & P. Kerswill (Eds.), *Dialect change* (pp. 196-220). Cambridge University Press.
- Kithulgodu, E. & Mendis, D. (2020). From analysis to pedagogy: Developing ESP materials for the Welcome Address in Sri Lanka. *English for Specific Purposes*, 60, 140-158
- Kirk, J. M., Kallen, J. L., Lowry, O., Rooney, A. & Mannion, M. (2007). International Corpus of English: Ireland Component. The ICE-Ireland Corpus: Version 1.2. Belfast: Queen’s University Belfast and Dublin: Trinity College Dublin. [version 1.2.1 released December 2009]

- Kirk, J. M. (2018). The Pragmatics of Well as a Discourse Marker in Broadcast Discussions. In S. Hoffmann, A. Sand, S. Arndt-Lappe & L. M. Dillmann (Eds.), *Corpora and Lexis: Vol. 81*. (pp. 140–172). https://doi.org/10.1163/9789004361133_007
- Koczogh, H. V. & Furkó B. P. (2011). Gender differences in the use of the discourse markers ‘you’ know and ‘I mean’. *Argumentum*, 7, 1-18.
- Koch, T., Romero, P., Stachl, C. (2022). Age and gender in language, emoji, and emoticon usage in instant messages, *Computers in Human Behavior*, 126, 1-12.
- Kohn, K. (2012). Pedagogic corpora for content and language integrated learning: Insights from the BACKBONE project. *The Eurocall Review*, 20(2), 1-20.
- Kraaz, M., & Bernaisch, T. (2020). Backchannels and the pragmatics of South Asian Englishes. *World Englishes*, 41(2), 224–243. <https://doi.org/10.1111/weng.12522>
- Kubota, R., (2015). Inequalities of Englishes, English speakers, and languages: A critical perspective on pluralistic approaches to English. In R. Tupas (Ed.), *Unequal Englishes: the politics of Englishes today* (pp. 1-19). Palgrave Macmillan.
- Labov, W. (1966). *The social stratification of English in New York City*. Center for Applied Linguistics.
- Labov, W. (1994). Principles of linguistic change: Vol. 1: Internal Factors. Blackwell.
- Labov, W. (2001). Principles of Linguistic Change: Vol. 2: Social Factors. Blackwell.
- Lamiroy, B. (1994). Pragmatic connectives and L2 acquisition: The case of French and Dutch. *Pragmatics*, 4, 183–201.
- Lange, C. (2009). “Where’s the party jaar!”: Discourse particles in Indian English. In T. Hoffmann & L. Siebers (Eds.), *World Englishes: Problems, Properties and Prospects*. (1st ed., pp. 207–225). John Benjamins Publishing Company.
- Lange, C. & Leuckert, S. (2021). Tag questions and gender in Indian English. In T. Bernaisch (Ed.), *Gender in World Englishes* (pp. 69–93). Cambridge: Cambridge University Press. [Doi: 10.1017/9781108696739.004](https://doi.org/10.1017/9781108696739.004)
- Laserna, C., Seih, Y., & Pennebaker, J. (2014). *Um . . . who like says you know*: Filler word use as a

- function of age, gender, and personality. *Journal of Language and Social Psychology*, 33(3), 328 - 338.
- Lee-Goldman, R. (2011). No as a discourse marker. *Journal of Pragmatics*, 43(10), 2627–2649.
<https://doi.org/10.1016/j.pragma.2011.03.011>
- Leimgruber, J. R. E., Lim, J.J., & Hiramoto, M. (2021). Ethnic and gender variation in the use of colloquial Singapore English discourse particles. *English Language and Linguistics*, 25(3), 601-620.
- Lenk, U. (1998). Discourse markers and global coherence in conversation. *Journal of Pragmatics*, 30(2), 245–257. [https://doi.org/10.1016/S0378-2166\(98\)00027-7](https://doi.org/10.1016/S0378-2166(98)00027-7)
- Levey, S. (2006). The sociolinguistic distribution of discourse marker like in preadolescent speech. *Multilingua*, 25(4), 413–441. <https://doi.org/10.1515/MULTI.2006.022>
- Leuckert, S., & Rüdiger, S. (2021). Discourse markers and world Englishes. *World Englishes*, 40(4), 482–487. <https://doi.org/10.1111/weng.12535>
- Li, L., Lorenz, E., & Siemund, P. (2023). The ages of pragmatic particles in Colloquial Singapore English: A corpus study based on oral history interviews. *English World-Wide*, 44(1), 91–117.
<https://doi.org/10.1075/eww.21016.li>
- Lim, Lisa. (2007). Mergers and Acquisitions: On the Ages and Origins of Singapore English Particles. *World Englishes*, 26(4), 446–473. <https://doi.org/10.1111/j.1467-971X.2007.00522.x>
- Lim, Lisa. (2013). Kaduva of privileged power, instrument of rural empowerment? The politics of English (and Sinhala and Tamil) in Sri Lanka. In L. Wee, R. B. H. Goh, & L. Lim (Eds.), *The politics of English: Vol 4. South Asia, Southeast Asian and the Asia Pacific* (1st ed., pp. 61-80). John Benjamins.
- Lohmann, A. & Koops, C. (2016). Aspects of discourse marker sequencing. In G. Kaltenböck, E. Keizer & A. Lohmann (Eds.), *Outside the clause: Form and function of extra-clausal constituents* (pp. 417-446). John Benjamins.
- Lohmann, A. & Koops, C. (2022). Pragmatic marker combinations: Introduction. *Journal of*

Pragmatics, 196, p. 1-5.

Lorenz, E. (2022). "We use English but not like all the time like"- Discourse marker like in UAE

English. *Frontiers in Communication*, 7. <https://doi.org/10.3389/fcomm.2022.778036>

Loureiro-Porto, L. (2017). ICE vs GloWbE: Big data and corpus compilation. *World Englishes*, 36(3),

448–470. <https://doi.org/10.1111/weng.12281>

Louwerse, M. M., & Mitchell, H. H. (2003). Toward a Taxonomy of a Set of Discourse Markers in

Dialog: A Theoretical and Computational Linguistic Account. *Discourse Processes*, 35(3), 199–

239. https://doi.org/10.1207/S15326950DP3503_1

Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The spoken BNC2014: Designing

and building a spoken Corpus of everyday conversations. *International Journal of Corpus*

Linguistics, 22(3), 319–344. <https://doi.org/10.1075/ijcl.22.3.02lov>

Love, R., Hawtin A., & Hardie, A. (2017). The British National Corpus 2014: User manual and

reference guide (version 1.0). Lancaster: ESRC Centre for Corpus Approaches to Social

Science. <http://corpora.lancs.ac.uk/bnc2014/doc/BNC2014manual.pdf>

Lowenberg, P. H. (1986). Non-native varieties of English: Nativization, norms, and implication. *Studies*

in Second Language Acquisition, 8, 1-18.

Luke, K. K. (1987, August 17- 22). The establishment of common ground in natural conversation: The

case of the utterance particle LA in Cantonese [Paper Presentation]. International Pragmatics

Conference, Antwerp.

Lutzky, U. (2012). *Discourse markers in early modern English* (1st ed.). John Benjamins Publishing

Company.

Macaulay, R. (1977). *Language, Social Class and Education*. Edinburgh University Press.

Macaulay, R. (2002). You know, it depends. *Journal of Pragmatics*, 34 (6). 749-767.

Maschler, Y. (1994). Metalanguaging and discourse markers in bilingual conversations. *Language in*

Society, 23 (3). 325-366

Matthews J. J. (1984). Good and mad women: The historical construction of femininity in twentieth

- century Australia. George Allen & Unwin Australia Pty Ltd.
- McCarthy, M.J. (1998). *Applied linguistics*. Cambridge University Press.
- McElhinny, B. S. (2021). How does water talk, and other hopeful questions about and beyond gender and language. *Gender and Language*, 15(4), 527-536. <https://doi.org/10.1558/genl.21520>
- Mendis, D. (2006). Situating SMS (Short Message Service) discourse. *The Sri Lanka Journal of the Humanities*, XXXII (1 & 2), 125-134.
- Mendis, D. and Rambukwella, H. (2021). Sri Lankan Englishes. In A. Kirkpatrick (Ed.), *The Routledge Handbook of world Englishes* (2nd ed., pp. 178-193). Routledge.
- Mesthrie R. & Bhatt R. M. (2008). *World Englishes: The study of new varieties*. Cambridge University Press.
- Meyler, M. (2007). *A dictionary of Sri Lankan English*. Colombo: Mirisgala.
- Miller, J & Weinert, R. (1995). The function of like in dialogue. *Journal of Pragmatics*, 23(4) 365–93. [doi:10.1016/0378-2166\(94\)00044-F](https://doi.org/10.1016/0378-2166(94)00044-F)
- Miller, J. & Weinert, R. (1998). *Spontaneous spoken language: Syntax and discourse*. Clarendon.
- Miller J. (2009). *Like and other discourse markers* in P. Peters, P. Collins, & A. Smith (Eds.), *Comparative studies in Australian and New Zealand English grammar and beyond* (1st ed. pp. 317-338). John Benjamins.
- Milroy, L. (1980). *Language and social networks*. University Park Press.
- Milroy, L. & Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Basil Blackwell.
- Miskovic-Lukovic, M. (2009). Is there a chance that I might kinda sort of take you out to dinner? The role of the pragmatic particles kind of and sort of in utterance interpretation. *Journal of Pragmatics*, 41(3), 602–25.
- Mondorf, B. (2004). Tag Questions According to Gender. In H. V. P. Blumenthal, K. V. Heusinger, & I. Plag, B. Primus & R. Wiese (Eds.), *Gender Differences in English Syntax: Vol. 491*. Walter de Gruyter GmbH.
- Mooney, A., & Evans, B. (2023). Language and Gender. In *Language, Society and Power* (6th ed., pp.

- 128–154). Routledge. <https://doi.org/10.4324/9781003120957-6>
- Mukherjee, J. (2008). "Sri Lankan English: evolutionary status and epicentral influence from Indian English". In K. Stierstorfer (Ed.), *Anglistentag 2007 Münster: Proceedings*. Trier: Wissenschaftlicher Verlag Trier, (pp. 359–68).
- Mullany, L. (2003). *Identity and role construction: a sociolinguistic study of gender and discourse in management*. (Publication No. 10290514) [Doctoral dissertation, Nottingham Trent University]. ProQuest Dissertations and Theses Global.
- Muller, C. (1994). *Yakada yaka*. Penguin Books: India.
- Müller, S. (2005). *Discourse markers in native and non- native discourse*. John Benjamins.
DOI: 10.1075/pbns.138.
- Murphy, B. (2010). *Corpus and Sociolinguistics: Investigating age and gender in female talk*. John Benjamins Publishing Company.
- Neary Sundquist, C. (2014). The use of pragmatic markers across proficiency levels in second language speech. *Studies in Second Language Learning and Teaching*, 4(4), 637--663.
- Nelson, G., Wallis, S. & Bas A. (2002). *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins.
- Newman M. L., Groom, C. J., Handelman, L. D. & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, 45(3), 211-236. DOI: [10.1080/01638530802073712](https://doi.org/10.1080/01638530802073712)
- Nguyen, H. (2023). Interactional sociolinguistics as a multi-pronged approach to office hour rapport management. *Journal of English for Academic Purposes*, 62, 1-11.
- Nikula, T. (1996). *Pragmatic force modifiers: A study in interlanguage pragmatics*. [Doctoral dissertation, University of Jyväskylä] <http://urn.fi/URN:ISBN:978-951-39-8072-6>
- Noor, M. (2021). Nativisation and variation of discourse marker in Pakistani English. *University of Chitral Journal of Linguistics and Literature*, 5(2), 112-136.

- Nordhoff, S. (2013). *The genesis of Sri Lanka Malay a case of extreme language contact*. In S. Nordhoff (Ed.) *Brill's Studies in South and Southwest Asian Languages: Vol. 3* (1st ed., pp. 3-12). Brill.
- Novotana, R. (2016). The role of discourse markers in spoken English [Diploma thesis, Masaryk University]. <https://is.muni.cz/th/ppa39/DiplomaThesis.pdf>
- Oates, S. L. (2000). Multiple discourse marker occurrence: Creating hierarchies for natural language generation. In D. Byron & P. Vanderheyden (Eds.), *Proceedings of the ANLP-NAACL 2000 Student Research Workshop* (pp. 41-45). Morgan Kaufmann Publishers Inc.
- Odato, C. V. (2013). The development of children's use of discourse like in peer interaction. *American Speech*, 88(2), 117–143. <https://doi.org/10.1215/00031283-2346825>
- Oladipupo, R. & Unuabonah, F. (2021). Extended discourse-pragmatic usage of now in Nigerian English. *World Englishes*, 40(3), 371-389.
- Onysko, A. (2021). Where are WEs heading to? In A Onysko (Eds.) *Research Developments in World Englishes* (1st ed., pp. 1-10). Bloomsbury Academic.
- Othman, Z. (2010). The use of okay, right and yeah in academic lectures by native speaker lecturers: Their 'anticipated' and 'real' meanings. *Discourse Studies*, 12(5), 665-681.
- Ostermann, A. C. (2021). Going South and zooming into what also matters in language, gender and sexuality. *Gender and Language*, 15(4), 611–624. <https://doi.org/10.1558/genl.21528>
- Östman, J. O. (1981). *You know: A discourse-functional study*. John Benjamins.
- Östman, J.O., (1982). The symbiotic relationship between pragmatic particles and impromptu speech. In N. E. Enkvist (Ed.), *Impromptu Speech: A Symposium* (pp. 147-177). Abo Akademi.
- Östman, J. O. (2006). Construction in cross-language research: Verbs as pragmatic particles in Solv. In K. Aijmer & A. M. Simon-Vandenberg (Eds.), *Pragmatic Markers in Contrast* (1 st ed., pp.237-257). Elsevier.

- Öztürk, Y., & Durmuşoğlu Köse, G. (2021). "Well (er) you Know ...": Discourse markers in native and non-native spoken English. *Corpus Pragmatics: International Journal of Corpus Linguistics and Pragmatics*, 5(2), 223–242. <https://doi.org/10.1007/s41701-020-00095-9>
- Palacios Martinez, I. (2015). Variation, development and pragmatic uses of innit in the language of British adults and teenagers. *English Language and Linguistics*, 19(3), 383–405. <https://doi.org/10.1017/S1360674314000288>
- Parviainen, H. (2016). The invariant tag isn't it in Asian Englishes. *World Englishes*, 35(1), 98-117.
- Parakrama, A. (1995). *De-hegemonizing language standards: Learning from (post) colonial Englishes about "English"*. London: Macmillan.
- Parakrama, A. (2010). "Naduth unge baduth unge" (Mistranslated as the rules & tools are theirs): Some thoughts on the language of privilege and the privilege of language. In S. Fernando, M. Guenesekera, A. Parakrama (Eds.), *English in Sri Lanka: Ceylon English, Lankan English, Sri Lankan English* (1st ed., pp. 78-102). Sri Lanka English Language Teachers' Association.
- Parakrama, A. (2012). The malchemy of english in sri lanka: Reinforcing inequality through imposing extra-linguistic value. In V. Rapatahana & P. Bunce (Eds.), *English Language as Hydra: Its Impacts on Non-English Language Cultures* (pp. 107–132). Multilingual Matters.
- Parakrama, A. (2016). Extra-linguistic value (of English in Sri Lanka) = Zuzwang (for non-elite users). *International Journal of Humanities and Applied Social Science*, 1(2), 23-30.
- Passé, H.A. (1948). The English Language in Ceylon. *University of Ceylon Review*, 3, 50-65.
- Passé, H.A. (1955). *The use and abuse of English*. Oxford University Press.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality & Social Psychology*, 77(6), 1296–1312.
- Perera, K., & Strauss, S. (2015). High-focus and time-immediate indexicals: A study of Sinhala discourse markers me: 'this' and dæn 'now.' *Journal of Pragmatics*, 85(2), 32–46. <https://doi.org/10.1016/j.pragma.2015.05.022>
- Petterson-Traba, D. (2018). Revisiting you know and I mean: some notes on the functions of the two

- pragmatic markers in contemporary spoken American English. *Research in Corpus Linguistics*, 67–81. <https://doi.org/10.32714/ricl.06.06>
- Pichler, H. (2013). *The structure of discourse-pragmatic variation*. John Benjamins.
- Preisler, B. (1986). *Linguistic sex roles in conversation: Social variation in the expression of tentativeness in English*. De Gruyter Mouton. <https://doi.org/10.1515/9783110862973>
- Quirk, R. Greenbaum, S., Leech, G. and Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.
- Rajapakse, A. (2008). A descriptive analysis of the language of the Burghers of Sri Lanka. In D. Fernando & D. Mendis (Eds.), *English for education, empowerment and equality* (pp. 49-58). International Conference of the Sri Lanka English Language Teachers' Association.
- Rajapakse, A. (2024). The impact of linguistic racism and coloniality on Sri Lankan English studies: the case of Burgher English. *International Multilingual Research Journal*. 18(1), 22-33. DOI: [10.1080/19313152.2023.2212938](https://doi.org/10.1080/19313152.2023.2212938)
- Ranaweera, M. C. (2007). Register variation in Sri Lankan English newspaper language [Unpublished Master's thesis]. Northern Arizona University.
- Ranwala, C. (2015). English, Lanka! In Sri Lanka, the English language remains a symbolic marker of privilege, refinement and class—as well as a means of upward mobility. *Economic and Political Weekly*, 50(18), 69–69.
- Ratwatte, H. (2012). Kaduva, karawila or giraya? Orientation towards English in 21st Century Sri Lanka. *Vistas: Journal of Humanities and Social Sciences*, 7(8), 180-203.
- Rayson, P., Leech, G., and Hodges, M. (1997). Social differentiation in the use of English vocabulary: some analyses of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics*, 2(1). 133 - 152. John Benjamins.
- Redeker, G. (2006). Discourse markers as attentional cues at discourse transitions. In K. Fischer (Ed.), *Approaches to discourse particles* (1st ed., pp. 339-358). Elsevier.
- Reichelt, S. (2021). Recent developments of the pragmatic markers kind of and sort of in spoken

- British English. *English Language and Linguistics*, 25(3), 563-580.
- Rees-Miller, J. (2011). Compliments revisited: Contemporary compliments and gender. *Journal of Pragmatics*, 43(11), 2673-2688.
- Reppen, R. (2010). Building a corpus: what are the key considerations? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed., pp. 31-37). Routledge.
- Reppen, R. (2011). Using corpora in the language classroom. In B. Tomlinson (Ed.) *Materials Development in Language Teaching* (pp. 35–50). Cambridge University Press.
<https://doi.org/10.1017/9781139042789.003>
- Revis, M. & Bernaisch, T. (2020). The pragmatic nativisation of pauses in Asian Englishes. *World Englishes*, 39(1), 135-153.
- Samarakkody, M., & Braine, G. (2014). Teaching English in Sri Lanka: From colonial roots to Lankan English. In G. Braine (Ed.), *Teaching English to the World: History, Curriculum, and Practice* (pp. 147–158). <https://doi.org/10.4324/9781410612861>
- Sankoff, G. (2006). Apparent time and real time. In K. Brown (Ed.), *Elsevier Encyclopedia of Language and Linguistics*. (2nd ed., pp. 110- 116). Elsevier.
- Sarangi, S. (2007). Editorial: The anatomy of interpretation: Coming to terms with the analyst’s paradox in professional discourse studies. *Text & Talk*, 27(5–6), 567–584.
<https://doi.org/10.1515/TEXT.2007.025>
- Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language (Baltimore)*, 79(2), 233–281. <https://doi.org/10.1353/lan.2003.0136>
- Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.
- Schneider, E. W. (2014). New reflections on the evolutionary dynamics of world Englishes. *World Englishes*, 33(1), 9-32.
- Schneider, E. W. (2016). World Englishes on YouTube: Treasure trove or nightmare? In E. Seoane & C. Suárez-Gómez (Eds.), *World Englishes: New theoretical and methodological considerations: Vol. G57*. John Benjamins Publishing Company.

- Schneider, K.P. (2010). Variational Pragmatics. In M. Fried, J. Östman & J. Verschueren (Eds.), *Handbook of Pragmatics Highlights: Vol 6 Variation and Change: Pragmatic Perspectives*, (1st ed., pp. 239-267) John Benjamins Publications.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press.
- Schleef, E. (2005). Gender, power, discipline, and context: On the sociolinguistic variation of okay, right, like, and you know in English academic discourse. In C. Sunakawa, T. Ikeda, S. Finch, & M. Shetty (Eds.), *Texas linguistic forum* (Vol. 48, pp. 177-186). Department of Linguistics.
- Schleef, E. (2009). A cross-cultural investigation of German and American academic style. *Journal of Pragmatics*, 41(6), 1104–1124. <https://doi.org/10.1016/j.pragma.2009.01.002>
- Schleef, E. & Mackay, B. (2022) Evaluation of Pragmatic Markers: The Case of You Know. In E. Peterson, T. Hiltunen & J. Kern (Eds.) *Discourse-pragmatic variation and change* (pp. 40-60). Cambridge University Press. <https://doi.org/10.1017/9781108864183.004>
- Schoning, C., Helder, J., & Diskin-Holdaway, C. (2023). Vi snakker sådan: An analysis of the Danish discourse-pragmatic marker sådan. *Nordic Journal of Linguistics*, 46(1), 46–82. <https://doi.org/10.1017/S0332586521000159>
- Schourup, L. (1985). *Common discourse particles in English conversation*. Garland.
- Schourup, L. (1999). Discourse markers. Tutorial overview. *Lingua*, 107, 227–65.
- Schweinberger, M. (2012). The discourse marker LIKE in Irish English. In M. Ní Chiosáin & B. Migge (Eds.), *Varieties of English around the world: Vol G44 New Perspectives on Irish English* (1st ed., pp. 179-201). John Benjamins Publishing Company.
- Schweinberger, M. (2018a). Swearing in Irish English: A corpus based quantitative analysis of the sociolinguistics of swearing. *Lingua*, 209, 1-20.
- Schweinberger, M. (2018b). The discourse particle eh in New Zealand English. *Australian Journal of Linguistics*, 38(3), 395–420. <https://doi.org/10.1080/07268602.2018.1470458>
- Scivoletto, G. (2023). *Discourse markers in Sicily: A synchronic, diachronic, and sociolinguistic analysis*. Brill.

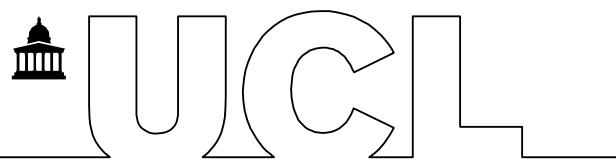
- Séguy, I, Courgeau, D., Caussin, H., Buchet, L., (2019). *Chronological age, social age and biological age*. <https://doi.org/10.13140/rg.2.2.28706.68801>
- Shakir, M., & Deuber, D. (2023). Compiling a corpus of South Asian online Englishes: A report, some reflections and a pilot study. *ICAME Journal*, 47(1), 119–139. <https://doi.org/10.2478/icame-2023-0007>
- Shakir, M., & Deuber, D. (2024). Code-switching in South Asian English CMC. *English World-Wide*. <https://doi.org/10.1075/eww.23068.sha>
- Shapiro, D. (1989). *Psychotherapy of neurotic character*. Basic Books.
- Sheydaei, I. (2021). Gender identity and nonbinary pronoun use: exploring reference strategies for referents of unknown gender. *Gender and Language*, 15(3), 369–393. <https://doi.org/10.1558/genl.18871>
- Siegel, M. E. A. (2002). Like: The discourse particle and semantics. *Journal of Semantics*, 19, 35–71.
- Slobe, T. (2018). Style, stance, and social meaning in mock white girl. *Language in Society*, 47(4), 541–567.
- Spender, D. (1980). *Man Made Language*. London: Routledge.
- Stenström, A. (2006). The Spanish discourse markers *o sea* and *pues* and their English correspondences. In K. Aijemer & A.M. Simon-Vendenbergen (Eds.), *Studies in Pragmatics: Vol 2 Pragmatic Markers in Contrast*. (1st ed., pp. 155-172). Elsevier.
- Stenström, A. (2014). *Teenage Talk from general characteristics to the use of pragmatic markers in a contrastive perspective* (1st ed.). Palgrave Macmillan UK. <https://doi.org/10.1057/9781137430380>
- Suárez-Gómez, C., & Seoane, E. (2021). The role of age and gender in grammatical variation in world Englishes. *World Englishes*, 42(2): 327–43. doi: <https://doi.org/10.1111/weng.12546>.
- Szudarski, P. (2022). Corpora and Teaching Vocabulary and Phraseology. In R. Jablonkai & E. Csomay (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning* (1st ed., Vol. 1, pp. 41–55). Routledge. <https://doi.org/10.4324/9781003002901-5>

- Symons, D. K. (2011). Psychological age. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child labour and development: Vol 3* (pp. 1180- 1180). Springer.
- Tagliamonte, S. A. & Hudson, R. (1999). Be like et al. beyond America: the quotative system in British and Canadian youth. *Journal of Sociolinguistics*, 3(2),147-72.
- Tagliamonte, S. A., & D'Arcy, A. (2007). Frequency and variation in the community grammar: Tracking A new change through the generations. *Language Variation and Change*, 19(2), 199–217.
<https://doi.org/10.1017/S095439450707007X>
- Tannen, D. (1994). *Gender and discourse*. Oxford University Press.
- Tao, H. (2003). Turn Initiators in spoken English: A corpus-based approach to interaction and grammar. In P. Leistyna & C. F. Meyer (Eds.), *Corpus analysis: Language structure and language use: Vol 46* (pp. 187–207). BRILL. https://doi.org/10.1163/9789004334410_011
- Tay, L. C., Chan, M. Y., Yap, N. T., & Wong, B. E. (2016). Discourse particles in Malaysian English: What do they mean? *Bijdragen Tot de Taal-, Land- En Volkenkunde*, 172(4), 479–509.
<https://doi.org/10.1163/22134379-17204002>
- Torgersen, E. N., Gabrielatos, C., Hoffmann, S., & Fox, S. (2011). A corpus-based study of pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory*, 7(1), 93–118.
<https://doi.org/10.1515/cllt.2011.005>
- Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge University Press.
- Tupas, R. & Rubdy, R. (2015). Introduction: From world Englishes to unequal Englishes. In R. Tupas (Ed.), *Unequal Englishes: the politics of Englishes today* (pp. 1-19). Palgrave Macmillan.
- Underhill, R. (1988). Like is, like, focus. *American Speech*, 63(3), 234–246.
<https://doi.org/10.2307/454820>
- Unger, C. (1996). The scope of discourse connectives: Implications for dis course organization. *Journal of Linguistics*, 32, 402–438.
- Unuabonah, F. O., & Muro, L. P. (2022). Borrowed Swahili discourse-pragmatic features in Kenyan

- and Tanzanian Englishes. *Intercultural Pragmatics*, 19(4), 489–512.
<https://doi.org/10.1515/ip-2022-4003>
- Unuabonah, F. O., & Oladipupo, R. O. (2021). Bilingual pragmatic markers in Nigerian English. *World Englishes*, 40(3), 390–406. <https://doi.org/10.1111/weng.12453>
- Vaish, V. (2005). A peripherist view of English as a language of decolonization in post-colonial India. *Language Policy*, 4(2), 187–206.
- Varga, M. (2020). The Role of Linguistic Impression Management: The Case of Hungarian Courtroom Discourses. *Corpus Pragmatics: International Journal of Corpus Linguistics and Pragmatics*, 4(3), 381–400. <https://doi.org/10.1007/s41701-020-00087-9>
- Verschueren, J. (1999). *Understanding Pragmatics*. London: Arnold Publishers.
- Vessey, R (2023). Language and gender in Canadian Chief Medical Officers’ tweets during the COVID-19 Pandemic. *Critical Discourse Studies*, 20(2), 200–217.
[DOI: 10.1080/17405904.2022.2050772](https://doi.org/10.1080/17405904.2022.2050772)
- Vincent, D. (2005). The journey of non-standard discourse markers in Quebec French: Networks based on exemplification. *Journal of Historical Pragmatics*, 6(1), 188–210.
- Vine, B. & Holmes, J. (2023). Doing leadership in style: Pragmatic markers in New Zealand workplace interaction. *Intercultural Pragmatics*, 20(1), 1–27.
- Walisundara, D. & Hettiarachchi, S. (2016). English language policy and planning in Sri Lanka: A critical overview. In R. Kirkpatrick (Ed.), *English language education policy in Asia: Vol. 11* (pp. 301–332). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-22464-0_14
- Wang, Q. (2020). A corpus-based contrastive analysis of I think in spoken Hong Kong English: Research from the International Corpus of English (ICE). *Australian Journal of Linguistics*, 40(3), 319–345. <https://doi.org/10.1080/07268602.2020.1823817>
- Watts, R. J. (1988). A relevance-theoretic approach to commentary pragmatic markers: The case of actually, really and basically. *Acta Linguistica Hungarica*, 38, 235–260.

- Wenzlhuemer, R. (2008). *From Coffee to Tea Cultivation in Ceylon, 1880-1990*. Leiden-Boston: Brill.
- West, C. & Zimmerman, D. (2002). Doing gender. In S. Fenstermaker and C. West (Eds.), *Doing Gender, Doing Difference* (pp. 3-23). Routledge.
- Westphal, M. (2024). Eh Across Englishes: A Corpus-Pragmatic Analysis of the Corpus of Global Web-Based English. *Corpus Pragmatics: International Journal of Corpus Linguistics and Pragmatics*, 8(1), 53–75. <https://doi.org/10.1007/s41701-023-00159-6>
- Weydt, H. (2006). ‘What are particles good for?’ In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 205-218). Brill.
- Wierzbicka, A. (2003). *Cross-cultural pragmatics: The semantics of human interaction* (2nd ed.). De Gruyter, Inc.
- Wijetunga, S.N. (2008) The stigma of “Not Pot English” in Sri Lanka: A study of production of /o/ and /ɔ/ and implications for instructions [Master’s thesis, Georgia State University]. https://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1000&context=alesl_theses
- Wilson, G., Westphal, M., Hartmann, J., & Deuber, D. (2017). *The use of question tags in different text types of Trinidadian English*. *World Englishes*, 36(4), 726–743. <https://doi.org/10.1111/weng.12247>
- Wiltschko, M., Denis, D., & D’Arcy, A. (2018). Deconstructing variation in pragmatic function: A transdisciplinary case study. *Language in Society*, 47(4), 569-599. <https://doi.org/10.1017/S004740451800057X>
- Winford, D. (1993). Variability in the use of perfect have in Trinidadian English: A problem of categorial and semantic mismatch. *Language Variation and Change*, 5(2), 141–188.
- Wolf, H.-G., & Polzenhagen, F. (2009). *Applications of cognitive linguistics: Vol 8. World Englishes: A cognitive sociolinguistic approach*. Mouton de Gruyter. <https://doi.org/10.1515/9783110199222>
- Wolfram, W. (1969). *A Sociolinguistic Description of Detroit Negro Speech*, Washington D.C.: Centre for Applied Linguistics.

Zandon, N. (2018, August 01). *How to stop saying “um”, “ah”, and “you know”*, Harvard Business Review. <https://hbr.org/2018/08/how-to-stop-saying-um-ah-and-you-know>



Zhang, J. (2024). Variational pragmatics in Chinese discourse markers zhege and nage: The influence of region and gender. *Journal of Pragmatics*, 230, 76–88.

<https://doi.org/10.1016/j.pragma.2024.07.001>

Zheng, Q. (2015). Revisiting *you know* using the BNC web query system: a Sociopragmatic analysis. In Beeching, K., Woodfield, H. (Eds.), *Researching sociopragmatic variability* (pp. 94-118).

Palgrave Macmillan. https://doi.org/10.1057/9781137373953_5

Zimmerman, D. & West, C. (1975) Sex roles, interruptions and silences in conversation. In B. Thorne & N. Henley (Eds.) *Language and Sex: Difference and Dominance* (pp. 105-129). Newbury House.

APPENDIX A

UCL Department

DEPARTMENT OF RESEARCH ETHICS APPLICATION FORM

Guidance notes

UCL expects all staff and students to follow its Research Ethics regulations (<http://ethics.grad.ucl.ac.uk/>).

All staff & PGR students* embarking on research with human participants must complete this form and submit it to the Chair of the DIS Research Ethics Committee (REC), **before** they start their research.

**(MA & MSc students registered for the Dissertation module are generally presumed to be doing research exempt from requirement for full UCL ethical*

clearance (i.e. their research does not involve vulnerable subjects and is non-sensitive, anonymised and non-invasive or interactive, see [UCL guidance on research exempt from requiring more than just Departmental level approval](#), see [UCL guidance on research exempt from full UCL research ethics committee consideration at <https://ethics.grad.ucl.ac.uk/exemptions.php>](#)).

However **all students** whose research involves human subjects must follow the process outlined on the relevant Moodle page and discuss their research approach with their personal tutors and then supervisors at their first meeting. Further advice and guidance on the exemptions and further research ethics processes are available from the [Chair of the Dept Research Ethics Committee \(REC\)](#)

The DIS REC Chair will review the form and decide whether:

- the proposed research is **exempt** from further UCL Research Ethics Committee review; **Or**
- the proposed research requires **further information / clarification or full approval by the A&H Local Research Committee (low risk) or the full UCL Research Ethics Committee (high risk)** - the REC Chair will notify you (and if appropriate your supervisor) and advise you on how to proceed.

Changes to previously exempt research projects: if you are planning to change your research project or methodology, you **MUST** contact the Dept REC Chair, as soon as possible and provide relevant details as your project may now no longer be exempt. Please also note that in addition to research ethics (whether exempt or not), researchers collecting and processing personally identifiable data of any kind (e.g. email addresses of interviewees or IP addresses of survey participants, personal data from interviews or surveys, etc.) in the course of their research will need to [apply for data protection registration](https://www.ucl.ac.uk/dataprotection/guidance-staff-students-and-researchers/research/research-registrationsguidance) (<https://www.ucl.ac.uk/dataprotection/guidance-staff-students-and-researchers/research/research-registrationsguidance>)

UCL researchers (staff and doctoral students), who are planning to conduct fieldwork, should consult the [Framework for starting or resuming fieldwork at nonUCL settings](#) and apply for an [approval through RiskNET separately](#) to this form.

DEPARTMENT

RESEARCH ETHICS APPLICATION FORM**1. Personal Details**

First Name	Surname	Email
Mahishi	Ranaweera	ranaweera.ranaweera.19@ucl.ac.uk

2. For UCL PGR students only

Programme of study (MA, MSc, MRes, MPhil, PhD, etc.)	Title (provisional) of dissertation and name of Principal Supervisor
PhD	A corpus based study on the use of Pragmatic Markers in Sri Lankan English Dr. Rachele de Felice

3. All applicants

Brief description of proposed & suggested research methodology (including details of topic, human participants and plans for anonymity, procedures to acquire and document informed consent from participants etc.)
--

Brief Description of Research topic

The main objective of the current study is to survey the pragmatic markers in Standard Sri Lankan English (SSLE). Additionally, it looks at the use and functions of the Pragmatic Markers (PMs) in SSLE. It also examines gender specific PM use. Further, the study examines whether there are nativized PMs in the data.

Outline Research Methods / Data Collection practices

The study uses YouTube videos which featured interviews with SSLE speakers as its data. A total of 84 online semi formal to formal interviews published between 2012 to 2021 are selected. The online YouTube videos will be downloaded and transcribed. The speakers at speech turn are given a code as their ID.

A corpus of the transcriptions will be created. Each transcription is restricted to 2750 words. Only the transcriptions will be analysed to understand the functions and the use of PMs in SSLE. The data will be manually and electronically

analysed. AntConc software will be used to analyse the data electronically.

Brief details of participants (age, approx. number, characteristics)

The participants are public figures/well known personalities in Sri Lanka such as actors, entrepreneurs and academics. The ages approximately range from 24 to 80.

Will research be collected / reported anonymously or pseudoanonymously? If not, why not?

The research will be reported pseudo-anonymously. The speech turns in the transcript uses a coded ID for the speakers. However, the names of participants are not anonymised within the text as the participants are well known figures in the country and they are aware that they are part of videos in a public domain.

Brief outline of how the informed consent of research participants is to be obtained and documented? (if you have draft Information sheets / consent forms please attach)

Consent has not been obtained from the research participants as the study uses videos in a public domain that does not require permission from the researcher to access.

What personal identifiable data do you anticipate collecting and how will you be storing this?

No personal identifiable data will be collected by contacting the participants. Only the speech samples are analysed at word level.

Briefly outline any ethic issues you anticipate arising out of this research and how you will address them? (including consideration of any possible harm or distress to participants or researchers and how these risks will be addressed and minimised, what rights do participants have (withdraw, etc))

The proposed research does not anticipate any ethics issue as the findings do not profess any positive or negative claim in the individual participants' personalities. It is merely a comment on the use of a language feature in English in Sri Lanka.

- 4. Exemptions. If you believe your research is exempt from requiring further approval from the full UCL Research Ethics Committee (see list of exemptions and circumstances of exempt research <https://ethics.grad.ucl.ac.uk/exemptions.php>) please**

Reason for exemption:

1. Research involving information freely available in the public domain. For example, published biographies, newspaper accounts of an individual's activities and published minutes of a meeting, whilst still personal data under the Data Protection Act would not require ethics review.

The study explores published content which features adults on YouTube videos which is a public domain. The videos are professional productions for public viewing.

YouTube allows three types of videos to be uploaded according to their Ts and Cs. (https://support.google.com/youtube/answer/157177?hl=en-GB&ref_topic=9386940#zippy=%2Cunlisted-videos%2Cpublic-videos)

They are:

Private videos

Public videos

Unlisted videos

The study only uses public videos. YouTube Ts and Cs state the following with regard to public videos:

Anyone at YouTube can see public videos. They can also be shared with anyone using YouTube. They're posted on your channel when you upload them and appear in search results and related video lists.

It must be reiterated that the proposed research does not concern the individuals in the videos. It uses only speech samples which will contribute to the aggregate data. Only the aggregate data will be analysed. The data will be analysed for function, frequency and abstraction of patterns in the use of Pragmatic Markers. Therefore, it can be assured that the data analysis will not be harmful, offensive, intrusive or ethically dubious to the people projected in the videos.

briefly say why and indicate which exemption/s you are claiming:

5. Primary Investigator (Supervisor/Personal Tutor for PG Students)

Name:

Signature:

6. Date form submitted to DIS REC Chair

Date: 06/07/2022

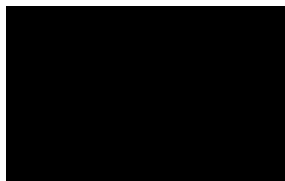
FOR OFFICE USE ONLY (to be completed by REC Chair)

1. Proposed research is **exempt for requiring further approval**
2. Proposed research requires **approval by the A&H Local Research Committee (low risk) or the full UCL Research Ethics Committee (high risk)**
[delete as applicable]

If (2) Applicant notified on:

Chair UCL DEPT REC Name: Julia Jordan

Signature:



(Chair UCL DEPT REC)

APPENDIX B

List of the videos with the YouTube Links, publisher information and license type.

Speaker ID	Youtube video ID	Published Year	Published by	YouTube license type
<01AF>	https://www.youtube.com/watch?v=cuJxbbsJNgA	27 Apr 2018	Pulse.lk	YouTube Standard License
<02AF>	https://www.youtube.com/watch?v=33CTMwYDwg4	19 Jun 2020	Mind Adventures Theatre Co.	YouTube Standard License
<03AF>	https://www.youtube.com/watch?v=gFfOAMJIsyE	12 Apr 2019	Pulse.lk	YouTube Standard License
<04AF>	https://www.youtube.com/watch?v=pn-s1lCxo4	27 May 2020	Hi!! Online	YouTube Standard License
<05AF>	https://www.youtube.com/watch?v=T_shrCB-zKM	6 Apr 2018	Pulse.lk	YouTube Standard License
<06AF>	https://www.youtube.com/watch?v=T9xaQwPnSbw	23 Feb 2018	Pulse.lk	YouTube Standard License
<07AF>	https://www.youtube.com/watch?v=RE38NvUj5Mo	8 Feb 2019	Pulse.lk	YouTube Standard License
<08AF>	https://www.youtube.com/watch?v=rEEJeNoK_Y	27 Dec 2019	Pulse.lk	YouTube Standard License
<09AF>	https://www.youtube.com/watch?v=270nutvEeb0	14 Feb 2020	Pulse.lk	YouTube Standard License
<10AF>	https://www.youtube.com/watch?v=Eds7PsbZqk	22 Sept 2018	Pulse.lk	YouTube Standard License
<11AF>	https://www.youtube.com/watch?v=lwoiW0Uvu9E	23 Aug 2019	Pulse.lk	YouTube Standard License
<12AF>	https://www.youtube.com/watch?v=V_UUQsJOyy0	24 Jan 2020	Pulse.lk	YouTube Standard License
<01AM>	https://www.youtube.com/watch?v=ZfuRh5t6NPA	1 Feb 2019	Pulse.lk	YouTube Standard License
<02AM>	https://www.youtube.com/watch?v=JvVCwUqR6j4	7 Dec 2018	Pulse.lk	YouTube Standard License

<03AM>	https://www.youtube.com/watch?v=QtsFBhInBdo	9 Nov 2018	Pulse.lk	YouTube Standard License
<04AM>	https://www.youtube.com/watch?v=YGJRZZ7BNxM	23 May 2020	Pulse.lk	YouTube Standard License
<05AM>	https://www.youtube.com/watch?v=3kPPOegoMdo	22 Jun 2018	Pulse.lk	YouTube Standard License
<06AM>	https://www.youtube.com/watch?v=EkKCVye_LBc	24 Apr 2018	Pulse.lk	YouTube Standard License
<07AM>	https://www.youtube.com/watch?v=UC1cDNylLMg	9 Mar 2018	Pulse.lk	YouTube Standard License
<08AM>	https://www.youtube.com/watch?v=f1RtAlAJlk8	10 Aug 2018	Pulse.lk	YouTube Standard License
<09AM>	https://www.youtube.com/watch?v=aabc-MSQ0jM	19 Oct 2018	Pulse.lk	YouTube Standard License
<10AM>	https://www.youtube.com/watch?v=svLRRy1MEws	9 Aug 2019	Pulse.lk	YouTube Standard License
<11AM>	https://www.youtube.com/watch?v=SobxQtt8P4U	31 Aug 2018	Pulse.lk	YouTube Standard License
<12AM>	https://www.youtube.com/watch?v=oFiKMOOyIRU	24 May 2019	Pulse.lk	YouTube Standard License
<01AcF >	https://www.youtube.com/watch?v=Pwm8rtCMyns	8 May 2021	Roar LK	YouTube Standard License
<02AcF >	https://www.youtube.com/watch?v=kJ739qsg4rE	5 Apr 2019	Pulse.lk	YouTube Standard License
<03AcF >	https://www.youtube.com/watch?v=nVp3d-mHFK4	7 Aug 2021	Roar LK	YouTube Standard License
<04AcF >	https://www.youtube.com/watch?v=zyhOSk8AGH0	24 May 2019	Sri Lanka Science Channel	YouTube Standard License
<05AcF >	https://www.youtube.com/watch?v=VPPoBaubue8	4 Apr 2018	Chanakya Jayadeva	YouTube Standard License
<06AcF >	https://www.youtube.com/watch?v=licuecDEVBM	10 Apr 2021	Daily Mirror Online	YouTube Standard License

<07AcF >	https://www.youtube.com/watch?v=MASigQJTgaQ	23 Oct 2012	Young Asia Television	YouTube Standard License
<08AcF >	https://www.youtube.com/watch?v=YZWlZ4fQsPE	16 Jun 2021	Hi!! Online	YouTube Standard License
<09AcF >	https://www.youtube.com/watch?v=Ur-kNESA0u8	20 Jul 2020	Retired United	YouTube Standard License
<10AcF >	https://www.youtube.com/watch?v=U2Jve83j6Gk	18 Sept 2020	Tonguesonfire	YouTube Standard License
<11AcF >	https://www.youtube.com/watch?v=bnkjLxTrgl4	9 Dec 2021	Hi!! Online	YouTube Standard License
<12AcF >	https://www.youtube.com/watch?v=Gg0y-XDIYMg	22 Apr 2021	Hi!! Online	YouTube Standard License
<01AcM >	https://www.youtube.com/watch?v=6L_xfZqNlp8	28 Jul 2021	Hi!! Online	YouTube Standard License
<02AcM >	https://www.youtube.com/watch?v=vf3AJqjhz0	4 Mar 2016	Postgraduate Institute of Management	YouTube Standard License
<03AcM >	https://www.youtube.com/watch?v=Y9Ag-9FgdY	30 Jun 2015	Social Scientists' Association	YouTube Standard License
<04AcM >	https://www.youtube.com/watch?v=2uGyTbuMdCA	2 Apr 2019	Newsfirst Sri Lanka	YouTube Standard License
<05AcM >	https://www.youtube.com/watch?v=axA5LjSG-gE	25 Mar 2021	Hi!! Online	YouTube Standard License
<06AcM >	https://vimeo.com/40021456	11 Mar 2012	Young Asia Television	don't need the copyright holder's permission. Fair use: Nonprofit educational uses
<07AcM >	https://www.youtube.com/watch?v=CJdVNZy3ykl	6 Jul 2021	Newsfirst Sri Lanka	YouTube Standard License
<08AcM >	https://www.youtube.com/watch?v=jVUoPW5T28	24 Mar 2021	Newsfirst Sri Lanka	YouTube Standard License

<09AcM >	https://www.youtube.com/watch?v=3v6Bede470k	6 Mar 2018	Newsfirst Sri Lanka	YouTube Standard License
<10AcM >	https://www.youtube.com/watch?v=-AXHINx7MCQ	23 Jan 2022	The In-between Space	YouTube Standard License
<11AcM >	https://www.youtube.com/watch?v=dueaF3jx5M	25 Aug 2021	Hi!! Online	YouTube Standard License
<12AcM >	https://www.youtube.com/watch?v=sTfo1MEuELY	9 Jul 2021	Newsfirst Sri Lanka	YouTube Standard License
<01EF>	https://www.youtube.com/watch?v=X2IXYUfMtKc	22 Jan 2018	Pulse.lk	YouTube Standard License
<02EF>	https://www.youtube.com/watch?v=8kFW67gZxCw	30 Dec 2019	Hi!! Online	YouTube Standard License
<03EF>	https://www.youtube.com/watch?v=7v5m-xLHJ5s	23 Nov 2016	Pulse.lk	YouTube Standard License
<04EF>	https://www.youtube.com/watch?v=yM3z3SLBrEw	16 Oct 2020	Hi!! Online	YouTube Standard License
<05EF>	https://www.youtube.com/watch?v=8p3FMOUHBwE	1 May 2021	Hi!! Online	YouTube Standard License
<06EF>	https://www.youtube.com/watch?v=Kfs9DXKzpKE	21 Nov 2019	Hi!! Online	YouTube Standard License
<07EF>	https://www.youtube.com/watch?v=2MM0sLBj7gk	4 Sept 2020	Hi!! Online	YouTube Standard License
<08EF>	https://www.youtube.com/watch?v=rz7FpBpH_S8	8 May 2021	Hi!! Online	YouTube Standard License
<09EF>	https://www.youtube.com/watch?v=1NLwRGc1AhA	3 Dec 2019	Hi!! Online	YouTube Standard License
<10EF>	https://www.youtube.com/watch?v=0G0a8TYDTAo	13 Mar 2021	Hi!! Online	YouTube Standard License
<11EF>	https://www.youtube.com/watch?v=Ud2LDQjHjCk	6 Mar 2021	Hi!! Online	YouTube Standard License
<12EF>	https://www.youtube.com/watch?v=20zC080GYbl	30 Oct 2021	Hi!! Online	YouTube Standard License

<01EM>	https://www.youtube.com/watch?v=l8YVElu7yCk	15 Jul 2020	Hi!! Online	YouTube Standard License
<02EM>	https://www.youtube.com/watch?v=shER2o2A9k	10 Jul 2017	Pulse.lk	YouTube Standard License
<03EM>	https://www.youtube.com/watch?v=1kLGNO19ZN0	10 Aug 2017	Pulse.lk	YouTube Standard License
<04EM>	https://www.youtube.com/watch?v=aOzEkJh4Sxl	31 Jul 2016	Pulse.lk	YouTube Standard License
<05EM>	https://www.youtube.com/watch?v=PPOeFdTpUoY	22 Feb 2017	Pulse.lk	YouTube Standard License
<06EM>	https://www.youtube.com/watch?v=0VOZO7i3XJM	8 Jun 2018	Pulse.lk	YouTube Standard License
<07EM>	https://www.youtube.com/watch?v=ttqJFRoPbYo	10 Aug 2016	Pulse.lk	YouTube Standard License
<08EM>	https://www.youtube.com/watch?v=RTdo2QICrBU	24 Dec 2016	Pulse.lk	YouTube Standard License
<09EM>	https://www.youtube.com/watch?v=KYIFL3uWL9l	26 Jul 2017	Pulse.lk	YouTube Standard License
<10EM>	https://www.youtube.com/watch?v=8uGK9AGsKu8	21 Dec 2020	Hi!! Online	YouTube Standard License
<11EM>	https://www.youtube.com/watch?v=qp1G4pWcwVU	10 Feb 2020	Hi!! Online	YouTube Standard License
<12EM>	https://www.youtube.com/watch?v=AlNiuTCoYGM	7 Aug 2021	Hi!! Online	YouTube Standard License

APPENDIX C

Length of each video

Speaker ID	Youtube video ID	Length of the video (Mins)
<01AF>	https://www.youtube.com/watch?v=cuJxbbsJNgA and part 2 (6:32 mins of) <u>Gymshark - Phys Ed (youtube.com)</u>	
<02AF>	https://www.youtube.com/watch?v=33CTMwYDwg4	
<03AF>	https://www.youtube.com/watch?v=gFfOAMJlsyE and part 2 (<u>Aseka Wijewardena on Anything But - Sequence Z (youtube.com)</u>)	
<04AF>	https://www.youtube.com/watch?v=pn-s1lCxo4	
<05AF>	https://www.youtube.com/watch?v=TshrCB-zKM and part 2 <u>Menaka De Fonseka on Anything But - Sequence Z (youtube.com)</u>	
<06AF>	https://www.youtube.com/watch?v=T9xaQwPnSbw and part 2 (4:51 mins of) <u>Soundarie David Rodrigo on Anything But - Sequence Z (youtube.com)</u>	
<07AF>	https://www.youtube.com/watch?v=RE38NvUj5Mo and part 2 (5:07 mins of) <u>Thaji Dias on Anything But - Sequence Y (youtube.com)</u>	
<08AF>	https://www.youtube.com/watch?v=rEEJeNoK-Y	
<09AF>	https://www.youtube.com/watch?v=270nutvEeb0 and part 2 (1:33 mins of) <u>Mariazelle Goonetilleke on Anything But - Sequence Y (youtube.com)</u>	
<10AF>	https://www.youtube.com/watch?v=EdS7PsbZqk and part 2 (4:16 mins of) <u>Ameena Hussein on Anything But - Sequence Y (youtube.com)</u>	
<11AF>	https://www.youtube.com/watch?v=lwoiWOUvu9E and part 2 (4:40 mins of) <u>Dinakshie Priyasad on Anything But - Sequence Y (youtube.com)</u>	
<12AF>	https://www.youtube.com/watch?v=VUUQsJOyy0 and part 2 (2:08 mins of) <u>Corrine Almeida on Anything But - Sequence Y (youtube.com)</u>	
<01AM>	https://www.youtube.com/watch?v=ZfuRh5t6NPA and part 2 (4:07 mins of) <u>Feroze Kamardeen on Anything But - Sequence Y (youtube.com)</u>	
<02AM>	https://www.youtube.com/watch?v=JvVCwUqR6j4 and part 2 (4:58 mins of) <u>Santhush Weeraman on Anything But - Sequence Z (youtube.com)</u>	
<03AM>	https://www.youtube.com/watch?v=QtsFBhInBdo and part 2 (5:33 mins of) <u>Senaka De Silva on Anything But - Sequence Y (youtube.com)</u>	
<04AM>	https://www.youtube.com/watch?v=YGJRZZ7BNxM	
<05AM>	https://www.youtube.com/watch?v=3kPPOegoMdo and part 2 (6:34 mins of) <u>Ravindra Randeniya on Anything But - Sequence Z (youtube.com)</u>	

<06AM>	https://www.youtube.com/watch?v=EkKCVye_LBc and part 2 (4:53 mins of) <u>Sashane Perera on Anything But - Sequence Z (youtube.com)</u>	
<07AM>	https://www.youtube.com/watch?v=UC1cDNyILMg and part 2 (5:43 mins of) <u>Ashan Dias on Anything But - Sequence Z (youtube.com)</u>	
<08AM>	https://www.youtube.com/watch?v=f1RtAIAJlk8 and part 2 (3:37 mins of) <u>Jehan Aloysius on Anything But - Sequence Y (youtube.com)</u>	
<09AM>	https://www.youtube.com/watch?v=aabc-MSQ0jM and part 2 (3:11 mins of) <u>Chandran Rutnam on Anything But - Sequence Y (youtube.com)</u>	
<10AM>	https://www.youtube.com/watch?v=svLRRy1MEws and part 2 (5:12 mins of) <u>Lakshman Joseph de Saram on Anything But - Sequence Y (youtube.com)</u>	
<11AM>	https://www.youtube.com/watch?v=SobxQt8P4U and part 2 (4:07 mins of) <u>(19) Kevin Cruze on Anything But - Sequence Y - YouTube</u>	
<12AM>	https://www.youtube.com/watch?v=oFiKMOOyIRU and part 2 (8:06 mins of) <u>Damith Fonseka on Anything But - Sequence Z (youtube.com)</u>	
<01AcF>	https://www.youtube.com/watch?v=Pwm8rtCMyns	
<02AcF>	https://www.youtube.com/watch?v=kj739qsg4rE and part 2 (9.12 mins of) <u>Madhubhashini Disanayaka Ratnayake on Anything But - Sequence Y (youtube.com)</u>	
<03AcF>	https://www.youtube.com/watch?v=nVp3d-mHFK4	
<04AcF>	https://www.youtube.com/watch?v=zyhOSk8AGH0	
<05AcF>	https://www.youtube.com/watch?v=VPPoBaubue8	
<06AcF>	https://www.youtube.com/watch?v=licuecDEVBM	
<07AcF>	https://www.youtube.com/watch?v=MASigQJTgaQ	
<08AcF>	https://www.youtube.com/watch?v=YZWIz4fQsPE	
<09AcF>	https://www.youtube.com/watch?v=Ur-kNESA0u8	
<10AcF>	https://www.youtube.com/watch?v=U2Jve83j6Gk	
<11AcF>	https://www.youtube.com/watch?v=bnkjLxTrgl4	
<12AcF>	https://www.youtube.com/watch?v=Gg0y-XDIYMg	
<01AcM>	https://www.youtube.com/watch?v=6L_xfZqNlp8	
<02AcM>	https://www.youtube.com/watch?v=vf3AJqjhAz0	
<03AcM>	https://www.youtube.com/watch?v=Y9Ag-9Fg-dY	
<04AcM>	https://www.youtube.com/watch?v=2uGyTbuMdCA	
<05AcM>	https://www.youtube.com/watch?v=axA5LjSG-gE	
<06AcM>	https://vimeo.com/40021456	
<07AcM>	https://www.youtube.com/watch?v=CJdVNzy3yki	
<08AcM>	https://www.youtube.com/watch?v=jVUoPW_5T28	
<09AcM>	https://www.youtube.com/watch?v=3v6Bede47Ok	
<10AcM>	https://www.youtube.com/watch?v=-AXHINx7MCQ	
<11AcM>	https://www.youtube.com/watch?v=dueaF3jx_5M	
<12AcM>	https://www.youtube.com/watch?v=sTfo1MEuELY	

<01EF>	https://www.youtube.com/watch?v=X2IXYUfMtKc	
<02EF>	https://www.youtube.com/watch?v=8kFW67gZxCw	
<03EF>	https://www.youtube.com/watch?v=7v5m-xLHJ5s	
<04EF>	https://www.youtube.com/watch?v=yM3z3SLBrEw	
<05EF>	https://www.youtube.com/watch?v=8p3FMOUHBwE	
<06EF>	https://www.youtube.com/watch?v=Kfs9DXKzpKE	
<07EF>	https://www.youtube.com/watch?v=2MM0sLBJ7gk	
<08EF>	https://www.youtube.com/watch?v=rz7FpBpH_S8	
<09EF>	https://www.youtube.com/watch?v=1NLwRGc1AhA	
<10EF>	https://www.youtube.com/watch?v=0G0a8TYDTAo	
<11EF>	https://www.youtube.com/watch?v=Ud2LDQjHjCk	
<12EF>	https://www.youtube.com/watch?v=20zC080GYbl	
<01EM>	https://www.youtube.com/watch?v=l8YVElu7yCk	
<02EM>	https://www.youtube.com/watch?v=_shER2o2A9k	
<03EM>	https://www.youtube.com/watch?v=1kLGNO19ZN0	
<04EM>	https://www.youtube.com/watch?v=aOzEkJh4Sxl	
<05EM>	https://www.youtube.com/watch?v=PPOeFdTpUoY	
<06EM>	https://www.youtube.com/watch?v=0VOZO7i3XJM and part 2 (min 5.43 of) Anything But with Kolu - Sequence Y (youtube.com)	
<07EM>	https://www.youtube.com/watch?v=ttqJFRoPbYo	
<08EM>	https://www.youtube.com/watch?v=RTdo2QICrBU	
<09EM>	https://www.youtube.com/watch?v=KYIFL3uWL9I	
<10EM>	https://www.youtube.com/watch?v=8uGK9AGsKu8	
<11EM>	https://www.youtube.com/watch?v=qp1G4pWcwVU	
<12EM>	https://www.youtube.com/watch?v=AlNiuTCoyGM	

APPENDIX D

Metadata with regard to ethnicity of the speaker

The speaker ID (F/M) refers to the gender of the speaker: female or male. The codes used such as <01AF>, <02AF> and so on refer to the speaker IDs in the corpus.

Speaker ID (F/M)	Ethnicity	Speaker ID (F/M)	Ethnicity	Speaker ID (F/M)	Ethnicity
<01AF>	Sinhalese	<01AcF>	Sinhalese	<01EF>	Tamil
<02AF>	Sinhalese	<02AcF>	Sinhalese	<02EF>	Sinhalese
<03AF>	Sinhalese	<03AcF>	Sinhalese	<03EF>	Sinhalese
<04AF>	Burgher	<04AcF>	Burgher	<04EF>	Burgher
<05AF>	Sinhalese	<05AcF>	Sinhalese	<05EF>	Sinhalese
<06AF>	Tamil	<06AcF>	Sinhalese	<06EF>	Sinhalese
<07AF>	Sinhalese	<07AcF>	Muslim	<07EF>	Sinhalese
<08AF>	Sinhalese	<08AcF>	Sinhalese	<08EF>	Sinhalese
<09AF>	Sinhalese	<09AcF>	Sinhalese	<09EF>	Burgher
<10AF>	Muslim	<10AcF>	Tamil	<10EF>	Muslim
<11AF>	Sinhalese	<11AcF>	Sinhalese	<11EF>	Sinhalese
<12AF>	Sinhalese	<12AcF>	Sinhalese	<12EF>	Muslim
<01AM>	Muslim	<01AcM>	Tamil	<01EM>	Muslim
<02AM>	Sinhalese	<02AcM>	Sinhalese	<02EM>	Sinhalese
<03AM>	Sinhalese	<03AcM>	Sinhalese	<03EM>	Sinhalese
<04AM>	Sinhalese	<04AcM>	Sinhalese	<04EM>	Sinhalese
<05AM>	Sinhalese	<05AcM>	Sinhalese	<05EM>	Sinhalese
<06AM>	Sinhalese	<06AcM>	Sinhalese	<06EM>	Sinhalese
<07AM>	Sinhalese	<07AcM>	Sinhalese	<07EM>	Sinhalese

<08AM>	Tamil	<08AcM>	Sinhalese	<08EM>	Sinhalese
<09AM>	Tamil	<09AcM>	Sinhalese	<09EM>	Sinhalese
<10AM>	Sinhalese	<10AcM>	Sinhalese	<10EM>	Muslim
<11AM>	Tamil	<11AcM>	Sinhalese	<11EM>	Muslim
<12AM>	Sinhalese	<12AcM>	Tamil	<12EM>	Tamil

Table 1: Speaker distribution according to ethnicity

APPENDIX E

The Checklist: Phonological features of SLE

Feature	Attested by
The absence of aspiration of word initial voiceless plosives such as /p/ and /k/	Fernando, 1985
Retroflex sounds where Std. E. would use alveolar plosives -eg: /t/ and /d/	Fernando, 1985
Devoicing of /z/ in word initial, word final and intervocalic positions.	Fernando, 1985
Use of dental plosives /t̪/ and /d̪/ in place of Std. E. /θ/ and /ð/	Fernando, 1985
Use of a flap articulation of /r/ than the frictionless continuant of Std. E. for initial	Fernando, 1985
The use of a clear /l/ in the final position where Std. E. would use a dark /l/	Fernando, 1985
The use of an approximant or a labio-dental frictionless continuant /ʋ/ for both Std. E. /v/ and /w/ initially	Fernando, 1985
The degree of lip rounding in the labial sounds generally, but particularly in /f/ and /w/. The degree is dependent on the scale of formality, although these consonants are never accompanied in SLE with as much labialisation as in Std. E..	Fernando, 1985
Doubling of a final consonant in a stressed syllable when it is followed by an initial vowel in the next syllable, accompanied by an absence of juncture over word boundaries. This is usually found in informal, friendly conversations. Eg: Come up- /kʌmmʌp/	Fernando, 1985

The long vowels in the Sri Lankan pronunciation are shorter than those heard in Received Pronunciation	Passe, 1948 (cited in Fernando, 1985)
Long vowels in SLE in place of diphthongs in Std. E.	Fernando, 1985
The use of /o:/ in place of /ɔ:/	Fernando, 1985; Ekanayaka, 2020
The use of short back rounded half-close vowel /ɒ/ in words like 'omit' where Std. E. would use /ou/	Fernando, 1985
The difference in the quality of diphthongs. Std. E. uses falling diphthongs, whereas in SLE the first element is only slightly more prominent than the second element. SLE diphthongs are usually shorter than the corresponding Std. E. sounds.	Fernando, 1985
The use of the diphthongs /ea/ where Std. E. uses /eə/ as in 'there'	Fernando, 1985
There are differences in the diphthongization of the triphthong /aue/ between Std. E. and SLE	Fernando, 1985
In the diphthongs /ai/, /ɔi/, /au/ the final element comes fairly close to the frictionless continuants /j/ and /w/ in casual colloquial speech, i.e. on the furthest point of the scale of formality. E.g. 'so how how?'	Fernando, 1985
A slight tendency to replace the middle element of triphthongs /aue/ and /aiə/ with bilabial or palatal frictionless continuants in casual colloquial style e.g. 'power'- /pavə/	Fernando, 1985
The use of /a/ for final a, ah in unstressed syllables where Std. E. uses /e/	Fernando, 1985

The non-use of the neutral vowel in weak forms of words like at, for, of, to, do etc.	Fernando, 1985
The tendency to use the neutral vowel /ə/, or a sound intermediate between the full vowel /e/ and /ə/ in all unstressed vowels in final syllables.	Fernando, 1985
Primary stress is placed on the first syllable of the word in SLE. Stress would typically be placed on the second syllable in BrE.	Meyler, 2007
Use of the diphthong /ai/ in SLE in lieu of /i/ or /I/ in BrE	Meyler, 2007

APPENDIX F

Age of the speakers

Table 1 shows speaker distribution according to age in this study. The ages that were guessed by what is mentioned as life experiences in their interviews and by their various profiles online are marked with an * on their birth year in the table 1 below.

Speaker ID (F/M)	(Approximate*) DoB and Age as of 2024 ()	Speaker ID (F/M)	(Approximate*) DoB and Age as of 2024 ()	Speaker ID (F/M)	(Approximate*) DoB and Age as of 2024 ()
<01AF>	1963 (61)	<01AcF>	1986* (38)	<01EF>	1969 (55)
<02AF>	1975 (49)	<02AcF>	1969 (55)	<02EF>	1988 (36)
<03AF>	1992 (32)	<03AcF>	1990* (34)	<03EF>	1964 (60)
<04AF>	1997 (27)	<04AcF>	1979 (45)	<04EF>	1965* (59)
<05AF>	1969 (55)	<05AcF>	1958 (66)	<05EF>	1994* (30)
<06AF>	1975 (49)	<06AcF>	1965 (59)	<06EF>	1975* (49)

<07AF>	1988 (36)	<07AcF>	1970 (54)	<07EF>	1997 (27)
<08AF>	1943 (81)	<08AcF>	1973*(51)	<08EF>	1972* (50)
<09AF>	1960* (64)	<09AcF>	1958 (66)	<09EF>	1985 (39)
<10AF>	1964 (60)	<10AcF>	1959 (65)	<10EF>	1960* (64)
<11AF>	1990 (34)	<11AcF>	1970* (54)	<11EF>	1995* (29)
<12AF>	1964 (60)	<12AcF>	1980 (44)	<12EF>	1989* (35)
<01AM>	1972 (52)	<01AcM>	1978 (46)	<01EM>	1985* (39)
<02AM>	1977 (47)	<02AcM>	1967 (57)	<02EM>	1949 (75)
<03AM>	1961* (63)	<03AcM>	1950 (74)	<03EM>	1976 (48)
<04AM>	1958 (66)	<04AcM>	1958 (66)	<04EM>	1964 (60)
<05AM>	1945 (79)	<05AcM>	1983 (41)	<05EM>	1961 (63)
<06AM>	1988 (36)	<06AcM>	1959 (65)	<06EM>	1961 (63)
<07AM>	1981 (43)	<07AcM>	1980 (44)	<07EM>	1982 (42)
<08AM>	1976 (48)	<08AcM>	1958 (66)	<08EM>	1965 (58)
<09AM>	1948 (76)	<09AcM>	1967* (57)	<09EM>	1960 (64)
<10AM>	1962 (62)	<10AcM>	1974 (50)	<10EM>	1963 (61)
<11AM>	1985 (39)	<11AcM>	1962* (62)	<11EM>	1965* (59)
<12AM>	1969 (55)	<12AcM>	1958 (66)	<12EM>	1991* (33)

Table 1: Speaker distribution according to age

All the participants distributed according to the social generation is given in Appendix A.

The following speakers are included in each social generation.

Baby Boomers (Birth year)		Generation X (Birth year)		Generation Y (Birth year)	
1946-1964		1965-1980		1981-1997	
<01AF>	<09AM>	<02AF>	<02AM>	<03AF>	<11AM>
<05AF>	<10AM>	<06AF>	<08AM>	<04AF>	<01EM>

<08AF>	<03AcM>	<02AcF>	<12AM>	<07AF>	<07EM>
<09AF>	<04AcM>	<04AcF>	<01AcM>	<11AF>	<12EM>
<10AF>	<06AcM>	<06AcF>	<02AcM>	<01AcF>	
<12AF>	<08AcM>	<07AcF>	<05AcM>	<03AcF>	
<05AcF>	<11AcM>	<08AcF>	<07AcM>	<02EF>	
<09AcF>	<12AcM>	<11AcF>	<09AcM>	<05EF>	
<10AcF>	<02EM>	<12AcF>	<10AcM>	<07EF>	
<03EF>	<04EM>	<01EF>	<03EM>	<09EF>	
<10EF>	<05EM>	<04EF>	<08EM>	<12EF>	
<03AM>	<06EM>	<06EF>	<11EM>	<11EF>	
<04AM>	<09EM>	<08EF>		<06AM>	
<05AM>	<10EM>	<01AM>		<07AM>	

Table 3: The list of speakers in each social generation

APPENDIX G

<C06EF> Shanila de Livera <06EF> / Tashya Segel <17HF>>

<17HF> welcome back to the Founder and today we are here once again this beautiful space the Colombo Cooperative and today I have with me Shanila from Foto Design and we are going to get to know a little bit about her story Shanila how are you today

<06EF> good and thank you for having me

<17HF> awesome and how has your morning been so far

<06EF> good thank you

<17HF> good okay awesome so tell me a little bit more about your company

<06EF> okay so erm Foto Design is the name of my company and we specialize in er custom picture framing so erm let me be brief and tell you my tag line which is art becomes personal once defined by a frame

<17HF> wow that's amazing

<06EF> so be it er painting a photograph an object whatever it is if you like it frame it

<17HF> if you like it frame it that's pretty awesome Shanila tell us when your company started

<06EF> okay so Foto Design started in nineteen eighty six er way before my time but erm it was started by my dad and erm it was actually a hobby of his which he I guess didn't have much time to run erm started with only laminating so when I took over it was two thousand and two I didn't really intend to take over my erm dad wanted me to close things up and erm before I erm pursued my line of study which was completely different to what erm I'm doing right now so erm I could say I erm started picture framing picture framing segment of Foto Design which is its main product now

<17HF> yes

<06EF> in two thousand and two

<17HF> okay that's amazing that you carried on the legacy now what would you say actually brought about the inspiration for this product

<06EF> okay so to be honest I wasn't I don't I didn't think of myself as a person who would own my business one day or have something of my own I always felt I had the attributes to be led by somebody always and be guided erm so there was one story though that caught my eye and that was the story of Anita Roddick and how she transferred her interest and her passion into her business er which is now known as the Body Shop

<17HF> Body Shop yes

<06EF> erm so one little story that she says is about she=how she visited Sri Lanka and she was travelling and she saw this village ladies use pineapple on their face and that got her thinking to do these natural products and that was an inspiration for her to start the Body Shop so that story excited me but then again I didn't really think about er er having my own place but when erm my dad told me to close this business down and I was looking into erm what I could do and suddenly out

of nowhere picture framing part came in to my head the fact erm that a picture frame can transform a painting or some whatever is valuable to the customer can transform it and give it new life erm that became my inspiration and a fun thing to do transferred into my passion so

<17HF> yeah what a beautiful inspirational er thing that you picked up from a a person like Anita Roddick right I believe one thing that a lot of people have a lot of difficulty in you know er they have ideas, everyone has ideas but but er putting it into practice is the challenge and what would you say erm brings about ideas for you

<06EF> okay so erm a lot of people have ideas but looking in what I feel is looking into the finer detail of implementing those ideas are probably not looked into because you get excited when you have a new idea you know and you just want to go and do it

<17HF> yeah that's true

<06EF> but erm factors such as er er the capital involved cost er finding the right people to carry on your idea erm knowing and studying your target audience all of that needs to be accounted into in order to pursue that idea that was born so erm I also believe in having your heart your total heart in it because when tough times come that is what will erm help you to carry on without giving up

<17HF> yeah

<06EF> so when you think of an idea I think just need to look into the big picture

<17HF> yeah and anybody can take an example out of that as well

<06EF> yeah

<17HF> so er tell me about you know one of the things that you had to do is that take over

<06EF> mhm

<17HF> your parents company

<06EF> mhm

<17HF> right

<06EF> mhm

<17HF> in two thousand and two I believe

<06EF> mhm

<17HF> and er obviously there would have been challenges

<06EF> mhm

<17HF> so what was the toughest challenge that you had to face

<06EF> okay so when I took the company over my dad wasn't really expecting me to take it forward

<17HF> okay

<06EF> cause we were just was thinking of closing it down because it was not his main business you know but I got excited when I saw this picture when I thought about this picture framing thing so it was a company that already had it's staff so changing that mind set of the staff they were quite rigid erm probably perceived me as a small brat

<17HF> okay

<06EF> maybe

<17HF> [laughter]

<06EF> erm also erm the company had a loan er so I was stepping into kind of a liability with this little dream of mine erm yeah and this constant thing in my mind am I doing the right thing er studied and I've come back to do to get into the corporate world erm so that balance was there those would be the challenges

<17HF> so challenges yeah we spoke about that but now tell me about what marketing strategies that you had to come you know face practically

<06EF> okay so my main marketing strategy is word of mouth erm this means I had to ensure that the product is of very good quality erm I must say I don't always get it right erm but at least it pushes me and er it makes me keep trying until I get the product right erm I always think if I if the client were to invite me over for a meal or for erm tea or something to their house if I sit there and I look at the wall and see that picture and I'm disappointed with it

<17HF> mm

<06EF> that's not going to make me happy

<17HF> [laughter]

<06EF> so I always think of it like that erm and er see you see when you do something that you really like and you are passionate about this is it's not a pain to do something like this it's actually very challenging and a lovely thing to do

<17HF> yeah what would you say has been the hardest decision you've had to make so far

<06EF> mm whether I should do this interview or not

<17HF> [laughter]

<06EF> no I was just joking er okay so okay so erm when Foto Design was quite new erm and I was trying to get my name out there we got approached by a high profile client

<17HF> okay

<06EF> okay I was so excited because I thought okay this is my chance to make it you know get out there and then he will go and tell everybody about me and then you know

<17HF> yeah

<06EF> you are quite sorted then so erm they they had a very short deadline it was a big challenge and erm we took it over we did it but however when we handed over the finished product I myself

noticed a small mistake done on our part but the client didn't notice it and neither did his team and they accepted it

<17HF> ok

<06EF> but this was bothering me

<17HF> [laughter]

<06EF> so er I erm I would I was toying with the idea of should I tell him erm and lose this opportunity of making it you know

<17HF> okay

<06EF> or should I just let it go unnoticed

<17HF> yeah

<06EF> kind of thing but then when I really thought about it I wanted I always wanted the foundation of my business to be built on honesty

<17HF> yeah yeah

<06EF> so I erm went ahead and I told him what happened

<17HF> okay

<06EF> er but he was so appreciative of it

<17HF> okay

<06EF> that in the end he erm continued coming to me got much more orders and encouraged all his colleagues to come to me as well so I got what I wanted out of the job

<17HF> okay

<06EF> but through a hard decision

<17HF> what a rewarding moment I would say

<06EF> it

<17HF> for being honest at the same time

<06EF> it really was a de=defining moment for me

<17HF> [laughter]

<06EF> in my work experience

<17HF> that's amazing

<06EF> yeah

<17HF> and er one of the things that would have been frustrating right so obviously frustration comes with any startups

<06EF> mhm

<17HF> what would you say has been your frustrating moment having a start up like this

<06EF> mhm frustrations

<17HF> [laughter]

<06EF> [laughter] okay so erm looking back I would say the most frustrating part for me was when I started

<17HF> okay

<06EF> and er because it's a new thing for me I studied the product and you know erm had been reading into my competition I know what they are doing and I feel like we are offering the same product for probably less cost and more personalized than anything but always the client would want to go to the more established er places because of I guess reliability

<17HF> mm

<06EF> I can't blame them for it

<17HF> yeah [laughter]

<06EF> but you know that was a bit frustrating that we could offer the same product erm but er people would always go to the bigger er and the long=the person who has been in the industry longer

<17HF> mm

<06EF> we were quite new at the time so those were those were little

<17HF> little frustrations

<06EF> yeah mm apart from that there is always frustrations like erm handling staff you know erm clients who don't who can't really erm communicate what they want so I have to kind of imagine you know but erm if I look at the big picture again I couldn't do without my staff I've a great team and or from my clients I learn so much so yeah

<17HF> obviously when it comes to er you know having your own start up there are a few challenges that you have to face like for example financing

<06EF> mm

<17HF> was financing for you er any easier

<06EF> erm being a small and medium enterprise I would erm I would say it's best to always avoid taking loans

<17HF> yeah

<06EF> but if you are pushed to do that I feel it's very important that you have a backup plan about how you are going to pay back

<17HF> mhm

<06EF> erm and be responsible about that just stick to that erm yeah

<17HF> yeah were the banks helpful for you at all

<06EF> erm banks are helpful

<17HF> [laughter]

<06EF> er if you cooperate with them and er work with them again I would say you have to have that payback plan

<17HF> yeah

<06EF> and er stick to it I think that would be

<17HF> so tell me Shanila what has er been something that you failed before you succeeded have you ever had a moment where you failed obviously

<06EF> erm when it comes to failure I actually think now because I've been running this business for almost er for seventeen years erm failure is something you have to go through in order to keep running a successful business erm it helps you to erm learn how to handle crisis to handle a cross section of clients to handle staff problems erm so er there have been many instances where I have failed and I thought should I just stop and give up but now looking back in retrospect I think that's what has made me and just to give a small example when I was during my pregnancy it's li=er it would be so difficult er because sometimes clients or you can't expect clients to think that you are reliable when you have so much going on in your personal life

<17HF> that's true yeah

<06EF> but erm all of those have just added to

<17HF> yeah

<06EF> my experience

<17HF> <u=?> yeah so what really gets you up in the morning

<06EF> what gets me up in the morning is the morning school run [laughter]

<17HF> [laughter]

<06EF> no but on a serious note erm I know I might sound like a cliché

<17HF> [laughter]

<06EF> er philosophical person but er

<17HF> that's no problem with that [laughter]

<06EF> [laughter] I believe in er I believe that everyone has a purpose in life

<17HF> that's right

<06EF> er and I'm very grateful that there is life

<17HF> mhm

<06EF> so I always think that every day is a new day to erm a new opportunity to erm be of some kind of influence to the people that surround you on a daily basis you meet your I meet my staff I meet even my family my children my er husband my friends erm when I drop my kids I meet parents and I think it's an opportunity to share something that we've been blessed with

<17HF> yeah

<06EF> or something good with someone else

<17HF> alright

<06EF> so

<17HF> yeah so stepping away from the seriousness I have a fun question for you [laughter]

<06EF> mhm

<17HF> what's the craziest thing you've done at office

<06EF> I don't think I can reveal

<17HF> [laughter]

<06EF> all the crazy things I've done in office

<17HF> that's what they all say [laughter]

<06EF> [laughter] but my friends know this story and they love it

<17HF> okay

<06EF> so I'll share it erm there was once a client who came in and I just assumed he couldn't hear and that he couldn't speak

<17HF> okay

<06EF> so I spoke to him in sign language the best=not that I know the sign language but in the best

<17HF> okay okay

<06EF> possible way that I could and he thought that I couldn't hear and I couldn't speak

<17HF> oh

<06EF> so he started talking to me in sign language as well

<17HF> [laughter]

<06EF> so we both had this five minute conversation in sign language until he resorted to say oh is it and then I got such a shock and I ran out of my office it was very embarrassed when my staff questioned me about it yeah

<17HF> but it is actually crazy it's more awkward than crazy

<06EF> mhm

<17HF> I would say but fun story what would you say is your definition of success

<06EF> erm I believe it's piece of mind

<17HF> piece of mind okay that's it short and simple

<06EF> mhm

<17HF> piece of mind is is very important

<06EF> that's it

<17HF> actually a lot of people forget mental health is also important very important er another thing is for the fact that you have your company now if you were to take over a company or a brand er other than what you are doing now what would it be

<06EF> okay it would definitely be barefoot [laughter]

<17HF> okay [laughter]

<06EF> It's my favourite brand in the whole wide world

<17HF> can I ask why

<06EF> I just love I love the concept

<17HF> mhm

<06EF> I love that they use local crafts people I love the colour but I wouldn't really want to take it over because I don't do I don't think I would do such a good job with it

<17HF> [laughter]

<06EF> so yeah

<17HF> so now considering that er you don't have

APPENDIX H

Speaker distribution according to age in CSSLE

Speaker ID (F/M)	(Approximate*)Date of Birth and Age as of 2024 ()	Speaker ID (F/M)	(Approximate*)Date of Birth and Age as of 2024 ()	Speaker ID (F/M)	(Approximate*)Date of Birth and Age as of 2024 ()
<01AF>	1963 (61)	<01AcF>	1986* (38)	<01EF>	1969 (55)
<02AF>	1975 (49)	<02AcF>	1972* (52)	<02EF>	1996* (28)
<03AF>	1992 (32)	<03AcF>	1990*	<03EF>	1964 (60)
<04AF>	1997 (27)	<04AcF>	1979 (45)	<04EF>	1965* (59)
<05AF>	1969* (55)	<05AcF>	1958*(66)	<05EF>	1994* (30)
<06AF>	1975* (49)	<06AcF>	1965 (59)	<06EF>	1975* (49)
<07AF>	1988 (36)	<07AcF>	1970 (54)	<07EF>	1977* (47)
<08AF>	1943 (81)	<08AcF>	1973*(51)	<08EF>	1972* (50)
<09AF>	1960* (64)	<09AcF>	1958 (66)	<09EF>	1985* (39)
<10AF>	1964 (60)	<10AcF>	1959 (65)	<10EF>	1960* (64)
<11AF>	1990 (34)	<11AcF>	1970* (54)	<11EF>	1964* (60)
<12AF>	1964 (60)	<12AcF>	1980 (44)	<12EF>	1989* (35)
<01AM>	1972 (52)	<01AcM>	1978*(46)	<01EM>	1985* (39)
<02AM>	1977 (47)	<02AcM>	1967 (57)	<02EM>	1949 (75)
<03AM>	1961* (63)	<03AcM>	1950 (74)	<03EM>	1976 (48)
<04AM>	1958 (66)	<04AcM>	1958* (66)	<04EM>	1964 (60)
<05AM>	1945 (79)	<05AcM>	1983* (41)	<05EM>	1961 (63)
<06AM>	1988 (36)	<06AcM>	1959* (65)	<06EM>	1961 (63)
<07AM>	1981 (43)	<07AcM>	1980 (44)	<07EM>	1982 (42)
<08AM>	1976 (48)	<08AcM>	1958 (66)	<08EM>	1966 (58)

<09AM>	1948 (76)	<09AcM>	1967* (57)	<09EM>	1960 (64)
<10AM>	1962 (62)	<10AcM>	1974 (50)	<10EM>	1963 (61)
<11AM>	1985 (39)	<11AcM>	1962* (62)	<11EM>	1965* (59)
<12AM>	1969 (55)	<12AcM>	1958 (66)	<12EM>	1991* (33)

