

Faculté de philosophie, arts et lettres

Discourse marker *Like* across the varieties of English in South and Southeast Asia: A comparative corpus-based study

Auteur : Siqi Xu

Promotrice : Prof. Dr. Gaëtanelle Gilquin

Année académique 2023-2024

Master [120] en linguistique, à finalité spécialisée: Linguistics
and its Applications for a Multilingual Society

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Gaëtanelle Gilquin, for providing valuable guidance and feedback for my thesis. Moreover, I extend my appreciation to all the professors at UCLouvain for their support and inspiration.

I wish to express my heartfelt thanks to my parents for their unwavering support and encouragement throughout my studies.

Finally, I would like to thank all my friends for their encouragement and companionship during the completion of this thesis.

Table of Contents

1. INTRODUCTION.....	5
2. THEORETICAL BACKGROUNDS.....	6
2.1 INTRODUCTION	6
2.2 DISCOURSE MARKERS.....	7
2.2.1 Introduction	7
2.2.2 Definition	8
2.2.3 Major approaches to discourse markers.....	10
2.2.4 Features.....	13
2.2.5 Functions.....	16
2.3 DISCOURSE MARKER <i>LIKE</i>	21
2.3.1 The history of like.....	21
2.3.2 The positions of like	27
2.3.3 The scope of like	29
2.3.4 The prosody of like.....	31
2.3.5 The functions of like.....	31
2.4 VARIETIES OF ENGLISH	38
2.4.1 Hong Kong English.....	39
2.4.2 Indian English	41
2.4.3 Singapore English	43
2.4.4 Philippine English.....	44
3. DATA AND METHODOLOGY	47
3.1 INTRODUCTION	47
3.2 CORPUS AND DATA.....	47
3.2.1 Introduction and justification for using ICE corpus.....	47
3.2.2 Description of data.....	49
3.3 CODING METHOD.....	50
3.3.1 Introduction	50
3.3.2 The selection of discourse marker like	51
3.3.3 The positions of discourse marker like.....	53
3.3.4 The functions of discourse marker like	54
3.4 ANALYSIS METHOD	58
3.4.1 Proportion of DM use of like and non-DM uses of like across varieties of English.....	58
3.4.2 Frequency of DM use of like across varieties of English.....	58
3.4.3 Distribution of DM use of like in different positions across varieties of English.....	59
3.4.4 Distribution of DM use of like in different functions across varieties of English	59

4. RESULTS AND DISCUSSION OF PROPORTION AND FREQUENCY OF DM USE OF <i>LIKE</i>	59
4.1 RESULTS OF PROPORTION OF DM USE OF <i>LIKE</i> AND NON-DM USES OF <i>LIKE</i> ACROSS VARIETIES OF ENGLISH	60
4.2 RESULTS OF FREQUENCY OF DM USE OF <i>LIKE</i> ACROSS VARIETIES OF ENGLISH	63
4.3 DISCUSSION	65
5. RESULTS AND DISCUSSION OF POSITIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i> ACROSS VARIETIES OF ENGLISH	67
5.1 DESCRIPTION OF POSITIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i> ACROSS VARIETIES	67
5.2 RESULTS OF THE DIFFERENCES ACROSS VARIETIES IN POSITIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i>	70
5.2.1 <i>Clause-initial like across varieties</i>	71
5.2.2 <i>Clause-medial like across varieties</i>	75
5.2.3 <i>Clause-final like across varieties</i>	78
5.2.4 <i>Non-clausal like across varieties</i>	82
5.3 DISCUSSION	85
6. RESULTS AND DISCUSSION OF FUNCTIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i> ACROSS VARIETIES OF ENGLISH	88
6.1 DESCRIPTION OF FUNCTIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i> ACROSS VARIETIES	88
6.2 RESULTS OF THE DIFFERENCES ACROSS VARIETIES IN FUNCTIONAL DISTRIBUTION OF DM USE OF <i>LIKE</i>	93
6.2.1 <i>Focus marker usage of like across varieties of English</i>	94
6.2.2 <i>Explanation marker usage of like across varieties of English</i>	97
6.2.3 <i>Repair marker usage of like across varieties of English</i>	101
6.2.4 <i>Exemplifier usage of like across varieties of English</i>	104
6.2.5 <i>Approximator usage of like across varieties of English</i>	106
6.2.6 <i>Linkage usage of like across varieties of English</i>	109
6.3 DISCUSSION	112
7. CONCLUSION	116
REFERENCES	118

1. Introduction

The discourse marker *like* has been extensively studied in various Inner Circle English varieties. Notable research includes Schweinberger's (2015) investigation of the discourse marker *like* in Irish English and southeastern British English, Tagliamonte's (2005) study in Canadian English, and Miller's (2009) work on New Zealand and Australian English.

Valentine (1991) highlights the need for more research on the pragmatics of Outer Circle varieties, which she refers to as "new English varieties." Such studies are crucial for understanding the discourse strategies and cultural norms within these English varieties. By treating these varieties as independent linguistic systems with their own unique features, we can gain deeper insights into their pragmatic aspects (Valentine, 1991: 333).

Historically, the discourse marker *like* in Outer Circle varieties has received less attention compared to its study in Inner Circle varieties. Research on its usage in Expanding Circle varieties is even scarcer. Recent studies have begun to address this gap, such as Rüdiger's (2021) examination of *like* in Korean English and Diskin's (2017) analysis of its use among Chinese and Polish migrants in Ireland. Despite this growing interest, there remains a limited amount of research on the use of the discourse marker *like* across different Outer Circle varieties.

This study aims to fill this gap by comparing the usage of the discourse marker *like* across four South and Southeast Asian English varieties: Hong Kong English, Indian English, Singapore English, and Philippine English. The following research questions are addressed:

1. To what extent do the four varieties differ in the proportion of the discourse marker uses of *like* versus non-discourse uses of *like*?
2. To what extent do the four varieties differ in the frequency of the discourse marker uses

of *like*?

3. To what extent do the four varieties differ in the positional distribution of the discourse marker uses of *like*?

4. To what extent do the four varieties differ in the functional distribution of the discourse marker uses of *like*?

The structure of this thesis is as follows: Chapter 2 reviews existing literature on discourse markers in general and provides an overview of the discourse marker *like* and the various English varieties. Chapter 3 describes the data and methodology used in the study. Chapter 4 presents an analysis of the proportion of *like* used as a discourse marker versus its non-discourse marker uses, along with a comparison of its frequency across the four varieties. Chapter 5 examines the positional distribution of the discourse marker *like*, while Chapter 6 analyzes its functional distribution. Finally, Chapter 7 concludes the thesis and offers suggestions for future research.

2. Theoretical backgrounds

2.1 Introduction

The following chapter aims to present an overview of the theoretical backgrounds concerning discourse markers, with a particular focus on discourse marker *like* and on varieties of English in south and southeast Asia. Section 2.2 is primarily concerned with the issues surrounding discourse markers. Section 2.3 mainly deals with the discourse marker *like*, including its historical development, its various positions, its scope, its prosody and, finally, its functions. Section 2.4 primarily presents an overview of the existing literature on the varieties of English in south and southeast Asia. The four varieties selected for analysis in the current study are Hong Kong English, Indian English, Singapore English, and Philippine English. This section

introduces the situation and development of the four varieties separately. Additionally, the current research and key findings on the use of discourse marker *like* in the four target varieties is demonstrated in this section.

2.2 Discourse markers

2.2.1 Introduction

Prior to undertaking a detailed examination of the discourse marker *like*, it is first necessary to define what the term ‘discourse marker’ refers to. First of all, the intricate nature of discourse markers gives rise to the different terminology observed in the literature on the subject. Indeed, different scholars employ plenty of different terms that are either closely or more distantly related to the terminology of the same concept. It is challenging to achieve consensus on this topic, as different terms are associated with disparate theoretical perspectives. Given their status as a nexus of different linguistic aspects, discourse markers present a challenging case for classification into a single, clearly delineated category. In addition to the conceptual ambiguity and lack of consensus surrounding this concept, discourse markers nevertheless exhibit certain common features from both the syntactic and semantic perspectives. Furthermore, the multifunctionality of discourse markers in spoken discourses cannot be overlooked. It is therefore inevitable that the debate surrounding discourse markers will continue, with a substantial body of literature emerging that addresses this topic from a multitude of perspectives.

The following section sets out to define and characterise discourse markers. Section 2.2.2 provides an initial overview of the terminology associated with discourse markers. Subsequently, it differentiates between one of the most contentious terms, ‘discourse particles’, and the term employed in the present study, namely ‘discourse markers’. Furthermore, this section examines the various definitions of discourse markers presented in the literature. Section 2.2.3 presents some major approaches applied to the analysis of discourse markers. Section 2.2.4 provides an overview of the principal features of discourse markers. The

aforementioned features will be employed as criteria for determining whether an instance is a discourse marker or not in the current study. Section 2.2.5 presents the principal functions that discourse markers can fulfil from both a textual and an interpersonal perspective.

2.2.2 Definition

The terminology used to describe discourse markers is one of the most controversial aspects of the field. Brinton highlights that there are more than twenty terms that have been coined to refer to it. (Brinton 1996, as cited in Schourup, 1999, p.228). The following terms are used to describe the same concept: ‘discourse operator’ (Redeker, 1991), ‘semantic conjuncts’ (Quirk et al., 1985), ‘discourse particle’ (Schourup, 1983; Siegel, 2002), ‘discourse connective’ (Blakemore, 1987), ‘pragmatic expression’ (Erman, 1987), ‘pragmatic marker’ (Fraser, 1988, 1990, 1996; Brinton, 1996; Erman, 2001; Aijmer, 2005), ‘pragmatic particles’ (Östman, 1995), ‘phatic connectives’ (Bazanella, 1990).

The occurrence of these different terms reflects the authors’ different research interests and emphasis. One of the most competitive terms for discourse markers is ‘discourse particles’. Many scholars differentiate between the two terms for various reasons, even though both terms are employed to denote elements with functions within discourse. According to Schourup (1999), firstly, ‘particle’ is categorized as a syntactic class. By contrast, the term discourse marker is more a functional class which comprises several syntactic classes (see section 2.2.4.6). Secondly, the referential scope of the term ‘particle’ itself is unclear. Regarding the use of the term ‘discourse particle’, ‘particle’ refers to “items that do not fit easily into any well-established word class” (Schourup, 1999: 229). Zwicky (as cited in Schourup 1999) notes that ‘particle’ is frequently “used to identify items that have idiosyncratic distributions and peculiar semantics” (Schourup, 1999: 229). In contrast, ‘discourse marker’ has a more restricted referential range. It is possible to find some commonalities such as “optionality” (see section 2.2.4.1), “non-truth conditionality” (see section 2.2.4.2) on nearly all these items that are regarded as ‘discourse markers’ (Schourup, 1999: 229). Thirdly, ‘particle’ predominantly

refers to scalar particles and modal particles (Schourup, 1999: 229; Risselada & Spooren, 1998: 131). Finally, the choice between ‘discourse particle’ and ‘discourse marker’ is often related to a different choice of approach (Risselada & Spooren, 1998: 131). Studies on particles tend to pay more attention to semantics, while studies on markers more often focus on discourse analysis, benefiting from corpus-based data.

Moreover, the definition of discourse markers is another issue that has not yet achieved consensus. Some definitions overlap, while others are controversial across different perspectives. Schourup (1999: 234) uses the three most common features of DMs to define them: “a syntactically optional expression that does not affect the truth-conditions associated with an utterance it introduces and is used to relate this utterance to the immediately preceding utterance.” Additionally, definitions can be made from either a grammatical or functional point of view. From the grammatical perspective, Zwicky (1985: 302) states that DMs comprise “a grammatically significant class of items, in English and other languages generally.” Another author who deals with DMs from a grammatical aspect is Fraser. Fraser (1988, 1990) defines discourse marker as: “(i) has a core meaning which can be enriched by the context; and (ii) signals the relationship that the speaker intends between the utterance the DM introduces and the foregoing utterance (rather than only illuminating the relationship, as Schiffrin suggests).” From a functional perspective, Schiffrin (1987) defines discourse markers as “sequentially dependent elements which bracket units of talk” (Schiffrin, 1987: 31). Redeker (1991) proposes a definition of DMs based on coherence as Schiffrin but criticizes a pragmatic bias in the term DM. She defines discourse markers as “linguistic signals of textual coherence links” (Redeker, 1991:1139). Generally, the terms and definitions from each framework are based on different selection of theories, and the desire to include or exclude certain elements. Moreover, due to such uncertainties, the debate surrounding discourse markers will continue, making it a domain with theoretical aspirations.

2.2.3 Major approaches to discourse markers

From the 1980s until the present day, discourse markers have been a topic of debate, attracting interest from various linguistic perspectives. As mentioned above, the choice of term and definition is based on different theories. Different approaches are used to analyze discourse markers. The two most common are coherence-based theory and relevance theory, which emerged in the 1980s. Subsequently, with the development of this field, grammatical-pragmatic theory and indexical theory began to be adopted. Finally, the use of corpus data to analyze discourse markers has become popular in recent years, and it will be used in the current study too. In the following section, these major approaches will be reviewed.

2.2.3.1 Coherence-based theory

Schiffrin (1987: 24) proposes that coherence is “constructed through relations between adjacent units in discourse”. According to Schiffrin, discourse markers aid in creating coherence by simultaneously establishing multiple contextual coordinates, which facilitate the incorporation of different elements of discourse (Schourup, 1999: 237). She also proposes a model called the five planes of talk: ideational structure, action structure, exchange structure, participation framework, and information state. Discourse markers are used to locate the utterance within one or several of these planes. For instance, *oh* signals the transitions of information state, it is thus firstly be related to the information state plane. Secondly, it is related to participation framework, as it indicates speakers and hearers in specific productive and receptive roles. Finally, it is deemed to be related to action structure, marking actions such as clarifications (Schiffrin, 1987: 316).

Schiffrin’s theory was later modified and developed by Redeker (1991). She revises Schiffrin’s five planes of talk to three components: ideational structure, rhetorical structure and sequential structure. This alternation aims at setting clearer boundaries of discourse markers. For instance, “non-anaphoric uses of *now* are excluded, and adverbial uses of *now* are included” (Schourup, 1999: 238).

2.2.3.2 Relevance theory

In the relevance theory proposed by Blakemore (1987), DMs are referred to as ‘discourse connectives’ on the grounds that the term DM conflates two discrete categories of expression: “those which are non-truth conditional but conceptual, and those which are both non-truth conditional and nonconceptual” (Schourup, 1999: 240). This approach to DMs is related to the Relevance Theory framework proposed by Sperber and Wilson (1986). Blakemore considers discourse connectives as conventional implicatures proposed by Grice (Fraser, 1989: 936). In this sense, the principal function of DMs is not to maintain coherence between utterances but to express an inferential connection. The emphasis here is on how relevance is fulfilled in the utterance.

2.2.3.3 Grammatical-pragmatic theory

Fraser is the founder of this grammatical-pragmatic theory. Fraser (1987) describes DMs as a subclass of ‘pragmatic markers’. Pragmatic markers are described as “linguistically encoded clues which signal the speaker’s potential communicative intentions” (Fraser, 1996: 168), instead of solely indicating the relationship as Schiffrin (1987) proposed. Moreover, Fraser (1987) defines DM as a linguistic expression that excludes non-verbal expressions such as *ah*. Besides, Fraser’s framework makes a strict distinction between content meaning and pragmatic meaning (Maschler & Schiffrin, 2015: 193). According to Schourup (1999: 239), Fraser’s approach sets a good standard for deciding which items to include or exclude, and also allows those excluded items to fit into other classes. Nevertheless, there are still opposing voices regarding his approach. Schourup (1999: 239) points out that, due to the limitation to connections between successive discourse segments, criticisms aimed at DM functions based on inter-utterance coherence have also influenced Fraser’s definition.

2.2.3.4 Grammaticalization theory

In grammaticalization theory, emphasis is placed on the diachronic development of a discourse marker from its lexical meaning to finally losing its semantic content and gaining pragmatic

meaning. In this process, grammatical changes also occur. Such notions are proposed by Aijmer (2002). Traugott (1995: 1) proposes that grammaticalization is “the process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts becomes grammatical.” Grammaticalization provides discourse markers with the capacity for multifunctionality (Aijmer & Simon-Vandenberg, 2011: 229). The term ‘pragmaticalization’ is used to describe the same phenomenon but focuses on the “semantic bleaching and pragmatic enrichment characteristic of the changes to pragmatic markers” (Erman & Kotsinas, 1993; Aijmer, 1997; Dostie, 2004, as cited in Aijmer & Simon-Vandenberg, 2011: 230).

2.2.3.5 Corpus linguistic approach

Thanks to the rapid development of computer technology, the use of corpora to analyze discourse markers has become increasingly popular in recent years. With the help of corpora, quantitative and qualitative analysis can be conducted at the same time to investigate discourse markers. According to Aijmer (2002:3), “corpora represent actual performance and provide the opportunity to study the distribution and function of particles in extensive text extracts representing different registers”. Aijmer (2002) uses the London-Lund Corpus to study the distribution of English discourse markers in different registers.

Moreover, comparisons can be made across languages and varieties using corpora. Müller (2005) compares the native and non-native use of English discourse markers. Fung & Carter’s (2007) comparison between Hong Kong English and British English use of discourse markers shows that Hong Kong learners use more discourse markers with a referential function, while British speakers use discourse markers for a wider range of functions. In Schweinberger (2015), the study compares Irish English and British English use of the specific discourse marker *like*. The results suggest that differences exist in frequency, social profile and positioning preference between these two groups. All these studies demonstrate the advantages of using a corpus-based approach to investigate discourse markers.

2.2.4 Features

Although there are different perspectives concerning the definition and terminology of discourse markers, they do share some common features. Some of these features are more predominant, while others may be more marginal or non-defining. The following section will review some of the main features related to discourse markers. Some of these features will later be adopted as standards to select the use of *like* as a discourse marker.

2.2.4.1 Optionality

Most studies regard discourse markers as optional in the syntax and semantics. That is to say, the removal of discourse markers will not affect the grammaticality of the utterance (Fraser, 1988, as cited in Schourup, 1999). Moreover, their use will not add extra meaning to the elements they link (Schourup, 1999: 231; see also Hölker, 1991). As Schiffrin (1987) points out, discourse markers are “syntactically detachable” from a sentence. The omission of discourse markers will not make the sentence “ungrammatical or unintelligible” (Brinton, 1996). However, this does not mean that discourse markers are useless. They serve as hints for listeners to interpret the utterance appropriately (Schourup, 1999). As Fraser (1988: 21) claims, discourse markers signal the discourse relationship between the two utterances they connect. Therefore, despite the syntactic optionality of discourse markers, they still play a role in implicit interpretation.

2.2.4.2 Non-truth conditionality

Non-truth conditionality is another important feature of discourse markers, which means that the occurrence of the discourse marker does not influence the truth condition of a proposition in the utterance. Levinson (1983: 87-8) characterises discourse markers as “words have at least a component of meaning that resists truth-conditional treatment” (see also Hölker, 1991; Schourup, 1999). Blakemore (2006) points out that there are other discourse indicators such as discourse adverbials like *frankly*. Therefore, she claims that non-truth conditionality is an important feature to distinguish discourse markers from other discourse indicators. She then

concludes that “their non-truth-conditionality has meant that they play a role in discussions of the non-unitary nature of linguistic meaning and the relationship between semantics and pragmatics” (Blakemore, 2006:222).

2.2.4.3 Weak clause association

Weak clause association is a feature that can be related to the syntactic optionality of discourse markers, as Brinton (1996) lists both syntactic optionality and weak clause association as syntactic features. According to Brinton (1996: 34), discourse markers are “either outside the syntactic structure or loosely attached to it.” Quirk et al. (1985) consider discourse markers (which they call conjuncts) to have a “detached and superordinate” relation with the clause (as cited in Schourup, 1999: 235).

It is worth noting that weak clause association often correlates with phonological independence (Schourup, 1999: 233). Brinton (1996) points out that they form a separate tone group, which indicates that they are phonologically weakly associated with the rest of the clause. The discourse markers selected in Hansen (1997) all fulfill the features of “remain external to the structure of the clause” and “tend to constitute independent tone units” (Hansen, 1997: 156).

2.2.4.4 Initiality

The consideration that discourse markers often occur in discourse initial positions has long been confirmed by many scholars, although it is not a defining feature. Schourup (1999: 233) proposes that “the tendency of DMs to appear initially is probably related to their ‘superordinate’ use to restrict the contextual interpretation of an utterance.” More specifically, it is discourse markers that decide the direction of the interpretation of an utterance. Hansen (1997:156) points out that discourse markers “prototypically introduce the discourse segments they mark.” Brinton (1996) claims that discourse markers are restricted to sentence-initial positions. Moreover, the initial position does not always mean the first word of the utterance; instead, it indicates “the position of DMs in relation to the central clause elements” (Schourup, 1999: 233) because discourse markers often gather at the beginning of the utterance and in

other places. Additionally, this initial position is syntactically loosely attached and an appealing position for interactional and textual planning (Aijmer, 2002: 29).

2.2.4.5 Orality

Brinton (1996) points out that discourse markers are an oral feature rather than written feature and are allied with informality. Nevertheless, they are not solely restricted to oral discourse; they do occur in written discourse with different forms depending on formality/informality (Fraser, 1990: 389; Redeker, 1990:379; Schourup, 1999: 234). For example, Schourup (1999: 234) points out that *also* is usually used in spoken discourse as a discourse marker, while *moreover* is more often used in written discourse as a discourse marker.

The reasons why discourse markers are used more often in oral discourses may have two possible explanations. Firstly, they are regarded as “a sign of the dysfluency and carelessness” (Brinton, 1996:33). They exhibit “a high degree of utterance planning” (Schourup, 1999: 234). That is to say, they serve as a method to give the speaker more time to think and plan the utterance. Secondly, the meaning of discourse markers “presupposes a familiarity with the addressee not typical of impersonally addressed writing” (Schourup, 1999:234). For instance, *you see* encodes that the listener has access to the background of the topic which the speaker is going to discuss.

2.2.4.6 Syntactic diversity

Syntactic diversity means that discourse markers could belong to diverse syntactic classes. This feature is related to their grammatical status, so it can also be named as multigrammaticality (Fung & Carter, 2007) or multi-categoriality (Schourup, 1999). According to Schiffrrin (1987: 64), they could be items that belong to various classes: adverbs (e.g. *now*, *anyway*), interjections (e.g. *oh*, *ah*, *gosh*), verbs (e.g. *look*, *see*), subordinate conjunctions (e.g. *since*, *so*), coordinate conjunctions (e.g. *and*, *but*), clauses (e.g. *you know*, *I mean*), literal phrases (e.g. *as a result*) (Schourup, 1999; Fraser, 1988).

The feature that discourse markers come from several different syntactic classes may be

evidence of their grammaticalization process (see section 2.2.3.4). In other words, discourse markers are initially elements from different word classes, and due to the unidirectionality of grammaticalization, they eventually all become discourse markers. Moreover, Heine & Kaltenböck (2021) focuses on the change from comment clause (e.g. *I think, I suppose, I expect*) to discourse markers and finds that grammaticalization is a process that accompanies such changes from their beginning to their final stage.

2.2.5 Functions

Discourse markers are usually said to be multifunctional, functioning at several linguistic levels simultaneously (Brinton: 1996:33). Due to such multifunctionality and the various approaches used to analyze them, a general agreement of their functions has not been reached until now. Jucker & Ziv (1998) suggest that discourse markers could function at textual, attitudinal, cognitive and interactional levels. Aijmer (2002:39), who distinguishes between textual and interpersonal functions, points out that “textual and interpersonal domains are not mutually exclusive.”

The following section will review the principal functions mentioned in former studies respectively from a textual and an interpersonal perspective. The first three functions from section 2.2.5.1 to section 2.2.5.3 are concerned with textual functions: discourse markers serve as connectors, topic switching devices, and repair markers. The latter four functions from section 2.2.5.4 to section 2.2.5.7 are related to interpersonal functions: discourse markers used to show attitudes, work as turn-taking devices, mark shared knowledge and raise participants’ attention.

2.2.5.1 Connectors

Connectivity has long been a typical feature of discourse markers. They are referred to as “connectors” (Jucker & Ziv, 1998:1), primarily indicating connections between utterances and discourse units. In Hansen (1977:160), discourse markers are defined as “linguistic items of variable scope, whose primary function is connective.” Similarly, Levinson (1983) emphasizes

that discourse markers “indicate, often in very complex ways, just how the utterance that contains them is a response to, or a continuation of, some portion of the prior discourse” (1983:197-8). Fraser (1996) also considers that the discourse marker “signals the relationship of the basic message to the foregoing discourse” (1996: 186).

Nevertheless, the connectivity of DMs is conceived from different aspects (Schourup, 1999: 230). Some may perceive this function from a “global” level, while others analyze it from a “local” level. For example, Blakemore (1987), who takes a global stance, believes that there is no need to distinguish between the discourse-internal uses and the discourse-initial uses of the same discourse marker. It is preferable to consider certain DMs as “relating the propositional content expressed by the current utterance to assumptions that may or may not have been communicated by a prior utterance” rather than as “relating two segments of text” (Schourup, 1999: 230-231). In contrast, from a local level, as Fraser (1988) points out, discourse markers may signal whether the current message aligns with or diverges from the context immediately preceding it. More specifically, Fraser’s definitions indicate that DMs establish a connection between two units of text, thereby enhancing inter-utterance coherence (Schourup, 1999: 230).

2.2.5.2 Topic switching devices

Discourse markers can serve as topic-switching devices from a textual perspective. They “signal links and transitions between topics” (Fung & Carter, 2007: 415). For example, in the following utterance:

- (1) but only one of them seems to cross the ankle joint and into the dorsal arch of the foot.
Okay. *So* now the next thing we really want to know is is there any way we could treat that stenosis? (CANCODE) (Fung & Carter, 2007: 422)

In (1), *so* is used to mark the end of the former topic and the tendency to shift to another topic (Fung & Carter, 2007: 421). In example (2), *Now* signifies the shift in topic to addressing how the previously raised questions can be answered:

- (2) OK so you're all happy with it. *Now* how are we going to approach it would anyone like to suggest a method? (CANCODE) (Fung & Carter, 2007: 422).

Moreover, discourse markers can signpost the beginning and ending of the talk, setting boundaries in discourse (e.g. *OK, right, well, let's start*), which is why sometimes they are called boundary markers (Brinton, 1996). Sinclair and Coulthard (1975) conduct a study on classroom discourse and finds that *well* is a frequently occurring marker that serves as a "frame" to indicate boundaries.

Finally, under this function, discourse markers can be used to indicate sequential relationships (Aijmer, 2002; Fung & Carter, 2007). This marks the logical sequence of the utterance, thereby facilitating comprehension. Expressions such as *firstly, secondly, then*, etc are employed in this way. In example (3), these markers are used to show the organised procedure and order of a debate.

- (3) So this is this is what we'll do.

Firstly introduce the speakers.

Yes.

Then introduce the topics of the debate and the main the main topics.

Er *thirdly* we'll give the reasons for actually having the debate in the first place (Fung & Carter, 2007: 421).

2.2.5.3 Repair markers

Discourse markers are sometimes used as so-called repair markers (Jucker & Ziv, 1998) to mark self-correction in the utterance. This is a typical function mentioned in Brinton (1996) and Aijmer (2002), although the former treats it at a local level while the latter treats it at a global level. Besides, it is often analyzed in a cognitive category as a reflection of the speaker's cognitive state (Fung & Carter, 2007: 415). Speakers sometimes may not express themselves clearly in unplanned speech or due to time constraints. Hence, DMs are tools for them "to

reformulate, rephrase, self-correct or repair their utterance” (Fung & Carter, 2007: 424).

It can be observed that there is always a frequency difference of DMs with such function between learner’s speech and native speaker’s speech. One typical example of such a function is *I mean* (Schiffrin, 1987). This function of DM is especially useful in an English as a lingua franca context as disfluencies or mistakes are obstacles for the speaker to correctly convey his/her thoughts (Kaur, 2011: 2711-2712). The use of *I mean* can promote the “clarity and explicitness in discourse organization” (Mauranen, 2007: 258, as cited in Kaur, 2011: 2712). Similar to *I mean*, so is another DM which provides “closure to an infelicitous part of the discourse” (Buysse, 2012: 1777). It is used more frequently in learner’s speech than in native speaker’s speech (Buysse, 2012).

2.2.5.4 Showing attitudes

This is a function of DMs at the interpersonal level. With this function, DMs usually indicate the speaker’s attitudes (e.g. *I think, sort of, like, just*) or show the speaker’s stance towards the truth of the proposition conveyed or the object being discussed (e.g. *actually, exactly, obviously*) (Anderson, 1998; Fung & Carter, 2007: 415). Speakers may use adverbs such as *actually, absolutely* to show their attitudes, as in the following example (4):

(4) I mean her main work’s on Pinter *really*. (Fung & Carter, 2007:419).

These are also called “commentary pragmatic markers”, which is a hypernym of DM in Fraser (1990:386).

In addition, to avoid face-threatening situations, speakers may use *sort of* as a softener of the utterance (Aijmer, 2002; Fung & Carter, 2007). Listeners may use DMs such as *yeah, oh, really?* to show their understanding or attitudes towards the speaker’s proposition (Jucker & Smith, 1998). Jucker & Smith (1998) points out that these reception markers aim to “provide differential feedback about the ease with which that information was integrated into the

receiver's state of knowledge" (Jucker & Smith, 1998: 197). The discourse marker *anyway* can be used to show both the attitude of the speaker and hearer (Takahara, 1998: 334).

2.2.5.5 Turn-taking

Discourse markers may influence the distribution of turn-taking or floor-holding (e.g. *now*) (Aijmer, 2002). For example, the DM *and* is used to signal the beginning of a conversational turn and always occur in the initial position (Stenström, 1994). Similar to Stenström, Tannen (1993) thinks that the turn-taking use of *and* aims at starting a new thought. *Cos* is usually a DM used to indicate one's attention to grasp the chance to speak (Fung & Carter, 2007: 423). Degand and Van Bergen (2018) conducts a study of instant messaging and shows that "the discourse markers in utterance-final position can function as a turn-transition mechanism, signaling that the turn is over, and the floor is open to the hearer" (Degand & Van Bergen, 2018: 47). Even in Kyratzis and Ervin-Tripp's study of young children's interactions, the turn-marking use of DM *but* is observed (Kyratzis & Ervin-Tripp, 1999: 1322).

2.2.5.6 Marking shared knowledge

Marking shared knowledge is an important function of DMs at the interpersonal level. Some similar terms for this function are "securing comprehension" (Erman, 2001) or "confirmation seekers" (Jucker and Ziv, 1998). *You know* is a typical marker of such a function. In Erman (1987), it is mentioned that *you know* "tends to be used in the presentation of given information or of what the speaker assumes to be known to the addressee" (Erman, 1987: 201). Similarly, in Fung & Carter (2007), *you know* "is often used to appeal to the assumed shared knowledge or experience of the speaker for the acceptance of information" (Fung & Carter, 2007: 419). Moreover, Jucker and Smith (1998) point out that *you know* can also be used when the information is shared not explicitly but implicitly. The speaker hopes the listener can construct meaning together by understanding the implication behind the explicit information. That is why it is considered as a confirmation-seeking device (Jucker & Ziv, 1998).

2.2.5.7 Raising participant's attention

In Redeker (1990), the main purpose of DMs is to draw the listener's attention to a specific connection between the forthcoming utterance and the current discourse context. In Bolden (2006), the discourse marker *so* is described as a preface of "other-attentive" topics or "recipient-attentive" topics. Its aim is to draw the listener's attention to the status of the discussing issues (Bolden, 2006: 670). Such DMs always focus on the operative process of the interaction and are thus called "operative markers", as pointed out by Trillo (2002). For example, *look* and *listen* are DMs used to raise the participant's attention to ensure the conversation progresses smoothly and without interruption (Trillo, 1997, as cited in Trillo, 2002: 777). Like Trillo, the study of Fung & Carter (2007: 419) uses *see* as example to illustrate such a function as "an utterance launcher". Moreover, they emphasize that *see* often "occurs in turn-initial position and signals that what follows is an explanation of what has preceded" (Fung & Carter, 2007: 419).

2.3 Discourse marker *like*

2.3.1 The history of *like*

Like is a frequently used discourse marker in everyday language known for its multifunctionality and informality (Buchstaller, 2001). This multifunctionality indicates that the grammaticalization process of *like* is ongoing (Romaine & Lange, 1991, as cited in Buchstaller, 2001). This section will review the history of *like* including its grammaticalization process and various meanings.

2.3.1.1 Grammaticalization of *like*

Like has undergone a long evolutionary history until now. Its origins can be dated back to 14th century according to the Oxford English Dictionary (OED) (Meehan, 1991). Numerous researchers have conducted studies focusing on the development of *like*. For instance, Meehan (1991) discusses the grammaticalization of *like* based on Lehmann's (1982) notion. Meehan's

model of the grammaticalization of *like* is entirely linear and unidirectional: from a “lexical item” to a “free grammatical morpheme”, eventually becoming an “affix” (Meehan, 1991: 37). The following examples will further explain Meehan’s (1991) model:

- (5) They stick’em right to the glass *like* Garfield cats (Meehan, 1991: 40).
- (6) Yeah... and *like* he’s really gonna be drownin’ his sorrows tonight (Meehan, 1991:44).

In example (5), *like* functions as a preposition and still has a strong lexical meaning. In example (6), *like* serves as a grammatical item to mark the focus (Meehan, 1991:45). Its semantic meaning is bleached enough in this usage (ibid.).

In contrast to Meehan’s model, Romaine and Lange (1991: 262) propose that “[a] simple linear model of grammaticalization is inadequate to account for these developments”. Their model for the grammaticalization of *like* is based on Traugott’s (1982) model. They posit *like* is initially used as a preposition, then as a conjunction, followed by its role as a discourse marker, and finally as a quotative complement (Romaine & Lange, 1991: 261). For example:

- (7) She looks *like* her father.
- (8) Winston tastes good *like* a cigarette should.
- (9) He brought along things for the picnic, food, drinks, and such-*like*
- (10) And there were *like* people blocking, you know?
- (11) Maya’s *like*, “Kim come over here and be with me and Brett.”

(Romaine & Lange, 1991: 244)

In example (7), *like* is used as a proposition. In example (8), *like* functions as a conjunction like “as”. However, such a use of *like* was rejected by The American Heritage Dictionary (1969) in 1950s. Nevertheless, the use of *like* as a conjunction is very common now in colloquial English (Romaine & Lange, 1991: 244). In example (9), *like* serves as a “denominal adjectivalizing suffix” (ibid.). It has a meaning of “having the qualities of” (ibid.). Originally, words ending in

-like was deemed as compounds. Gradually, they became independent suffix added after nouns (Jespersen, 1942, as cited in Romaine & Lange, 1991: 245). And this suffix use of *like* is regarded as the source of its discourse marker use (Jespersen, 1942, as cited in Romaine & Lange, 1991: 246). Example (10) illustrates the discourse marker use of *like*. Its function is to focus on discourse structure (Romaine & Lange, 1991: 245). In example (11), it is a quotative use of *like*. Romaine & Lange (1991) argues that these stages occur synchronically and do not happen sequentially (Romaine & Lange, 1991: 261-262).

However, Buchstaller (2001) critiques Romaine and Lange's branching model for its inability to "show the semantic-pragmatic link between functions at different ends of the channel" (Buchstaller, 2001: 22). Instead, she proposes a model that illustrates the multifunctionality of *like* and the simultaneous overlap of its functions (ibid). Her model in Figure 1, based on Lakoff's (1987) radial structure model, positions the comparative and approximative meanings at the center (Buchstaller, 2001: 32). The functional extensions, including quotative and focus marker functions, are placed to the left of the center, while epistemic hedge, pragmatic hedge, and filler functions are placed to the right. All these functions are interconnected with the center as well as with each other (ibid.). For instance:

(12) A: If someone slips on the ice outside a building, they could almost sue the architect for not having put ice-melting equipment in the sidewalk.

B: Yeah, into the sidewalk *like* heaters. But then... (Buchstaller, 2001: 24)

In example (12), *like* is used as an epistemic hedge to show the speaker B's uncertainty of whether *heaters* is an appropriate representation of the ice-melting equipment. This use of *like* is sourced from its central comparative or approximative meaning of "similar to" (Buchstaller, 2001:24). This is precisely the relationship between "a prototype and its less prototypical, specific instantiation of the category" (ibid.).

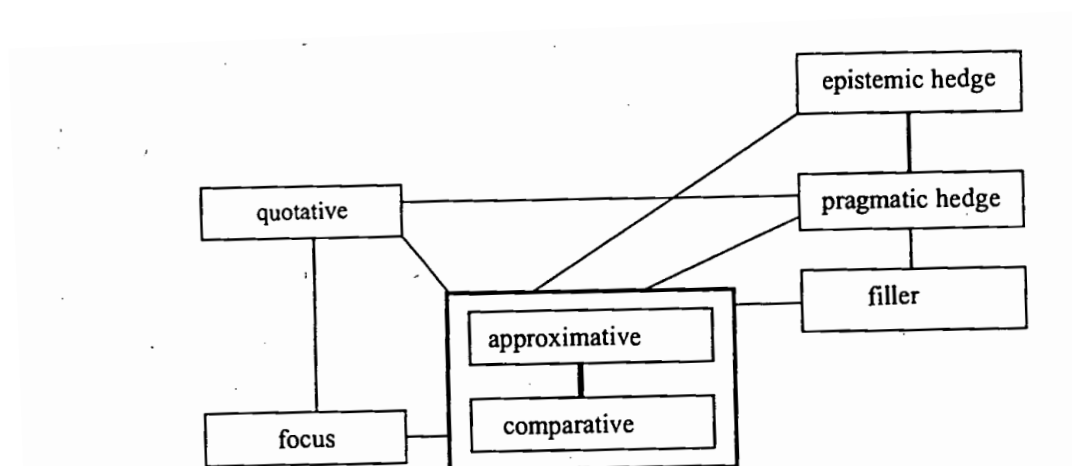


Figure 1. Buchstaller's model of *like* (Buchstaller, 2001:32)

In addition, D'Arcy (2005) also argues that the grammaticalization of *like* is unidirectional and proposes her model based on Traugott (1995) and Brinton (1996). She critiques Romaine & Lange's (1991) model for entailing significant shift in the constraints on this lexeme. Initially *like* has highly constrained syntactic positions as a preposition or conjunction. Then it suddenly has a great extent of freedom in syntactic positions and develops pragmatic function as a discourse marker without any intermediate stage (D'Arcy, 2005: 207). She supplements an essential intermediate stage of *like* as a sentential adverb. This is also what Traugott (1997) and Traugott & Dasher (2002) proposed. D'Arcy's (2005) unidirectional model begins with *like* serving as a preposition or conjunction, then progressing to a sentence adverb, followed by a discourse marker and ultimately evolving into a discourse particle (D'Arcy, 2005: 218-219). The following examples will illustrate this model:

- (13) It felt *like* everything had dropped away.
- (14) We need to smarten it up a bit *like*.
- (15) *Like* the week had already gone by.

(D'Arcy, 2005: 206)

In example (13), *like* is used as a conjunction. In example (14), *like* is used as a sentential adverb. In example (15), *like* functions as a discourse marker. The linkage between the

sentential adverb and the discourse marker is supported by “the adverb pre-dates the discourse marker” and “the discourse marker takes wider scope than the adverb” (D’Arcy’s, 2005: 207). Most importantly, this model connects what is typically deemed as British patterns (example 14) and American patterns (example 15) together (D’Arcy’s, 2005: 206).

2.3.1.2 Different meanings of *like*

A commonly accepted core meaning of *like* is comparison (Romaine & Lange, 1991; Buchstaller, 2001). Schourup (1983) proposes that the core meaning of *like* is “to express a possible unspecified minor non-equivalence of what is said and what is meant” (Schourup, 1983:31). In other words, *like* provides speakers with the opportunity to indicate a discrepancy between the actual discourse and their intended meaning.

Meehan (1991: 39) first proposes a “similar to” sense of *like* as per the OED, which “is an adverbial extension of the adjective use” and such a use can be found even in the 14th century. Then, she provides a more specific explanation of the “similar to” meaning, which is “approximately” (Meehan, 1991: 40). This specific meaning is often used in quantity phrases in modern English. Moreover, Buchstaller (2001) points out that this “similar to” meaning of *like* does not add much information to the speaker’s utterance and is therefore the most ideal particle “to fill a pause, and to hold the floor” (Buchstaller, 2001:25). Another specific explanation derived from the “similar to” sense is the “as if” meaning of *like*, which also demonstrates a sense of comparison (Meehan, 1991: 41). It has a broader scope than the “approximately” meaning (see Section 2.3.3).

Another meaning of *like* as “for example” first occurred in the 19th century, and this meaning equals the phrase “such as” (Meehan, 1991: 42). The meaning of *like* is become more generalized since then (Meehan, 1991:43).

After that, in constructions with *like* + adjective constructions or focus constructions, *like* functions more as a grammatical item that emphasizes the focus and can be completely omitted

(Meehan, 1991:45). Its influence on the meaning of the utterance is limited to placing less focus on a particular constituent of the utterance (ibid.). For example, consider the following two sentences:

(16) She's *like* really pregnant.

(17) She's really pregnant.

(Meehan, 1991: 46)

In example (16), the role of *like* is as a focus marker, serving as a free grammatical morpheme. Without it, as shown in example (17), the focus of the sentence shifts from the state of the subject to the subject herself (ibid.).

Another use of *like* mentioned in Meehan (1991) is highly associated with the quotations. This is also referred to as quotative *like*. The question of whether it is a discourse marker or not is a matter of contention. The following example serves to illustrate such a use of *like*:

(18) It's *like* 'How am I supposed to know?' (ibid.).

In this usage, *like* has a fixed position before an embedded clause and is therefore "more grammaticalized than the 'focus *like*'" (ibid.). A grammatical morpheme with a fixed position "is more likely to become an affix as it passes through the stages of grammaticalization" (Meehan, 1991:47). Indeed, in this context, *like* still originates from its "similar to" meaning due to the inaccuracy of the quoted information (ibid.). Romaine and Lange (1991: 227) offer a similar explanation, suggesting that the information following the quotative *like* reflects the speaker's thoughts or talk to himself/herself/themselves rather than an actual utterance. They regard such a use of *like* as a discourse marker because it shares some similarities with Schiffrin's (1987) definition for discourse markers (see section 2.2.2). They emphasize that such quotative usage of *like* enables the speaker to "retain the vividness of direct speech and thought while preserving the pragmatic force of indirect speech" (Romaine & Lange, 1991:

228).

However, this usage of *like* cannot be regarded as a discourse marker in the context of the current study. As optionality is one of the fundamental criteria employed in the present study for the identification of discourse markers (see section 3.3.2). Moreover, Buchstaller (2001: 26) points out that such a use of *like* does not fulfill the feature of being “syntactically detachable” as proposed by Schiffrin (1987) (see section 2.2.4.1). Romaine and Lange (1991) concede that the omission of the entire quotation is the only circumstance under which *like* can be omitted from the discourse (Romaine & Lange 1991: 246).

2.3.2 The positions of *like*

According to Romaine and Lange (1991: 261), *like* exhibits a “positional mobility”. However, Meehan (1991) points out that “the positioning of *like* in the sentence is never random”, as demonstrated in each of her examples where “the scope of *like* includes a post-verbal noun phrase” (Meehan, 1991: 40) (see section 2.3.3). Similarly, several scholars emphasize that the use of *like* is not free but occurs in specific positions with different pragmatic functions (Miller & Weinert, 1995; Underhill, 1988; D’Arcy, 2005; Tagliamonte, 2005). More specifically, Anderson (2001) distinguishes between clause-internal *like* and clause-external *like* (Anderson, 2001: 272-274, as cited in Schweinberger, 2015: 119). In Miller’s (2009) study, the position of *like* is divided into three different types: clause-initial, clause-medial and clause-final *like*. Therefore, this review will be based on a combination of Anderson’s (2001) and Miller’s (2009) categorizations, namely clause-initial *like*, clause-medial *like*, clause-final *like* and non-clausal *like* (Schweinberger, 2015: 121).

2.3.2.1 Clause-initial *like*

The clause-initial *like* connects the sequences of a dialogue (Fraser, 1988, 1990). This is one of the most frequent positions for *like* (Schourup, 1983; Underhill, 1988; Miller, 2009), and it is not surprising, as initiality is a typical feature of discourse markers (see section 2.2.4.4).

Clause-initial *like* is used to highlight new information (Miller & Weinert, 1995; Underhill, 1988). Moreover, the clause-initial *like* often serves a specification function (Miller, 2009; Schweinberger, 2015: 121), which is to further explain the preceding statement.

2.3.2.2 Clause-medial *like*

The clause-medial *like* also serves to exemplify the previous statement, similar to the clause-initial *like* (Miller & Weinert, 1995; Underhill, 1988). It can function as a focus marker (see section 2.3.1.2), indicating a focus on the elements immediately following it (Schweinberger, 2015: 122). In other words, it attracts the listener's attention to the subject within its scope. Additionally, it can also serve as a hedge, as proposed by Schourup (1983) (see section 2.3.1.2). Here, *like* functions more as a face-saving device rather than reflecting the actual discourse, as described by Buchstaller (2001) as an "epistemic hedge" (Schweinberger, 2015: 122).

2.3.2.3 Clause-final *like*

It is less common to find *like* in a clause-final position (Romaine & Lange, 1991; Miller, 2009). Miller (2009) suggests that this is probably due to its special function: not only does it highlight the information to its left, but also to indicate "the supply of and requests for explanations" (Miller, 2009: 328-329). Therefore, the clause-final *like* serves a particular explanatory function, which aims to avoid or revise incorrect interpretations and conclusions (ibid.). Besides, it is perhaps the oldest construction of *like* as it is attested in the novels of 1815 in OED (Miller, 2009), and D'Arcy (2005) and Schweinberger (2012) have shown that the grammaticalization of clause-final *like* occurred much earlier than the clause-initial and clause-medial *like* (Schweinberger, 2014: 72).

2.3.2.4 Non-clausal *like*

Non-clausal *like* refers to instances where *like* is "not integrated is not integrated into the underlying clausal structure and is motivated by psychological rather than pragmatic factors" (Schweinberger, 2015: 123). However, it still serves the pragmatic function of a floor-holding

device (ibid.). In this context, *like* functions similarly to other typical fillers such as *um* and *well*, giving the speaker more time to find the accurate word and plan their discourse (Siegel, 2002: 10). Therefore, this kind of *like* usually occurs alongside pauses, restarts, or repairs (Schweinberger, 2015: 123). Schourup notes that *like* occurs most frequently when there is lexical indecision on the part of the speaker (as cited in Siegel, 2002).

2.3.3 The scope of *like*

2.3.3.1 Scope related with the meaning of *like*

In Meehan (1991), the scope of *like* is defined as “the domain of the operation of the semantic elements” (Meehan, 1991: 40). The “approximately” meaning of *like* typically precedes a quantitative noun phrase, emphasizing the new information. The “as if” meaning of *like* is derived from the “similar to” sense, has a broader scope than the “approximately” meaning, in that the “as if” interpretation can encompass an entire clause. For instance:

(19) it cut *like* a hundred-mile path. (Meehan, 1991: 41).

(20) it was *like* I was watching someone else do it. (ibid.).

In example (19), *like* has the meaning of “approximately”, emphasizing the distance. The scope of *like* is the noun phrase. In example (20), *like* has the meaning of “as if”, compared with example (19), its scope includes the whole clause rather than a noun phrase.

The scope of the “for example” meaning of *like* is broader than that of the “similar to” meaning, as it can include a noun phrase, a prepositional phrase, or an entire clause. Here, the scope of *like* contains new information to which listeners should pay attention. For example:

(21) Do you have *like* a mint or something?

(22) It might be difficult if you start *like* in the fall.

(23) There are some other examples of past habituais *like... like* when we use...

(Meehan, 1991: 42)

In example (21), the scope of *like* covers a noun phrase; in example (22), the scope of *like* encompasses a prepositional phrase; in example (23), the scope of *like* covers an entire clause. All these examples exhibit the “for example” meaning of *like*. Moreover, the broader the scope, the more general the meaning becomes (Meehan, 1991: 43).

The scope of a focus marker *like* can range from a post-verbal noun phrase to an entire clause, and such a use of *like* is therefore considered as a free grammatical morpheme (Meehan, 1991: 46-47). The following examples serve to illustrate the scope of the focus marker *like*:

(24) We could do *like* sport comic books.

(25) *Like* it was 25 years ago when I had him as an instructor.

(Meehan, 1991: 44-45)

The scope of *like* in example (24) encompasses a post-verbal noun phrase and the scope of *like* in example (25) covers an entire clause. In these two examples, the use of *like* serves to “focus the listeners’ attention on specific information” (Meehan, 1991: 44).

It can be concluded that the newer meanings of *like*, such as the focus marker *like*, have a broader scope than the older meanings of *like*, such as the “approximately” meaning of *like*. In addition, the newer meanings of *like* are closer to a more general meaning (Meehan, 1991: 49).

2.3.3.2 Scope related with the position of *like*

The clause-initial *like* takes the scope over the discourse (D’Arcy, 2005) and extends to the entire clause or sentence (Siegel, 2002: 11), which is broader than the clause-medial *like* whose scope encompasses the elements to its right (Schweinberger, 2015: 122). In contrast, the clause-final *like* has an inverse scope compared with the clause-initial and clause-medial *like*, encompassing the elements to its left (ibid.) or having a scope over the entire preceding discourse but not beyond the sentence in which it occurs (Schweinberger, 2015: 132).

2.3.4 The prosody of *like*

The prosodic patterns associated with *like* functioning as a discourse marker has gained attention in recent years. When *like* is identified as a discourse marker, it exhibits a distinct prosodic pattern, akin to other discourse markers that “have a range of prosodic contours” (Schiffrin, 1987: 328). Drager (2011) highlights the existence of phonetic variations corresponding to changes in the function of *like*, i.e. thereby supporting the differentiation between discourse marker *like* and quotative *like*. Similarly, as discussed in Schleef and Turton (2018: 64), quotative *like* may gradually become less dependent on current prosodic and syntactic limitations. The consistent divergence in the speaker’s realizations of *like* based on its function implies that “these items are stored in the mind in a way that maintains the distinction” (Drager, 2011: 706).

The usefulness of prosodic cues in detecting the discourse marker *like* is also confirmed by Zufferey & Popescu-Belis (2004). They propose that the duration preceding *like* is another prosodic feature to label *like* as a DM, i.e. there is a typically relatively long pause before a discourse marker *like* (ibid.). As mentioned in Diskin (2017: 149), brief speaker silence often precedes clause-initial *like*, which is typically integrated into the same intonational phrase that follows it. Besides, a pause or silence following *like* usually signals its use as a DM (Zufferey & Popescu-Belis, 2004), predominantly observed in clause-final instances of *like* (Diskin, 2017). Clause-medial *like* is integrated “intonationally and semantically in the discourse” (ibid.). Non-clausal *like* (she calls it clause-between *like*) is frequently “accompanied by silence on both the left and right periphery” (ibid.), lacks semantic connections to the preceding or following discourse, and sometimes co-occurs with filled pauses (e.g. *um*) (ibid.).

2.3.5 The functions of *like*

Because of the multifunctionality of *like*, *like* occurs very frequently in colloquial English. This section will review the typical functions of *like* as a discourse marker discussed in previous

studies.

2.3.5.1 Highlighting focus

The highlighting focus function of discourse marker *like* has attracted the attention of numerous scholars. This highlighting function presents two possibilities. Firstly, it can emphasize the elements immediately following it as discussed in Meehan (1991) and Schweinberger (2015) (see section 2.3.2.2). A very common form is *like+adjective* (see example (16) in section 2.3.1.2). Secondly, it is recognized as marking focus on new information (Miller & Weinert, 1995; Underhill, 1988) (cf. section 2.3.2.1 above). Underhill identifies the primary function of *like* as marking “the most significant new information in a sentence” (Underhill, 1988: 238). The following examples serve to illustrate such a function of *like* as a discourse marker:

- (26) Student coming in for help on a homework assignment: I had problems *like* on the second question.
- (27) In an interchange between two students walking down the sidewalk: She had a research proposal *like* submitted to SIH.
- (Underhill, 1988: 238).

In example (26), the student had problems on the assignment is an existed information, *like* focuses on the concrete aspect of the problem. In example (27), the focus is left on *submitted to SIH* by discourse marker *like*, rather than on the research proposal itself.

According to Miller & Weinert (1995), *like* can be paraphrased using wh- or it- clefts, which “lend salience to specific constituents” (Miller & Weinert, 1995: 379). However, compared to wh- or it clefts, *like* has a weaker degree of focusing power due to its non-deictic nature and syntactic flexibility. Buchstaller (2001) notes that purely focusing uses of *like* do not exist, they are always accompanied by approximative or comparative implication.

2.3.5.2 Supplying and requesting explanation

Like is sometimes used to provide explanations to avoid ambiguities in certain circumstances. It has been observed that *like* is produced by instruction-givers to clarify misunderstandings and by instruction-followers to request “further information or for confirmation of an instruction” (Miller & Weinert, 1995: 379). Miller proposes that *like* is used to “anticipate an objection or to counter a proposition already expressed (and also to request an explanation)” (Miller, 2009: 336). In other words, when speakers perceive that their listener has drawn a wrong conclusion, they may offer further explanations to prevent misunderstandings (Miller, 2009). For instance:

(28) A: sorry I’m late my father’s in the Royal Edinburgh Royal Infirmary again

[The two listeners look at each other, knowing that A’s father has had serious heart surgery.]

A: it’s OK he’s just in for observation *like*. (Miller, 2009: 329)

In example (28), speaker A supply a further explanation ended by discourse marker *like*, indicating that he could know from the listeners’ body language that they had misinterpreted the situation.

This function of *like* is also frequently observed in Müller’s (2005: 217) study, where she identifies two sub-types. The first one is *like* appearing before a simple repetition form with almost the same words as the preceding utterance (ibid.). The second one is *like* introducing an explanatory reformulation with additional details to aid the listener better understand the meaning (ibid.). The following examples serve to explain these two sub-types:

(29) It said like poor and hungry later or whatever,

Mhm,

Like that he was poor and hungry.

(30) ...I I’d rather... say it was neutral for me *like* not good, not not really bad,

(Müller, 2005: 217-218)

In example (29), at first the speaker attempted to retell the precise title of a film but then failed. Therefore, the occurrence of *like* as a discourse marker serves to give an explanation that whatever the exact words are, the title of the film conveyed that the protagonist is “*poor and hungry*.” A simple repetition is turned into an explanation by the use of the discourse marker *like* (Müller, 2005: 217). In example (30), the speaker said “*neutral*” at first, then he or she changed the expression into “*not good, not not really bad*.” The use of discourse marker *like* serves to give an explanation which tends to help the listener to understand what the speaker wanted to express by “*neutral*.” The speaker does not repeat the same words before and after *like*, however, the content expressed before and after *like* are the same. Thus, *like* supplies an explanation to aid the listener understand the meaning more easily.

Another situation where discourse marker *like* is used is to request an explanation. For example, consider the utterance:

(31) Oh no I was going to develop that but um, ih I was just wondering what if you can give me a guide on what month *like* uh [ART ABCe1] (Miller, 2009: 334).

It is evident that in example (31), the speaker desires the interlocutor to provide more information on the target topic.

2.3.5.3 Restarting/Repairing discourse

In Schourup (1983), it is noted that there are many examples of *like* before a restart. For example:

(32) What we do is *like*- We did a uh (Schourup, 1983: 40).

In this case, the speaker interrupts the ongoing construction and begins again (ibid.). Schweinberger (2015) also highlights the restarting use of *like* in a non-clausal position (see

section 2.3.2.4).

Anderson (2001: 255) categorizes four more detailed usages of *like* as accompanied by false starts and self-repairs. Firstly, *like* functions as a restart where the elements preceding *like* are “syntactically unrelated to what follows it” (ibid.). For example:

- (33) But like it’s different if you’ve got a really bad cold and sometimes you have to, you can’t *like* ... sometimes you can hide it but I don’t go in front of someone. (Anderson, 2001: 255)

In example (33), a shifted opinion of the speaker is shown by the different structures before (*you can’t*) and after *like* (*you can*). The speaker restarts the cuts-off sentence by using the discourse marker *like*.

Secondly, *like* signals a self-repair when the speaker makes a minor correction, and the elements preceding *like* are syntactically related to what follows it (ibid.). Here, Anderson proposes that a new head of the verb phrase replaces the old one, while Müller (2005: 224) observes that there are cases where the structure is interrupted and reconstructed by the same lexical verb. The following examples serve to elucidate the distinctions with greater clarity:

- (34) it might be better to use *like*, just wait on the edge of like a jam or something like that, just let the ball come straight through (Anderson, 2001: 255).
- (35) The waiter keeps ta- *like* keeps tapping him (Müller. 2005: 224).

In example (34), the new head of the verb phrase (*wait*) after discourse marker *like* replaces the old head of the verb phrase (*use*). Nevertheless, in example (35), the utterance is cut off and repaired by the same lexical verb. Moreover, the example (34) can be distinguished from example (33) by the fact that the old and new lexical verbs are the “constituents of the same syntactic structure” (Anderson, 2001: 255).

Thirdly, *like* may occur when the speaker stops the utterance without further repairing or completing it. In this case, *like* marks a terminated utterance (ibid.). This is a typical clause-final use of *like* (ibid.). For instance:

- (36) Grace: Just tape conversations for school. Teacher wants to know about conversations,
like Dawn: Is it still running?
Grace: Yeah. (Anderson, 2001: 255)

Finally, *like* may connect two syntactically and thematically unrelated structures simply as a discourse link or a floor-holding device (see section 2.3.5.6) (ibid.).

2.3.5.4 Exemplifier

The exemplifier use of *like* has been proposed by many scholars. In this circumstance, one possible substitution for *like* is “for example” or “for instance” (Schourup, 1983; Meehan, 1991; Miller & Weinert, 1995; Jucker & Smith, 1998; Miller, 2009). Such a use of *like* can be illustrated by the following example:

- (37) A: My Dad’s very strict.
B: Is he?
C: like are there strict rules you know about coming in and all this?
(Miller & Weinert, 1995: 388)

In example (37), the speaker C requests for an example of the strictness of A’s Dad.

When used as a discourse marker, *like* is “an accurate but selective representation” (Schourup, 1983: 37). This use of *like* is to link the elements of prior discourse to what follows *like* (ibid.). Jucker & Smith (1998) also argue that the elements following *like* represent a typical situation rather than an actual situation. It shows a discrepancy between the situated utterance and the intended thought of this utterance (ibid.). This is the same when it functions as an approximator (see section 2.3.5.5) or hedge (see section 2.3.5.6).

2.3.5.5 Approximator

Like can be used as an approximator preceding a numerical expression. In this context, it can be substituted by “approximately” or “about” (Schourup, 1983; Buchstaller, 2001). For example:

- (38) so it was that was *like* three hundred dollars. [discourse marker; approximator] (Jucker & Smith, 1998: 183).

Buchstaller (2001) proposes that *like* indicates the approximate lexical choice and reduces the speaker’s responsibility. Moreover, *like* can also function as an approximator before other measurable units (Anderson, 1998: 153-154). For example:

- (39) and there’s *like* → [that much] gap between the earth and the top of the thing (ibid.).

In example (39), *like* occurs before an abstract measurable unit and still functions as an approximator. Besides, the approximator use of *like* can be deemed as a special case of hedge (see section 2.3.5.6 below) because it indicates a similarity between the reality and the intended thought (Jucker & Smith, 1998).

2.3.5.6 Planning discourse

Like can be used to plan the discourse from several aspects. On the one hand, it can be used to give the speaker enough time to search for the appropriate word, as in some unprepared speeches (Schourup, 1983; Buchstaller, 2001; Siegel, 2002). For example:

- (40) I have I take, um I’m taking math, and it’s *like* it’s called” (Buchstaller, 2001: 25).

The use of *like* in example (40) enables the speaker to fill the potential silence and imply there is more to come (ibid.). Therefore, it is a strategy used by speakers to hold the floor, it is a sign that the speaker wants to continue the talk (Schweinberger, 2015). In this use of *like*, it often appears before modifications or additions made by speakers who have difficulties organizing

their language but wish to continue their statement (Buchstaller, 2001). However, it is regarded as “hesitative” or “procrastinative” because of its informality (Schourup, 1983: 40).

On the other hand, *like* can be used as a hedge (Underhill, 1988; Buchstaller, 2001; Siegel, 2002; Schweinberger, 2015). In this case, *like* always occurs simultaneously with *kind of* or *sort of*. For instance:

- (41) Hockey is kind of *like* interesting too, ice-hockey. [discourse marker; hedge] (Jucker & Smith, 1998: 183).

In example (41), the speaker thinks hockey is an interesting sport but apparently not everyone has the same thought as the speaker. Therefore, both *kind of* and *like* are used as hedges to qualify the adjective *interesting*.

In Underhill (1988), this function of *like* is an extension from its numeric “approximately” meaning to non-numerical “sort of” meaning. This hedging use of *like* “leaves the statement slightly open” and “give an effect of politeness” (Underhill, 1988: 241). Moreover, Anderson argues that *like* is a “loose use of language” (Anderson, 1998: 152). Similarly, Jucker & Smith (1998) agree that *like* exhibits a looser closeness of fit between the actual discourse and the thought it represents than the listener may have expected. Therefore, they conclude that *like* is an “information-centered presentation marker” (Jucker & Smith, 1998: 191). In this case, the element modified by *like* may not be a perfect literal representation but is “good enough for current purposes” (ibid.).

2.4 Varieties of English

Kachru (1985) proposes a model for world Englishes with three circles: the Inner Circle, the Outer Circle, and the Expanding Circle. The Inner Circle Englishes include traditional native English varieties such as American English and British English. The Outer Circle Englishes comprise varieties of English used as a second language, including Singapore English and

Indian English. The Expanding Circle Englishes comprise varieties in which English is employed as a foreign language.

Schneider (2003) proposes a dynamic model for new varieties of English. He points out that “there is a shared underlying process which drives the formation of New Englishes, accounts for many similarities between them, and appears to operate whenever a language is transplanted” (Schneider, 2003: 241). He outlines five phases in the evolutionary process of new Englishes in the dynamic model: foundation, exonormative stabilization, nativization, endonormative stabilization and differentiation (Schneider, 2003). The four selected varieties in the current study, namely Hong Kong English, Indian English, Singapore English and Philippine English can be situated in different phases of Schneider’s (2003) dynamic model.

The following sections will commence with a concise overview of preceding studies examining the varieties of English spoken in South and Southeast Asia. Secondly, an overview will be provided of the existing research concerning the use of the word *like* as a discourse marker in the four target varieties included in the present study.

2.4.1 Hong Kong English

Different from Singapore English, Indian English and Philippine English, Hong Kong English has fewer “native speakers” (Hung, 2012; Schneider, 2007). More precisely, there are few people who use Hong Kong English in their daily communication. However, it is proposed that Hong Kong English has reached a more advanced stage, that is, phase 3 in Schneider’s (2003) model, with some phase 2 features still observable (Schneider, 2007: 133). This is because of the British has colonized Hong Kong tightly for a long time (ibid.).

Due to the relatively small number of “native speakers”, Hong Kong English has restricted distinctive lexical and grammatical features than other varieties of English (Hung, 2012). From the grammatical point of view, Hong Kong English exhibits features of over-generalisation, simplification (Hung, 2001a, as cited in Hung, 2012: 127). Incorporating the speaker’s first

language (Cantonese) and features of target language (English) is also a typical characteristic of Hong Kong English (Hung, 2012: 126). It is noteworthy that not only speakers with lower proficiency make a lot of mistakes, but also those who have higher proficiency exhibit plenty of non-standard features (Hung, 2012: 127). Sometimes, Chinese grammar can be used to explain some non-standard grammatical features in Hong Kong English (ibid.).

2.4.1.1 *Like*/Discourse markers in Hong Kong English

The study by Fung & Carter (2007) compares the usage of discourse markers among speakers of Hong Kong English and speakers of Native British English. The results show that in both varieties, discourse markers display their interpersonal, referential, structural and cognitive functions in oral communications.

In both Hong Kong English and British English, there are frequent use of discourse markers to “structure and organise speech on interpersonal, referential, structural, and cognitive levels” (Fung & Carter 2007). Furthermore, Fung & Carter (2007) point out that Hong Kong English speakers tend to use discourse markers with referential functions, such as *and*, *but*, *so* etc. In contrast, British speakers have a more balanced usage of discourse markers with the four functions.

In their study, the main functions of *like* as a discourse marker are interpersonal and cognitive usages. More precisely, the interpersonal function includes indicating the attitudes of the speaker. The cognitive usage serves to elaborate on the utterance by giving more details. The following is an example of *like* acts as a discourse marker from Fung & Carter’s (2007) data of Hong Kong secondary school student corpus:

(42) A: What what type of game

B: *Like* the Mono Monopoly game is popular in Hong Kong nowadays.

A: do you...

In example (42), the discourse marker *like* is used as an exemplifier (see section 2.3.5.4) to express speaker B's suggestion. The usage of the discourse marker *like* is situated in the sixth position among the 23 selected DMs in the Hong Kong data. The frequency of *like* as a discourse marker is situated in the ninth position in the British data. Based on their contrastive standard, the result shows a comparable use of *like* between the two groups. That is to say, the frequency of *like* in Hong Kong English is similar to its frequency in British English.

2.4.2 Indian English

Concerning the speaker numbers, "Indian English is one of the most important varieties on the globe" (Schneider, 2007: 161). Indian English has been viewed as a phase 4 variety in Schneider's (2003) dynamic model, and still with some features of phase 3 (Mukherjee, 2007). However, Indian English is only mastered by a minority of local people (Schneider, 2007: 161). Therefore, the application of the dynamic model is only to those parts of society that has been infiltrated by English, not the whole society of India (ibid.).

Indian English is strongly localized (ibid.). This is because in the case of learners of English in India, the linguistic models and teachers that they have typically interacted with have been other Indians, rather than native speakers. (Gupta 2001: 154; Bhatt 2004:1017-1018; Gargesh 2004: 994, as cited in Schneider, 2007: 167). In addition, language mixing is very common in India: a mixture of English and Hindi can be found in daily life, such as the advertisements of McDonald's (Schneider, 2007: 170-171). Therefore, It is not surprising that "nativized interactional markers play an important role in the structure of discourse" (Valentine, 1991: 331). Sahgal (1991: 303) finds that "English is treated not as a foreign language but as part of the cultural identity of India" (as cited in Schneider, 2007: 171). Moreover, a Standard Indian English has been accepted by lots of people (Gupta, 2001: 155, as cited in Schneider, 2007: 171).

2.4.2.1 *Like*/Discourse markers in Indian English

There are relatively fewer studies concerning discourse marker *like* in Indian English. Valentine (1991), Siemund et al. (2009) have noticed this vernacular use of English. Valentine (1991) finds that Indianization is a typical feature in the spoken English of Indian, not only concerning its grammatical and syntactic features but also can be observed in its pragmatics as discourse markers.

Siemund et al.'s (2009) study classifies the Outer Circle varieties of English (they call them non-standard varieties) into group A and group B according to the positional distribution of the DM *like* within the varieties. Group A indicates that occurrences of *like* appear more frequently in clause-marginal positions, that is, clause-initial and clause-final positions. Group B refers to *like* occurring more often in clause-medial position. In their study, the situation of *like* in Indian English belongs to Group A. It is worth noticing that another variety of English which also belongs to group A is Irish English (they call it Hiberno English) (ibid.). Therefore, concerning the positional distribution of *like* in Indian English, it has a similarity to Irish English.

Parviainen (2012) points out the preference of clause-final position for discourse markers in Indian English. Moreover, this trend has been a local innovation in India which has spread to Philippine, Singapore and probably Hong Kong English (ibid.). This usage is positively correlated with the informality of the discourse (ibid.).

Like is one example of a discourse marker that seems not to have a nativized source in Indian English (Valentine, 1991: 332). Because it is neither a discourse marker from Hindi, nor a discourse marker with innovative usage in Indian English. Although *like* in Indian English is neither as frequent as its use in American English nor has as many functions as in American English (ibid.). It is a discourse marker used by all Indian speakers in her study, exhibiting frequent occurrences in Indian English (ibid.). For example:

(43) *Like* for example even you have it like surname

- (44) *Like* if *like* an *like* suppose uh these days there are too many conditions regarding my marriage
(Valentine, 1991: 333)

Discourse marker *like* in example (43) is a redundancy of the expression *for example*. While in example (44), *like* functions as a filler to hold the floor. From a functional perspective, *like* in Indian English reveals “some of the meaning sense of superficiality” (Valentine, 1991: 332).

2.4.3 Singapore English

Singapore English has reached phase 4 and is likely to finish the whole process in Schneider’s (2003) dynamic model. As Schneider (2007) points out, Singapore English is far advanced due to its unique language policy, which is called “English-based bilingualism” (Tickoo, 1996: 438, as cited in Schneider, 2007: 153). Four official languages are admitted in Singapore: Mandarin Chinese, English, Tamil and Malay. English is the first language taught to the children, and his/her ethnic language among the three other official languages is taught to the children as a “mother-tongue Second language” (Foley, 1998: 130, as cited in Schneider, 2007: 156). Therefore, a lot of people use Singapore English in their daily communication (Schneider, 2007: 157).

Ansaldo (2009) points out that Malay has a great influence on the formation of Singapore English due to the “founder effect” (Mufwene, 2005). Nowadays, English has a special status in Singapore, as it is a language transcend the ethnic boundaries (Schneider, 2007: 156). Rubdy (2001: 334) points out that “English educated Singaporeans are seen as being potentially less ethnocentric and demonstrating greater loyalty to Singapore” (as cited in Schneider, 2007: 156). These processes lead to the occurrences of those unique forms in Singapore English from informal uses to formal uses (Schneider, 2007: 157).

2.4.3.1 *Like*/Discourse markers in Singapore English

“The spoken data indicate specific preferences and dispreferences for discourse markers in

Singaporean English and suggest sociolinguistic variation and potentially diachronic change” (Deuber et al., 2018: 80).

Sand (2013) conducted a study to investigate whether the use of computer-mediated communication, such as weblogs, promotes the written use of Singapore English. The results reveal that *like* does occur in the contents of bloggers but has a relatively lower frequencies than in conversations. According to the data from Corpus of Singapore weblogs, *like* as a discourse marker and with the function of hedging or focus marker accounts for 81.7%, while its quotative use has a proportion of 18.3%. The following examples serve to illustrate the usage of discourse marker *like* in Singapore English:

(45) I was *like* very shocked like he really went to the room and take out stuff and *like* showed us *like* eh

(46) ..., the volcano in Iceland erupted and air travel was disrupted for *like* forever?!!!

(Sand, 2013)

In example (45), the occurrences of discourse marker *like* are as typical hedges or fillers. In example (46), it is obvious that *like* serves as highlighting focus on the constituent after it.

Another unexpected result is that no significant differences occur when taking the factors of sex, age and level of education into account (ibid.). The only significant predictor is the home language (ibid.). These results need more investigations in future studies, as pointed out by the author. Moreover, in Singapore spoken English, discourse marker *like* is common, at least in teenagers’ discourses (Deuber et al., 2018: 80).

2.4.4 Philippine English

In comparison to the aforementioned varieties of English, Philippine English exhibits slight differences, having been shaped not by British colonialism but by American expansion (Schneider, 2007: 140). Despite its relatively brief history, Philippine English has been spread

at an astonishingly rapid rate (ibid.). Nevertheless, as a consequence of the language policy that designates the indigenous language as the national language, the status of English appears to shift from that of a second language to a more marginal position (ibid.).

2.4.4.1 *Like*/Discourse markers in Philippine English

Some studies of the use of discourse marker *like* in Philippine English have been conducted, mainly focusing on its functions (Walrod, 2006; Malicsi, 2010; Palacio & Gustilo, 2016). Siemund et al. (2009), as mentioned above (see section 2.4.2) they conducted a study mainly investigates the positional distribution of the discourse marker *like* within certain varieties. They point out that the use of discourse marker *like* in Philippine English belongs to group B, where the predominant position is clause medial. Moreover, Walrod (2006: 8) points out that discourse marker *like* often occurs before a verb or between clauses in Philippine English.

Palacio & Gustilo (2016) conducted a study of discourse markers in Philippine Computer Mediated Communication reveals the functions of discourse marker *like* in Philippine English. It serves as a tool for linking sentences and discourses, ranking tenth among the twenty most frequently used discourse markers in the relational category. The functional categorization of *like* is similar to what is reported in Fung & Carter (2007), where *like* can be categorized in the interpersonal category to show the speaker's attitudes in Hong Kong English. It is observed that *like* is used to "relate or show solidarity with the interlocutors" (Palacio & Gustilo, 2016: 13).

However, the cognitive use of *like* as an exemplifier in Fung & Carter (2007) is not found in Palacio & Gustilo's data. Malicsi (2010) conducted a study of Philippine English to prove it is a case of language drift. He proposes that the term "language drift" is used to describe the "random changes in forms and rules that diffuse throughout a speech community through cultural transmission" (Malicsi, 2010: 29). These changes "become regular and systematic", especially if the diffusers responsible for the diffusion of these changes are deemed as English

exemplars within the community (ibid.). Some of the data presented in this study has been derived from the transcription of participants at a National Convention. The participants are all individuals who are widely regarded as “reputable professionals in the English language” in Philippine. Their speech can be seemed as a diffusion of the linguistic forms and rules that they have produced (Malicsi, 2010: 39). It is noteworthy that two of the three instances of non-standard rhetoric are discourse marker *like* used as a redundancy with “for example” and “for instance” (Malicsi, 2010: 41). It is therefore necessary to conduct further studies in order to ascertain whether discourse markers *like* serves the function of exemplifiers in Philippine English.

Moreover, *like* can function as a hedging device in Philippine English, as observed in Palacio & Gustilo (2016). They gloss this usage of *like* as enabling the speaker to “sound less direct in addressing frustration and annoyance to the other interlocutor” (Palacio & Gustilo, 2016: 11). This function of *like* is also attested in Walrod (2006), whose study focuses on the influence of discourse markers on textual meaning. The data of this study come from speakers from northern Philippines. In this study, *like* is said to have the role as “a slight disclaimer as to the exact accuracy or exact reliability of the information” (Walrod, 2006: 6). It is reported that *like* can even be used up to three times per sentence, for example:

- (47) And we had a totally great time and reconnected and everything and when I’d left, *like*, unbeknownst to me, *like*, he was, *like*, really sure he wanted to get married, but we hadn’t talked about it at all or anything; I was still transitioning from Africa (Walrod, 2006: 6).

In example (47), when there are some critical turning points, *like* occurs very frequently as a hedging device.

3. Data and methodology

3.1 Introduction

This chapter introduces the data and methodology employed in the current study. Section 3.2 introduces the corpus and data selected for analysis. Section 3.2.1 gives a comprehensive overview of the International Corpus of English and justifies the reason for its selection. Section 3.2.2 presents an overview of the data utilised in the study. Section 3.3 outlines the encoding methodology. Firstly, section 3.3.1 gives a brief introduction of the software used to extract *like* automatically from the corpus, and clarifies the non-pragmatic uses of “like” that are excluded from the analysis. Secondly, the criteria for identifying the discourse marker use of *like* are clarified in section 3.3.2. Thirdly, the classification criteria for the positions (section 3.3.3) and the functions (section 3.3.4) of *like* are discussed in detail. Finally, section 3.4 presents the further quantitative and qualitative analysis methods.

3.2 Corpus and data

3.2.1 Introduction and justification for using ICE corpus

Data used in the current study are derived from the International Corpus of English (ICE). The data collection of ICE began in 1990 with the goal of compiling data for comparative studies of global English varieties. Both spoken and written English were collected by twenty-six research teams for their own national or regional variety of English. Each team follows the same corpus design criteria and the same annotation scheme to ensure compatibility across varieties.

Each sub-corpus contains 500 texts, of which 300 texts are spoken data and 200 texts are written data. Each text contains approximately 2000 words, thus a total of 600,000 words of spoken data could be obtained. The spoken corpora comprise two parts: dialogues and monologues, with 180 texts in dialogues and 120 texts in monologues. The 180 dialogues consist of 100 private discourses (90 face-to-face conversations and 10 phone calls) and 80

public discourses (20 classroom lessons, 20 broadcast discussions, 10 broadcast interviews, 10 parliamentary debates, 10 legal cross-examinations and 10 business transactions). Since the discourse marker *like* is most likely to occur in the informal conversations (cf. section 2.3.1). The focus of the current study is on the dialogues, specifically the face-to-face conversations within the private section of dialogues. Due to the time-consuming nature of manual data coding, it was not feasible to code all the texts. Therefore, only the top fifty texts from the face-to-face conversations in the four sub-corpora were selected for the data encoding process. The results indicated that this sample size was sufficient to yield reliable findings.

The decision to utilise the ICE corpus in the current study is based on two primary considerations. First, it provides a large amount of authentic data for the study of English varieties. Since it comprises inner circle and outer circle varieties and thus is valuable for studies aiming to compare varieties. Moreover, it includes spoken data, which is especially relevant for the current study, as orality (cf. section 2.2.4.5) is an important feature of discourse markers. As mentioned before, all the sub-corpora are compiled with the same criteria, which makes it suitable for comparative studies.

The second reason for using ICE corpus is its accessibility. It can be obtained from the Zurich University platform (<https://www.ice-corpora.uzh.ch/en/access.html>). After the completed license is sent to ice@es.uzh.ch, the corpora can be downloaded and unpacked within 10 days. So far, hundreds of studies have used the ICE corpus for their research. Several studies have used it in the comparative study of discourse markers. For example, Miller (2009) used ICE corpus to compare the usage of *like* and other discourse markers in Australian and New Zealand English. Wang (2020) used ICE-Hong Kong and ICE-Great Britain to investigate the use of *I think* by native and non-native speakers. These various studies demonstrate the suitability of the ICE corpus for comparative studies of discourse markers. However, research comparing outer circle varieties within this corpus remains relatively limited.

3.2.2 Description of data

Table 1. Basic Description of Texts

	HKE	IndE	PhiE	SinE
Number of texts	50	50	50	50
Total number of words	121076	105166	108125	107408

Table 1 provides a basic information of the texts used in the current study. It can be observed that the number of texts of each variety is the same, which makes it possible to compare the usage across different English varieties. Each text is produced by different people, therefore, the data among different groups are balanced. As mentioned in the markup manual for spoken texts in ICE, utterances produced by non-corpus speakers are considered extra-corpus material. For instance, utterances from Americans in the Hong Kong corpus. These utterances were included in the corpus when they were essential for the context. All those extra-corpus texts were marked between the two symbols: <X> signals the beginning of the extra-corpus content and </X> signals the ending of the extra-corpus content. The extra-corpus speaker was identified as <\$Z>. However, these utterances were not tagged or parsed in the corpus thus they were discarded in the total number of words. Moreover, the occurrences of *like* produced by these non-corpus speakers were excluded from the current study because they could not represent the real situation in those selected varieties. After loading the data into the #Lancasbox v.6.0 software, the total number of words for each variety was displayed automatically by the software.

3.3 Coding method

3.3.1 Introduction

The #Lancasbox v.6.0 software was used to extract data from the four corpora. Firstly, all the selected data were loaded into the software. Then, the number of files, tokens, types and lemmas were automatically calculated by the software. Secondly, using the Key Words in Context (KWIC) function, all occurrences of *like* were extracted in concordance with its context on both the left and right. The lists of all instances of *like* were then downloaded into Excel files. However, not all returns were relevant. The automatic results from the software included all instances of “like”, and not all of these instances function as discourse markers. The following examples illustrate how “like” is used in contexts other than as a discourse marker:

a. *Like* as a verb

(48) I I have no reson to to *like* this place (ICE Hong Kong: S1A-001:19)

b. *Like* as a preposition

(49) I I think uh she really looks *like* a Chinese rather than a Japanese she (ICE Hong Kong: S1A-006: 94)

c. *Like* as a noun

(50) So that was your *like* (ICE Philippines: S1A-019:357)

d. Quotative *like*

(51) When we were really stuck then I’d go to him *like* sorry can’t find it etcetera etcetera (ICE Philippines: S1A-009: 181)

Therefore, a manual disambiguation is conducted to ensure that only the instances of *like* functioning as a discourse marker are included in the data, which is the focus of the current study. Section 3.3.2 outlines the criteria for selecting the discourse marker *like*, while Sections 3.3.3 and 3.3.4 describe the encoding process for the positions and functions of discourse

marker *like* separately.

3.3.2 The selection of discourse marker *like*

The first step of coding involved selecting only the discourse marker *like*. Each instance of *like* was examined to determine whether it functioned as a discourse marker. It is important to note that this examination was conducted within the limited context automatically displayed by the software. In cases where there was insufficient context to determine whether *like* was a discourse marker, additional context was consulted to ascertain its usage in the text. If a decision regarding the classification of *like* as a discourse marker remained uncertain, that instance was excluded from the data set. For example:

(52) A: Oh I don't like, to have, different stuff inside

I just *like*, you know

B: Plain

A: Plain (ICE Hong Kong: S1A-009: 146)

In this example, *like* can function either as a discourse marker to gain additional time to organise the discourse, or as a verb to express the preference of the speaker.

The selection of discourse marker *like* was based on characteristics of discourse markers mentioned in section 2.2.4, namely optionality and non-truth conditionality. First, *like* was removed from the discourse, and then the discourse was subjected to a syntactic assessment to ascertain its continued acceptability. If the correctness of the sentence was affected, the instance was rejected. For example:

(53) It's it's it's something *like* powder Do you drink it Yes I (ICE Hong Kong: S1A-001: 7).

The omission of *like* in example (53) renders the sentence syntactically incorrect. It can therefore be concluded that the term *like* is not considered to be a discourse marker in this context.

Secondly, an examination of the non-truth conditionality of *like* in the given instance was conducted. In the event that the removal of *like* had an impact on the truth conditionality of the discourse, the instance in question was excluded from the data set. For example:

- (54) Enough drama so it is it is *like* a documentary you see how they uh (ICE Singapore: S1A-049: 679).

In example (54), the removal of *like* does not affect the syntactic correctness of the sentence, the optionality matches. Nevertheless, the absence of the *like* does affect the truth conditionality of the sentence. For example, “it is a documentary” differs from “it is *like* a documentary”. The former confirms the subject is a documentary while the latter indicates just a similarity between the subject and the documentary. The use of *like* here contributes to the truth conditionality of the sentence and the instance is therefore excluded from the data. The discourse marker use of *like* is illustrated in example (55) and (56):

- (55) It’s *like* it’s almost the same as well (ICE Hong Kong: S1A-005: 50).
 (56) No we I don’t say they are *like* uh they are for what you call (ICE India: S1A-008: 67).

After the manual selection of discourse marker *like*, Table 2 below illustrates the total number of *like* retrieved automatically by the KWIC function and the number of *like* instances identified as discourse markers.

Table 2. Total number of *like* and number of DM *like*

	ICE-HK	ICE-Ind	ICE-Phi	ICE-Sin
Total <i>like</i>	760	646	987	688
DM <i>like</i>	164	180	363	212

3.3.3 The positions of discourse marker *like*

The second step of coding was to determine the positions of *like*. In concordance with the literature review, these were coded in four different positions: clause-initial, clause-medial, clause-final and non-clausal *like* based on Anderson (2001), Miller (2009) and Schweinberger (2015).

Each position is explained with an example:

1. Clause-initial *like*, when *like* occurs before the subject of the or the beginning of the clause. Other pauses, fillers, interjections, conjunctions, interruptions or discourse markers immediately before *like* do not influence the clause-initial status of *like*, as these elements did not influence the meaning of the discourse. For example:

(57) *Like* Hong Kong is a colony of British (ICE Hong Kong: S1A-005: 49).

2. Clause-medial *like*, when *like* occurs after the subject of the clause but before other constituents of the clause and is not at the end of the clause. For example:

(58) No the one where they were uhm they were *like* worshipping that golden cow or something that they have made (ICE Philippine: S1A-007: 124).

3. Clause-final *like*, when *like* occurs at the end of an entire clause and is preceded by other constituents of the clause. If *like* is followed by pauses, fillers, interjections, interruptions, other discourse markers, the position of *like* is still considered as clause-final, since these elements are regarded as symbols of the segmentation to distinguish it from another clause and do not influence its clause-final position. For example:

(59) when you are in in your home *like*
oh my home Yeah uhm Oh just (ICE Hong Kong: S1A-005: 51).

4. Non-clausal *like*, when *like* occurs between two separate clauses or *like* is surrounded by pauses, fillers, interjections, interruptions, or other discourse markers. Additionally, when *like* appears in the first word of a clause and is followed by a comma is also considered non-

clausal. Moreover, when *like* is used to introduce phrasal units but does not constitute a complete clausal unit, it is also deemed as non-clausal.

(60) a little bit insulting people uh uh *like* uh Oh right I think Yeah that (ICE Hong Kong: S1A-009: 137).

3.3.4 The functions of discourse marker *like*

The final step of coding was to determine the functions of the discourse marker *like*. Based on the functions described in section 2.3.5, six different functions are classified in the current study in as presented in Table 3.

Table 3. Functions of DM *like*

- | |
|---|
| <ol style="list-style-type: none"> 1. Highlighting focus (Focus) 2. Supplying/requesting explanation (Explanation) 3. Restarting/repairing discourse (Repair) 4. Exemplifier 5. Approximator 6. Planning discourse (Link) |
|---|

The first function of discourse marker *like* is highlighting focus, when *like* emphasizes the element immediately following it, a common form is *like+adjective*. In addition, when *like* emphasizes the most important new element in the clause, it is also considered to have this function. For example:

(61) I'm assuming you go for four days

so the fourth

because three days is *like* too short (ICE Singapore: S1A-011: 127).

(62) Then shoes they have so many pairs she *like* must match match them (ICE Singapore: S1A-048: 673).

In example (61), compared with “three days is too short”, the role of *like* is as a focus marker. Without it, the focus of the sentence shifts from the predicate to the subject. In example (62), the necessity of the subject’s activity is the significant and special information in the sentence.

The second function of discourse marker *like* is supplying or requesting explanation. Firstly, when *like* occurs before a repetition element with nearly the same words as the elements precede *like*. Secondly, when *like* introduces a reformulation with more details. Thirdly, when *like* occurs before or after a question that seeks for more information about the target topic. The following examples serve to illustrate the various types of this function of *like*:

(63) I I asked for something much higher

Uhm

Like much higher means uh it’s about six more than five thousand dollars higher (ICE Singapore: S1A-036: 480).

(64) I’m a pure vegetarian

Sometimes we take eggs

Like my family all the people they are taking non-veg but only me the pure vegetarian one (ICE Indian: S1A-007: 53).

(65) Well so uhm the uh the thing is in the States everybody drives when they are really young and

Like how old

Well I think I was driving by the time I was fourteen (ICE Hong Kong: S1A-039: 569).

In example (63), the speaker first mentions a request for something much higher. The listener answers with a filler, indicating confusion. *Like* occurs before a repetition of the speaker’s former expression. In addition, *like* introduces an explanation with a more concrete numeric expression which turns the mere repetition into an explanation. In example (64), the speaker first mentions that *I’m a pure vegetarian*, which then becomes his or her family *all the people*

they are taking non-veg but only me I'm the pure vegetarian. *Like* marks the reformulation with more backgrounds as an explanation, aiming at giving more information to the listener. In example (65), the use of *like* before a question indicates the intention of the interlocutor is to request a more precise age to illustrate the statement of the speaker.

The third function of discourse marker *like* is restarting or repairing discourse. Firstly, when *like* occurs in case where a structure is cut off and is restarted by the same words. Secondly, when *like* occurs before a minor correction which is syntactically related to the elements precede *like*. Thirdly, when *like* occurs at the end of the sentence without more information and marks the termination. The following examples are provided to illustrate the various types of this function of *like*:

(66) It's a it's a it's *like* it's a seminary (ICE Singapore: S1A-040: 534).

(67) Uh uhm yeah unlike for us we have to *like* wait uh-huh (ICE Philippines: S1A-013: 233).

(68) Yeah but I think it's more, it's more of a case to case basis *like*

Uh uhm

For you it came earlier (ICE Philippines: S1A-040: 748).

In example (66), *like* functions as a marker indicating a restart of the discourse, because the speaker cuts off the utterance and restarts the utterance with the same words before *like*. In example (67), *like* serves as a symbol of self-repair of the speaker. The speaker resumes the utterance with an element that is a constituent of the syntactic structure preceding *like*. The syntactic coherence is still maintained. In example (68), *like* appears when the utterance is cut off by the speaker without resuming it.

The fourth function is the discourse marker *like* is to serve as an exemplifier. When *like* occurs before a first-occurring item which is a typical situation of the element before *like* not necessarily an actual situation. In example (69), Hyatt is a representation of typical expensive hotels and is introduced by *like* which functions as an exemplifier. Furthermore, instances in

which *like* is followed by “*for example*” or “*for instance*” are also regarded as belonging to this category. This is exemplified in example (70):

- (69) place except for the very expensive hotel *like* uhm Hyatt uhm Hyatt Sheraton Uhm How (ICE Hong Kong: S1A-002: 23).
- (70) He always goes against these teachers who are promoted *like* for example, I was nominated before to be promoted to be given the full-load status together with other four English teachers (ICE Philippine: S1A-008: 156).

The fifth function of discourse marker *like* is to serve as an approximator. This occurs when *like* is used before a numerical expression or other measurable units. The following two examples are provided to illustrate this function:

- (71) uhm you haven’t studied in here for *like* te ten years Oh okay Yah that (ICE Hong Kong: S1A-010: 183).
- (72) But on other hand there are also uhm uh uh scripts about uh Hong Kong situation uh political or on a more *like* daily level about life in Hong Kong (ICE Hong Kong: S1A-013: 232).

In example (71), *like* is an approximator to show the uncertainty of the speaker about the exact period. Moreover, in example (72), *like* occurs before an abstract measurable unit “level” to illustrate there are different aspects of life in Hong Kong.

The final function of *like* as a discourse marker is to plan the discourse, with the objective of linking the discourse and maintaining its coherence, as well as softening the tone. Primarily, the function of *like* is to serve as a floor-holding device. Secondly, it is employed as a hedging device. The following two examples serve to illustrate this function:

- (73) Uhm uhm *like* even *like* uh now I get the set of drawings from structure from contractor drawing there’s slight difference (ICE Singapore: S1A-015: 187).

(74) Aiyoh then I suddenly look at myself from outside sort of
like distance myself from (ICE Singapore: S1A-028: 327).

In example (73), the discourse marker *like* is surrounded by fillers and is thus considered as a floor-holding device, which aims to gain more time to organise the utterance. In example (74), the use of discourse marker *like* serves to soften the tone, as evidenced by the hedging expression “*sort of*”, which occurs prior to it. This mitigates the absoluteness of the discourse.

3.4 Analysis method

3.4.1 Proportion of DM use of *like* and non-DM uses of *like* across varieties of English

The first research question aims to gain insight into the proportion of discourse marker use of *like* in relation to the other non-pragmatic functions that can be fulfilled by *like*. To do so, the relative frequency of discourse marker use of *like* and non-discourse marker uses of *like* per 100,000 words were calculated due to the different number of words in the four sub-corpora. After that, the proportion of discourse marker use of *like* and non-discourse marker uses of *like* in each of the four varieties were calculated. A comparison was conducted among them subsequently.

After that, statistical tests were adopted to check to what extent the four varieties differ in their use of *like* as a discourse marker and non-discourse marker uses of *like*. Firstly, the chi-square test was adopted to check whether there was a significant difference among the four varieties. Secondly, a post hoc test was conducted to explore whether there were significant differences between each two of the four varieties. The Holm method was applied in the case of multiple comparisons to avoid errors. Only the results from the Holm method meet the significance threshold of 0.05 were deemed as significant differences.

3.4.2 Frequency of DM use of *like* across varieties of English

The second research question aims to analyze the frequency of discourse marker use of *like*

across varieties. Firstly, the relative frequency of discourse marker use of *like* per 100,000 words of the four varieties was compared. Secondly, the chi-square test was used to check whether there were significant differences in the frequency of discourse marker use of *like* across groups. After that, a post hoc test was used to investigate whether there were significant differences between each pair of the four varieties. The Holm method was applied to adjust for multiple comparisons.

3.4.3 Distribution of DM use of *like* in different positions across varieties of English

The third research question aims to explore to what extent the four varieties differ in the distribution of discourse marker use of *like* in different positions. Firstly, the distribution of *like* in four different positions was calculated in each of the four varieties. Secondly, each variety was analysed separately to account for the positional distribution within the variety. Thirdly, the chi-square test was used to check whether there was a significant difference in the distribution of positions across varieties. Finally, a post hoc test was applied to explore whether there were statistical significances in each position between each pair of the four varieties. Similarly, Holm method was used to correct the results of multiple comparisons. Qualitative analysis was combined with the results of the quantitative analysis.

3.4.4 Distribution of DM use of *like* in different functions across varieties of English

The fourth research question aims to investigate to what extent the four varieties differ in the distribution of discourse marker use of *like* in different functions. The general process was the same as the analysis of positional distribution of the discourse marker *like* across varieties.

4. Results and discussion of proportion and frequency of DM use of *like*

This chapter firstly describes the results of the proportion of the discourse marker use of *like* and non-discourse marker uses of *like* across varieties of English, as detailed in section 4.1.

Secondly, section 4.2 primarily concerns the frequency of the discourse marker use of *like* across varieties of English. Finally, section 4.3 provides further discussion concerning the frequency of discourse marker use of *like*.

4.1 Results of proportion of DM use of *like* and non-DM uses of *like* across varieties of English

Table 4. Relative frequency of DM use of *like* and non-DM uses of *like*

	Relative frequency DM <i>like</i>	Relative frequency non-DM <i>like</i>
ICE-Hong Kong	135.45	492.25
ICE-India	171.16	443.11
ICE-Singapore	197.38	443.17
ICE-Philippines	335.72	577.11

Table 4. presents the relative frequency of discourse marker use of *like* and non-discourse marker uses of *like* per 100,000 words in each of the four varieties. It can be observed that the relative frequency of *like* as a discourse marker is lower than the relative frequency of its non-pragmatic functions in all the four varieties. This suggests that the non-discourse marker uses of *like* are more prevalent in each of the varieties. In addition to this general similar tendency among these varieties, there are differences across varieties.

In Hong Kong English, the relative frequency of the use of the discourse marker use of *like* is markedly lower than that of non-discourse marker uses of *like*. The ratio of the discourse marker use of *like* to its non-discourse marker uses is 0.275. The use of the discourse marker *like* is over 3.5 times less frequent than its non-pragmatic functions in Hong Kong English. The ratio of the discourse marker use of *like* to non-discourse marker uses of *like* in Indian English is relatively low (0.386), indicating a preference for non-pragmatic functions of *like*. In both

Indian English and Singapore English, the relative frequency of non-discourse marker *like* is almost identical, while Singapore English speakers use discourse marker *like* slightly more often than Indian English speakers. Despite Philippine English having the highest frequency of discourse marker use of *like*, Philippine speakers still use *like* more frequently for its non-pragmatic functions.

Table 5. Results of comparison of DM use of *like* and non-DM uses of *like* between each pair of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	6.30e-03	9.45e-03
HkE vs. PhiE	6.39e-12	4.07e-11
HkE vs. SinE	6.27e-05	1.88e-04
IndE vs. PhiE	1.85e-04	3.70e-04
IndE vs. SinE	2.37e-01	2.37e-01
PhiE vs. SinE	1.14e-02	1.37e-02

The results of the chi-square test indicate that there is a statistically significant difference in the frequency of discourse marker use of *like* and the frequency of non-discourse marker uses of *like* across the four varieties ($\chi^2 = 48.996$, $df = 3$, $p < 0.001$). The results of the post hoc tests, as presented in Table 5, indicate that, with the exception of the difference between Indian English and Singapore English, which is not statistically significant ($p = 0.237$), there are statistically significant differences between each pair of the varieties.

The proportion of discourse marker use of *like* and non-discourse marker uses of *like* in each of the four varieties is presented in Figure 2.

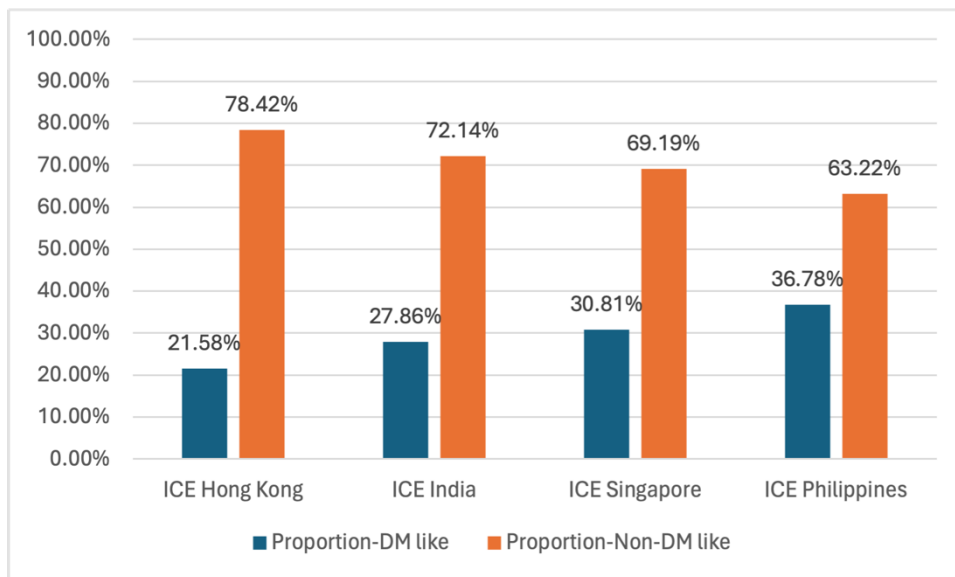
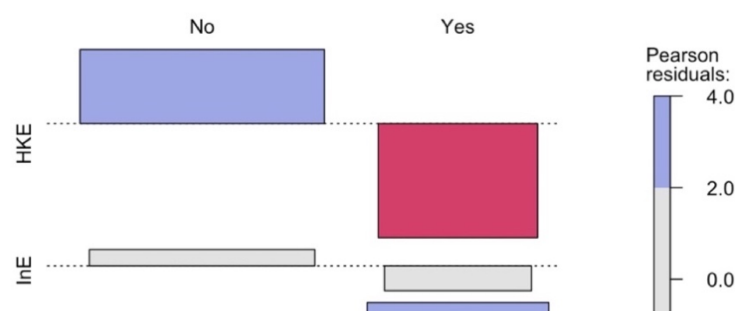


Figure 2. Proportion of DM *like* and non-DM *like*

Figure 2 exhibits a comparable pattern to that pattern observed in Table 4. Figure 2 illustrates that the discourse marker *like* is used in Hong Kong English at a rate of 21.58%. In other words, the discourse marker *like* in Hong Kong English accounts for only one-fifth of the total occurrences of *like*. As with the findings presented in Table 4, the data for Indian English indicate that the proportion of discourse marker use of *like* is less than 30% of the total occurrences of *like*. The proportion of discourse marker use of *like* in Singapore English has reached 30%. Finally, Philippine English displays the highest proportion of discourse marker use of *like*. The usage pattern of discourse markers and non-discourse markers is more balanced in this variety.

Figure 3 presents an association plot of discourse marker use of *like* and non-discourse marker uses of *like*. The term “No” is used to refer to non-discourse marker uses of *like*, while the term “Yes” is used to refer to discourse marker uses of *like*. From the association plot, it can be observed that Hong Kong English is significantly underrepresented in the use of the discourse marker *like*, while Philippine English is significantly overrepresented in the use of the discourse marker *like*.



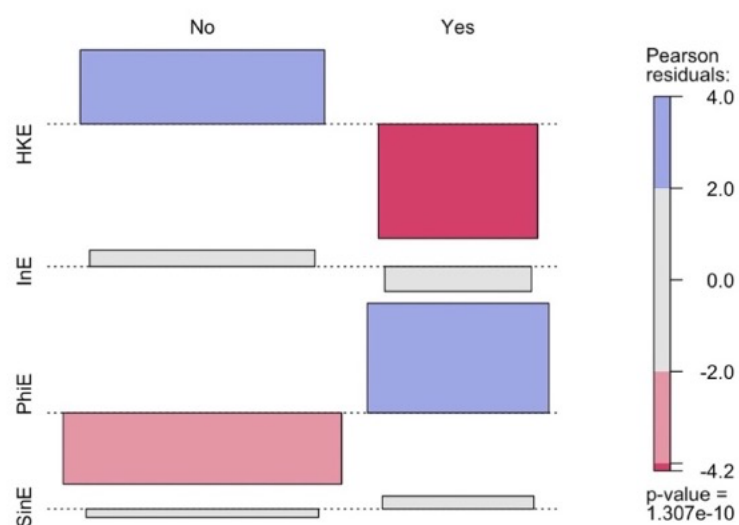


Figure 3. Association plot of DM use of *like* and non-DM uses of *like*

4.2 Results of frequency of DM use of *like* across varieties of English

Figure 4. presents the relative frequency of *like* functions as a discourse marker in each of the four varieties.

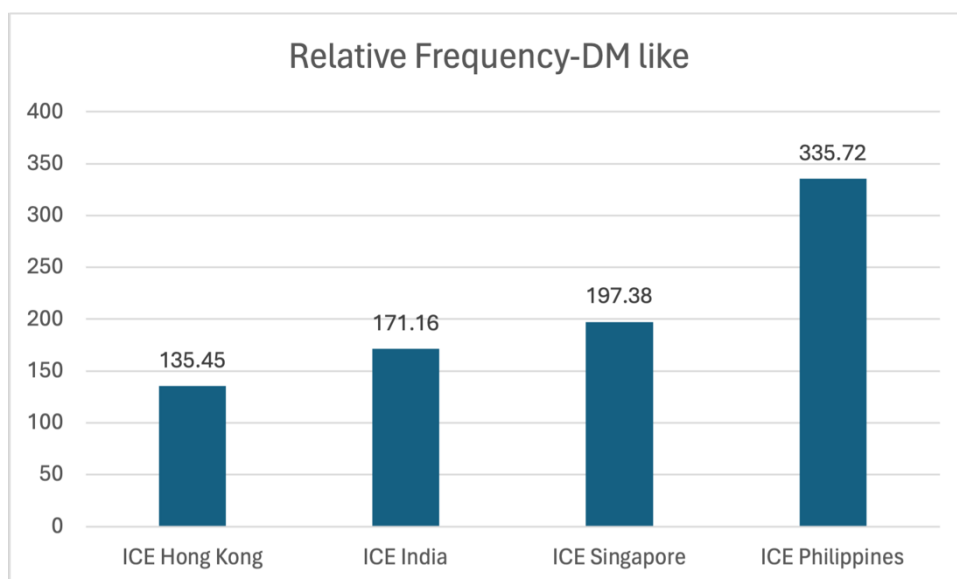


Figure 4. Relative frequency of DM *like*

The variety with the lowest frequency of the discourse marker use of *like* is Hong Kong English, with a relative frequency of 135.45 occurrences per 100,000 words. Subsequently, Indian English and Singapore English demonstrate an increasing frequency, with 171.16 and 197.38 occurrences per 100,000 words respectively. Finally, Philippine English exhibits the highest frequency of the discourse marker use of *like*, with a relative frequency of 335.72 occurrences per 100,000 words. The frequency of use of *like* as a discourse marker among Philippine English speakers is approximately 2.5 times that observed among Hong Kong English speakers.

The frequency of the discourse marker use of *like* differs significantly across the four varieties ($\chi^2 = 108.24$, $df = 3$, $p < 0.001$). Moreover, the results of the post hoc test in Table 6 indicate that the frequency of the discourse marker use of *like*, the differences between Hong Kong English and Indian English ($p = 0.388$), and between Singapore English and Indian English ($p = 0.212$) are not statistically significant. In contrast, the results for the remaining pairs of varieties demonstrate statistically significant differences.

Table 6. Results of frequency of DM use of *like* between each pair of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	3.88e-01	3.88e-01
HkE vs. PhiE	4.38e-18	2.63e-17
HkE vs. SinE	1.33e-02	3.99e-02
IndE vs. PhiE	4.05e-15	2.02e-14
IndE vs. SinE	1.06e-01	2.12e-01
PhiE vs. SinE	3.03e-10	1.21e-09

4.3 Discussion

The results of the quantitative analysis demonstrate that the varieties of English spoken in South and Southeast Asia continue to employ the word *like* with greater frequency for its non-pragmatic functions than as a discourse marker. This is feasible due to the relatively recent emergence of the use of *like* as a discourse marker. As can be observed from the grammaticalization process of *like*, the discourse marker use of *like* usually occurs in a relatively latter stage with almost no semantic meaning (see section 2.3.1.1). The current study's participants typically acquired English as a second language and do not use vernacular forms with the same frequency as native speakers of English. It is therefore unsurprising that individuals from Outer Circle varieties employ *like* with its conventional lexical meaning and non-pragmatic functions more frequently than they utilise it as a discourse marker.

With regard to the frequency of use of the discourse marker *like*, Philippine English exhibits the highest incidence. This may be attributed to the observation made by Schneider (2007), namely that Philippine English is a consequence of American colonialism, whereas the other varieties selected for analysis in the present study can be attributed to British colonialism (cf. section 2.4.4). Moreover, Siegel (2002) indicates that it was the Valley Girls in Californian who initially employed the use of *like* as a discourse marker, thereby facilitating its spread as a vernacular usage. In other words, the use of the discourse marker *like* in Philippine English can be traced back to American English. This provides an explanation as to why Philippine English speakers utilise the discourse marker with greater frequency than the other three varieties.

The frequency of use of the discourse marker *like* in Singapore English is the second highest. This may be attributed to the fact that it is a more advanced variety, a consequence of the country's "English-based bilingualism" policy. As previously stated, English holds a unique position in Singapore, serving as the sole first language taught to all children (see section 2.4.3). Consequently, they commence their English education at an early age, with English becoming

an integral part of their daily lives. In addition to acquiring knowledge of standard English expressions, learners also gain familiarity with the vernacular uses of the language, such as the discourse marker *like*. The data employed in the present study were obtained from face-to-face, private conversations (see section 3.2.1). As Siegel (2002) notes, the discourse marker *like* is often used to show intimacy and thus occur more frequently in the conversations between friends. This offers an explanation as to why, in the current study, Singapore English exhibits a relatively high frequency in using *like* as a discourse marker.

Indian English is the second least frequent variety in using *like* as a discourse marker. This may be attributed to the fact that *like* is a discourse marker which has no nativized sources in Indian English, as highlighted by Valentine (1991) (cf. section 2.4.2.1). Nevertheless, Indian English speakers tend to utilise those discourse markers with a nativized source more frequently, as these elements sometimes originate from other native languages in India, such as Hindi.

Discourse marker *like* is not used so frequent as other discourse markers such as *so* in Hong Kong English (cf. 2.3.1.1). As documented in Fung & Carter (2007), the usage of the discourse marker *like* is situated in the sixth position among the 23 selected discourse markers in the Hong Kong data (cf. section 2.3.1.1). This provides a partial explanation for the observation that Hong Kong English speakers exhibit the lowest frequency of use of the discourse marker *like*. Moreover, Schweinberger (2014) conducted a study of the discourse marker *like* across several varieties of English using the ICE corpus. In the study, the British English was identified as the variety with the lowest frequency of the discourse marker *like*. In the present study, Hong Kong English is identified as the variety exhibiting the lowest frequency of the discourse marker *like*. This finding is consistent with the results of Fung & Carter (2007), which indicate that the frequency of the discourse marker *like* in Hong Kong English is similar to its frequency in British English (cf. section 2.4.1.1).

5. Results and discussion of positional distribution of DM use of *like* across varieties of English

This chapter presents the results of the positional distribution of discourse marker use of *like* across varieties of English. Initially, section 5.1 demonstrates and provides clarification on the results of descriptive statistical analysis. Subsequently, the results of statistical inference are presented in section 5.2. Finally, section 5.3 presents a further discussion and interpretation of the findings of the preceding two sections.

5.1 Description of positional distribution of DM use of *like* across varieties

Table 7. depicts the overview of the positional distribution of discourse marker use of *like* in each of the varieties. From Table 7, it can be found that the clause-medial position is the most prevalent in the positional distribution of discourse marker *like* among these four varieties in general. Moreover, it seems that discourse marker *like* has a relatively low frequency in the clause-initial position across these four varieties, with a relative frequency of 176.56 per 100,000 words. The frequency of occurrence of the discourse marker *like* in clause-final and non-clausal positions is almost identical, with 198.34 and 191.59 occurrences per 100,000 words, respectively.

Table 7. Relative frequency of positional distribution of DM use of *like* across varieties

	Clause-initial per 100,000 words	Clause-medial per 100,000 words	Clause-final per 100,000 words	Non-clausal per 100,000 words	Total
ICE-Hong Kong	35.51	45.43	32.21	22.30	135.45
ICE-India	44.69	24.72	45.64	56.10	171.16
ICE-Singapore	25.14	80.07	39.10	53.07	197.38

ICE-Philippines	71.21	123.01	81.39	60.12	335.72
Total	176.56	273.22	198.34	191.59	839.71

Nevertheless, the various varieties exhibit disparate patterns of positional distribution when the *like* is employed as a discourse marker.

In Hong Kong English, the most frequent position of *like* as a discourse marker is clause-medial whereas the least frequent position is non-clausal. The clause-initial and clause-final positions of discourse marker *like* in Hong Kong English are almost identical. Although Hong Kong English has the lowest frequency in the usage of the discourse marker *like*, regarding its usage on clause-medial position, the frequency of the discourse marker *like* in Hong Kong English is higher than its frequency in Indian English. With regard to the clause-initial position, Hong Kong English uses the discourse marker *like* more frequently than Singapore English.

In contrast to the other three varieties, in Indian English, the clause-medial is the least frequent position of the discourse marker *like*. The most common position for the discourse marker *like* is non-clausal position in Indian English, followed by the clause-final and clause-initial positions.

Singapore English also exhibits a preference for the clause-medial position of the discourse marker *like* while the least common position is clause-initial. The second most popular position for discourse marker use of *like* in Singapore English is the non-clausal one.

In Philippine English, similar to Hong Kong English and Singapore English, the clause-medial position of discourse marker *like* is the most frequent one. In contrast to Indian English, the least frequent position of discourse marker *like* in Philippine English is the non-clausal one. Interestingly, although the frequency of discourse marker *like* in Hong Kong English is strikingly less than its occurrences in Philippine English, these two varieties have similar

patterns in the positional distribution of discourse marker *like*.

Figure 5. provides a more detailed illustration of the positional distribution of discourse marker *like* across the varieties.

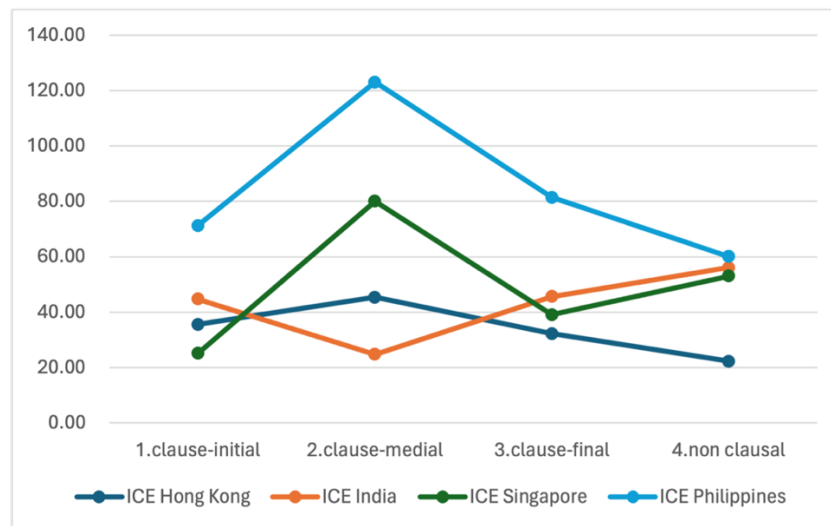


Figure 5. Positional distribution of DM *like* across varieties of English

It can be found from Figure 5 that Philippine English scores the highest frequency for all four positions of the discourse marker *like*. In clause-initial and clause-final position, except Philippine English, Indian English is slightly higher than Hong Kong English and Singapore English.

In clause-medial position, except Philippine English, it is Singapore English that uses *like* as a discourse marker the most frequently, which is significantly higher than the frequency in Hong Kong English and Indian English. Moreover, the distance between Hong Kong English and Indian English in the usage of clause-medial *like* is still large.

In the non-clausal position of the discourse marker *like*, Philippine English, Indian English and Singapore English have a similar frequency and are strikingly higher than that of Hong Kong

English. The association plot in Figure 6 further depicts the positional distribution pattern across varieties of English.

It can be observed from Figure 6 that, Indian English is highly underrepresented in clause-medial position of discourse marker use of *like* and is highly overrepresented in the non-clausal position of this vernacular use of *like*. In contrast to Indian English, Singapore English is highly overrepresented in the clause-medial position when using discourse marker *like*. However, Singapore English is highly underrepresented concerning the usage of clause-initial discourse marker *like*.

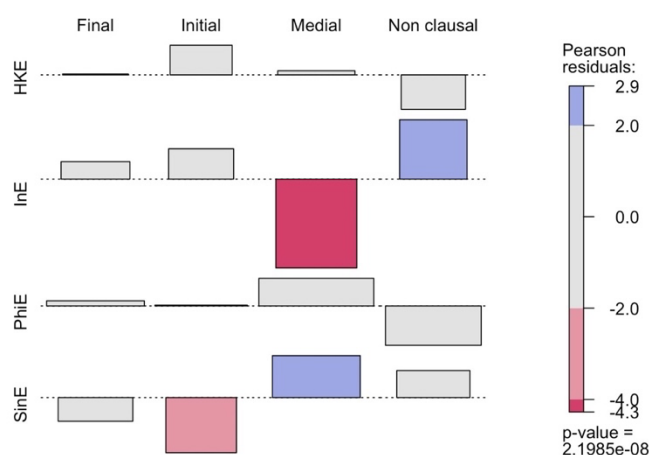


Figure 6. Association plot of positional distribution of DM *like* across varieties

5.2 Results of the differences across varieties in positional distribution of DM use of *like*

To provide further evidence of the correlation between positional distributions of the discourse marker *like* and varieties, the results of the statistical tests are subjected to a rigorous examination.

The result of the chi-squared test indicates a statistically significant difference in the positional distribution of discourse marker use of *like* across the four varieties ($\chi^2 = 53.652$, $df = 9$, $p < 0.01$). The results of the post hoc test in Table 8 indicate that, with the exception of the difference between Hong Kong English and Philippine English, as well as the difference

between Hong Kong English and Singapore English, the differences between each pair of varieties in the positional distribution of the vernacular use of *like* are statistically significant. The results serve to test and verify one of the findings of the descriptive statistics, which posits that Hong Kong English speakers and Philippine English speakers exhibit similar usage patterns with respect to the positional distribution of this vernacular use of *like* (section 5.1).

Table 8. Results of positional distribution of DM *like* between each two of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	4.66e-05	9.32e-05
HkE vs. PhiE	6.38e-01	6.38e-01
HkE vs. SinE	1.49e-03	2.24e-01
IndE vs. PhiE	2.60e-07	7.80e-07
IndE vs. SinE	9.06e-08	5.44e-07
PhiE vs. SinE	7.02e-03	8.42e-03

5.2.1 Clause-initial *like* across varieties

The result of the chi-squared test for the clause-initial position of the discourse marker *like* exhibits that there are statistically significant differences across varieties ($\chi^2 = 26.948$, $df = 3$, $p < 0.001$). The results of the post hoc test in Table 9 reveal that the usage of clause-initial *like* exhibits statistically significant differences when comparing Philippine English with each of the other three varieties. No statistically significant difference is detected between any two of Hong Kong English, Indian English and Singapore English.

Table 9. Results of clause-initial *like* between each two of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
---------------------------	---------------	-------------------------

HkE vs. IndE	6.73e-01	6.73e-01
HkE vs. PhiE	1.91e-03	9.55e-03
HkE vs. SinE	5.58e-02	1.12e-01
IndE vs. PhiE	7.06e-03	2.82e-02
IndE vs. SinE	2.01e-02	6.03e-02
PhiE vs. SinE	9.44e-07	5.66e-06

The results from Table 9 confirms the findings of the descriptive statistics. Philippine English speakers use *like* as a discourse marker in the clause-initial position strikingly more frequently than speakers from other three varieties of English (section 5.1).

5.2.1.1 Qualitative analysis within the varieties of English

In Hong Kong English, the clause-initial position of discourse marker *like* is the second most frequent position. As previously stated in section 3.3.3, the initial position of *like* is not affected by the presence of other discourse markers. The following examples serve to illustrate the usage of clause-initial *like* in Hong Kong English:

(75) Z: Lots of other things shows and

A: Show yeah

B: Uh

A: *Like* the shows in the casino are fabulous (ICE Hong Kong: S1A-024: 381).

(76) Look at the dog

It's really scared

I mean *like* you can put this huge thing in front of him, you know (ICE Hong Kong: S1A-047: 705).

In example (75), the discourse marker *like* occurs in the clause-initial position and has a scope

over the entire noun phrase following it. In example (76), another discourse marker *I mean* occurs before discourse marker *like*, however, *like* still has a scope encompassing the entire proposition that follows it. Furthermore, it is noteworthy that, in the dataset of Hong Kong English, *you know* and *I mean* are the only two discourse markers that occur before the clause-initial *like*.

In Indian English, the frequency of the clause-initial position of the discourse marker *like* is third among the four positions. This position of *like* includes the situations when *like* is literally the first word of the clause, as well as situations in which *like* is preceded by other discourse marker or fillers (cf. section 3.3.3). The following examples are used to illustrate its usage:

(77) So I've got my old grandmother and making for us you know

So *like* if we have to make we'll naturally prefer a pressure cooker (ICE India: S1A- 007: 57).

(78) A: Accha how many minutes I must do like that

B: Uh *like* I will tell you once (ICE India: S1A- 043: 502).

In example (77), *like* occurs in the clause-initial position and is preceded by another discourse marker *so*. It has a scope over the conditional clause modified by *if*. In example (78), a filler *Uh* occurs before *like*, however, *like* is still considered as in the clause-initial position. It has a scope over the entire sentence following it. In general, in Indian English, *so* is the most common discourse marker which occurs before the clause-initial *like*, other fillers such as *Uh* and interjections like *Oh* are also frequent elements which occur before the clause-initial *like*. However, unlike Hong Kong English, discourse marker *you know* only occurs one time before clause-initial *like* while another discourse marker *I mean* is not detected before clause-initial *like* in the Indian English dataset.

In Singapore English, the clause-initial position is the least frequent among the four positions of discourse marker *like*. Moreover, Singapore English also has the lowest frequency of clause-

initial *like* across the four varieties of English. As in Indian English, sometimes another discourse marker *so* precedes the clause-initial discourse marker *like*, as illustrated in (81). However, it has a distinctive usage of discourse marker *like* in this position. The following examples serve to illustrate the clause-initial *like* in Singapore English:

- (79) *Like* aiyah why are we here in the first place (ICE Singapore: S1A-012: 142).
(80) Ya but *like* if she step on her toes than why the hell does
she we're all excreted to (ICE Singapore: S1A-047: 653).
(81) So *like like* I was saying they came to conduct (ICE Singapore: S1A-032: 435).

It can be observed from the two examples that some distinctive discourse markers co-occur frequently with the clause-initial *like* in Singapore English, such as *aiyah* in (79) and *Ya* in (80). In both examples, *like* has a forward scope encompassing the entire higher-level construction following it. Moreover, similar to Indian English, discourse marker *so* also frequently occurs before the clause-initial *like* in Singapore English.

In Philippine English, the clause-initial *like* occupies the third among the four positions of discourse marker *like*. However, Philippine English has the highest frequency of clause-initial *like* across the four varieties. This usage can be illustrated by the following examples:

- (82) I mean they're they're old people I mean *like* their ages range from late thirties (ICE Philippines: S1A-012: 209).
(83) Yeah *like* C P said that (ICE Philippines: S1A-042: 762).

In both examples, the clause-initial *like* is preceded by another discourse marker or interjection, namely *I mean* in (82) and *Yeah* in (83). Moreover, all those other typical discourse markers, fillers or interjections that are detected to occur before clause-initial *like* in the former three varieties, such as *so*, *you know*, *uh*, can be found in the dataset of Philippine English, with the exception of those distinctive discourse markers in Singapore English. This phenomenon

probably explains the reason why Philippine has such a high frequency of clause-initial *like* compared to the other three varieties, as *like* can co-occur with diverse other constituents, not being restricted to several of them.

5.2.2 Clause-medial *like* across varieties

Table 10. Results of clause-medial *like* between each two of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	1.27e-03	3.81e-03
HkE vs. PhiE	1.28e-08	6.40e-08
HkE vs. SinE	9.04e-03	9.04e-03
IndE vs. PhiE	2.15e-17	1.29e-16
IndE vs. SinE	1.43e-08	6.40e-08
PhiE vs. SinE	1.49e-03	3.81e-03

According to the result of the chi-squared test for the clause-medial *like*, the statistically significant difference exists across the four varieties ($\chi^2 = 83.813$, $df = 3$, $p < 0.001$). The results of the post hoc test in Table 10 verify that there are statistically significant differences between each pair of the varieties when using discourse marker *like* in a clause-medial position.

5.2.2.1 Qualitative analysis within varieties of English

The total frequency of clause-medial *like* is the highest among the four varieties of English. However, the four varieties employ this kind of *like* in different ways.

In Hong Kong English, this is the most popular position for *like* as a discourse marker. On this position, *like* is often situated between an auxiliary and the lexical verb, or between a prepositional verb and the noun phrase following, as illustrated in example (84) and (85) respectively. Moreover, *like* can occur between a copula and the predicative adjective, as

illustrated in example (86):

- (84) And then she is *like* using all her own fruit to make the trifle (ICE Hong Kong: S1A-010: 140).
- (85) What what do you think about *like* the difference between Hong Kong and Singapore (ICE Hong Kong: S1A-010: 156).
- (86) And there you know he's *like* so mature you know (ICE Hong Kong: S1A-041: 606).

In Indian English, however, in contrast to the other three varieties, this is the least frequent position among the four positions of discourse marker *like*. Moreover, Indian English has the lowest frequency of clause-medial *like* across the four varieties of English. In Indian English, clause-medial *like* usually occurs between a transitive verb and its direct object, or between the head of a prepositional phrase and its complement, as illustrated in example (87) and (88), respectively:

- (87) He was doing in driving *like* a lorry (ICE India: S1A- 017: 227).
- (88) I'm uh during the mosquito season in *like* Amritsar (ICE India: S1A- 013: 149).

In Singapore English, the clause-medial position of *like* is the most frequent one. Moreover, this usage of discourse marker *like* in Singapore English occupies the second position across the four varieties of English. In Singapore English, clause-medial *like* can occur between all the constituents mentioned in the former two varieties. Besides, *like* can modify a diverse range of constituents, such as a noun phrase, as illustrated in (89); a verb phrase, as illustrated in (90); or a prepositional phrase, as illustrated in (91):

- (89) The other side is *like* car-park (ICE Singapore: S1A-001: 4).
- (90) Uhm uhm so they must *like* open themselves to (ICE Singapore: S1A-008: 86).
- (91) That doctor has a very long lunch-break *like* from eleven to three (ICE Singapore: S1A-027: 290).

It is noteworthy that in Singapore English, clause-medial *like* frequently occur after a hedging device *sort of* and before some numerical expressions, as presented in (92) and (93) respectively. Moreover, *like* can occur between the subject and the verb in Singapore English, as demonstrated in (94):

(92) The clock just sort of *like* fades away (ICE Singapore: S1A-021: 242).

(93) It's not necessary to do a *like* three-year Arts Admin course (ICE Singapore: S1A-042: 562).

(94) So she *like* forgets some of the things she learnt (ICE Singapore: S1A-039: 506).

In general, the use of the clause-medial *like* in Singapore English exhibits greater variability, as it can occur in a variety of positions, modifying a range of constituents within the clause in which it occurs.

In Philippine English, the clause-medial *like* has the highest frequency among the four positions. Moreover, the clause-medial *like* in Philippine English also occupies the first place across the four varieties of English. Similar to Singapore English, the clause-medial *like* in Philippine English can modify a wide range of constituents, such as a noun phrase, a verb phrase or a prepositional phrase, as presented in (95), (96) and (97):

(95) I thought we were staying by *like* a few steps (ICE Philippines: S1A-007: 109).

(96) You're a fourth year can you just *like* walk through a crowd of third years (ICE Philippines: S1A-029: 539).

(97) The scariest part of the movie was *like* in the beginning (ICE Philippines: S1A-012: 197).

Moreover, the use of clause-medial *like* that occurs before an adverbial phrase or before a pronoun is first time detected across the four varieties of English. The following examples (98) and (99) illustrate these two cases:

(98) And then after but he said *like* little by little that's when he's gonna (ICE Philippines:

S1A-015: 250).

(99) You know it it really scares like everyone (ICE Philippines: S1A-048: 925).

Finally, in Philippine English, it is *kind of/kinda* which frequently occurs before a clause-medial *like*, instead of the hedging device *sort of* in Singapore English. Furthermore, a great number of numerical expressions can be identified after clause-medial *like* in Philippine English. The following examples (100), (101) and (102) illustrate these cases:

(100) It's kinda *like* so fun (ICE Philippines: S1A-018:337).

(101) It's good It's kind of *like* mixing Knorr (ICE Philippines: S1A-023: 395).

(102) She wasn't able to sleep for *like* three nights (ICE Philippines: S1A-012: 198).

5.2.3 Clause-final *like* across varieties

The result of the chi-squared test for the clause-final position of discourse marker *like* indicates the presence of statistically significant differences across varieties ($\chi^2 = 28.77$, $df = 3$, $p < 0.001$). The results of the post hoc test in Table 11 indicate that there are statistically significant differences in the usage of clause-final *like* between Philippine English and each of the other three varieties. No statistically significant differences are detected between any two of the other three varieties. These results are consistent with those observed for clause-initial *like*. Therefore, it can be observed that there are similarities in the clause-marginal use of discourse marker *like* in the four varieties of English spoken in South and Southeast Asia.

Table 11. Results of clause-final *like* between each two of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	3.35e-01	1.00
HkE vs. PhiE	1.37e-05	8.22e-05
HkE vs. SinE	7.39e-01	1.00
IndE vs. PhiE	6.04e-04	2.42e-03

IndE vs. SinE	5.27e-01	1.00
PhiE vs. SinE	5.47e-05	2.74e-04

5.2.3.1 Qualitative analysis within varieties of English

Hong Kong English occupies the least frequent one across the four varieties of English in using the clause-final discourse marker *like*. Most of the time, this usage of *like* is situated in the end of a main clause and is used to introduce the dependent clause. Sometimes, the clause-final *like* is used to mark an incomplete discourse. Moreover, as mentioned in section 3.3.3, other discourse markers or fillers that follow *like* will not influence its clause-final position. This is a common situation in the dataset of Hong Kong English. The following examples illustrate the cases:

- (103) But the requirement is *like* uhm you haven't studied in here (ICE Hong Kong: S1A-010: 182).
- (104) May may maybe I'll I'll put some symbols *like*, you know oh this is not my, my experience is from, from my my friend (ICE Hong Kong: S1A-040: 581).

In example (103), *like* occurs in the last slot of the main clause; thus, it is considered as in a clause-final position. It has a forward scope to introduce the dependent clause. In example (104), since the clause is terminated by the discourse marker *like*, the speaker changes the focus of the topic to other things.

In Indian English, the frequency of the clause-final *like* occupies the second most frequent among the four positions. In most of the cases, the clause-final *like* marks the termination of the clause. It can modify an adjective phrase, a noun phrase or the whole clause in which it occurs. It is noteworthy that there are cases in which discourse marker *like* is situated at the end of a question. The following examples illustrate these cases:

(105) No I can teach English, uh I don't know what kind of English is this, but that's sometimes, there are Englishes which are spoken by local people which are quite intolerable *like* (ICE India: S1A- 028: 302).

(106) I was accompanying him *like* (ICE India: S1A- 040: 445).

(107) Ah I thought you have bought some ordinary thing *like* (ICE India: S1A- 040: 446).

(108) You didn't go anywhere *like*? (ICE India: S1A- 021: 256).

In example (105) and (107), the clause-final discourse marker *like* has a backward scope to modify the adjective phrase and the noun phrase, separately. In example (106) and (108), *like* has a backward scope over the entire clause in which it occurs. Additionally, the clause-final *like* is combined with the question marker in (108).

In Singapore English, the frequency of the clause-final *like* occupies the third most common one among the four positions. The clause-final *like* can modify an adjective phrase or a prepositional phrase. Similar to Hong Kong English, in most cases, the clause-final *like* is used between the main clause and the dependent clause or to mark the incomplete clause. Interestingly, this usage is usually presented in a pattern such as: *I mean like/I mean it's like*. The following examples serve to illustrate these cases:

(109) If it's something really serious *like* (ICE Singapore: S1A-005: 61).

(110) I did tell the lady at the SBS *like* (ICE Singapore: S1A-021: 232).

(111) I mean it's like you know he's very well-trimmed white beard and wearing beige-coloured shirt (ICE Singapore: S1A-034: 469).

(112) A: Ya it's different also I mean it's *like*

B: I think it depends on who you hang out with also (ICE Singapore: S1A-034: 471).

(113) A: But wish but that's not my primary concern right now I mean *like*

B: You never you would never consider going say joining say SDT or it's very lifestyle (ICE Singapore: S1A-044: 599).

In example (109) and (110), the clause-final *like* has a backward scope to modify the elements to its left, namely an adjective phrase and a prepositional phrase, respectively. In example (111), (112) and (113), *I mean* frequently co-occurs with discourse marker *like*. However, in example (111), *like* has a forward scope and introduce the dependent clause, while in (112) and (113), *like* is scopeless and marks the incomplete clause.

In Philippine English, the clause-final discourse marker *like* is the second most frequent among the four positions, according to frequency of occurrence. Furthermore, Philippine English exhibits the highest frequency of clause-final *like* across the four varieties. Similarly, *like* can be used to modify a variety of phrases, including adjective phrases and noun phrases. Additionally, *like* can be used to modify an indefinite adverb or an indefinite pronoun, as illustrated in example (114) and (115). When *like* is used between the main clause and the independent clause, further patterns are observed, such as *they're like/he's like*, which are not limited to the structure *it's like* as seen in Singapore English. Finally, the structure *kinda/kind of* co-occurs with the clause-final *like*, as seen in clause-medial position in Philippine English. The following examples serve to illustrate the clause-final *like* in Philippine English:

- (114) People can go anywhere and practically anywhere *like* if if they can they can access (ICE Philippines: S1A-012: 200).
- (115) Especially if there's somebody else *like* uhm (ICE Philippines: S1A-038: 684).
- (116) Let's get out of here and they're *like* you know they didn't make me pansin (ICE Philippines: S1A-007: 132).
- (117) Their coach maybe the coach is he's *like* you know they don't have a lot (ICE Philippines: S1A-025: 425).
- (118) It's kinda *like* we volt in (ICE Philippines: S1A-033: 593).
- (119) A: And then you know everything kind of *like*
 B: So what so he smiles (ICE Philippines: S1A-046: 829).

In example (116) and (117), *like* is still considered as in a clause-final position although it is followed by another discourse marker *you know* (section 3.3.3). In (118) and (119), the hedging device *kinda/kind of* occurs together with the clause-final *like*. In example (118), *like* has a forward scope encompassing the embedded clause. While in (119), the scope is missing.

5.2.4 Non-clausal *like* across varieties

Table 12. Results of non-clausal *like* between each two of the varieties

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	5.59e-04	0.002800
HkE vs. PhiE	7.44e-05	0.000446
HkE vs. SinE	1.06e-03	0.0042240
IndE vs. PhiE	5.90e-01	1.00
IndE vs. SinE	8.53e-01	1.00
PhiE vs. SinE	4.69e-01	1.00

The result of the chi-squared test for the non-clausal position of discourse marker *like* indicates the presence of statistically significant differences across varieties ($\chi^2 = 26.692$, $df = 3$, $p < 0.001$). The results of the post hoc test in Table 12 indicate that there are statistically significant differences in the usage of clause-final *like* between Hong Kong English and each of the other three varieties. No statistically significant differences are detected between each two of the other three varieties.

5.2.4.1 Qualitative analysis within varieties of English

In Hong Kong English, the frequency of the non-clausal position of discourse marker *like* is the least frequent one among the four positions. Moreover, Hong Kong English has the lowest frequency of using the non-clausal discourse marker *like* across the four varieties of English.

The non-clausal *like* in Hong Kong English is often accompanied by fillers. Sometimes, *like* occurs in the beginning of a clause and is separated from the clause by a comma, as illustrated in (121). The following examples illustrate its usage in Hong Kong English:

(120) A: *Like* uhm

Z: Dessert

A: Yeah desserts uhm, some dim sum with seafood, I think uhm, etcetera (ICE Hong Kong: S1A-032: 473).

(121) No I mean not exactly one month but a little bit more

Like, a month and ten days or so, I try to I try to make more time on it (ICE Hong Kong: S1A-007: 98).

(122) You have everything all mixed in on one page, so uh *like like like like like* uh just is I think the the situation is very unique because you you don't really say that (ICE Hong Kong: S1A-013: 244-248).

In (120), discourse marker *like* is followed by a filler in a separate line in the transcription and without more information, thus is considered in a non-clausal position. In a similar situation in (122), *like* is surrounded by fillers, and interestingly, here the speaker uses *like* five times between two independent clauses.

In Indian English, the non-clausal position of discourse marker *like* is the most prevalent among the four positions. Furthermore, Indian English is the second most frequent variety of English for the use of non-clausal discourse marker *like*. As is the case in Hong Kong English, the non-clausal *like* in Indian English is frequently accompanied by fillers, as illustrated in (123). Sometimes, a non-clausal *like* is employed to modify a noun phrase, yet it fails to constitute a clausal unit due to the absence of a predicate, as exemplified in (124). Moreover, in (125), it can be observed that *like* is separated from the preceding clause by a comma, which indicates that there is a pause. The discourse marker *like* does not belong to the latter clause, as evidenced

by the prosodic break indicated in the transcription:

(123) A: Prime Minister uh sorry, Karna, Chief Minister relief fund

B: Karna

B: Uh *like* uh

A: Uh as well as hear in the news that uh, uh Deve Gawda also (ICE India: S1A- 002: 21).

(124) A: Right

B: *Like* uh District Industries Corporation, industries, there is uh one corporation called uh Karnataka State Financial Corporation

A: Industries (ICE India: S1A- 020: 241).

(125) You have the matriarchal system na, *like* where do your, the woman is the more dominant in the family (ICE India: S1A- 039: 440).

In Singapore English, the frequency of non-clausal discourse marker *like* occupies the second most prevalent among the four positions. Different from the former two varieties, the situation in which *like* is surrounded by fillers is rare in Singapore English. In most cases, non-clausal *like* is used to introduce phrasal units which do not constitute a complete clause, such as noun phrases, as illustrated in (126), or a prepositional phrase, as illustrated in (127). Moreover, a hedging device *sort of* sometimes occurs before the non-clausal discourse marker *like*, which introduces a verb phrase, as illustrate by (128):

(126) Small little egg you know

Like those pigeon or what egg uh (ICE Singapore: S1A-013: 168).

(127) A: There were few plays every night

B: Uhm

A: *Like* at seven forty-five (ICE Singapore: S1A-025: 274).

(128) A: Aiyoh then I suddenly look at myself from outside sort of *like* distance myself from

B: Uh huh (ICE Singapore: S1A-028: 327).

Philippine English is the most prevalent variety of English in the use for non-clausal discourse marker *like*. However, in the Philippine English dataset, the frequency of non-clausal *like* is the lowest among the four positions. In Philippine English, non-clausal *like* is surrounded by fillers but not as frequent as the cases in Hong Kong and Indian English. Interestingly, non-clausal *like* in Philippine English is often combined with the expression *for example*, as illustrated in (130). Furthermore, in Philippine English, the discourse marker *like* frequently occurs between two similar clauses, wherein the clause after *like* constitutes a restart of the clause preceding *like*, as illustrated in example (131). It is regarded as a non-clausal *like* because *like* serves as a repair marker, operating as an independent entity and not belonging to either of the two clauses. The following examples serve to illustrate non-clausal *like* in Philippine English:

(129) A: And not if not the last at least they will if ever the relationship will end it will not end

B: *Like* uh (ICE Philippines: S1A-042: 766).

(130) Uh he, he always *like* for example for me I've I was never a lawyer (ICE Philippines: S1A-009: 180).

(131) uh he would promise *like* he would you know get up on certain days (ICE Philippines: S1A-010: 188).

In example (129), the discourse marker *like* is followed by a filler without further discourse, and thus is considered to be in a non-clausal position. In example (130), *like* is in a non-clausal position and serves to reiterate the expression *for example*.

5.3 Discussion

The total occurrences of clause-initial *like* are the least frequent position of the discourse marker *like* among the four varieties. This result contradicts the findings of previous studies which posit that this is one of the most prevalent positions for *like* as a discourse marker (cf.

Schourup, 1983; Underhill, 1988; Miller, 2009). In contrast, the clause-final position of the discourse marker *like* is more popular among the four varieties. This result also contradicts the findings of previous studies that have identified this position as less common (cf. section 2.3.2.3). It is possible that this discrepancy can be attributed to the fact that the aforementioned studies primarily concentrated on the positional distribution of the discourse marker *like* among Inner Circle varieties, whereas the four selected varieties in the present study are classified as Outer Circle varieties.

In comparison to the other three varieties, Philippine English exhibits a greater preference for utilising the clause-initial discourse marker *like* with considerably higher frequency. This may be attributed to the fact that Philippine English has its roots in American English. Consequently, it adheres to the established pattern observed in American English, which has been documented to use the discourse marker *like* with greater frequency in a clause-initial position (Underhill, 1988). Philippine English speakers use the discourse marker *like* in a clause-medial position with the highest frequency among the four positions. This finding is consistent with the literature of Siemund et al. (2009), who also reported a similar result (see section 2.4.4.1). Additionally, the observation that the discourse marker *like* often occurs before a verb in clause-medial position is in alignment with the findings of Walrod (2006) (see section 2.4.4.1).

From a qualitative perspective, Philippine English exhibits a preference for combining a diverse range of other discourse markers and fillers with the clause-initial *like*, such as *I mean*, *you know*, *so*, *yeah*, *uh*, etc., rather than just several of them (cf. section 5.2.1.1). Furthermore, a hedging device *kinda/kind of* is often combined with the discourse marker *like* at both the clause-medial and clause-final positions. This may be attributed to the influence of American English, in which these two expressions are more prevalent. Furthermore, the non-clausal *like* in Philippine English is frequently accompanied by the expression *for example* (cf. section 5.2.4.1). This finding is in contrast to that of Palacio & Gustilo (2016), yet is in alignment with

the findings of Malicsi (2010) (cf. section 2.4.4.1).

Singapore English speakers tend to use their distinctive discourse markers, such as *ya* (of Malay origin) and *aiyah* (of Mandarin Chinese origin), before the clause-initial discourse marker *like* (cf. section 5.2.1.1). This finding is consistent with the observation that four official languages exist simultaneously in Singapore (cf. section 2.4.3). Speakers of different ethnic languages are able to recognise each other through the use of distinct discourse markers. Additionally, another discourse marker *I mean* often occurs before a clause-final discourse marker *like* in Singapore English. Moreover, the hedging device *sort of* is often observed before the clause-medial and non-clausal discourse marker *like*, as seen in Singapore English, which exhibits a distinct pattern from Philippine English.

Hong Kong English speakers use other discourse markers such as *I mean*, *you know*, with greater frequency before a clause-initial discourse marker *like*. In Hong Kong English, the clause-medial *like* is the most frequent position and is often situated between an auxiliary and the lexical verb (cf. 5.2.2.1). Furthermore, it is notable that fillers frequently occur after the clause-final and non-clausal discourse marker *like*.

Indian English speakers employ the clause-medial *like* with the lowest frequency across the four varieties, whereas they exhibit a preference for the clause-final *like*. These findings align with those of Siemund et al. (2009) and Parviainen (2012) (cf. section 2.4.2.1). In contrast to the usage observed in Hong Kong English, Indian English speakers employ other discourse markers and fillers, such as *so* and *uh* with greater frequency before a clause-initial *like*. Furthermore, non-clausal *like* in Indian English is frequently accompanied by fillers and is employed to introduce certain phrasal units.

6. Results and discussion of functional distribution of DM use of *like* across varieties of English

This chapter places particular emphasis on the functional distribution of discourse marker *like* across varieties of English. Section 6.1 describes the results of the descriptive statistical analysis of the functional distribution across varieties of English. Section 6.2 presents the statistical results pertaining to the extent of differences across varieties of English. Finally, section 6.3 presents further discussions and interpretations of the findings presented in sections 6.1 and 6.2.

6.1 Description of functional distribution of DM use of *like* across varieties

Table (13) illustrates the relative frequency of the functional distribution of discourse marker use of *like* across the four different varieties. Table (13) reveals that the primary function of discourse marker use of *like* is its linkage function, with a relative frequency of 242.69 occurrences per 100,000 words. Moreover, the discourse marker *like* is observed to be used as an approximator to a lesser extent across the four varieties, with a relative frequency of 74.52 occurrences per 100,000 words. The second most frequent function of the discourse marker *like* is its focus function, while the third most frequent function is its explanation function. The relative frequency of the repair function of discourse marker *like* and of its usage as an exemplifier are almost identical, with 85.98 and 86.79 occurrences per 100,000 words, respectively.

	Focus	Explanation	Repair	Exemplifier	Approximator	Linkage	Total
ICE-HK	25.60	25.60	14.04	22.30	12.39	35.51	135.45
ICE-Ind	28.53	43.74	22.82	21.87	1.90	52.30	171.16

ICE-Sin	58.65	31.65	14.90	11.17	16.76	64.24	197.38
ICE-Phi	90.64	45.32	34.22	31.45	43.47	90.64	335.72
Total	203.42	146.32	85.98	86.79	74.52	242.69	839.71

Table 13. Relative frequency of functional distribution of DM use of *like* across varieties

Nevertheless, each variety has its own pattern of the functional distribution of discourse marker *like*.

In Hong Kong English, the most common function of *like* as a discourse marker is to create linkage. Furthermore, the focus function of the discourse marker *like* is the second most frequent function, occurring with the same relative frequency of 25.60 occurrences as the explanation function. The exemplifier function is slightly less prevalent than its focus and explanation functions, with a relative frequency of 22.30 occurrences per 100,000 words. The repair function of the discourse marker *like* is the second least frequent function in Hong Kong English. Finally, the approximator function is the least frequent of the six functions.

In Indian English, the linkage function of discourse marker *like* is the most frequently occurring of the six functions, with a relative frequency of 52.30 occurrences. The explanation function occupies the second position, with a relative frequency of 43.74, which is not significantly lower than that of the linkage function. The focus function occupies the third position, with a relative frequency of 28.53 occurrences, which represents a relatively lower frequency in comparison to its explanation function. There is a slight difference in the frequency of the repair function and the exemplifier function of discourses marker *like*, with 22.82 and 21.87 occurrences, respectively, in Indian English. Finally, discourse marker *like* used as an approximator still occupies the lowest position among the six functions, with an extremely low relative frequency of 1.9 occurrences per 100,000 words.

In Singapore English, as in Hong Kong English and in Indian English, the linkage function of discourse marker *like* occupies the first position among the six functions. The focus function of *like* as a discourse marker occupies the second position, with a relative frequency of 58.65 occurrences. The explanation function is situated in the third position, exhibiting a relatively greater difference to its focus function, with a frequency of 31.65 occurrences per 100,000 words. It should be noted that, in contrast to the findings in Hong Kong English and Indian English, the approximator function is not the least frequent in Singapore English. This function occupies the fourth position in Singapore English, with a relative frequency of 16.76 occurrences. The two least frequent functions of the discourse marker *like* in Singapore English are its repair and exemplifier functions, with a frequency of 14.90 and 11.17 occurrences, respectively.

In Philippine English, the linkage function of discourse marker *like* remains the most frequent, accompanied by the focus function. These two functions exhibit the same frequency, with an occurrences rate of 90.64 occurrences per 100,000 words. Subsequently, the explanation function of the discourse marker *like* occupies the third position, with a frequency of 45.32 occurrences. It can be observed that the approximator frequency is not the least frequent in Philippine English. Furthermore, it has an almost identical frequency of 43.47 occurrences with the explanation function. As observed in Singapore English, the two least frequent functions of the discourse marker *like* in Philippine English are also the repair and the exemplifier functions, with a frequency of 34.22 and 31.45, respectively. In conclusion, the functional distribution of the discourse marker *like* in Philippine English exhibits a pattern of similarity to that observed in Singapore English.

Figure 7. provides a more comprehensive overview of the functional distribution of discourse marker *like* across the four varieties.

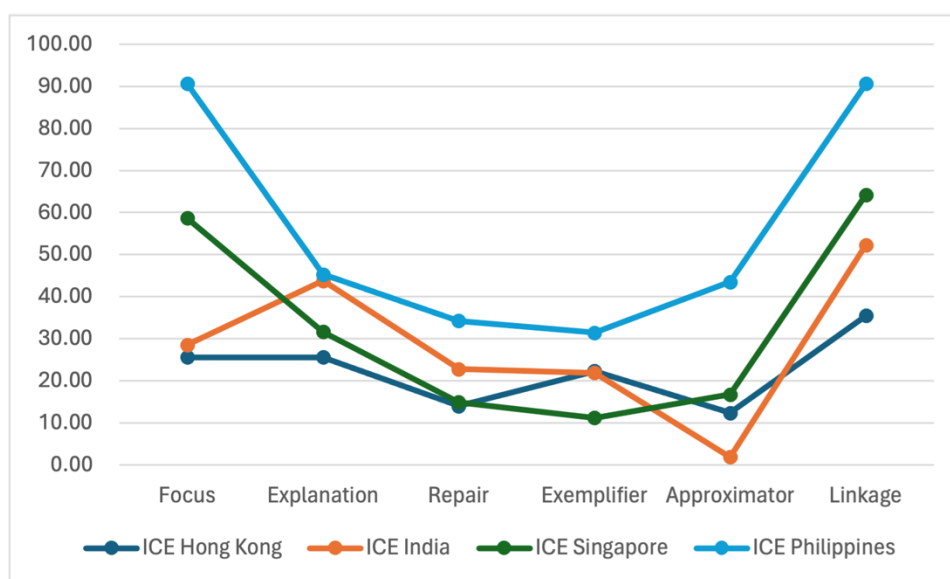


Figure 7. Functional distribution of DM *like* across varieties of English

Figure 7 illustrates that, in each function of the discourse marker *like*, Philippine English exhibits the highest frequency of occurrence, surpassing the other three varieties. The functional distribution of the utilisation of discourse marker *like* in Singapore English is similar to that observed in Philippine English, albeit with a relatively lower frequency in each function.

With regard to the focus function of the discourse marker *like*, Philippine English displays a remarkably higher frequency than the other three varieties. Subsequently, the frequency of this function in Singapore English is strikingly higher than in Indian English and Hong Kong English. In this function, Indian English and Hong Kong English have a similar frequency.

In the explanation function of discourse marker *like*, it is noteworthy that Philippine English and Indian English have a similar frequency and occupy the first and the second places. Subsequently, Singapore English occupies the third position. Finally, Hong Kong English speakers use this function the least frequently. However, the frequency of use of this function in Singapore English is not significantly higher than that observed in Hong Kong English.

The repair function of discourse marker *like* is employed with the greatest infrequency by

speakers of Singapore English and Hong Kong English. The second most frequent position within this function is occupied by Indian English speakers. Philippine English retains the top position in this function.

In terms of the exemplifier function of discourse marker *like*, Philippine English occupies the foremost position. The frequency of use of this function is almost identical in India English and Hong Kong English. However, the frequency of use of this function in Philippine English is not significantly higher than that observed in India English and Hong Kong English. Speakers of Singapore English employ this function of discourse marker *like* with the lowest frequency, despite exhibiting a relatively high frequency of overall usage of discourse marker *like*.

The use of discourse marker *like* as an approximator is most prevalent in Philippine English, with a strikingly higher frequency than in the other three varieties. It is noteworthy that this function is almost undetectable in Indian English, with only two documented instances among the total 171 instances of discourse marker *like* in Indian English. Similarly, this function is not particularly prevalent in Singaporean and Hong Kong English. The use of this function is slightly more prevalent among Singapore English speakers than among those in Hong Kong.

Regarding the linkage function of the discourse marker *like*, Philippine English exhibits a remarkably high frequency of usage in comparison to the other three varieties. Singapore English occupies the second position, while Indian English and Hong Kong English are in the third and last positions, respectively. The frequency of this function across the four varieties is consistent with the overall frequency of *like* as a discourse marker across the four varieties. In other words, the frequency is displayed in a descending order, with Philippine English, Singapore English, Indian English and finally Hong Kong English occupying the respective positions.

The association plot in Figure 8 further explains the functional distribution of discourse marker *like* across varieties of English.

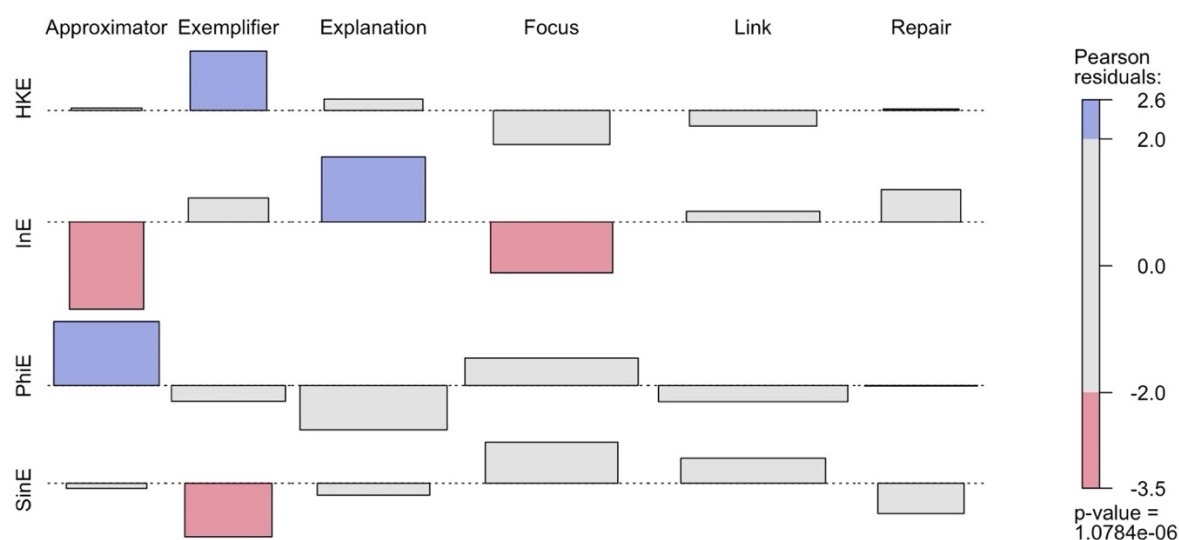


Figure 8. Association plot of functional distribution of DM *like* across varieties

As illustrated in Figure 8, the exemplifier function of *like* as a discourse marker is significantly overrepresented in Hong Kong English. In Indian English, the explanation function of the discourse marker *like* is highly overrepresented, whereas the focus and approximator functions are markedly underrepresented. In contrast, the approximator function of the discourse marker *like* is observed to be highly overrepresented in Philippine English. In Singapore English, the exemplifier function of *like* as a discourse marker is significantly underrepresented.

6.2 Results of the differences across varieties in functional distribution of DM use of *like*

The following sections seek to provide further verification of the correlation between the functional distributions of discourse marker *like* and varieties. The result of the chi-squared test indicates there are significant difference in the functional distribution of discourse marker *like* across varieties of English ($\chi^2 = 56.3$, $df = 15$, $p < 0.001$). The results of the post hoc test, as presented in Table 14, with the exception of the results between Hong Kong English and Philippine English and between Philippine English and Singapore English, statistically

significant differences are attested in each pair of varieties in the functional distribution of discourse marker use of *like*. It is worthy of note that the result of the difference between Hong Kong English and Philippine English is just slightly above the threshold of statistical significance ($p = 0.0502$). The results of the post hoc test serve to further validate one of the findings of the descriptive statistics, namely that Philippine English and Singapore English exhibit similarities in the functional distribution of the discourse marker *like* (section 6.1).

Table 14. Results of the functional distribution of DM *like* between each two of the varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	9.73e-03	1.46e-02
HkE vs. PhiE	4.18e-02	5.02e-02
HkE vs. SinE	3.97e-03	7.94e-03
IndE vs. PhiE	5.83e-07	3.50e-06
IndE vs. SinE	1.16e-05	3.48e-05
PhiE vs. SinE	1.54e-01	1.54e-01

6.2.1 Focus marker usage of *like* across varieties of English

The result of the chi-squared test for the focus function of discourse marker *like* indicates that there are statistically significant differences across the four varieties of English ($\chi^2 = 56.09$, $df = 3$, $p < 0.001$). The results of the post hoc test, as presented in Table 15, indicate that, with the exception of the difference between Hong Kong English and Indian English, statistically significant differences are detected between each pair of varieties.

Table 15. Results of focus function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	8.98e-01	8.98e-01
HkE vs. PhiE	3.66e-09	1.83e-08
HkE vs. SinE	9.65e-04	2.90e-03
IndE vs. PhiE	1.85e-09	1.11e-08
IndE vs. SinE	6.22e-04	2.49e-03
PhiE vs. SinE	5.81e-03	1.16e-02

The results presented in Table 15 corroborate the findings of the descriptive statistics, indicating that speakers of Philippine English utilise discourse marker *like* with greater frequency than the other three varieties to highlight focus. Furthermore, there is no significant quantitative difference between the usage of this function by speakers of Hong Kong English and Indian English.

6.2.1.1 Qualitative analysis within varieties of English

In Hong Kong English, the focus function of discourse marker *like* is the second most prevalent among the six functions. Interestingly, a relation between the position and this function can be observed in the dataset of Hong Kong English: over half of the instances occur in a clause-medial position (18 instances out of the 31 instances). In most of the cases, *like* occurs before a noun phrase, as illustrated in (132), or an adjective, as illustrated in (133). It is noteworthy that in (133), the focus function of the discourse marker *like* is further reinforced by the intensifier *so*:

(132) Why you can take some *like* motion movement but I can't (ICE Hong Kong: S1A-011: 184).

(133) She was *like* so shock (ICE Hong Kong: S1A-041: 602).

In Indian English, the focus function of the discourse marker *like* is the third most frequent of

the six functions. As in Hong Kong English, one-third of the instances occur in a clause-medial position (11 instances out of the 30 instances). In contrast to Hong Kong English, in Indian English, the focus function of the discourse marker *like* is observed to occur before a verb phrase, as illustrated in (134), or a noun phrase, as illustrated in (135). It is notable that no instances were identified in which the discourse marker *like* was used before an adjective, despite this being a common occurrence in the focus function of the marker *like* (cf. section 3.3.4).

(134) But then nobody *like* motivated me yes I know (ICE India: S1A-021: 242).

(135) It's pouring *like* cats and dog whole day (ICE India: S1A-041: 468).

In Singapore English, the focus function of the discourse marker *like* represents the second most frequent of the six identified functions. Surprisingly, this function of *like* exhibits a strong relation with the clause-medial position (53 instances out of the 63 instances). The focus function of *like* can be observed modifying a variety of constituents, including noun phrases (136), verb phrases (137), adjective phrases (138 and 141), prepositional phrases (139), and adverbs (140). Note that in (141), the focus function of the discourse marker *like* is further reinforced by the intensifier *very*:

(136) Can you imagine *like* countdown two days to your wedding (ICE Singapore: S1A-002: 15).

(137) No but you don't have to *like* give him (ICE Singapore: S1A-031: 390).

(138) Why is this one *like* much louder than anyone else's spirituality (ICE Singapore: S1A-008: 77).

(139) I heard a voice from *like* out of nowhere (ICE Singapore: S1A-031: 377).

(140) Belvedere is *like* enough (ICE Singapore: S1A-016: 196).

(141) He is *like* very slow but he can recognise places (ICE Singapore: S1A-048: 663).

In Philippine English, the focus function of discourse marker *like* is the most frequent among

the six functions. Philippine English also exhibits a strong correlation between the clause-medial position and this function of discourse marker *like* (74 instances out of the 98 instances). The focus function of *like* has the capacity to emphasise a range of constituents, including noun phrases (142), adjective phrases (145 and 146), prepositional phrases (143), and verb phrases (144). It is noteworthy that in (145) and (146), the focus function of the discourse marker *like* is further reinforced by the intensifiers *really* and *so*, respectively:

- (142) If if they can they can access *like* different offices just through that bar code (ICE Singapore: S1A-012: 201).
- (143) Yeah everybody was *like* in war point and everybody started creaming (ICE Singapore: S1A-034: 621).
- (144) So I have to to *like* leave the house early to (ICE Singapore: S1A-039: 724).
- (145) Uh she's *like* really nice (ICE Singapore: S1A-048: 928).
- (146) They were *like* so stupid and mindless (ICE Singapore: S1A-041: 756).

6.2.2 Explanation marker usage of *like* across varieties of English

The result of the chi-squared test for the explanation function of discourse marker *like* indicates that statistically significant differences do not exist across the four varieties of English ($\chi^2 = 5.85$, $df = 3$, $p = 0.1191$). The results of the post hoc test, as presented in Table 16, indicate that no statistically significant differences are detected between each pair of varieties.

The results presented in Table 16 confirm the findings of the descriptive statistics, namely that the difference between Hong Kong English and Singapore English and the difference between Indian English and Philippine English are inconsequential. Furthermore, minor quantitative differences exist between the remaining pairs of varieties in their utilisation of this function of discourse marker *like*.

Table 16. Results of explanation function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	0.0874	0.437
HkE vs. PhiE	0.0442	0.265
HkE vs. SinE	0.7100	1.000
IndE vs. PhiE	0.7580	1.000
IndE vs. SinE	0.1800	0.540
PhiE vs. SinE	0.0997	0.437

6.2.2.1 Qualitative analysis within varieties of English

In Hong Kong English, the explanation function of the discourse marker *like* is also the second most frequent of the six functions, occurring with the same frequency as the focus function. This function also has a correlation with the positional distribution: over one-third of the instances occur in a clause-initial position (13 instances out of the 30 instances), aiming to supply more information on the preceding discourse. In some cases, *like* occurs before a question, serving to request more information about the target topic, as illustrated in (147). In other cases, the occurrence of discourse marker *like* points out that this is a shared knowledge as a further explanation. Furthermore, another discourse marker *you know* frequently occurs before the discourse marker *like* which serves to supply an explanation. The following examples serve to illustrate this function of the discourse marker *like* in Hong Kong English:

(147) And *like* have you ever read some books about the ritual (ICE Hong Kong: S1A-025: 389).

(148) I I think so

Like you told the different two style (ICE Hong Kong: S1A-027: 412).

(149) You know *like* if they give me a tip or (ICE Hong Kong: S1A-039: 556).

In example (148), the speaker uses the discourse marker *like* to introduce a given information provided by the listener in the preceding conversations. In example (149), the occurrence of the discourse marker *you know* indicates that the following information will be an explanation of the preceding discourse, the discourse marker *like* has the same function as *you know*.

In Indian English, the explanation function of the discourse marker *like* is the second most frequent of the six functions. Similarly, the clause-initial position is correlated with this function: 19 instances out of the 46 instances with the explanation function occurring in this position. In some cases, the discourse marker *like* introduce a more specific reformulation of the elements before *like*, as illustrated in (150). Similarly, as in Hong Kong English, *like* can introduce a question, serving to request for more information, as illustrated in (151). Furthermore, *like* can introduce a possibility which serves as a supplementary information of the preceding proposition, as illustrated in (152):

(150) Latest we had, uh *like* the latest movie this Lamhe of Sridevi, that, they don't, the whole thing is uh picturised out there (ICE India: S1A-008: 81).

(151) And so what do you think about their, this one future or, uh
like will it be good or, uh will there be any, politics or what do you think about it (ICE India: S1A-050: 644).

(152) A: There is no, indepth knowledge given about any particular subject

B: And

A: *Like* if the person is from a higher background like, we find that that knowledge is not (ICE India: S1A-042: 478).

In Singapore English, the explanation function of the discourse marker *like* is the third most frequent of the six functions. In contrast to the former two varieties, in Singapore English, nearly half of the instances of this function occur in a non-clausal position (15 instances out of

the 34 instances). There are a lot of cases in which the conjunction *because* occurs before the discourse marker *like*, as *because* is a causal conjunction, the information after *like* is no doubt an explanation, as illustrated in (153). Furthermore, other discourse markers *I mean* and *so* often occurs before the discourse marker *like* which serves to supply additional information or present the background to the preceding discourse, as illustrated in (154) and (155), respectively:

(153) I may have to do it because *like* one teacher told me (ICE Singapore: S1A-011: 123).

(154) A: I drive there

B: Huh

A: Walk there

B: No I mean *like* other places definitely you'll you'll drive (ICE Singapore: S1A-023: 265).

(155) So *like* when both of us just stand and just look at him you know he said hello (ICE Singapore: S1A-031: 375).

In Philippine English, the explanation function of the discourse marker *like* is the third most frequent of the six functions. This function of the discourse marker *like* mainly occurs in two positions: the clause-initial position (20 instances out of the 49 instances) and followed by the non-clausal position (18 instances out of the 49 instances). With this function, as in Singapore English, other discourse marker *I mean* frequently occurs before the discourse marker *like*, as illustrated in (156). Moreover, also as in Singapore English, the conjunction *because* is used before *like* in a truncated form *'cause*, as illustrated in (158). As in Hong Kong and Indian English, the discourse marker *like* with the explanation function is frequently used to introduce a question, aiming at requesting for more details, as illustrated in (157). Finally, as in Hong Kong English, another discourse marker *you know* frequently preceding the discourse marker *like* with the explanation function, as illustrated in (159):

(156) A: Why don't you tell me about yo

B: Okay you ask me something anything ask me I mean, *like*

what do we guys do in Manila for fun (ICE Philippine: S1A-029: 530).

(157) A: However my Dad, uhm is still deciding is still deciding actually which is very expensive

B: *Like* how much is it

A: Three thousand two hundred (ICE Philippine: S1A-041: 753).

(158) It's it could really get frustrating 'cause *like* you have no one to talk to (ICE Philippine: S1A-012: 210).

(159) Some of them involve uh you know *like* one said this slanderous remarks (ICE Philippine: S1A-006: 83).

6.2.3 Repair marker usage of *like* across varieties of English

The result of the chi-squared test for the repair function of discourse marker *like* indicates that statistically significant differences are detected across the four varieties of English ($\chi^2=11.957$, $df = 3$, $p < 0.01$). As illustrated in Table 17, the results of the post hoc test demonstrate that statistically significant differences are attested between Hong Kong English and Philippine English, as well as between Singapore English and Philippine English. No statistically significant differences are identified between the remaining pairs of varieties.

Table 17. Results of repair function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	0.27400	0.6180
HkE vs. PhiE	0.00650	0.0325
HkE vs. SinE	0.86200	0.8620
IndE vs. PhiE	0.09600	0.3840

IndE vs. SinE	0.20600	0.6180
PhiE vs. SinE	0.00392	0.0235

The results presented in Table 17 verify the findings of the descriptive statistics. In other words, the significant differences are only attested between the top variety, namely Philippine English, and the two varieties with the lowest frequency of this function of discourse marker *like*.

6.2.3.1 Qualitative analysis within varieties of English

In Hong Kong English, the repair function of the discourse marker *like* occupies the second least frequent of the six functions. This function of the discourse marker *like* has a correlation with the clause-final position, as evidenced by 12 out of the 17 instances. In the majority of instances, the discourse marker *like* functions as a marker to indicate a restart of the discourse (cf. section 3.3.4), as illustrated in (160) and (161):

- (160) That's okay just, just like me to *like* to speak English yah maybe a bit shy but, only more practice yeah (ICE Hong Kong: S1A-010: 162).
- (161) The just fix it *like* uh you you you have to fix the phone number (ICE Hong Kong: S1A-027: 409).

In Indian English, the repair function of the discourse marker *like* occupies the third least frequent of the six functions. As in Hong Kong English, this function of the discourse marker *like* is correlated with the clause-final position, as evidenced by 19 out of the 24 instances. In most instances of Indian English, the discourse marker *like* functions as a self-repair marker, rectifying a minor error that has occurred in the preceding utterance. This is exemplified in (162). Furthermore, another common occurrence is that the discourse marker *like* occurs in a clause-final position, indicating the termination of the current discourse (cf. section 3.3.4), as illustrated in (163). Finally, the occurrence of the discourse marker *like* in this situation is less frequent than in Hong Kong English. It can, however, function as a marker to indicate a restart of the discourse, as illustrated in (164):

(162) we feel there is nothing *like*

something like God but there's something some superpower (ICE India: S1A-012: 136).

(163) If you come out uhm, if you go and come out

Uhm, that means uh, you are a very pure person in heart *like* (ICE India: S1A-012: 144).

(164) Uhn they don't uh *like* they don't want to understand the topic (ICE India: S1A-042: 485).

In Singapore English, the repair function of the discourse marker *like* occupies the second least frequent of the six functions. As in Hong Kong and Indian English, this function of the discourse marker *like* is correlated with the clause-final position, as evidenced by 12 out of the 16 instances. Most of the time, *like* is used to indicate a restart of the discourse and to mark the termination of the current discourse, as illustrated in (165) and (166), respectively:

(165) We're going to do after the trip *like* after we've (ICE Singapore: S1A-009: 98).

(166) A: Throw it out

B: and you can put in other ingredients *like*

A: So as long as it's frothing just scoop off all the froth (ICE Singapore: S1A-019: 207).

In Philippine English, the repair function of the discourse marker *like* occupies the second least frequent of the six functions. As in earlier varieties of English, this function of the discourse marker *like* also exhibits a correlation with the clause-final position, as evidenced by 26 out of the 37 instances. Similarly to Indian English, the discourse marker *like* can introduce a self-repair, as illustrated in (167). Moreover, the discourse marker *like* can indicate a restart, as is the case in all former varieties of English. This is demonstrated by the following example (168). Similarly to Indian and Singapore English, the discourse marker *like* can indicate the end of the current utterance, as illustrated in (169):

(167) Uhm actually whe when you're a bit creative you can actually think of ways to *like* keep you keep yourself busy (ICE Philippine: S1A-034: 636).

(168) He gives me *like* he gives me work every day (ICE Philippine: S1A-012: 204).

- (169) A: You know what, what we should do, we should *like*,
 B: What man
 A: Clean ourselves up (ICE Philippine: S1A-048: 904).

6.2.4 Exemplifier usage of *like* across varieties of English

Table 18. Results of exemplifier function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	0.57200	0.74000
HkE vs. PhiE	0.37000	0.74000
HkE vs. SinE	0.01630	0.08150
IndE vs. PhiE	0.14500	0.43500
IndE vs. SinE	0.06300	0.25200
PhiE vs. SinE	0.00118	0.00708

The result of the chi-squared test for the exemplifier function of discourse marker *like* reveals that there are statistically significant differences across the four varieties of English ($\chi^2 = 10.583$, $df = 3$, $p < 0.05$). As evidenced by the post hoc test results presented in Table 18, the sole statistically significant difference is observed between Philippine English and Singapore English. No statistically significant differences are attested between the remaining pairs of varieties.

The results presented in Table 18 echoes the findings of the descriptive statistics, indicating that the significant difference is only identified between the variety with the highest frequency in using this function of discourse marker *like*, namely Philippine English and the variety with the lowest frequency in using this function of discourse marker *like*, namely Singapore English.

6.2.4.1 Qualitative analysis within varieties of English

In Hong Kong English, the exemplifier function of the discourse marker *like* occupies the third least frequent of the six functions. This function of the discourse marker *like* exhibits a correlation with the clause-medial position, as evidenced by 26 out of the 27 instances. The discourse marker *like* with this function is used to introduce a typical representation rather than an actual situation of the elements occurring before *like*, as illustrated in (170):

- (170) Why don't why don't she cook uh the Western Western food *like* steak (ICE Hong Kong: S1A-003: 33).

In Indian English, the exemplifier function of the discourse marker *like* occupies the second least frequent of the six functions. A correlation can be observed between this function of the discourse marker *like* and the positional distribution. Specifically, 9 out of the 23 instances occur in a clause-medial position, while 7 out of the 23 instances occur in a non-clausal position. Similarly to Hong Kong English, the discourse marker *like* introduces a typical situation described in the preceding utterance. This is illustrated in (171). Moreover, the discourse marker *like* serves to reinforce the expression *for example*, as evidenced in (172):

- (171) You have a very famous uh feminist, poet, *like* Jayprabha (ICE India: S1A-014: 168).

- (172) A: So I like to do a lot more things, before getting married

B: So *like* for example Diploma (ICE India: S1A-024: 278).

In Singapore English, the exemplifier function of the discourse marker *like* occupies the least frequent of the six functions. In 50% of cases (6 out of the 12 instances), the discourse marker *like* with this function is employed in a non-clausal position. Similarly, as with the preceding two varieties of English, the discourse marker *like* introduces a typical example that is sufficient for the listener to comprehend the speaker's meaning. This is demonstrated in (174). Moreover, the discourse marker *like* serves to reinforce the expression *for instance*, as illustrated in (173):

(173) What I've asked him to do so far is *like* for instance the
sewage unit uh once the reply came back (ICE Singapore: S1A-015: 182).

(174) A: You mean that kind of thing including the three major departments

B: Which three

A: *Like* Bursar's and so on

B: Ya ya ya (ICE Singapore: S1A-045: 604).

In Philippine English, the exemplifier function of the discourse marker *like* is the least frequent of the six functions. In 50% of cases (17 out of the 34 instances), the discourse marker *like* with this function is employed in a clause-initial position. Similarly to the former varieties, the discourse marker *like* is primarily employed to introduce a typical example, thereby facilitating the listener's comprehension of the speaker's thought. This is exemplified in (175). Moreover, in Philippine English, the instances in which the discourse marker *like* serves to reinforce the expression *for example* are more frequent than the former varieties of English, as demonstrated in (176). Additionally, there is one instance in which the discourse marker *like* serves to reinforce the expression *for instance*, as demonstrated in (177), respectively:

(175) You want to hold a major position *like* secretary auditor (ICE Philippine: S1A-008: 165).

(176) He also, was uh I don't know a very responsible person uh in this sense 'no *like* for example we the staff I mean I would make him laugh with booboos (ICE Philippine: S1A-009: 176).

(177) I mean in the government positions because uhm *like* for instance no matter if we keep on electing officials we keep on choosing from them (ICE Philippine: S1A-021: 368).

6.2.5 Approximator usage of *like* across varieties of English

The result of the chi-squared test for the focus function of discourse marker *like* indicates that there are statistically significant differences across the four varieties of English ($\chi^2 = 52.732$, $df = 3$, $p < 0.001$). As illustrated in Table 19, the post hoc test results indicate that, with the

exception of the difference between Hong Kong English and Singapore English, statistically significant differences are observed between each pair of varieties.

Table 19. Results of approximator function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	1.62e-03	3.24e-03
HkE vs. PhiE	4.82e-05	2.41e-04
HkE vs. SinE	6.02e-01	6.02e-01
IndE vs. PhiE	1.29e-10	7.74e-10
IndE vs. SinE	3.47e-04	1.29e-03
PhiE vs. SinE	3.22e-04	1.29e-03

The results presented in Table 19 corroborate the findings of the descriptive statistics, namely that the usage of this function of discourse marker *like* is markedly more prevalent in Philippine English than in the other three varieties. Moreover, its usage in Singapore English is slightly more frequent than in Hong Kong English (section 6.1).

6.2.5.1 Qualitative analysis within varieties of English

In Hong Kong English, the approximator function of the discourse marker *like* occupies the least frequent of the six functions. A correlation can be observed between this function of the discourse marker *like* and the positional distribution. Specifically, 7 out of the 15 instances occur in a clause-medial position, while 6 out of the 15 instances occur in a clause-initial position. In most of the cases, the discourse marker *like* occurs before a numerical expression, as illustrated in (178). Moreover, as illustrated in example (72) (cf. section 3.4.4), the discourse marker *like* can be used to introduce other abstract measurable units (cf. section 3.4.4).

(178) For example on one page they will have *like* uh twenty to thirty columns so so (ICE Hong

Kong: S1A-013: 241).

In Indian English, the approximator function of the discourse marker *like* occupies the least frequent of the six functions. The approximator function of the discourse marker *like* was observed on only two occasions, with an extremely low frequency. In example (179), the discourse marker *like* is also considered an approximator due to the adverb *almost*, indicating an approximate time period. In example (180), the discourse marker *like* is used to indicate the approximate start date of the holiday, thereby relieving the speaker of the obligation to provide an exact date.

(179) I went when I was very young, *like* a child almost, so I don't remember anything clearly
(ICE India: S1A-014: 159).

(180) A: So, have your holidays started Rashmi,

B: Yes aunty from, *like*, from last, Monday (ICE India: S1A-044: 507).

In Singapore English, the approximator function of the discourse marker *like* occupies the third least frequent of the six functions. It is notable that the majority of the instances occur in a clause-medial position, as evidenced by 14 out of the 18 instances. In most of the cases, the discourse marker *like* is positioned before a numerical expression, as illustrated in (181). Moreover, the discourse marker *like* can occur before a time period, as illustrate in (182):

(181) It's about *like* six seven hundred (ICE Singapore: S1A-003: 24).

(182) It's *like* uh only about a few months (ICE Singapore: S1A-042: 564).

In Philippine English, the approximator function of the discourse marker *like* occupies the third least frequent of the six functions. Similarly to the preceding varieties of English, the majority of the discourse marker *like* in this function occur in a clause-medial position, as evidenced by 37 out of the 47 instances. The two most prevalent pattern is still the discourse marker *like* introduces a time point or the amount of money, as illustrated in (183) and (184), respectively:

- (183) And I have to wake up *like* five o'clock in the morning (ICE Philippine: S1A-018: 345).
- (184) The food is okay quite expensive 'cause a meal would cost *like* a hundred fifty bucks (ICE Philippine: S1A-012: 208).

6.2.6 Linkage usage of *like* across varieties of English

The result of the chi-squared test for the repair function of discourse marker *like* indicates that statistically significant differences are detected across the four varieties of English ($\chi^2 = 25.4$, $df = 3$, $p < 0.001$). As illustrated in Table 20, the results of the post hoc test demonstrate that statistically significant differences are attested between Hong Kong English and Philippine English, as well as between Indian English and Philippine English. No statistically significant differences are identified between the remaining pairs of varieties.

The results presented in Table 20 provide further clarification regarding the findings of the descriptive statistics. In other words, despite the fact that this function of discourse marker *like* is used far more frequently in Philippine English than in the other three varieties, the statistically significant differences are only attested between two pairs of the varieties. Furthermore, it is notable that the difference between Hong Kong English and Singapore English in the utilisation of this function of the discourse marker *like* is marginally above the threshold of statistical significance ($p = 0.0560$).

Table 20. Results of linkage function of DM *like* across varieties of English

Comparison between groups	p. chi-square	p. adjusted. chi-square
HkE vs. IndE	2.25e-01	4.18e-01
HkE vs. PhiE	3.62e-06	2.17e-05
HkE vs. SinE	1.40e-02	5.60e-02
IndE vs. PhiE	5.08e-04	2.54e-03
IndE vs. SinE	2.09e-01	4.18e-01

6.2.6.1 Qualitative analysis within varieties of English

In Hong Kong English, the linkage function of the discourse marker *like* is observed to occupy the most frequent of the six functions. A correlation can be observed between this function of the discourse marker *like* and the positional distribution. In particular, 17 out of the 43 instances occur in a non-clausal position, while 13 out of the 43 instances occur in a clause-final position. In most of the cases, the discourse marker *like* is employed as a floor-holding device, enabling the speaker to gain additional process time. This is exemplified in (122), as discussed in section 5.2.4.1. The speaker uses *like* five times in order to gain additional time to organise their speech. Additionally, the discourse marker *like* serves to reinforce the hedging device *kind of*, as illustrated in (185):

(185) A: Yeah we buy it in uh you know uh in in, in the street

A: They are

B: Kind of *like* a burger stand (ICE Hong Kong: S1A-009: 144).

In Indian English, the linkage function of the discourse marker *like* occupies the most frequent of the six functions. It is noteworthy that the majority of the instances occur in a non-clausal position, as evidenced by 31 out of the 55 instances. Similarly, as observed in Hong Kong English, the discourse marker *like* is surrounded by fillers, which appear to be used as a means of maintaining control of the conversation, as illustrated in (186). Moreover, the discourse marker *like* serves to reinforce the hedging device *sort of*, as illustrated in (187):

(186) They still follow those uh *like like* uh you call it when we go (ICE India: S1A-008: 82-83).

(187) You know you know you get the creep sort of *like* you see, my God what some ghost

some haunt, so I always (ICE India: S1A-040: 454).

In Singapore English, the linkage function of the discourse marker *like* occupies the most frequent of the six functions. A correlation can be observed between this function of the discourse marker *like* and the positional distribution. Specifically, 28 out of the 69 instances occur in a non-clausal position, while 23 out of the 69 instances occur in a clause-final position. In Singapore English, the discourse marker *like* is surrounded by not only fillers, as illustrated in (188), but also by other discourse markers, as illustrated in (189), in order to gain more time to maintain control of the utterance. Moreover, the instances in which the discourse marker *like* serves to reinforce the hedging device *sort of* are more frequent than the former two varieties of English. This usage is illustrated in (190):

(188) He said hello uh *like* uh

I heard a voice from like out of nowhere (ICE Singapore: S1A-031: 376).

(189) You know *like like* I think you you should have given it to him (ICE Singapore: S1A-031: 370).

(190) When we came back we sort of *like* everybody was fighting for that that check (ICE Singapore: S1A-016: 195).

In Philippine English, the linkage function of the discourse marker *like* occupies the most frequent of the six functions, occurring with the same frequency as the focus function. A correlation can be observed between this function of the discourse marker *like* and the positional distribution. Specifically, 43 out of the 98 instances occur in a clause-final position, while 27 out of the 98 instances occur in a non-clausal position. In Philippine English, the discourse marker *like* is surrounded by fillers and other discourse markers simultaneously, as illustrated in (191). Additionally, the discourse marker *like* can serve to reinforce the hedging device *kinda/kind of*, as illustrated in (100) and (101) in section 5.2.2.1, as well as in (118) and (119) in section 5.2.3.1. Moreover, an additional hedging device, *more of*, may be used in

conjunction with the discourse marker *like*, as illustrated in (192):

(191) So *like* uhm so when I heard that uhm (ICE Philippine: S1A-047: 840).

(192) But but I you know I tend to think that it's more of *like* a rebound thing for for Joey (ICE Philippine: S1A-006: 88).

6.3 Discussion

In general, the four varieties of English demonstrate both similarities and differences in the functional distribution of the discourse marker *like*. Firstly, the four varieties of English exhibit a general preference for the linkage function of the discourse marker *like*. In other words, the linkage function is the most prevalent of the six functions in each variety. This is consistent with the findings of Hansen (1977), who notes that the primary function of discourse markers is to facilitate connectivity (see section 2.2.5.1). Furthermore, while Philippine English employs the linkage function of *like* in a clause-final position with greater frequency, the other three varieties of English tend to utilise the function of *like* in a non-clausal position.

From a qualitative perspective, all four varieties demonstrate a tendency to utilise the linkage function of *like* in conjunction with fillers. These findings corroborate Valentine's (1991) assertion regarding Indian English that the linkage function of the discourse marker *like* serves to maintain coherence in discourse (cf. section 2.4.2.1). Furthermore, the Singapore English and Indian English varieties demonstrate a preference for utilising the hedging device *sort of* in conjunction with the linkage function of *like*. This finding is in accordance with Sand's (2013) assertion regarding Singapore English (cf. 2.4.3.1). Additionally, both Hong Kong English and Philippine English employ the hedging device *kind of* before the linkage function of *like*. It is noteworthy that Philippine English also employs another hedging device, *more of*, in combination of this function of the discourse marker *like*. These findings align with those of Palacio & Gustilo (2016) and Walrod (2006) for Philippine English (cf. section 2.4.4.1). They both point out the popularity of employing the discourse marker *like* to reinforce the hedging

devices in Philippine English.

Secondly, while the explanation function of the discourse marker *like* occupies the third most frequent of the six functions in Singapore English, the other three varieties of English utilise this function of the discourse marker *like* with a frequency that is the second most frequent of the six functions. This is consistent with the findings of Jucker & Ziv (1998), who indicate that a frequent function of discourse markers is that they can be used as “confirmation seekers” (cf. section 2.2.5.6). Furthermore, while Singapore English displays a preference for employing this function of the discourse marker *like* in a non-clausal position, the other three varieties exhibit a tendency to utilise this function in a clause-initial position.

From a qualitative perspective, with the exception of Singapore English, the other three varieties tend to use the explanation function of the discourse marker *like* to introduce a question, with the aim of requesting further explanation of the topic under discussion. However, there are similarities between Philippine English and Singapore English in the use of this function of the discourse marker *like*. Both varieties employ another discourse marker, namely *I mean* or a causal conjunction, *because*, preceding the discourse marker *like* with an explanatory function (cf. section 6.2.2.1). The sole distinction is that Philippine English speakers tend to utilise a shortened form of *because*, namely *'cause*. This may be attributed to the fact that Philippine English is derived from American English.

The function of the discourse marker *like* to highlight the focus plays an important role in all four varieties of English. In Hong Kong English, the discourse marker *like* is employed in this function with the second highest frequency, equaling that of the explanation function. In the case of Indian English, the function is employed the third most frequently of the six functions. In Singapore English, it is the second most frequent of the six functions. This finding is consistent with the results of Sand's (2013) study on Singapore English, as discussed in section 2.4.3.1. In Philippine English, the discourse marker *like* is employed with the greatest

frequency to highlight the focus, ranking as the most frequent of the six functions and occurring at the same frequency as the linkage function.

From a qualitative perspective, with the exception of Indian English, the remaining three varieties frequently utilise the pattern *like+adjective* to indicate the focus function of the discourse marker *like*. This finding is in accordance with that of Meehan (1991) (cf. section 2.3.1.2). Furthermore, with regard to the *like+adjective* construction in this function, different varieties demonstrate somewhat disparate patterns. An adverb frequently occurs between the discourse marker *like* and the adjective in this form, and the adverb varies between varieties (cf. section 6.2.1.1). In Hong Kong English, the adverb is *so*. In Singapore English, the adverb is *very*. In Philippine English, the adverb is *really*. However, Indian English does not utilise the focus function of the discourse marker *like* in this manner.

The discourse marker *like* is employed by Hong Kong English speakers with the exemplifier function with the third highest frequency of the six functions. This finding corroborates that of Fung & Carter (2007), who also report this function of the discourse marker *like* in their dataset. Nevertheless, the exemplifier function of the discourse marker *like* is employed the least frequently in Philippine English and Singapore English. The exemplifier function of the discourse marker *like* is employed with the second lowest frequency of the six functions in Indian English.

From a qualitative perspective, all four varieties employ this function of the discourse marker *like* to introduce a typical representation of the elements preceding *like*, rather than an actual representation. This result is in accordance with Schourup (1983) (cf. 2.3.5.4). With the exception of Hong Kong English, all the other three varieties tend to use this function of the discourse marker *like* in conjunction with the expressions *for example* or *for instance*. This finding aligns with that of Valentine's (1991) study for Indian English (cf. section 2.4.2.1), and aligns with that of Malicsi's (2010) study for Philippine English (cf. section 5.3).

In Indian English, the discourse marker *like* is often used in a clause-final position to perform the function of repair or restart. It is notable that only two pairs of varieties demonstrate a significant difference in the frequency of utilisation of this function of the discourse marker *like*. In other words, the four varieties exhibit a similar frequency of use with regard to this function. Parviainen (2012) posits that Indian English exhibits a preference for the use of discourse markers at the clause-final position. This phenomenon has manifested as a local innovation in India and has subsequently spread to Philippine, Singaporean, and potentially Hong Kong English (cf. section 2.4.2.1). This can partially explain why the four varieties do not differ significantly in the frequency of using this function.

From a qualitative perspective, the Philippine English and Indian English speakers employ this function of the discourse marker *like* to self-repair, to restart the discourse, and to terminate the discourse. In contrast, Hong Kong English speakers only utilise this function to restart the discourse. Singapore English speakers employ this function to restart and terminate the discourse.

In examining the approximator function of the discourse marker *like*, it can be observed that this function is the least frequent of the six functions in both Hong Kong English and Indian English. It is the second lowest frequency of the six functions in Singapore English. The function is employed with the third lowest frequency of the six functions in Philippine English, yet its occurrences in Philippine English are significantly higher than in the other three varieties. This may be attributed to the total instances of the discourse marker *like* in Philippine English being the highest across the four varieties.

From a qualitative perspective, Hong Kong English speakers employ this function of the discourse marker *like* to introduce a general numerical expression. Indian English speakers utilise this approximator function of the discourse marker *like* to introduce a time period. Philippine and Singapore English speakers employ this function of the discourse marker *like*

to introduce both the amount of money and a time period.

7. Conclusion

This study utilised the ICE corpora to examine the usage of the discourse marker *like* across varieties of English in South and Southeast Asia. Specifically, the study investigated the use of *like* across four English varieties: Hong Kong English, Indian English, Singapore English, and Philippine English. The findings reveal both similarities and significant differences in the use of *like* across these varieties. These differences include the proportion of *like* used as a discourse marker versus its non-discourse marker uses, its frequency, positional distribution, and functional distribution. Several explanations have been proposed to account for these significant differences.

Firstly, all four varieties of English show a preference for the non-pragmatic functions of *like*, such as its use as a verb or preposition, over its role as a discourse marker.

Secondly, there are notable differences in the frequency of *like* as a discourse marker across the four varieties. Philippine English exhibits the highest frequency of *like* as a discourse marker. This may be due to its origins in American English, which uses *like* more frequently in this capacity compared to British English, the origin of the other three varieties. Singapore English ranks second in frequency, possibly due to the “English-based bilingualism” policy, which increases the use of *like* in this way. Indian English ranks second lowest, which could be because *like* does not have a nativized source in Indian English, leading speakers to use other, more nativized markers. Hong Kong English shows the lowest frequency, likely due to its historical British influence, reflecting the lower usage of *like* seen in British English.

Thirdly, regarding positional distribution, the study found both similarities and significant differences across the four varieties. Contrary to previous studies that highlight the predominance of clause-initial *like*, the varieties in this study show a greater preference for

clause-final *like*. This difference might be due to previous research focusing on Inner Circle varieties, whereas this study examines Outer Circle varieties. The clause-final position may also be a local innovation in Indian English that has influenced other varieties. Except for Indian English, the clause-medial position is the most frequent across the remaining three varieties. The non-clausal position is relatively frequent in all but Hong Kong English. These varieties also exhibit distinct qualitative patterns in the positional distribution of *like*.

Finally, regarding functional distribution, there are both similarities and significant differences. Consistent with previous studies, the linkage function of *like* is the most prevalent across the four varieties. The explanation function is the second most frequent overall, except in Singapore English. In Singapore English, the focus function is the second most prevalent, while in Indian English, it ranks third. In Hong Kong English, the focus function is tied for second with the explanation function, and in Philippine English, it is the most frequent with the linkage function. The remaining functions—repair, exemplifier, and approximator—are relatively marginal across all varieties. Notably, each variety displays distinct qualitative usage patterns in the functional distribution of *like*.

This study is limited to examining discourse markers in four Outer Circle varieties of English. Future research could explore the use of *like* in Expanding Circle Englishes. Additionally, this study focuses on face-to-face, private conversations. Future studies could examine *like* in other genres, such as public discourse, to gain further insights into its usage. Comparing results across different genres may provide a deeper understanding of the factors influencing the use of *like*. Future research could also investigate additional discourse markers within these four varieties to achieve a more comprehensive understanding of discourse marker usage patterns in these linguistic communities.

References

- Aijmer, K. (2002). English discourse particles. *English Discourse Particles*, 1-315.
- Aijmer, K. (2005). Evaluation and pragmatic markers. In *Strategies in Academic Discourse* (pp. 83-96). Amsterdam: John Benjamins.
- Aijmer, K., & Simon-Vandenberg, A. M. (2011). Pragmatic markers. *Discursive Pragmatics*, 8, 223-247.
- Andersen, G. (1998). The Pragmatic Marker like from a Relevance-theoretic Perspective. In A. H. Jucker & Y. Ziv (Eds.), *Pragmatics and Beyond New Series* (pp. 147-170). Amsterdam & New York: John Benjamins.
- Andersen, G. (2001). *Pragmatic Markers and Sociolinguistic Variation*. 1-366.
- Ansaldo, U. (2009). *Contact Languages: Ecology and Evolution in Asia*. Cambridge University Press.
- Bazzanella, C. (1990). Phatic connectives as interactional cues in contemporary spoken Italian. *Journal of Pragmatics*, 14(4), 629-647.
- Blakemore, D. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Blakemore, D. (2006). Discourse markers. In B. A. Porter & C. M. Davis (Eds.), *The Handbook of Pragmatics* (pp. 221-240). Blackwell.
- Blyth, C., Recktenwald, S., & Wang, J. (1990). I'm like, "say what?!": A new quotative in American oral narrative. *American Speech*, 65(3), 215-227.
- Bolden, G. B. (2006). Little words that matter: Discourse markers "so" and "oh" and the doing of other-attentiveness in social interaction. *Journal of Communication*, 56(4), 661-688.
- Brinton, L. J. (1996). *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin; New York: Mouton de Gruyter.
- Buchstaller, I. (2001). An alternative view of like: Its grammaticalisation in conversational American English. *Edinburgh Working Papers in Applied Linguistics*, 2, 21.
- Buysse, L. (2012). So as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, 44(13), 1764-1782.
- Cameron, R. (2005). Aging and gendering. *Language in Society*, 34(1), 23-61.
- Dailey-O'Cain, J. (2000). The sociolinguistic distribution of and attitudes toward focuser like and quotative like. *Journal of Sociolinguistics*, 4(1), 60-80.
- D'Arcy, A. (2005). Like: Syntax and development. Unpublished PhD dissertation, University of Toronto.

- Degand, L., & Van Bergen, G. (2018). Discourse markers as turn-transition devices: Evidence from speech and instant messaging. *Discourse Processes*, 55(1), 47-71.
- Deuber, D., Leimgruber, J. R., & Sand, A. (2018). Singaporean internet chat compared to informal spoken language: Linguistic variation and indexicality in a language contact situation. *Journal of Pidgin and Creole Languages*, 33(1), 48-91.
- 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.
- Drager, K. K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, 39(4), 694-707.
- Eckert, P. (2014). Language and gender in adolescence. In M. McElhinny (Ed.), *The Handbook of Language, Gender, and Sexuality* (pp. 529-545). Blackwell.
- Erman, B. (1987). Pragmatic expressions in English: A study of you know, you see and I mean in face-to-face conversation. *Stockholm Studies in English* 69. Stockholm: Almqvist & Wiksell.
- Erman, B. (2001). Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics*, 33(9), 1337-1359.
- 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.
- Fraser, B. (1987). Pragmatic formatives. In *The Pragmatic Perspective* (pp. 179). John Benjamins.
- Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica*, 38(1-4), 19-33.
- Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, 14(3), 383-398.
- Fraser, B. (1996). Pragmatic markers. *Pragmatics*, 6, 167-190.
- Fraser, B. (1999). What are discourse markers?. *Journal of Pragmatics*, 31(7), 931-952.
- Fung, L., & Carter, R. (2007). Discourse Markers and Spoken English: Native and Learner Use in Pedagogic Settings. *Applied Linguistics*, 28(3), 410-439.
- George, L. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Hansen, M. B. M. (1997). Alors and donc in spoken French: A reanalysis. *Journal of Pragmatics*, 28(2), 153-187.
- Heine, B., & Kaltenböck, G. (2021). From clause to discourse marker: On the development of comment clauses. *Language Sciences*, 87, 101400.

- Hung, T. T. (2012). Hong Kong English. In E. L. Low & A. Hashim (Eds.), *English in Southeast Asia: Features, Policy and Language in Use*. John Benjamins Publishing.
- Jucker, A. H., & Smith, S. W. (1998). And people just you know like ‘wow’: Discourse markers as negotiating strategies. In A. H. Jucker & Y. Ziv (Eds.), *Pragmatics and Beyond New Series* (pp. 171-202). Amsterdam & New York: John Benjamins.
- Jucker, A. H., & Ziv, Y. (1998). Discourse markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Pragmatics and Beyond New Series* (pp. 1-12). Amsterdam & New York: John Benjamins.
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism. In R. Quirk & H. G. Widdowson (Eds.), *World Englishes: Critical Concepts in Linguistics* (pp. 241-270).
- Kaur, J. (2011). Raising explicitness through self-repair in English as a lingua franca. *Journal of Pragmatics*, 43(11), 2704-2715.
- Kyrtziz, A., & Ervin-Tripp, S. (1999). The development of discourse markers in peer interaction. *Journal of Pragmatics*, 31(10), 1321-1338.
- Lakoff, R. (1973). *Language and Woman’s Place*. *Language in Society*, 2(1), 45-79.
- Leuckert, S., & Rüdiger, S. (2021). Discourse markers and world Englishes. *World Englishes*, 40(4), 482-487.
- Levey, S. (2006). The sociolinguistic distribution of discourse marker like in preadolescent speech.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Malicsi, J. (2010). Philippine English: A case of language drift. *立命館言語文化研究*, 22(1), 29-58.
- Maschler, Y., & Schiffrin, D. (2015). Discourse markers: Language, meaning, and context. In *The Handbook of Discourse Analysis* (pp. 189-221).
- Meehan, T. (1991). It’s like, ‘what’s happening in the evolution of like?’. *Kansas Working Papers in Linguistics*, 16, 37-51.
- Miller, J. (2009). Like and other discourse markers. In *Comparative Studies in Australian and New Zealand English* (pp. 317-337).
- Miller, J., & Weinert, R. (1995). The function of LIKE in dialogue. *Journal of Pragmatics*, 23(4), 365-393.
- Mufwene, S. S. (2005). Language evolution: The population genetics way. In G. Hauska (Ed.), *Gene, Sprachen, und ihre Evolution* (pp. 30–52). Universitätsverlag Regensburg.
- Mukherjee, J. (2007). Steady states in the evolution of New Englishes: Present-day Indian English as an equilibrium. *Journal of English Linguistics*, 35(2), 157-187.

- Müller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamins.
- Östman, J. O. (1995). Pragmatic particles twenty years after. In B. Wårvik, S.-K. Tanskanen, & R. Hiltunen (Eds.), *Organization in Discourse: Proceedings from the Turku Conference* (pp. 95-108). Anglicana Turkuensia; Vol. 14, University of Turku.
- Palacio, M. A., & Gustilo, L. (2016). A pragmatic analysis of discourse particles in Filipino computer-mediated communication. *Online Submission*, 16(3), 1-19.
- Parviainen, H. (2012). Focus particles in Indian English and other varieties. *World Englishes*, 31(2), 226-247.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14(3), 367-381.
- Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics*, 29(6), 1139-1172.
- Risselada, R., & Spooren, W. (1998). Introduction: Discourse markers and coherence relations. *Journal of Pragmatics*, 30(2), 131-133.
- Romaine, S., & Lange, D. (1991). The use of like as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech*, 66(3), 227-279.
- Rüdiger, S. (2021). Like in Korean English speech. *World Englishes*, 40(4), 548-561.
- Sand, A. (2013). Singapore weblogs: Between speech and writing. *Malay*, 91, 7-9.
- 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 *Twelfth Annual Symposium About Language and Society-Austin* (Vol. 48, pp. 177-186).
- Schleef, E., & Turton, D. (2018). Sociophonetic variation of like in British dialects: Effects of function, context and predictability. *English Language & Linguistics*, 22(1), 35-75.
- Schneider, E. W. (2003). The dynamics of New Englishes: From identity construction to dialect birth. *Language*, 79(2), 233-281.
- Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.
- Schourup, L. C. (1983). *Common Discourse Particles in English Conversation*. Routledge.
- Schourup, L. (1999). Discourse markers. *Lingua*, 107(3-4), 227-265.

- Schweinberger, M. (2014). *The Discourse Marker LIKE: A Corpus-Based Analysis of Selected Varieties of English* (Doctoral dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky).
- Schweinberger, M. (2015). A comparative study of the pragmatic marker like in Irish English and in south-eastern varieties of British English. In *Pragmatic Markers in Irish English* (pp. 114-134).
- Siegel, M. E. (2002). Like: The discourse particle and semantics. *Journal of Semantics*, 19(1), 35-71.
- Siemund, P., Maier, G., & Schweinberger, M. (2009). Towards a more fine-grained analysis of the areal distributions of non-standard features of English. In *Language Contacts Meet English Dialects* (pp. 19-46).
- Sinclair, J., & Coulthard, M. (2013). Towards an analysis of discourse. In *Advances in Spoken Discourse Analysis* (pp. 1-34). Routledge.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition* (Vol. 142). Cambridge, MA: Harvard University Press.
- Stenström, A.-B. (1994). *An Introduction to Spoken Interaction*. London: Longman.
- Tagliamonte, S. (2005). So who? Like how? Just what?: Discourse markers in the conversations of young Canadians. *Journal of Pragmatics*, 37(11), 1896-1915.
- Takahara, P. O. (1998). Pragmatic functions of the English discourse marker anyway and its corresponding contrastive Japanese discourse markers. In Andreas H. Jucker & Yael Ziv (eds.), *Pragmatics and beyond new series*, Amsterdam & New York: John Benjamins, 327
- Tannen, D. (Ed.). (1993). *Gender and Conversational Interaction*. Oxford University Press.
- Traugott, E. C. (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In *Perspectives on Historical Linguistics* (pp. 245-271). Amsterdam: Benjamins.
- Traugott, E. C. (1995). The role of the development of discourse markers in a theory of grammaticalization. In *ICHL XII* (pp. 123). Manchester.
- Trillo, J. R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34(6), 769-784.
- Underhill, R. (1988). Like is, like, focus. *American Speech*, 63(3), 234-246.
- Valentine, T. M. (1991). Getting the message across: Discourse markers in Indian English. *World Englishes*, 10(3), 325-334.
- Walrod, M. (2006). The marker is the message: The influence of discourse markers and particles on textual meaning. *Philippine Journal of Linguistics*, 37(2), 100-119.

- 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.
- Zufferey, S., & Popescu-Belis, A. (2004). Towards automatic identification of discourse markers in dialogs: The case of like. In *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue at HLT-NAACL 2004* (pp. 63-71).
- Zwicky, A. M. (1985). Clitics and particles. *Language*, 283-305.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté de philosophie, arts et lettres

Place Blaise Pascal, 1 bte L3.03.11, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/fial