



Cross-Varietal Study of Politeness Markers in British, Malaysian and Pakistani English: A Corpus Based Study

Saira Qasim^a, Behzad Anwar^b

- a. Assistant Lecturer, University of Gujrat
- b. Associate Professor, University of Gujrat

ARTICLE INFO

Received:

May 02, 2026

Revision Received:

May 16, 2026

Accepted:

June 01, 2026

Available Online:

June 20, 2026

Keywords:

politeness variation;
tentativeness; deference;
solidarity; Britain; Pakistan;
Malaysia; corpus
pragmatics

ABSTRACT

This study inside the framework of variational pragmatics, examines a cross-cultural variation in the use of politeness formulae among British(GB), Pakistani(PK), and Malaysian(MY) English varieties. Using the data from the GloWbE corpus, this study investigates three politeness types (tentativeness, deference, and solidarity) and three levels of formality i.e. formal, neutral, informal. From GB, PK, and MY subcorpora, statistical results showed significant differences between all the varieties. Pakistani English, illustrating its hierarchical and religious cultural backgrounds, is associated with the greatest degrees of deference politeness and formal language; British English aligned with more indirect modes of communication employs tentativeness politeness; while Malaysian English emphasizing a focus on community and rapport has comparatively higher levels of solidarity markers. These findings foreground how politeness behavior in World Englishes is shaped through sociocultural variables such as power distance, urbanization, and religio-cultural values.

1. Introduction

In pragmatics, politeness has long been a key feature of inquiry for indicating the structures which help language to encode social relationships, power hierarchies, and cultural values. A subfield *variational pragmatics* investigates variations in language use across geographic and social boundaries (Schneider & Barron, 2008). The comparative corpus-based studies also indicate that various varieties of English implement unique politeness strategies. This study finds itself in this practice, by investigating three new contexts (Britain Pakistan and Malaysia). Pakistan and Malaysia share histories as British colonies, but with very different cultural attitudes and religious beliefs; English is an official or institutional language in both the countries.

1.1 Politeness, Culture, and English Varieties

Emphasizing social distance and the mitigation of enforcement, British culture is often represented as defined by indirectness and negative politeness strategies (Blum-Kulka, 1987; Ogiermann, 2009), whereas, Pakistani and Malaysian societies contrasting to it, are shaped by socialistic and high-power-distance orientations, in which linguistic respect markers are closely tied to religion and hierarchy (Kachru, 2005; Rahman, 2020). In both regions, English coexists with rich indigenous politeness systems Urdu and Malay respectively that effect pragmatic transfer. Therefore, expressions such as *thank you*, *please*, or *sir*

¹ Correspondence to: Assistant Lecturer, University of Gujrat

E-mail address: Saira.qasim@uog.edu.pk (S.Qasim).



may have culturally deep meanings of respect, while endearments like *bro* or *dear* often work as solidarity markers showing communal identity (Hashim & Low, 2018).

1.2 From Politeness Theory to Corpus Pragmatics

Culpeper and Gillings (2018) by following Brown and Levinson's (1987) differentiation between positive and negative politeness, reframed the taxonomy into three theoretically retrievable types, i.e., tentativeness, deference, and solidarity. Non-imposition strategies such as hedging and modal mitigation, e.g., *could you*, *would you mind*, are related to tentativeness whereas hierarchy and gratitude, such as *sir*, *thank you*, *sorry*, are established by deference. Likewise, intimacy and in-group rapport like *hi*, *mate*, *cheers* are shown through solidarity. The quantitative examination of these formulae across large datasets, is made possible through corpus-based approaches, allowing pragmatic variation to be traced through frequency and distributional patterns (Jucker, 2020 & Stefanowitsch, 2020).

1.3 The Rationale for a Britain-Pakistan-Malaysia Comparison

Theoretically and sociolinguistically, this triangular comparison is important where the source variety with well-established norms of indirect politeness is visualized by while a South-Asian postcolonial variety is exemplified by Pakistan, where English indexes status, education, and religious authority, whereas a multilingual, and a multiethnic context is depicted by Malaysia where English is used to mediate intercultural harmony. This study, by sorting out how these societies place the same set of formulaic expressions describes how cultural schemata individualism versus collectivism, secularism versus religiosity, and egalitarianism versus hierarchy shape linguistic politeness.

Furthermore, empirical research on English use in Pakistan and Malaysia strengthens this belief that culturally defined politeness patterns appear according to the accepted contexts. For example, apology responses of Pakistani English speakers are strongly shaped by sociocultural and socio-religious hierarchies; speakers used courteous strategies such as acceptance or acknowledgement toward interlocutors of higher or equal status but tended toward rejection strategies with lower-status interlocutors (Khan & Jabeen, 2018). The present study by merging *corpus pragmatics* and *variational pragmatics* contributes to understanding how pragmatic norms develop in World Englishes. Revealing how postcolonial and socio-religious contexts reshape inherited English politeness conventions, it caters quantitative evidence for showing that politeness behaviors are culturally mediated. Moreover, how globalization and online interaction have hybridized politeness strategies is also highlighted by focusing on digital discourse (web texts in GloWbE).

1.4 Research Questions

This study aims to determine empirical differences in the frequency and formality of politeness formulae across the GB, PK, and MY components of the *GloWbE* corpus. The leading questions are:

- i. What are the differences in British, Pakistani, and Malaysian English, in the frequency of tentativeness, deference, and solidarity politeness?
- ii. How do formulaic expressions related to these politeness types differ in usage frequency among the three varieties?
- iii. What variations emerge in the levels of formality, formal, neutral, and informal in politeness formulae across the three varieties?

2. Literature Review

2.1 Variational Pragmatics and Corpus-Based Politeness

Variational pragmatics is used to explore how pragmatic phenomena diverge across regional and social varieties of the same language (Schneider & Barron, 2008). Corpus linguistics has become a significant methodological resource in this field, offering large-scale quantitative investigations of speech acts, formulaic sequences, and politeness strategies in naturally occurring data (Jucker, 2013; Stefanowitsch, 2020). Latest studies suggest that politeness is best approached through diagnostic forms i.e., recurrent expressions such as *please*, *thank you*, or *hi* whose distribution across corpora offers a verifiable perceptivity into regional pragmatic norms (Jucker, 2020; Terkourafi, 2015).

Culpeper and Gillings' (2018) study of politeness variation in England using the Spoken BNC2014 is a major advance in corpus-based politeness research. They showed systematic north-south differences in the use of conventionalized politeness formulae (e.g. *please, thank you, hi, bye*) by offering a triangular typology (tentativeness, deference, and solidarity) politeness. Drawing on this, a socio-pragmatic variation in Britain across eight social variables (gender, age, region, population density, social class, education, setting, and medium) was analyzed by van Dorst, Gillings, and Culpeper (2024), confirming that social and spatial factors significantly shape the distribution of polite formulae in British English.

Most recently, comparing Britain, Australia, and Hong Kong, Xie (2024) applied this triadic framework to web-based data from the GloWbE corpus. Xie by using twenty-one formulaic expressions, highlighted that Australian English displays the highest frequencies of deference and solidarity politeness, while Hong Kong English shows more tentativeness and neutral-formality formulae, with these patterns interpreted in relation to urbanization, population density, and pragmatic transfer from Cantonese. For present research, the study by Xie is direct methodological model.

2.2 Diachronic and Cross-Regional Perspectives on (Im)Politeness

Diachronic work beyond the synchronic regional comparisons, has also illustrated how politeness practices shift over time. Movie Corpus was used by Jucker (2023) for analyzing stereotypes about declining politeness in British and American English, showing that evaluations of "politeness loss" are historically situated rather than purely linguistic facts. This work along with van Dorst et al. (2024), promotes the idea that politeness is dynamic, shaped by larger socio-historical processes and mediated by media genres.

Regionally, the flourishing work on World Englishes shows that politeness patterns in postcolonial varieties frequently diverge from those of the metropolitan norm. Studies of Asian Englishes record that a pragmatic transfer from indigenous norms (e.g., Malay, Cantonese, Urdu) yielding distinctive usage of address forms, honorifics, and mitigators, hybrid politeness systems are emerging from the contact between English and local languages (Hashim, 2020; Wong, 2017). However, many of these studies focus on qualitative speech-act realizations (e.g., requests, apologies) rather than on the systematic distribution of specific formulaic expressions in large web corpora.

2.3 Politeness in Pakistani English

Politeness in Pakistani context has been investigated typically within Brown and Levinson's (1987) framework, in relation to both English and local languages. Early work on linguistic politeness among English-using Pakistanis was analyzed by Rahman (1998), noting that forms of address and indirectness are strongly influenced by social hierarchy, religion, and notions of *izzat* (honour). Specific genres are also explored in recent works, such as business correspondence and broadcast discourse. For example, exploring the use of English in Pakistani business letters and motivational speeches has demonstrated a strong preference for negative-politeness strategies (deference, honorifics, apologies) and formulaic respect markers such as *sir* and *dear* (Saleem, 2020; Zafar, 2023).

Studies are also available on televised discourse. The research on politeness strategies in Pakistani morning shows records a frequent use of positive politeness (compliments, inclusive pronouns) alongside religious invocations and deferential address terms, reflecting an interplay of entertainment, morality, and hierarchy in Pakistani media (Ahmed, 2024; Saleem, 2020). Other work has examined Pakistani learners' pragmatic competence in acts such as disagreement and apologizing, highlighting gaps between grammatical knowledge and sociopragmatic suitability (Ahmed, 2024).

Though there is an evolving literature on linguistic politeness in Pakistani context, yet there are some gaps. First, the focus of most Pakistani studies remains on speech acts (requests, apologies, disagreement) rather than on frequency patterns of conventionalized expressions (e.g., *please, thank you, hi*). Second, these studies use relatively small spoken or written datasets (classroom talk, TV shows, letters) rather than large web-based corpora. Third, usually British or American English is taken as comparative variety to contrast with Pakistani English rather than positioning it alongside multiple Asian or postcolonial varieties in a unified quantitative design (Saleem, 2021).

2.4 Politeness in Malaysian English and Malay-English Contact

Historical emergence, structural features, and functional range of Malaysian English (MalE) in a multilingual, multicultural setting has been observed (Hashim, 2020; Hashim, Tan, & Low, 2012). Studies from a pragmatic perspective, have highlighted how Malay politeness norms such as careful honorific systems and strong orientation to harmony and face-saving shift into English usage. For instance, how Malay forms of address and deference shape English greetings, invitations, and requests, often

producing hybrid politeness strategies in Malaysian English have been explored by Lim (2012) and later work on Malay-English politeness.

Malaysian ESL learners are now compared with other groups. A study by Alakrash and Bustan (2020) was conducted on Arab EFL and Malaysian ESL students' request strategies and it was found that Malaysian students tend to prefer more indirect, mitigated requests (query preparatory, hedged performatives), consistent with collectivist and high-context communication norms. Another paper written by Harun and Abdul Rahman (2024) on politeness strategies in Malay and English greeting and invitation expressions shows that though English and Malay differ structurally yet underlying socio-cultural rules of deference and warmth are remarkably similar, with Malaysian speakers using politeness to negotiate both respect and solidarity across languages.

Yet, like Pakistani English, most studies on Malaysian politeness are not corpus-based at scale and seldom focus on specific conventionalized English formulae in web discourse. Rather, they use elicited data, classroom interactions, or small conversational corpora. There is still a limited learning of how frequently expressions such as *please*, *thank you*, *bro*, or *hi* occur in Malaysian web texts, and how these frequencies compare quantitatively with other varieties.

2.5 Identified Gap and Contribution of the Present Study

Taken together, it is demonstrated through recent work in variational pragmatics that (a) formulaic politeness expressions can be consistently categorized into tentativeness, deference, and solidarity and used to observe regional and social variation (Culpeper & Gillings, 2018; van Dorst et al., 2024; Xie, 2024); and (b) Pakistani and Malaysian English show recognizable politeness norms shaped by religion, hierarchy, and collectivism (Rahman, 1998; Saleem, 2021; Hashim, 2020), but these have seldom been examined through large web-based corpora or in direct comparison with British English.

The present study using the same set of 21 politeness formulae but in a new tri-varietal design: Britain, Pakistan, and Malaysia, addresses this gap by following Xie's (2024) corpus-based diagnostic approach and applying it to the GB, PK, and MY components of GloWbE. This suggests not only a systematic, frequency-based comparison of tentativeness, deference, and solidarity politeness, but also the levels of formality, across one inner-circle and two postcolonial Asian English varieties, by extending both variational pragmatics and corpus pragmatics to under-researched contexts.

3. Theoretical Framework

3.1 From Face and Politeness Strategies to Politeness Types

According to Brown and Levinson's (1987) framework politeness is conceptualized as strategic redress of face-threatening acts through negative politeness (non-imposition, deference) and positive politeness (solidarity, approval). Though this difference remains effective, yet it has been criticized for both its universalist assumptions and its foggy evaluative labels "positive" and "negative" (Eelen, 2001; Jucker, 2020).

Culpeper and Gillings (2018) structuring on these critiques, declare three more transparent, empirically traceable politeness types: tentativeness, deference, and solidarity. In tentativeness politeness mitigation and non-imposition (e.g., *could you*, *would you mind*) is included, while in deference politeness respect and hierarchy (*sir*, *madam*, *thank you*) is foregrounded, and in solidarity politeness in-group bonding and friendliness (*hi*, *mate*, *cheers*) is accented. Using the Spoken BNC2014, their corpus-based study of politeness variation in England, presents that these three types can be systematically operationalized through diagnostic expressions and display clear regional patterns (e.g., a north-south divide in the use of informal solidarity markers).

This triadic typology was further extended by van Dorst et al. (2024) with a broader sociopragmatic design, showing that politeness formulae vary significantly across gender, age, social class, urbanization, and setting. It was confirmed by their large-scale analysis that the distribution of tentativeness, deference, and solidarity formulae is sensitive to both social and spatial variables in Britain.

This framework is adapted by Xie (2024) for web-based English varieties, comparing Britain, Australia, and Hong Kong using GloWbE data and a set of twenty-one conventionalized politeness expressions. Through Xie's findings it has been demonstrated that Australian English show higher frequencies of deference and solidarity politeness, whereas Hong Kong English displays more tentativeness and neutral-form expressions i.e., patterns interpreted in relation to urbanization and pragmatic transfer from Cantonese. Xie's methodological choices are directly followed in present study but are re-contextualized for Britain, Pakistan, and Malaysia, allowing a focus on two under-researched Asian World Englishes.

3.2 Relational Work theory and the Contextual Nature of (Im)Politeness

The relational work approach is another pillar of this framework (Locher & Watts, 2008), according to which politeness is not a fixed property of linguistic forms but an emergent from ongoing negotiation of interpersonal relationships. So, any utterance from this perspective may be evaluated as polite, impolite, appropriate, or over-polite depending on norms that are locally co-constructed in interaction. For instance, in multilingual, postcolonial contexts such as Pakistan and Malaysia, relational work is further shaped by the coexistence of several politeness systems i.e., in Pakistan Urdu, Punjabi, and other local languages and in Malaysia Malay, Chinese varieties, Tamil, and others. Thus, English politeness practices in such settings are expected to exhibit a pragmatic hybridity, with English formulae taking on functions influenced by local norms of respect, honour, and harmony.

At the same time, it is acknowledged by relational-work theorists that some expressions become conventionalized as indexically polite or informal such as recurrent items like *please*, *thank you*, *sir*, or *hi* are identifiable politeness resources, just because they are repeatedly used to manage face and relationships (Locher & Watts, 2008; Terkourafi, 2015). Through this perceptivity, using formulaic expressions as corpus-diagnostic tools seems to be legitimized, but only if they are interpreted against their socio-cultural backdrop rather than treated as inherently polite in all contexts.

3.3 Formulaic Expressions, (In)Formality, and Corpus Pragmatics

Further, operationalization is the third component of the framework that concerns how to turn an abstract notion like politeness into measurable corpus variables. Using large, computer-readable language collections to investigate pragmatic phenomena are favoured by corpus pragmatics (Jucker, 2013; Landert et al., 2023). Conventionalized formulaic expressions in this paradigm, work as “diagnostics” meaning that their frequency and distribution can show patterns of politeness behaviour across regions and social groups (Culpeper & Gillings, 2018; Xie, 2024).

This study following Xie (2024) also adopts a set of twenty-one expressions, each assigned to one of three politeness types i.e., tentativeness, deference, solidarity and one of three formality levels formal, neutral, and informal. Taking earlier work as a guide, expressions such as *sir*, *madam*, *would you mind*, and *thank you* are addressed as more formal diagnostics, whereas forms such as *hi*, *bye*, *thanks*, and *bro* indicate informality and familiarity. Therefore, this two-fold categorization offers for analyzing not only what type of politeness is favored in each variety but also whether these favorite options are realized through more formal or informal signals.

Previous studies in British English, show strong criterion of indirectness and non-imposition, with tentativeness strategies playing a prominent role (Blum-Kulka, 1987; Culpeper & Demmen, 2011; Culpeper & Gillings, 2018). While, as far as Pakistani English is concerned here politeness in institutional and semi-formal online discourse is closely intertwined with hierarchical respect, religion, and the culturally prominent concept of honour, advising a possible emphasis on deference markers such as *sir*, *madam*, and *thank you* (Rahman, 1998; Saleem, 2020). Whereas, in context of Malaysian English, socialistic orientations and Malay politeness measures encourage harmony, warmth, and community, which may be manifested in relatively high use of solidarity markers such as *hi* and *bro* and a combination of respect and friendliness (Hashim, 2020; Lim, 2012).

Hence, the study by considering these formulae as probabilistic indicators rather than absolute labels, fits with relational-work theory while still utilizing the strengths of large-scale corpus evidence. A suitable empirical basis is provided by the GB, PK, and MY components of GloWbE for implementing this framework in web-based discourse, sanctioning a systematic comparison of how tentativeness, deference, solidarity, and (in)formality are encoded in British, Pakistani, and Malaysian English.

4. Methodology

4.1 Research Design

This study by merging quantitative frequency analysis with qualitative interpretation employs a corpus-based comparative design within the framework of variational pragmatics. The study intends to investigate three politeness types: *tentativeness*, *deference*, and *solidarity* along with three levels of *formality*: formal, neutral, and informal using twenty-one conventionalized formulaic expressions. According to the approach politeness is operationalized as a function of recurrent linguistic forms that indicate social relationships in context (Culpeper & Gillings, 2018; van Dorst et al., 2024).

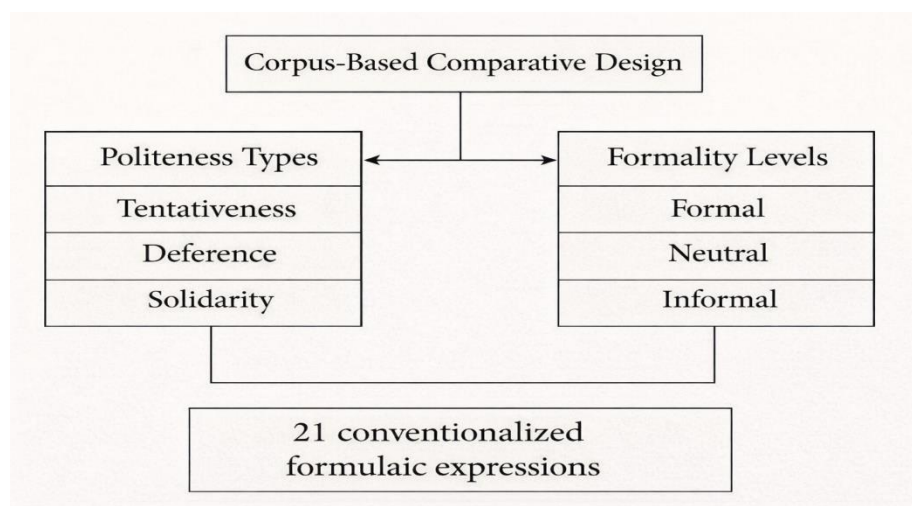


Fig.1: Corpus-based comparative design linking politeness types and formality levels through twenty-one formulaic expressions

4.2 Corpus Data

The Global Web-Based English (GloWbE) corpus, a 1.9-billion-word collection of web texts from twenty English varieties worldwide, was exclusively considered to extract the data (Davies, 2015) Only the following three national components are used in the present study:

Variety	Sub-corpus Code	Tokens	Contextual Description
British English	GB	387,615,074	Inner-circle, standard reference variety
Pakistani English	PK	51,367,152	Outer-circle, South-Asian, Islamicate society
Malaysian English	MY	41,643,730	Outer-circle, multilingual, collectivist context

Table.1: Total number of tokens in all the three varieties of English along with their contextual description

Using GloWbE's online interface, all data were accessed (updated 2024) and cleaned for duplicates or non-English text. To ensure consistency among corpora, only tokens from native or near-native English sources (excluding machine-translated material) were retained.

4.3 Diagnostic Expressions

Culpeper and Gillings (2018) and Xie's (2024) diagnostic criteria were followed and twenty-one formulaic expressions for retrieval were selected. These items occur frequently, are easily retrievable, and represent the three politeness types and formality levels.

Politeness Type	Expressions	Formality Level
Tentativeness	<i>could you, would you mind, please, perhaps, is it possible, can you, do you think</i>	Mostly neutral/formal
Deference	<i>sir, madam, thank you, thanks, sorry, excuse me, cheers</i>	Mixed (formal-neutral)
Solidarity	<i>hi, hello, bye, goodbye, love, mate, bro</i>	Mostly informal

Table.2: Twenty-one politeness expressions according to their types and formality level

Each retrieved token was analyzed in its original textual context to verify pragmatic aim. For instance, *cheers* working as “thanks” was counted as deference but *cheers* meaning “toast” was eliminated.

w_gb_g.txt	that he and Harry are linked and that Harry is	perhaps	the biggest threat to him ? <p> Jareth scowled , ? and
w_gb_g.txt	excuse in cases of VAW . In these non-theocratic cases ,	perhaps	the GBV framework is a useful one . However , the
w_gb_g.txt	specific requirements for looping violins than I ever thought possible .	Perhaps	the hardest part was simply explaining precisely what the
w_gb_g.txt	also passed to him on Count Geoffrey 's death in 1151 .	Perhaps	the most audacious acquisition was the strategically important duchy
w_gb_g.txt	n't agree with ' ethically , politically or even stylistically ' . So	perhaps	the obvious question is , what attracted her to the
w_pk_g.txt	breach of trust . They will slip from the bridge of '	Sir	t ' and drop into Hell-fire . <p> (al-K ?
w_pk_g.txt	and maintains trusts (Amanat) shall be able to cross the '	Sir	t ' easily and will reach Paradise . <p> None of
w_pk_g.txt	for diplomacy with the Imam , I said to him : " O	Sir ,	eminent teacher , what is the sin of the Muslim ,
w_pk_g.txt	p> In the Qur'anic terminology when the word Kh ?	sir (	loser) is used it denotes the one who is
w_pk_g.txt	the Relatives Will be Able to Cross the Bridge of	Sir ?	t Smoothly <p> ' Sir ? t (the bridge passing over
w_my_g.txt	I 'm lucky enough to get to do something I	love	for a living , it also reminds me to count
w_my_b.txt	film required coordination from all aspects . He looked back at	LOVE	FOR LIFE (JUI OI) that he spent all his
w_my_g.txt	call our own . <p> But , oh it was glorious . I	love	the beach . It 's my zen and my dream ,
w_my_b.txt	going to fire me ! <p> Mmmm ... these look delicious . I	love	the look of that crisp outside and almost choux-

Fig.2: Formality level of politeness expressions across three English varieties

4.4 Data Extraction and Coding Procedure

The process of retrieval was performed by using the search interface of GloWbE with regular-expression queries and to allow cross-corpus comparison, results were normalized to frequency per million words (pmw). For coding automatic retrieval and manual screening was done. for an automatic retrieval, GB, PK, and MY varieties of English were explored for each expression and manual screening was done by going through 1,000 random concordance lines for each variety were independently verified to confirm politeness usefulness. Each verified token was tagged for (a) politeness type, (b) formality level, and (c) discourse function e.g., request, greeting, and gratitude.

4.5 Software and Statistical Tool

The Global Web-Based English (GloWbE) corpus (Davies, 2013; updated 2024 interface) was selected to retrieve all the data and organized in Microsoft Excel 2021 for normalization and preliminary tabulation. For manually inspecting concordance lines and verifying the contextual validity of each politeness expression the detailed file view was thoroughly explored. Using Pearson's chi-square (χ^2) test, statistical testing was conducted in IBM SPSS Statistics Version 27 to identify significant cross-varietal differences in the frequency of politeness formulae. While ggplot2 was used to generate visualizations. Xie's (2024) corpus-based procedure has been reflected through this analytical setup, which relies on normalized frequency and chi-square significance testing to establish a cross-varietal pragmatic variation, without calculating effect-size.

4.6 Statistical Analysis

The following quantitative procedures were used to assess variation across varieties:

Normalization: To account for corpus size differences frequencies were converted to pmw.

$$\text{Normalized frequency} = \frac{\text{Raw hits}}{\text{Total words}} \times 1,000,000$$

Descriptive Statistics: For each politeness type and formality level, frequencies were tabulated and compared across GB, PK, and MY, mean and standard deviation were also calculated.

Chi-Square Tests: It was determined through Pearson's χ^2 tests whether frequency distributions of each politeness type differed significantly across the three varieties. Significance was established at $p < .05$.

4.7 Validity and Reliability

Methodological validity was guaranteed first through replication of the established diagnostic set of previous studies, second through coder triangulation, and third through quantitative normalization. Then, reliability was built through transparent operational definitions and manual verification of dubious cases.

5. Data Analysis

5.1 Overview of Analytical Approach

The retrieved data from the **GB**, **PK**, and **MY** portions of GloWbE were analyzed quantitatively to study variation in the frequency of twenty-one politeness formulae categorized under **tentativeness**, **deference**, and **solidarity** types. Using Pearson's χ^2 each occurrence was verified in context, normalized per million words "normalized frequency = (raw hits / total words) $\times$ 1,000,000", and tested statistically for cross-varietal significance

5.2 Frequency Distribution across Politeness Types

Distinct cross-varietal profiles emerged through the normalized counts per million words (pmw). The values reveal that Malaysian English maintains the highest overall density of politeness formulae (1739.30pmw), followed by British English (1682.30 pmw) and Pakistani English was at the lowest (1319.17 pmw). A statistically significant difference, χ^2 (40, N = 792 276) = 14456.59, $p < .001$, was pointed through a chi-square independence test using the full corpus counts (N = 792 276) showing that the three varieties differ markedly in their deployment of politeness formulae. The Likelihood Ratio test provides support to the Pearsons Chi-Square result, thus reinforcing the statistical significance of the varietal differences in distribution. The Effect sizes (Cramer's V = .096), indicates that difference is statistically significant but with small effect sizes.

Summary of Politeness Type Density (pmw)

Type	GB	PK	MY	Dominant Type
Tentativeness	624.02	515.35	601.17	GB highest; PK close to solidarity
Deference	474.97	277.71	445.33	GB highest; MY second
Solidarity	583.30	526.11	692.81	MY highest
Total	1,682.30	1,319.17	1,739.30	MY overall highest

Table.3: Normalized per million words, based on corpus sizes: GB = 387 615 074; PK = 51 367 152; MY = 41 643 730

5.3 Formality-Level Patterns

After calculation, formal politeness (e.g., *could you, would you mind, sir, madam*) was found least used in all three varieties: GB = 93.13pmw, PK = 69.38 pmw and MY = 44.28 pmw. Frequent use of informal and moderate expressions was also observed. The highest informal frequency was noted in MY(890.53 pmw) followed by GB(773.16pmw) and PK(629.66 pmw).

Distribution of Politeness Expressions by Formality Level Across Three Varieties (pmw)

	Representative Expressions	GB (pmw)	PK (pmw)	MY (pmw)	Dominant Variety	Pragmatic Orientation
Formality Level						
Formal	could you, would you mind, sir, madam, is it possible	93.13	69.38	44.28	GB	Institutional politeness, hierarchical deference
Moderate	please, thank you, sorry, perhaps, can you, do you think	816.00	620.12	804.49	GB (closely followed by MY)	Tentative cooperation, polite request strategies
Informal / Solidarity	thanks, love, hi, bro, mate, cheers, hello, bye, goodbye, excuse me	773.16	629.66	890.53	MY	Expressive friendliness, digital camaraderie
Total (ΣX)	—	1,682.30	1,319.17	1,739.30	MY overall	

Table.4: Per million frequencies of expressions according to their formality level and with the description of dominant varieties and pragmatic orientation

Chi-square Tests Results of Formality Levels

Chi-square tests were used to analyze whether the distribution of formulaic expressions used at three levels of formality (formal, moderate and informal) varies significantly across the three Englishes (GB, PK, MY). The results of the Chi-square test indeed indicate a very strong relationship between formality and English variety, with Pearson $\chi^2 (4, N = 792,276) = 1,607.21, p < .001$. This evidence demonstrates that formulaic expressions are systematically patterned across levels of formality and not randomly distributed by variety. While Cramer's $V = .032$ shows small association.

In more practical terms, this implies that each variety of English has specific flavor preferences for formulaic language at various points along the formality continuum. There are also particular expressions of feelings that are overrepresented in one variety (e.g., GB) when compared to the use of these same expressions in moderate or informal systems (PK and MY). The statistically significant result Validates the role of formality as an important pragmatic variable conditioning cross-varietal divergence in formulaicity.

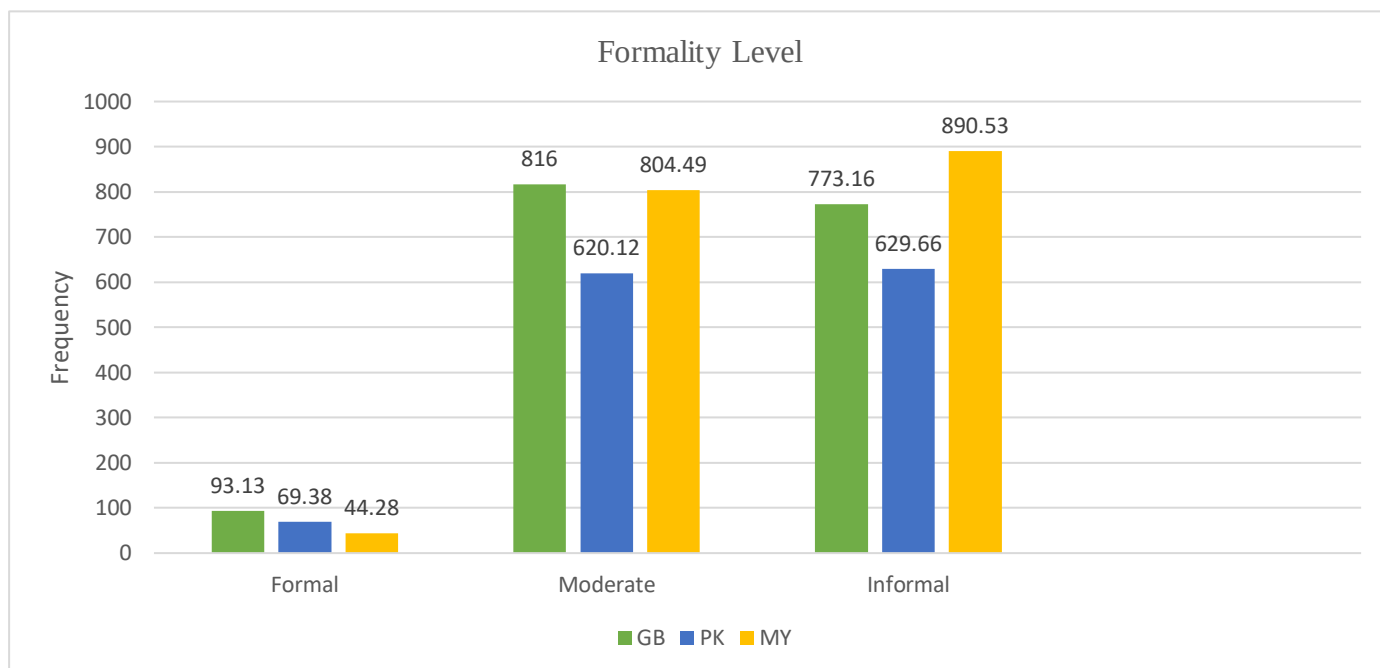


Fig. 3: Distribution of Politeness Formality Levels Across English Varieties (pmw)

It visualizes formal, moderate, and informal frequencies for British (GB), Pakistani (PK), and Malaysian (MY) English as reflected in Table 4.

Interpretation

GB shows steep increase outperforms MY and PK, with MY and PK exhibiting a gentler increase. This would appear to support the idea that moderated formal contexts promote formulaic expressions in all varieties, but more so in British English.

The biggest difference is seen on informal level, in which MY shows an increasing pattern of formulaic use and it surpasses GB and PK significantly. GB and PK also use increase over formal and medium but the effect is much less extreme. This trend suggests that formulaic use reflects family or ethnic influences on culturally embedded preferences for routinized and affiliative language even in informal conversation, as is the case within MY.

On the whole, the chart reveals a very transparent continuum of formality level on all three varieties: for all these traditions, there is an increase of formulaic expression as one moves from more to less formal context. However, the degree of enhancement is quite different and GB has the highest sensitivity towards formality. This result further supports the conclusion that formality level and English varieties are strongly connected, and that use of formulaic language is mediated by somewhat differing pragmatical norms in GB, PK, and MY.

5.4 Frequency of Key Expressions

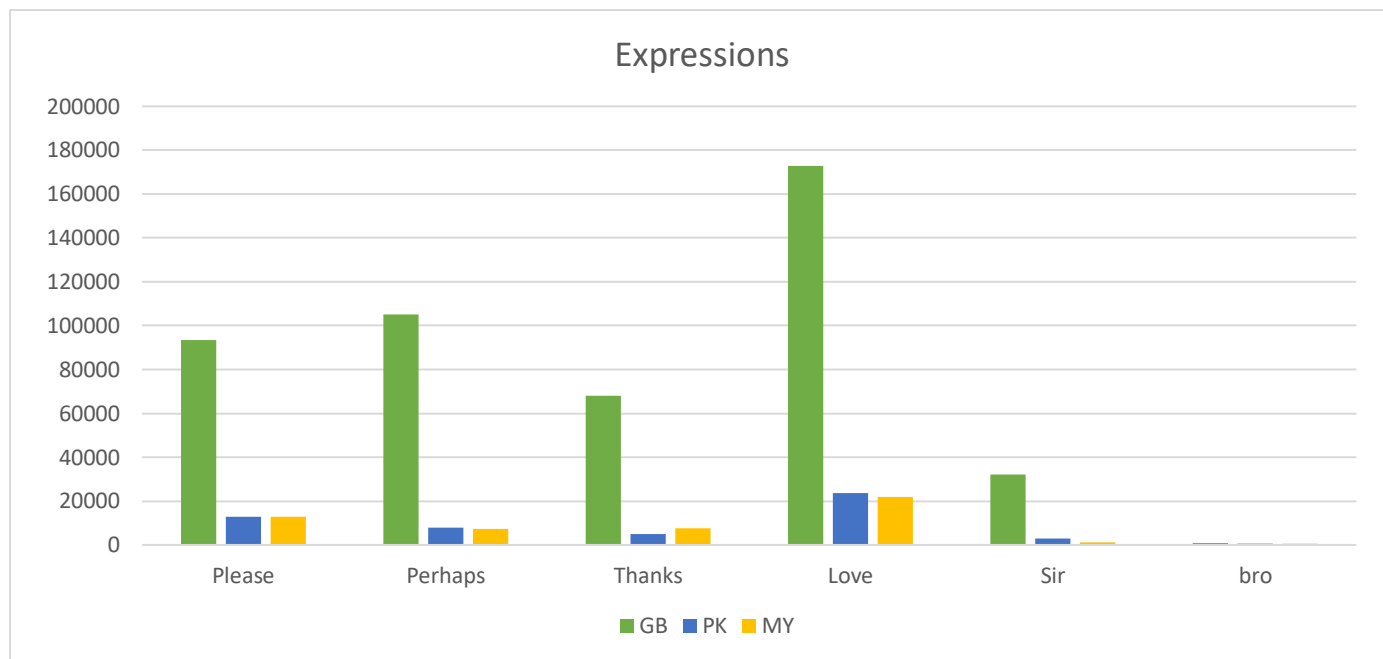


Fig. 4: Frequency of Key Politeness Expressions Across English Varieties

It compares the frequencies of *please*, *perhaps*, *thanks*, *love*, *sir*, and *bro* across British (GB), Pakistani (PK), and Malaysian (MY) English

Interpretation

British English as shown in figure 4, strongly shows an approval for *please* and *perhaps*, accenting on indirectness and mitigation strategies characteristic of the British politeness norm. Overall, most expressions are far more common in British English, especially *love*, *perhaps*, *please* and *thanks*. The very high frequency of *love* in GB points to its high degree of conventionality as a term of affiliation and interpersonality widely deployed not only in intimate contexts, but also ordinary informal interaction. Equally, the frequent use of *perhaps* and *please* indicates British English's indirectness and mitigation preference.

In contrast, what Pakistani English seems to be demonstrating is one where, both *sir* and *love* seem to coexist as high frequency markers of both hierarchical deference and interpersonal warmth, a duality that allows compatibility between levels of formal respect whilst not negating engaging with the emotions. Lower frequencies of all investigated expressions on the other hand are observed for Malaysian English suggesting a more indirect or context mediated use of politeness instead. The frequencies for most expressions are much lower in Pakistani and Malaysian English, but there is some variation.

For example, *sir* is salient in PK and MY but not GB, suggesting that the hierarchical respect is more prevalent in these two varieties with the use of honorific. It is this pattern which mirrors the larger sociocultural norms since the honorific markers are more pragmatically relevant. Words like *bro* show a very low frequency across all three varieties and are likely quite context-bound or limited to specific informal domains, or simply not as conventionalized formulaic expressions in the corpus genres covered.

Through this cross-varietal variation, it is foregrounded how cultural communicative preferences influence the selection of lexical politeness forms such as Britain tends toward tentative hedging, Pakistan balances power and solidarity, and Malaysia displays a pragmatic restraint and economy in overt politeness expression.

Normalized Frequencies (per million words, pmw)

No.	Expression	GB (raw)	PK (raw)	MY (raw)	GB (pmw)	PK (pmw)	MY (pmw)
1	could you	882	88	81	2.28	1.71	1.94
2	would you mind	226	26	21	0.58	0.506	0.504
3	Please	93,356	12,928	12,780	240.84	251.67	306.888
4	Perhaps	105,108	7,908	7,360	271.16	153.95	176.73
5	is it possible	2,060	452	294	5.31	8.79	7.06
6	can you	21,872	2,839	2,547	56.42	55.26	61.16
7	do you think	18,377	2,231	1,952	47.41	43.43	46.87
8	Sir	32,201	2,781	1,185	83.07	54.14	28.45
9	madam	731	217	263	1.88	4.22	6.31
10	thank you	37,100	2,900	5,205	95.71	56.45	124.98
11	thanks	67,927	4,919	7,540	175.24	95.76	181.05
12	sorry	40,480	3,048	3,658	104.43	59.33	87.84
13	excuse me	1,103	119	143	2.84	2.31	3.43
14	cheers	4,564	281	551	11.77	5.47	13.23
15	hi	24,660	1,216	3,212	63.61	23.67	77.13
16	hello	10,821	580	1,648	27.91	11.29	39.57
17	bye	1,785	177	310	4.60	3.44	7.44
18	goodbye	3,998	244	445	10.31	4.75	10.68
19	love	172,712	23,580	22,025	445.5	459.04	528.89
20	mate	11,314	555	486	29.18	10.80	11.67
21	bro	806	673	725	2.08	13.10	17.41
Total (Σ)		652,083	67,727	72,431.	1682.30	1,319.17	1739.30

Table. 6: Twenty-one politeness expressions, their raw frequencies extracted from GloWbE and normalized per million frequencies

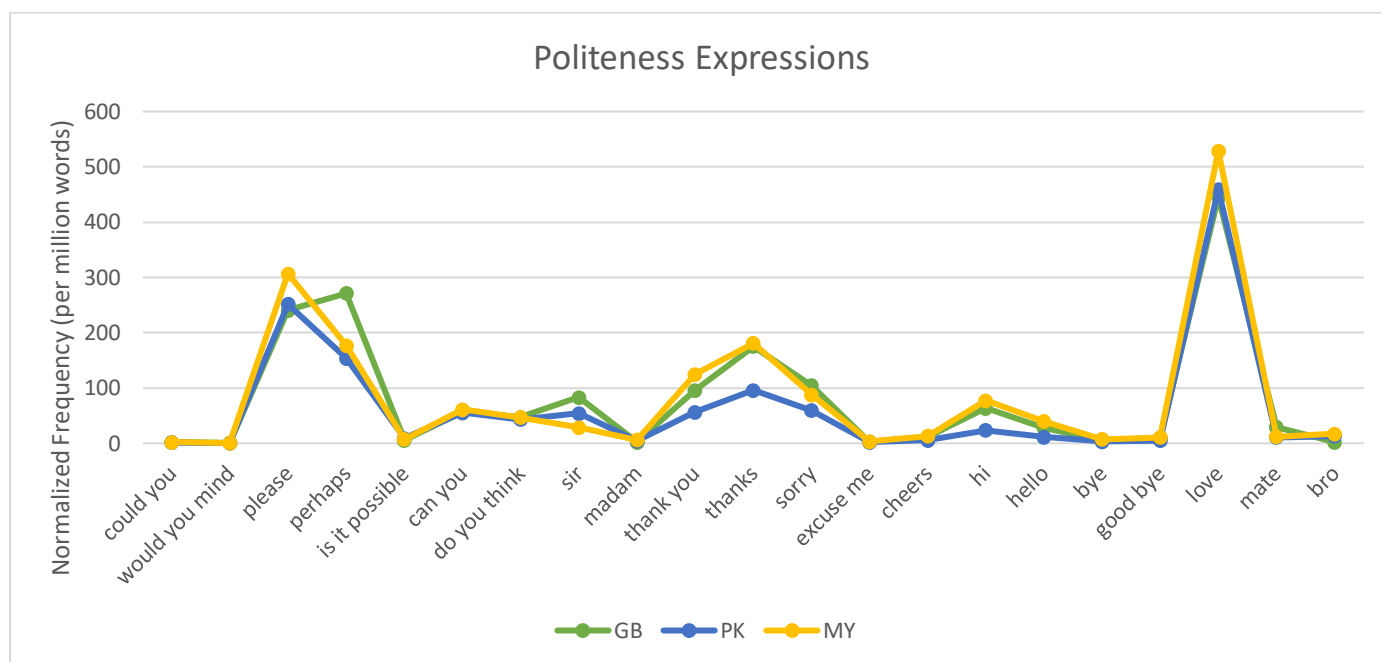


Fig 5: Comparative normalized frequencies (per million words) of 21 politeness expressions across British (GB), Pakistani (PK), and Malaysian (MY) English verified through GloWbE GB/PK/MY components (Davies, 2013; updated 2024 interface, accessed December 2025)

5.5 Interpretation of Trends

The frequencies differ clearly among the three varieties. The selected expressions scored the highest overall frequency in Malaysian English, closely followed by British English with Pakistani English showing the lowest total. British English largely favours indirectness and mitigation by expressing tentativeness via *perhaps*, *please* and *do you think*.

Pakistani English combines some hesitancy and compromise along with deferential forms like *sir*, *thank you*, *sorry*, etc. Solidarity markers for Malaysian English were more pronounced with the use of *love* and *thanks* in addition to *hi* and *bro*, as well as tentative and deferential expressions. This study indicated that British English is more towards tentativeness, Pakistani English more on warmth and respectful positioning meanwhile Malaysian English are strongly solidarity-person oriented. But words like *love*, *maybe* and *sir* as multifunctional must be read with an eye for context.

6. Results and Discussion

6.1 Britain: Tentativeness and the Legacy of Indirectness

The results indicate that British English exhibits the highest preference for tentative politeness compared to the other two varieties in this study. This makes this politeness type the most common in British data at around 624.02 per million words for tentativeness markers. Words like *perhaps*, *please*, *could you*, *can you*, *would you mind* and *do you think* are often settled to soften requests/suggestions/opinions/disagreement. These forms allow speakers to really get their point across without being too forward or blunt.

This predilection is rooted in the historical belief that British politeness has an indirect, mitigated character, one less imposing. Answer like A few hedges like *perhaps* serve to allow speakers to hedge their opinions, and clusters of modal expressions as in *could you* and *would you mind* carry out a request. Right, as well, *do you think* allows speakers to seek the listener's opinion rather than force a perspective. By the customarily frequent use of such formulas, we can conclude that speakers in British English are very often governed by a mitigation strategy that allows interlocutors to move freely and safely and without apparent showdown.

According to Locher and Watts' (2008) relational work theory, tentative expressions constitute important resources through which appropriate social relationships can be negotiated. They maintain interpersonal distance and demonstrate consideration of others and respect. This pattern is consistent with previous research linking British politeness to both restraint and indirectness. Still, words such as *perhaps* and *please* are multifunctional and not every occurrence has a politeness function necessarily. Which is why their pragmatic role needs to be verified by contextualised analysis of concordance lines.

6.2 Pakistan: Deference and Religious Respect

Pakistani English exhibits a multiple politeness profile whereby tentativeness and solidarity and deference cohere. Corrected normalized frequencies demonstrated that solidarity expressions occur at an approximate rate of 526.11 occurrences (pmw), tentativeness occurs at 515.35 pmw, and deference occurs at 277.71 pmw. These findings suggest that respect cannot solely define Pakistani English. Instead, politeness practices at its core have a mix of interpersonal warmth, mitigation and respectful positioning.

The focus on *sir* (54.14 pmw); *madam* (4.22 pmw), *thanks* (56.45 pmw) *sorry* (59.33pmw) which are cherished expressions in Pakistani English show how people remain respectable and polite while addressing literally anyone from strangers to even higher positions. Honorifics such as *sir*, and *madam* are no different; it is especially significant when the honorific utterances are within an institutional, educational, professional or hierarchy-structured interactions where literally a speaker may respect variations in status, age or deference. The local norms of *izzat*, or social honour where respect and appropriate behaviour in interactions are concerned, may also be a factor in the behavior patterns of such forms.

Meanwhile, there is extensive use of solidarity expressions in Pakistani English. The strong normalized frequency of *love* (459.04 pmw) and forms like *thanks* (95.76 pmw), *hi* (23.67 pmw), and *bro* (13.10 pmw) indicate that speakers also use formulas to construct friendliness, familiarity, and social closeness. Such harmony and consultation is emblematic of how users of Pakistani English recognize imbalances of power while also preserving friendships and facet.

Expressions of tentativeness are also noticeable in the Pakistani data. Frequent use of mitigation and non-imposition is shown through several forms, and they were *please* (251.67 pmw), *perhaps* (153.95 pmw), *can you* (55.26 pmw) and *do you think* (43.43 pmw). These phrases help speakers hedge requests, suggestions and opinions without compromising the collaborative spirit of communication. Pakistani courtesy thus comprises not only overt respect but also indirectness and relational awareness.

The occurrence of deferential forms might be understood in terms of Pakistan's hierarchical, collectivist and socio-religious communication codes. Common religious, familial, and cultural values emphasize respectful speech, humility (or high regard), gratitude, and honor for elders or authority figures. But the corpus data does not measure directly for religious belief or cultural attitudes. For this reason, the presentation of religious respect should be considered as a potential contextual factor rather than an established cause of the linguistic patterns identified above.

The Pakistani pattern illustrates speakers negotiating competing interpersonal demands in language through the lens of Locher and Watts' (2008) relational work theory. Deference to indicate status, solidarity to reduce distance, and tentativeness to lessen imposition. Overall, Pakistani English offers a more integrated and mixed politeness system whereby deference, friendliness and softening collaborate in controlling interpersonal relations.

6.3 Malaysia: Solidarity, Expressive Politeness and Collectivist Harmony

The frequent use of solidarity expressions in this study suggests that Malaysian English speakers tend to make overt linguistic choices to create warmth, familiarity and closeness. Instead of being silent or lacking explicit politeness, the results suggest an expressive and relational form as way of getting in contact. Thus, it can be suggested that greetings, informal address terms and gratitude formulae appearing very often in the transcripts might reflect Malaysian communicative environment emphasis on maintaining harmonious interpersonal relations mainly because of its multilingual and multicultural nature. Contemplative practices tend to be compatible with cultural rehearsals for social cooperative, mutual monitoring and attachment. But given that this study measures linguistic frequencies, rather than the underlying social attitudes or cultural values themselves, these cultural explanations are qualified as potential interpretations rather than causal conclusions.

This can also be seen in the high levels of tentativeness and deference exhibited by Malaysian English. As expected, as markers of tentativeness are naturally noticed, they occur at around 601.17 pmw which again means that Malaysian users depend on mitigation and non-imposition too. The presence of expressions such as *please* (306.89 pmw), *perhaps* (176.73 pmw), *can you*

(61.16 pmw) and *do you think* (46.87 pmw) indicate the use of solidarity strategy with strategies to soften propositions and requests in arguer conversations. Likewise, deference markers are present at 445.33 pmw, respectively, with much higher frequencies of *thank you* (124.98 pmw), *sorry* (87.84 pmw), and *madam* (6.31 pmw). Hence, Malaysian English should not be described as purely informal and solidarity-oriented. Instead, it has a politeness profile involving characteristics of both interpersonal warmth and mitigation alongside gratitude, apology, and respectful address.

In summary, this paper shows that Malaysian English presents a multi-dimensional but either solidarity-dominant or politeness-oriented pragmatic profile. The web-based discussion it leads seems to favour overt displays expression of thankfulness and alignment with others combined with a heavy reliance on hedging language and politeness strategies. Thus, the results suggest Malaysian English serves collectivist harmony not solely through implicit politeness but with a combination of explicit solidarity, mitigation, respect and relational engagement.

6.4 Cross-Varietal Interpretation

When all three varieties are compared proportionally the politeness hierarchy emerges as follows:

GB → Tentativeness, PK → Deference + Solidarity, MY → Solidarity-Dominant.

Through these distinctions it is validated that politeness expression correlates with each variety's socio-cultural and communicative context. Moreover, the statistical significance ($\chi^2(40) = 14,456.59$, $p < .001$) fosters that such variation is not random but systematically patterned according to pragmatic norms and audience expectations but at the same time this significance is small in magnitude (Cramer's $V = .096$).

7 Conclusion

Using twenty-one formulaic expressions from the GloWbE corpus, this comparative corpus-based study explored politeness variation in British, Pakistani, and Malaysian English. Considering Xie's (2024) triadic framework of tentativeness, deference, and solidarity as a guiding framework, the findings of this study reveal that politeness is a culturally dynamic rather than a universal linguistic system, as cross-varietal differences in politeness were discovered through statistical analysis $\chi^2(40, N = 792\ 276) = 14456.59$, $p < .001$, Cramer's $V = .096$. The findings of the study are aligned with Locher and Watts (2008) Relational work theory, which posits that politeness strategies are used in a strategic way to create, sustain or change social relationships and reflect interpersonal relevance power-based solidarity and macro-social context communication context. Here, British English was found leaning toward tentative politeness through hedging and modal verbs (*please, perhaps*). The selective use of softening mediated by hedges and modality in requests by British English speakers is a good illustration of relational work theory (Locher & Watts, 2008), reflecting cultural preferences for discipline and indirectness (Blum-Kulka, 1987; Culpeper & Demmen, 2011). In contrast, the Pakistani English version contained more joint usages of deference and solidarity strategies (sir, thanks, love), exemplifying simultaneous recourse to respect and camaraderie because they both co-occur within hierarchical collectivist norms (Rahman 2020; Saleem 2021). Additionally, Pakistanis combine their deference with solidarity strategies, creating a particularly interesting case of relational work theory (Locher & Watts, 2008), as it works within the framework that recognizes and respects hierarchy as means of expressing friendship and alliance in collectivist social systems. It is an example which serves as a signifier of the intricate give and take in different axes of relationships manifesting within the communications through this culture. The highest overall normalized frequencies of the selected expressions can be found in Malaysian English, which also shows an undoubtedly strong preference for solidarity markers over tentative and deferential forms, the mélange of solidarity, mitigation, gratitude and respect aligns with relational work because it is about maintaining relationships in the context. In general, the study has confirmed that World Englishes differ both pragmatically and linguistically. English varies to local moralities and communicative traditions through its reshaping of the practice of politeness in different contexts. These results can be viewed, from a didactic perspective, as reinforcing the need to raise awareness regarding regional Englishes and counterbalance uncritical use of native-speaker norms through the promotion of teaching pragmatics as culturally established.

References

- Ahmed, S. (2024). Politeness strategies in Pakistani television discourse: A pragmatic analysis. *Journal of Pragmatics*, 219, 45–60.
- Alakrash, H. M., & Bustan, E. S. (2020). Politeness strategies in requests by Arab EFL and Malaysian ESL learners. *International Journal of English Linguistics*, 10(3), 195–205.
- Barron, A., & Schneider, K. P. (Eds.). (2008). *Variational pragmatics: A focus on regional varieties in pluricentric languages*. John Benjamins. <https://doi.org/10.1075/pbns.178>
- Blum-Kulka, S. (1987). Indirectness and politeness in requests: Same or different? *Journal of Pragmatics*, 11(2), 131–146.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Culpeper, J., & Demmen, J. (2011). Nineteenth-century English politeness: Negative politeness, conventional indirect requests and the rise of the individual self. *Journal of Historical Pragmatics*, 12(1–2), 49–81.
- Culpeper, J., & Gillings, M. (2018). Politeness variation in England: The north–south divide revisited. *Journal of Pragmatics*, 138, 152–170.
- Davies, M. (2013). *Corpus of Global Web-Based English (GloWbE)*. Brigham Young University. <https://www.english-corpora.org/glowbe/>
- Davies, M. (2015). Corpora of Global Web-Based English: 20 English varieties, 1.9 billion words. *International Journal of Corpus Linguistics*, 20(1), 1–28.
- Eelen, G. (2001). *A critique of politeness theories*. St. Jerome Publishing.
- Harun, M., & Abdul Rahman, N. (2024). Politeness strategies in Malay and English greetings and invitations: A cross-linguistic study. *Asian Englishes*, 26(1), 34–51.
- Hashim, A. (2020). Politeness and communication in Malaysian English. *World Englishes*, 39(1), 87–101.
- Hashim, A., & Low, E. L. (2018). *English in Southeast Asia: Features, policy and language in use*. Routledge.
- Hashim, A., & Mahmud, M. M. (2020). Politeness strategies in refusals by Malay speakers of English. *Journal of Pragmatics*, 168, 38–54. <https://doi.org/10.1016/j.pragma.2020.06.004>
- Jucker, A. H. (2013). *Pragmatics and the history of English*. Edinburgh University Press.
- Jucker, A. H. (2020). Politeness in historical pragmatics. In A. H. Jucker & I. Taavitsainen (Eds.), *The handbook of historical pragmatics* (pp. 253–276). De Gruyter.
- Jucker, A. H. (2023). Stereotypes of declining politeness in British and American English. *Journal of Historical Pragmatics*, 24(1), 1–26. <https://doi.org/10.1075/jhp.22007.juc>
- Kachru, B. B. (2005). *Asian Englishes: Beyond the canon*. Hong Kong University Press.
- Khan, M. A., & Jabeen, F. (2018). Apology strategies in Pakistani English: A sociopragmatic study. *Journal of Pragmatics*, 135, 1–14.
- Landert, D., Jucker, A. H., & Haugh, M. (2023). *Corpus pragmatics*. Cambridge University Press.
- Lim, S. C. (2012). Politeness in Malaysian English: A pragmatic perspective. *Journal of Pragmatics*, 44(13), 1657–1670.
- Locher, M. A., & Watts, R. J. (2008). Relational work and impoliteness: Negotiating norms of linguistic behaviour. In D. Bousfield & M. A. Locher (Eds.), *Impoliteness in language* (pp. 77–99). Mouton de Gruyter.
- Ogiermann, E. (2009). *On apologising in negative and positive politeness cultures*. John Benjamins.
- Rahman, T. (1998). *Language and politics in Pakistan*. Oxford University Press.
- Rahman, T. (2020). *Language and politics in Pakistan* (2nd ed.). Oxford University Press.
- Saleem, S. (2020). Politeness and ideology in Pakistani broadcast discourse. *Discourse & Society*, 31(4), 417–435.
- Saleem, S. (2021). Politeness strategies in Pakistani English digital discourse. *Asian Englishes*, 23(3), 245–259.
- Schneider, K. P., & Barron, A. (2008). *Variational pragmatics: A focus on regional varieties in pluricentric languages*. John Benjamins.
- Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Language Science Press.
- Terkourafi, M. (2015). Conventionalization: From pragmatics to grammar. *Language and Linguistics Compass*, 9(10), 420–432.
- van Dorst, I., Gillings, M., & Culpeper, J. (2024a). Regional and social variation in English politeness formulae. *Journal of Pragmatics*, 209, 1–17.
- van Dorst, I., Gillings, M., & Culpeper, J. (2024b). Sociopragmatic variation in Britain: A corpus-based study of politeness. *Journal of Pragmatics*, 227, 37–56.
- Xie, F. (2024). Politeness variation in Britain, Australia, and Hong Kong: A corpus-pragmatic study. *Corpus Pragmatics*, 8(3), 313–333.
- Zafar, M. (2023). Politeness strategies in Pakistani business correspondence. *International Journal of Business Communication*, 60(2), 265–283.