UNIVERSITY OF BERGEN
Faculty of Humanities

Department of Linguistic, Literary and Aesthetic Studies

On the use of prepositional verbs by Pakistani ESL learners: A corpus study

Iram Nazir

Thesis for the degree of Master in Linguistics (LING350)
Fall 2023

Dedication

To my wonderful mother, Safia, whose constant support, infinite love, sacrifices, and prayers have been the foundation of my strength and determination to succeed.

In loving memory of my late father, Muhammad Nazir, whose steadfast trust in my abilities and the eternal blessings will always be cherished.

Acknowledgements

I would like to convey my heartfelt appreciation to everyone who helped me complete this MA thesis.

All glory be to Allah Almighty, whose mercy provided me with the strength to complete this difficult work. It would not have been feasible to give it its current shape without the help of many individuals around me. To begin with, I want to thank my supervisor, Prof. Koenraad De Smedt, for his essential assistance, experience, and valuable input. He also helped me write the R program for quantitative analysis. His guidance was vital to improving the quality of this research.

I am eternally thankful to my parents, especially my late father, for their constant encouragement, support, and faith in my skills, which have served as guidance for me throughout this academic journey. I want to thank my siblings and friends for offering emotional support during my challenging times. Despite personal challenges such as mental health concerns and the loss of my father, this accomplishment exhibits resilience and I recognize the courage required to keep going.

Iram Nazir

Abstract

English as a second language is widely used in Pakistan, but Pakistani students of ESL (English as a second language) still face difficulties in writing and speaking English. Errors are a natural part of second language learning, so teachers need to be aware of the causes, effects, and consequences of errors. A common challenge for ESL students is using prepositions correctly. Several studies have been conducted on ESL learners' use of English prepositions, but few on Pakistani ESL learners' errors using prepositional verbs. Therefore, it is important to investigate the hypothesis that Pakistani ESL learners may make relatively more errors in prepositional verbs than other Asian ESL learners. The main objectives of this study are to assess Pakistani ESL students' proficiency in the use of prepositional verbs, to investigate recurring errors, and to determine the role of transfer and L1 intervention in these errors. The current study employs a corpus linguistics approach with a mixed method that combines quantitative and qualitative data analysis. Qualitative methods identify and describe individual items as correct or incorrect, whereas quantitative methods group the correct and incorrect prepositional verbs to differentiate among Pakistani students compared to other Asian ESL students. This study focuses on identifying common prepositional verbs and an exploration of errors and challenges faced by Pakistani learners. The data comes from ICNALE (International Corpus Network of Asian Learners of English), which has collected data from ESL learners in ten Asian regions. Samples of prepositional verbs belonging to 28 lemmas in specified grammatical contexts were selected from the written essays module that has 5,600 entries by 2,800 participants. A total of 19027 observations were manually annotated as correct, incorrect, or irrelevant. Of these, 5106 relevant observations were subjected to a quantitative analysis, which suggests that Pakistani learners make relatively more errors (only 74.62% correct) than the other groups. Lemmas with high error rates among Pakistani learners are discussed, including an analysis of transfer and L1 interference as possible factors. There has been insufficient research on prepositional verbs used by Pakistani ESL learners, hence, this study will help English language teachers to identify these issues and rethink their teaching methods. It contributes to a pedagogical understanding of the challenges of second language acquisition. This study concludes by emphasizing the importance of addressing the challenges of prepositional verbs in ESL classes and provides insights for teachers, students and translators.

Keywords: Prepositional verbs, Pakistani ESL Learners, ICNALE, Interlanguage, Transfer, Urdu Interference.

Sammendrag

Engelsk som andrespråk (ESL) er mye brukt i Pakistan, men pakistanske ESL-studenter har fortsatt problemer med å skrive og snakke engelsk. Feil er en naturlig del av andrespråkstilegnelse, så lærere må være klar over årsakene, virkningene og konsekvensene av feil. En vanlig utfordring for ESL-studenter er korrekt bruk av preposisjoner. Det er utført flere studier på ESL-studenters bruk av engelske preposisjoner, men få studier på pakistanske ESL-studenters feil ved bruk av preposisjonsverb. Derfor er det viktig å undersøke hypotesen om at pakistanske ESL-studenter kan gjøre relativt flere feil i preposisjonsverb enn andre asiatiske ESL-studenter. Hovedmålene med denne studien er å vurdere pakistanske ESL-studenters ferdigheter i bruk av preposisjonsverb, undersøke tilbakevendende feil, og undersøke en mulig årsak relatert til L1-påvirkning. Den nåværende studien bruker en korpuslingvistisk tilnærming med en blandet metode som kombinerer kvantitativ og kvalitativ dataanalyse. Kvalitative metoder identifiserer og beskriver individuelle elementer som korrekte eller feil, mens kvantitative metoder grupperer de riktige og ukorrekte preposisjonsverbene for å skille mellom pakistanske studenter sammenlignet med andre asiatiske ESL-studenter. Denne studien fokuserer på å identifisere vanlige preposisjonsverb og analysere feil og utfordringer som møter pakistanske studenter. Dataene kommer fra ICNALE (International Corpus Network of Asian Learners of English), som har samlet inn data fra ESL-studenter i ti asiatiske regioner. Preposisjonsverb tilhørende 28 lemma i spesifiserte grammatiske sammenhenger ble selektert fra modulen med skriftlige stiler, som omfatter 5600 oppføringer av 2800 deltakere. Totalt 19027 observasjoner ble manuelt annotert som korrekte, ukorrekte eller irrelevante. Av disse ble 5106 relevante observasjoner gjenstand for en kvantitativ analyse, noe som tyder på at pakistanske studenter har en relativt høy feilrate (kun 74,62% riktig) enn de andre gruppene. Spesifikke lemma med en høy feilrate blant pakistanske studenter blir diskutert, inkludert en analyse av transfer og L1-interferens som mulige faktorer. Det har ikke vært tilstrekkelig mye forskning på preposisjonsverb hos pakistanske ESL-studenter, og derfor vil denne studien hjelpe engelsklærere til å identifisere disse problemene og revurdere undervisningsmetodene. Forhåpentlig bidrar dette til en pedagogisk forståelse av utfordringene ved andrespråkstilegnelse. Denne studien konkluderer ved å understreke viktigheten av å ta opp utfordringer knyttet til preposisjonsverb i ESL-klasser og skaffer innsikt til lærere, studenter og oversettere.

Contents

Chapter 1

Introduction

According to Ellis (1994, p. 4), "Language learning is a complex process that involves developing communicative competence, including knowledge of grammar, vocabulary, pragmatics, and discourse conventions." The necessity of second language learning is evident, as language is an essential aspect of human communication, and the ability to communicate in a second language has become a fundamental requirement in the globalized world.

This work focuses on the learning of English as a second language, and on the use of prepositional verbs by Pakistani learners as compared to other Asian learners. English has become the second most widely spoken language in the world and has undeniably gained international clout. English is the most commonly used language in foreign communication, and there is an increasing interest in studying it, also in Pakistan, where it is one of two official languages. The other official language is Urdu, an Indo-Aryan language which is the lingua franca of Pakistan, although a great number of other languages are also spoken by diverse ethno-linguistic groups. All educational levels in Pakistan offer English language instruction. Pakistani English is regarded by some as a subset of British English (Kachru and Nelson 2006).

Pakistan is a linguistically diverse country, with several regional languages spoken in various areas. While regional languages like Punjabi, Sindhi, Pashto, and Balochi are significant, Urdu serves as a common language that connects various linguistic diversities. It enables communication between people who speak various languages. Given Pakistan's linguistic diversity, Urdu provides a neutral and inclusive mode of communication. Individuals from diverse linguistic origins frequently converse in Urdu to overcome language hurdles. According to the country's constitution, Urdu is the national language of Pakistan. It is utilized in official papers, government communications, and national celebrations as a symbol of national unity. This formal acknowledgment leads to its dominance as the principal mode of communication. Urdu is used in everyday life, including family interactions, social events, and media, and is not confined to official situations. In addition, Urdu is frequently used as a medium of instruction in educational institutions throughout Pakistan, particularly at the primary and secondary levels. While English has grown increasingly popular in higher education, Urdu continues to be an important language for academic learning. This educational function reinforces its position as the predominant language of many Pakistanis.

As Masood et al. (2020, p. 111) stated, "Urdu along with Punjabi, Pashto, Sindhi, Saraiki, and Balochi is considered the first language of most of the speakers in Pakistan. English as a second language (ESL) is learned by the people due to the demand of educational institutions. Right from the beginning of schooling, almost every student in Pakistan starts learning English as a second language." Moreover, Farukh and Vulchanova (2015) discussed the reason Urdu is considered as a first language (L1) in Pakistan, namely that Urdu is the national language,

the national medium of instruction, and the language of the mass media. It is used for internal communication as well as writing. In the present study it is therefore considered as L1 due to its importance as a lingua franca and as the language in which all children learn to read.

Previous research on L1 interference for Pakistani ESL learners has focused on Urdu because it is the language spoken and understood by the majority of the population. Investigating the obstacles and patterns of language learning in Pakistan requires an awareness of how Urdu impacts the acquisition of other languages such as English. Few previous research works about Pakistani ESL learners have considered Urdu as L1 and investigated its interference in learning English (Farooq et al. 2012; Farukh and Vulchanova 2015; Ishaq and Bukhari 2016; Masood et al. 2020; Saeed et al. 2015; Zafar 2018). So, this present study considers Urdu as L1 and investigates its role in interlanguage errors under discussion.

Proficiency in one's native language does not guarantee proficiency in writing another language (Komba and Bosco 2015; Mahmood 2009; Sultana 2018). Having been a teacher of English in Pakistan for fifteen years, I have always found language learning concerns and difficulties to be of considerable interest. The errors that language learners make in their speech, writing, and other language skills, are the result of various factors, including an intricate interplay between the learners' first language and their exposure to the target language. Grammatical errors can have an impact on the clarity and interpretation of learners' speech and writing.

My goal in this study is to help people comprehend the learning of a very complicated and difficult category in English, namely prepositional verbs. As a native Urdu speaker and reasonably proficient speaker of English as a second language, I am constantly amazed by prepositional verbs and how they appear to convey so much of our most fundamental perspective of the world around us. The current project was prompted by my curiosity as to why their correct use is so difficult to learn, with specific emphasis on Pakistani learners of English, in comparison to other Asian learners.

Prepositional verbs are created by combining a verb with a preposition to provide a new meaning that cannot fully be deduced from the separate components. More generally, multiword sequences can be difficult to learn; unsurprisingly, learners use fewer multiword sequences than native speakers do (Christiansen and Arnon 2017). Due to the fact that the semantics of multiword verbs are not compositionally transparent, learners may find it difficult to use and understand them. Furthermore, multiword verbs from a learner's native language may conflict with the expression of the corresponding meanings in the target language (Christiansen and Arnon 2017; Gardner and Davies 2007).

It is a hypothesis of this thesis that an effect of the learner's own native language can be found in that learner's use of English prepositional verbs. To investigate this possibility in some detail, examples of such verbs were identified in a corpus of texts in English as a second language and a quantitative analysis was performed. The patterns of use of prepositional verbs by Pakistani learners, as apparent from corpus examples, were compared to that of other Asian learners.

The aim to identify and explain any deviations in the use of prepositional verbs is motivated by the belief that such insights will help both teachers and learners to be aware of the challenges for the correct use of prepositional verbs. It is hoped that this, in turn, may contribute to improved learning methods and materials that stimulate training in the correct use of English verb complementation, taking into account the specific linguistic background of Pakistani learners.

1.1   Aim and limitations of this study

Prepositional verbs have received little attention as a prominent category of multi-word verbs (Ella and Dita 2017), particularly in corpus-based analyses of Pakistani ESL learners. The goal

of this study is to address this void and make substantial contributions to the current literature on ESL in Pakistan. Based on materials from the ICNALE corpus (see Chapter 4), this study seeks to identify, count and compare errors in Pakistani and other Asian ESL learners' usage of prepositional verbs, and to provide possible explanations for specific errors by Pakistani learners.

This study has been limited to a corpus of essays written by a group of learners in ten regions. The possibility that spoken language might contain other kinds of errors was not taken into account. Only prepositional verb constructions were studied. No attention was given to any other syntactic and grammatical deviations in Pakistani ESL learners' writings. The purpose has solely been to annotate errors in the materials, to perform a quantitative analysis, and to provide grammatical descriptions and explanations for specific issues.

1.2 Problem statement

ESL learners in Pakistan, a multilingual community, encounter a variety of English-related challenges, which lead to errors. These errors might be made consciously or unintentionally as a result of ignorance or a lack of knowledge. These issues, if not handled swiftly, might lead to fossilization, making it difficult to quit later (Iqbal et al. 2019). This paper focuses on one such difficult area i.e., the correct use of prepositions, and more specifically, prepositional verbs. The current study focuses on the difficulties that Pakistani ESL learners have while utilizing prepositional verbs, as well as the impact of transfer and L1 interference in these mistakes. Prepositions appear to be simple, and students disregard them even though some prepositions are difficult (Mohamed et al. 2014). Pakistani ESL learners come from a variety of cultural and linguistic backgrounds, and they frequently struggle to create grammatically correct words. According to Sultana (2018), these errors may or may not be the result of first language intervention. Alternatively, pupils may attempt to translate the preposition in their L1, resulting in prepositional misreading (Abualzain 2017).

1.3 Significance

The corpus-based study will assist both English teachers who want to dedicate more time to prepositions and curriculum planners who want to uncover the best ways to improve with prepositions because there is no specific research on prepositional verb errors in the Pakistani ESL context. The current study aims to contribute to the existing literature on Pakistani ESL learners by providing a new perspective. The pedagogical implications will be useful to ESL instructors, students, and curriculum designers in developing remedial techniques. Thus, it will play a significant role in assisting the concerned stakeholders in taking remedial action and developing courses aligned with learners' needs and developmental stages.

1.4 Objectives

The main objectives of the present study are the following:

1. To utilize a qualitative method to identify and describe errors of prepositional verbs in the corpus.

2. To employ a quantitative method to determine the frequencies of relevant corpus items identified as correct or incorrect.

3. To investigate the possibility of L1 interference faced by Pakistani ESL learners acquiring English argument structure, as apparent from prepositional verb use.

1.5   Research Questions

1. What are the kinds of errors made by Pakistani learners in using English prepositional verbs?

2. Is there a significant difference between Pakistani and other Asian ESL learners in the frequency of their errors against prepositional verb use?

3. Can the errors by Pakistani learners be explained by L1 interference in the acquisition of argument structure?

1.6   Hypotheses

1. Pakistani ESL learners make errors in their use of prepositional verbs, including the use of wrong prepositions and omitting prepositions.

2. Pakistani ESL learners commit more errors in their use of prepositional verbs than other Asian ESL learners.

3. L1 argument structure can interfere with the use of prepositional verbs by Pakistani ESL learners.

The remainder of this thesis is organized as follows. Chapter 2 will present the theoretical background. Chapter 3 will focus on the particular constructions that have been researched. Chapter 4 will present the data and method, and Chapter 5 will describe the findings of the quantitative analysis. These will be further discussed in Chapter 6. The thesis will be concluded in Chapter 7.

Chapter 2

Theoretical background

2.1   Learner language research

Krashen (1982) distinguished between acquiring and learning as part of his Monitor Theory. Whereas the acquisition of a language is a natural process, learning a language is a conscious one. In the former, the learner partakes in natural communicative situations and does not necessarily get feedback. This is normally the case in learning the mother tongue or L1 at a young age. In the latter case, error correction is present and elements of grammar may be learned explicitly. This is often the case in learning a second or consecutive foreign language (L2) in classroom settings. Not all educators in foreign language agree to this distinction. Still, the study of how a second language is learned or acquired, often referred to as second-language acquisition (SLA), is a theoretical starting point of this thesis.

Second language acquisition (SLA) is a branch of linguistics that deals with the study of how people learn a second language, and it has become a subject of interest to researchers and educators in recent years. SLA research has shown that no matter what syllabus teachers use, learners of all languages have their own "built-in syllabus", or systematic developmental sequence (Corder 1967; Selinker 1972; Spada and Lightbown 2019; Tarone 1979). SLA is a complex and multifaceted process, and understanding the mechanisms that underlie it is essential for language teachers to design effective teaching methodologies (Krashen 1981). It is influenced by many factors, including age, motivation, aptitude, and social and cultural factors (Gass and Selinker 1994).

Learner language, also known as interlanguage, refers to the linguistic system that an individual develops during L2 learning. The concept was first introduced by Selinker (1972), who argued that L2 learners develop a systematic, but incomplete, language system that is partially influenced by their L1 and partially influenced by the target language (L2). It is a transitional system that is created by the learner as they attempt to make sense of the new language. This system is based on the learner's existing knowledge of L1, as well as their experience with L2. Interlanguage is a separate linguistic system with its own rules, structure, errors and deviations that are different from both L1 and L2. Learner language is characterized by both similarities and differences when compared to the target language, and it is often influenced by the learners' L1 and previously learned other languages, as well as their individual learning strategies and experiences. He argues that understanding the interlanguage of language learners is important for language teachers, as it can provide insights into the specific mistakes and learning difficulties that learners may experience. Interlanguage is dynamic and constantly evolving as learners gain more exposure to the target language and receive feedback on their use of it. It is a natural part of the language learning process and can serve as a valuable tool for language teachers to

help identify areas where learners may need additional support or instruction.

Interlanguage may be thought of as a transitional linguistic system at all levels (phonology, morphology, syntax, semantics, pragmatics) that is distinct from both the L1 and L2. It may be defined in terms of changing language patterns and norms, and it can also be explained in terms of learning processes in a cognitive and sociolinguistic context. According to Corder (1967) and Ellis (1994), understanding learners' interlanguage is critical for language teachers because it allows them to optimize language instruction and provide appropriate feedback suited to the learner's interlanguage needs.

There have been several studies on the topic of interlanguage (e.g., Block 2009; Corder 1971; Lyster and Ranta 1997; Selinker 1972; Tarone 2006). Studies have examined the role of transfer from the L1, the effects of instruction, the role of feedback, and the influence of social and cognitive factors. Lyster and Ranta (1997) examine the role of feedback in language learning, focusing on corrective feedback specifically. They suggest that while corrective feedback can be an effective tool for language learning, it must be provided in a way that is tailored to the learner's interlanguage. In other words, feedback that is too simplistic or too complex may not be as effective as feedback that is designed to target the particular needs of the learner's interlanguage. Similarly, Block (2009) explores the impact of context on interlanguage development, specifically looking at the role of social and cultural factors. He suggests that learners' interlanguage is not solely determined by individual factors, such as their first language or cognitive abilities, but is also influenced by the social and cultural environment in which they learn the new language. He argues that language teachers should be aware of these contextual factors and work to create a supportive and inclusive learning environment that considers the learners' social and cultural backgrounds. Corder (1971) suggests that errors made by L2 learners are not random but are rather systematic and reveal the learners' interlanguage system. Tarone (2006) proposes that interlanguage is dynamic and that learners' language production undergoes constant change as they acquire new input and integrate it into their interlanguage system.

2.2 Transfer theory and error analysis

Transfer theory, which is relevant to the current study's assumptions, argues that learners may transfer the rules and patterns of their L1 to the L2, resulting in mistakes and deviations from the target language norms. According to learner theory, second language learners frequently transfer linguistic traits from their first language to the target language. Transfer theory, a subset of learner theory, contends that learners are more likely to transfer structures that are similar in both languages, and that structure transfer might result in mistakes in the target language. In the context of prepositional verbs, the transfer of structures from L1 to English can lead to errors such as the inappropriate use of prepositions, the omission of prepositions, or the use of incorrect prepositions. These errors can affect the clarity and accuracy of communication, which can be a significant barrier to effective language learning.

Allen (2015) defined an argument as a noun phrase with a specific grammatical or semantic relationship to a verb, the existence of which is necessary for well-formedness in structures that incorporate that verb, whether explicit or indirect. She characterized arguments in terms of verb-related syntactic and semantic responsibilities. According to Akbarnezhad et al. (2020), argument structure is the arrangement of the quantity and types of arguments necessary for a verb in that structure to be well-formed. Jackendoff and Jackendoff (2002) state that "The problem of argument structure is central to any theory of grammar" (p.137). Understanding argument construction is essential for mastering predication. The verb detects predication, which may be used to understand how linguistic statements interpret events and circumstances (Akbarnezhad

et al. 2020). The key issue for EFL/ESL learners is determining which verbs will appear in which argument structures. According to Yazdi and Rezai (2015) and Akbarnezhad et al. (2020), the learning of English argument formation is difficult for ESL students; problematic interlanguage can be ascribed to proficiency level, L1 impact, and specific verbs.

Recent studies in SLA have focused on learners' errors to predict the difficulties involved in language learning. Alhamadi (2015) describes error analysis as the linguistic analysis of errors that learners produce. She discusses previous research on error analysis and its role in SLA. She argues that prior knowledge can be beneficial in learning a second language, but it can also lead to errors due to incorrect guessing. Errors are considered pedagogical grounds for defining the process of second language acquisition (SLA), as they are a fundamental aspect of language acquisition. They allow learners to expand their knowledge through self-correction. Teachers and researchers are responsible for determining the cause of errors, as they occur and are not acknowledged by learners. Alhamadi (2015) distinguishes between mistakes and errors. Errors indicate gaps in a learner's knowledge, while mistakes indicate occasional slips in performance. Error analysis aims to determine the learners' knowledge gaps and provide the teacher with appropriate information to help them form a more accurate concept in the target language (Brown 2007; Ellis 1994, 1997; Gass and Selinker 1994; Gass and Selinker 2008; Heydari and Bagheri 2012; Richards 1974; Spada and Lightbown 2019).

Furthermore, Alhamadi (2015) discusses two types of errors, one of which is transfer error, which occurs when a student uses his or her first language skill. This is also known as mother-tongue interference, and it is an issue with language transmission in that familiar patterns are favored over new, unfamiliar ones, and this preference becomes the basis for mother-tongue interference. Additionally, the influence of the first language can be positive if it assists in the acquisition of the second language, or it can be negative if inconsistencies between the two languages develop and cause learning issues and errors.

Bakken (2017) also discusses the role of interlanguage and transfer theory in SLA and errors. She argues that the behaviorist model of SLA emphasizes the importance of transfer in SLA. Positive transfer involves learners applying their native language habits to produce L2 structures, which are similar in the target language, whereas, Negative transfer occurs when learners are disturbed by their native language habits, producing erroneous L2 utterances. However, Generative accounts of SLA describe the competence of nonnative speakers in L2, based on Corder's hypothesis that L2 learners have their own developing systems (Corder 1975), which is why Selinker (1972) introduced the term Interlanguage (IL) to describe this system. Within generative approaches, learners' systems are understood as an unconscious mental representation of grammar, similar to native speakers. The present study investigates the role of L1 interference in using prepositional verbs by Pakistani vs. other Asian ESL learners, aiming to understand the impact of these transfer errors on their language development.

2.3   CEFR levels of language proficiency

The *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (CEFR) is an instrument put together by the Council of Europe as part of its promotion of new approaches to L2 teaching, notably the development of a communicative approach, while referring to increasing plurilingualism in Europe, in which knowledge and experience of several languages interrelate and interact (COE 2001). In order to harmonize teaching and certification requirements, the CEFR organizes language proficiency in six levels, A1, A2, B1, B2, C1 and C2, which can be regrouped into three broad levels: Basic User, Independent User and Proficient User, and that can be further subdivided according to the needs of the local context. Language

education programs and language assessments are often aligned with the CEFR levels to ensure that learners are making progress towards standardized language proficiency goals, thus allowing learners, teachers, and institutions to compare language ability across different contexts and languages. In the context of our study, it is relevant that B1 and B2 have been subdivided. The levels which are directly relevant in the present context are defined as follows.[1]

- A2 refers to the 'waystage' level or the second lowest level of language proficiency in the CEFR. It represents a basic level of competence in the language, where learners are able to grasp simple words and expressions linked to everyday situations and can communicate in basic conversations such as ordering meals or asking for directions.

- B1.1 is an intermediate level of language proficiency in the CEFR. It represents a 'threshold' level of competence in the language, where learners are able to understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc., and can produce simple connected text on topics that are familiar or of personal interest.

- B1.2 is also an intermediate ('threshold') level of language proficiency in the CEFR. It represents a higher level of competence in the language than B1.1, where learners are able to understand the main ideas of complex text on both concrete and abstract topics and can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers possible without strain for either party.

- B2+ refers to the 'vantage' or an advanced level of language proficiency in the CEFR, which represents a higher level of competence than B2. B2+ learners can understand complex texts on a wide range of topics of academic or professional contexts and can express themselves clearly and effectively in a variety of situations (see also EALTA 2016).

2.4 English as a second language (ESL)

ESL stands for English as a Second Language, and it refers to the process of learning and teaching English as a non-native language. The goal of ESL is to help non-native speakers of English develop their language skills in the four main areas: speaking, listening, reading, and writing (Kanno et al. 2017). The majority of speakers of English do not have this language as their mother tongue, but as an L2 they use for communication in a wider circle, such as in trade, higher education, administration and travel. English, like French, Spanish and Russian, are officially recognized in some multilingual countries as a means of public communication. Thus, people of such multilingual countries learn these languages in addition to their native language. For example, English is a second language in several South Asian countries like India, Bangladesh, Pakistan and Sri Lanka. Likewise, it is well known that French serves as an official L2 in several African countries.

ESL differs from other language learning approaches in several ways. Firstly, ESL learners often have specific needs and goals for learning English, such as academic or professional advancement. In contrast, foreign language learners may learn a language for personal interest or cultural enrichment. Secondly, ESL instruction often focuses on the development of communicative skills, whereas foreign language learning may emphasize reading and writing skills (Norton and Toohey 2011). Thirdly, ESL learners may be exposed to English in a variety of ways, either in academic, social, or professional settings, or in a setting of foreign language

---

[1]See also https://coe.int/en/web/common-european-framework-reference-languages/level-descriptions.

learning, with limited exposure to the language outside of the classroom (Cook 2016; Kanno et al. 2017).

Another key difference between ESL and other language learning is that ESL is often taught to adults who already have a basic understanding of their native language. This means that ESL teachers can build upon the student's existing language skills and focus on teaching the nuances of the English language. In contrast, other language learning may be taught to children who are still developing their language skills in general. ESL is typically associated with immigrant populations who have come to an English-speaking country and need to learn the language in order to communicate effectively and integrate into the new culture (Norton and Toohey, 2011).

Research has shown that effective ESL instruction should consider the unique needs and experiences of ESL learners, and should be tailored to their goals, interests, and cultural backgrounds (Norton and Toohey 2011). Additionally, ESL instruction should incorporate opportunities for authentic language use and interaction, as well as strategies for building learners' confidence and motivation (Cook 2016). One study found that intensive ESL instruction can improve the language proficiency of non-native speakers, particularly in the areas of speaking and listening (Baker and MacIntyre 2000). Another study found that ESL instruction that integrates language and content can be effective for promoting academic achievement in non-native speakers (Short and Fitzsimmons 2007).

2.5 Learner corpus research

The present study has a mixed method in the area of learner corpus research, as will be described in more detail in Chapter 4. Corpus linguistics has been one of the most exciting methodological developments in linguistics since its revival with digital methods from the 1960s. It is currently widely taught and its approaches and applications are amply described in the literature (e.g. McEnery and Wilson 2001). A corpus is essentially a large body of deliberately assembled machine-readable language materials. What must be included, in terms of size and types of materials, is dependent on the purpose of the corpus. A corpus may contain authentic texts in printed, oral, or multimedia forms representing a broad cross-section of a language or a specific linguistic variety, depending on the kind of research it is meant to support. Corpora should be documented and may contain linguistic annotation at various levels, most often with parts of speech (PoS).

With such data, corpus-based methods allow the study of language patterns and usage. Digital search and analysis methods allow researchers to quickly and efficiently filter large amounts of text and to identify and quantify patterns that are not as readily apparent with paper-based methods. Corpus studies can assist us in developing theories of language and describing its various characteristics (Hunston 2002). Corpora, as Hunston claims, objectively represent real-life natural language examples, and their findings can be applied in real-life situations. Because corpus studies can, in principle, be verified and replicated, they have become a dominant and frequently used method to study linguistic variation. Many corpus-based discourse studies have come up with diverse analyses of linguistic and discourse features (Jabeen et al. 2011).

The study of ESL and, more generally, SLA, has benefited greatly from the availability of learner language corpora. Indeed, many corpora have been constructed that contain authentic examples of learner language, in order to study the properties of interlingua in its various stages and as related to conditions of the learners and their learning contexts. Major conferences on Learner Corpus Research (LCR) are being held around the world, such as LCR 2013 in Bergen, Norway, to discuss the construction and use of learner corpora. Typically, such corpora are compilations of relatively small texts written by learners in the context of specific tasks, such as

essays written for assignments or tests. The subject of such essays can vary from relatively free or restricted. The genre varies between descriptive and argumentative. In order to investigate potentially relevant differences between learner groups, an array of properties of contributing writers are recorded, such as age, L1, CEFR level, etc. Finally, the texts themselves may be annotated in various ways, such as parts of speech (PoS), error types, and corrected text. A well-known example of a learner corpus is Andrespråkskorpuset (ASK) which has essays by various groups of learners of Norwegian (Tenfjord et al. 2006). ASK has been used in many lexical and grammatical studies of learner language, for instance, to study transfer in gender assignment by learners of Norwegian (Ragnhildstveit 2009).

The current study uses ICNALE, the International Corpus Network of Asian Learners of English, developed in Japan (Ishikawa 2023; Ishikawa 2011, 2013). This corpus contains a large number of topic-controlled speeches and essays produced by college students in ten countries/ regions in Asia, as well as English native speakers. It has four modules: spoken monologue, spoken dialogue, written essays, and edited essays. The corpus is available for online search at the ICNALE site and can also be downloaded from there.[2] More details on the data contained in ICNALE and the method used to explore it follow in Chapter 4.

ICNALE has previously been used in many studies which investigate the challenges which various groups in Asia are facing in learning ESL. The corpus has emerged as a key resource for a wide range of research projects examining the linguistic ability of English learners from different Asian nations. It has been vital to fostering comprehensive exploration into certain language domains among these students, and it addressed topics such as the complexity, accuracy, and fluency (CAF) in argumentative writing (Barrot and Gabinete 2021), phrasal verbs usage among Chinese and Japanese EFL learners (Haugh and Takeuchi 2022), the acquisition of articles by Chinese learners (Leroux and Kendall 2018; Park 2023), adjectives (Wang and Liu 2022), the comparison of collocation use by Turkish and Asian learners (Demirel and Kazazoğlu 2015), connectors (Cho Min 2020), gender differences in vocabulary (Ishikawa 2015), interview structures (Ishikawa 2019) and grammatical and lexical patterning of the word *make* in Asian learner writing (Lin and Lin 2019). Also, Schenck (2020) examines the influence of native and nonnative English-speaking teachers on Korean EFL writing using ICNALE data, while, Suzuki (2015) explores the uses of *get* in Japanese learner and native speaker writing, highlighting the complexities that Japanese learners face in using this verb.

A few Pakistani researchers have utilized ICNALE corpus analysis to explore linguistic features of Pakistani ESL learners. Aslam and Fayyaz (2021) investigate the use of nominalization in argumentative essays by Pakistani learners and native English speakers using ICNALE corpus-based comparative approach. Arslan et al. (2020) present a comparative study of derivational morphemes across native English speakers and ESL/EFL learners using the ICNALE corpus. Additionally, Alyas (2019) conducts a quantitative analysis of subject-verb agreement problems among Pakistani ESL students using ICNALE data. Anwar et al. (2020) focuses on lexical bundles in academic discourse, comparing native and non-native learner corpora from ICNALE. Similarly, Abdulaziz and Mahmood (2022) explore new dimensions in argumentative essay writing using multidimensional analysis techniques on a corpus containing essays by ENL, ESL, and EFL learners. Shah et al. (2021) conducts a corpus-driven analysis of conjunctions in academic writing among EFL learners, emphasizing the importance of mastering text cohesion.

These studies shed light on the language learning issues experienced by ESL and EFL learners, with a special focus on Asian students, by leveraging the enormous dataset inside ICNALE, and give significant insights for both pedagogical practices and future research attempts in En-

---

[2]http://language.sakura.ne.jp/icnale/

glish language education. These studies add to our understanding of language acquisition, writing skill, and the unique problems experienced by learners from various linguistic origins. They use corpus analysis to give significant insights into language usage trends and places for language learners to develop.

Chapter 3

Argument structure and prepositional use in ESL

3.1 Argument structure in English

Argument structure refers to the way in which arguments (subjects, objects, and complements) are combined with verbs to create meaningful sentences. In English, verbs have a particular argument structure, which includes the number and type of arguments that a verb requires. For example, the verb *eat* requires a subject and a nominal object, as in *John ate the sandwich*, while the verb *believe* can take a sentential complement, as in *I believe he is honest*. The argument structure of a verb determines how many arguments are needed, what their grammatical roles are, and the order in which they appear in the sentence.

An *argument* has been defined as a phrase having a specific grammatical or semantic association to a verb and whose explicit or indirect presence is required for well-formedness in structures that include that verb (Allen 2015). Arguments can be defined in terms of syntactic and semantic roles related to the verb. The configuration of the number and types of arguments required for a verb in that structure to be well-formed is known as *argument structure* (Akbarnezhad et al. 2020). Jackendoff and Jackendoff (2002, p. 137) state that "The problem of argument structure is central to any theory of grammar". Understanding argument structure is crucial for learning predication. As argued by Akbarnezhad et al. (2020, p. 2), "to understand how events and states are construed through linguistic expressions, the most typical kind of lexical item that supports predication is the verb.Thus, an analysis of how argument structure is realized in a given language is ultimately ananalysis of how verbs will behave in that language". The argument structure of verbs can vary depending on their semantic meaning and the syntactic context in which they are used. Some verbs may take different types of arguments depending on their transitivity, causativity, or aspect. Understanding argument structure is important for building grammatically correct sentences and for analyzing and interpreting meaning in written and spoken English.

The acquisition of English argument structure is a challenge for ESL students (Akbarnezhad et al. 2020; Yazdi and Rezai 2015). In particular, EFL learners must learn to determine which verbs will occur in which argument structures. Issues with argument structure can be explained through aspects of proficiency, L1 impact, and verb form.

3.2 Transfer of grammatical structure

In linguistics, *transfer* refers to the influence of a learner's first language (L1) on the acquisition of a second language (L2). In the context of argument structure, transfer can refer to the ways in which the argument structure of the learner's L1 can influence their production of argument structure in the L2. For example, if a learner's L1 has a different word order or requires different arguments in certain constructions, they may transfer these patterns into their L2. This can result in errors in their use of argument structure in the L2. On the other hand, transfer can also have positive effects on the acquisition of argument structure in the L2. For example, L1 knowledge of argument structure may facilitate the acquisition of certain aspects of the target language (Goad and White 2004). Similarly, studies have shown that transfer can facilitate the learning of cognate vocabulary items that are similar across languages (Ellis 1994; Odlin et al. 1989).

Transfer can occur at various levels of language, including phonology, syntax, and semantics. In terms of argument structure, transfer may lead to errors in the production of sentences that do not match the argument structure of the target language. Studies have shown that speakers of Japanese and Chinese, which have different argument structure patterns from English, may have difficulty acquiring the English argument structure system (Li and Shirai 2000; Tsimpli and Dimitrakopoulou 2007). The transfer of argument structure from learners' first language (L1) to the target language (L2) has been a topic of interest in the field of second language acquisition (SLA). Here are some common problems that learners face when transferring argument structure from L1 to L2:

1. Different word order: L2 learners may face difficulty when L2 requires a different word order than their L1. For example, if the L1 places the subject before the verb but L2 places the verb before the subject, the learner may produce non-target-like sentences (Jiang and Forster 2011).

2. Different case marking: L2 learners may have difficulty with case marking, which can affect the argument structure of sentences. For example, in German, the accusative case is used to mark the direct object, while the dative case is used for the indirect object. L2 learners may transfer the case marking system from their L1, resulting in incorrect use of cases in L2 (Ionin and Montrul 2010).

3. Different valency: L2 learners may struggle with L2 verbs that require different valency patterns than those in their L1. For example, if L1 verbs take an obligatory object but L2 verbs do not, L2 learners may produce sentences with an extra object (Sorace and Filiaci 2006).

4. L1 interference: L1 structures may interfere with L2 argument structure. For example, if the L1 has a transitive verb that is paired with a specific preposition, the learner may transfer this structure to L2, resulting in a non-target-like sentence (Gass and Selinker 1994).

Research has also highlighted the role of factors such as language proficiency, language use, and context of learning in the transfer of L1 knowledge to the target language. For instance, learners' proficiency in the target language can affect the extent and nature of transfer from the L1 (Jarvis and Pavlenko 2008; Ringbom 1987). Other studies have demonstrated that transfer can vary depending on the context of language use, such as whether the learner is using the target language for communication or for formal study (Cook 2001; Lyster and Ranta 1997).

3.3   Prepositional Verbs in English

In English, prepositions are closed-class words that "initiate NPs to mark the relationship between the NP and other parts of the sentence" (Dekeyser et al. 1979), thus forming prepositional phrases (PPs). There are however several possible functions that prepositional phrases can take. Prepositional phrases that function as objects to verbs are prepositional objects, such as *after her sister* in example (3.1). If PPs that do not function as objects may also function as adjuncts (or adverbial complement), as *after class* in the same example. A crucial characteristic of an object is the possibility of the passive transformation (Dekeyser et al. 1979, p. 294), as in (3.2), whereas adjuncts cannot become the subject of the passive, as in (3.3).

Another characteristic is that prepositional objects are not easily separable from their verbs, whereas adjuncts are. This difference is illustrated in (3.4), which is maybe not impossible, but more marked than (3.1).

(3.1) She was looking after her sister after class.

(3.2) Her sister was looked after.

(3.3) * Class was looked after.

(3.4) ? She was looking after class after her sister.

"Prepositional objects pattern with prepositional verbs," state Dekeyser et al. (1979, p. 294), which means they are in a fixed relationship where the preposition is selected by the verb so as to jointly constitute a particular semantics. Thus, *look after* has the meaning 'take care of' whereas *look for* has the quite different meaning 'search' and *look at* means 'watch'. In contrast, adjuncts like *after class* or *before class* do not affect the basic meaning of the verb, but impart circumstantial information about the event. Moreover, the preposition initiating an NP functioning as adjunct is not selected by the verb, but by the NP. Thus, the choice between *on* and *at* in (3.5) – (3.8) is not dependent on the verb *come*, but is determined by the nouns *Thursday* versus *night*.

(3.5) He came on Thursday.

(3.6) * He came at Thursday.

(3.7) He came at night.

(3.8) * He came on night.

Furthermore, prepositional verbs may sometimes be confused with particle verbs (also called phrasal verbs) because some words may be prepositions or adverbial particles, and in both cases, they may clearly affect the semantics of the verb they combine with. English phrasal verbs are frequent and therefore quite important (Aldukhayel 2014; Gardner and Davies 2007).

An adverbial particle does however not initiate an NP. In example (3.9) the particle verb does not have any object. In (3.10) the verb takes a direct object *her paper*, but it is not initiated by the particle *in*. The phrase *in her paper* may superficially look like a PP, but it is not. An important difference is that particles can be separated from the verb and placed after the noun phrase, as in (3.11). Such movement is not possible with prepositional verbs, cf. (3.12) – (3.13).

(3.9) He turned up.

(3.10) She turned in her paper.

(3.11) She turned her paper in.

(3.12) She applied for the job.

(3.13) * She applied the job for.

Clearly, many verbs may be used with different complementation patterns. In example (3.14) the verb *apply* is a prepositional verb, whereas in (3.15) it is a transitive verb taking an NP as direct object. In both cases, incorrect usage in learner language is possible, as in (3.16), where a preposition is omitted, or (3.17), where the wrong preposition is used. Thus, there are four cases that will interest us in learner language: correct or incorrect use of a PP, and correct or incorrect of an NP.

(3.14) I am applying for the job.

(3.15) I am applying two principles.

(3.16) * I am applying the job

(3.17) * I am applying with two principles.

3.4 Prepositions in Urdu

Urdu is an Indo-Aryan language spoken in Pakistan and India. Although there are many other languages spoken by ethno-linguistic groups in Pakistan, the focus of the present work will be on Urdu, which is the lingua franca in the country and an official language besides English. Urdu has a complex system of prepositions called *Harf-e-Jar*. Like English, Urdu has simple prepositions and multi-word (compound) prepositions. The following lists some common ones, with transliterations[1] and approximate English translations.

پر *par* 'at/on'

میں *mein* 'in'

ساتھ *saath* 'with'

کے ساتھ *ke saath* 'with'

کا / کی / کے *ka/ki/kay* 'of'

کو *ko* 'to'

سے *se/say* 'from/of'

کے لئے *kay liye* 'for'

کےدرمیان *kay darmiyan* 'between'

کے بارے میں *ke barey mein* 'about'

کے علاوہ *ke ilawa* 'apart from'

---

[1]Urdu is read from right to left, but the transliterations are from left to right, for ease of reading.

کے باوجود *ke bawajood* 'despite'

کے بغیر *ke baghair* 'without'

کے پاس *ke paas* 'near'

کے بعد *ke baad* 'after'

کے سامنے *ke samne* 'in front of'

Placement of prepositions in Urdu is different from placement in English. The prepositions in Urdu are usually placed after the word they modify. In that sense, they are postpositions, but for simplicity's sake we call them prepositions. For example, the Urdu sentence in (3.18) has a preposition کو *ko* 'to/towards' that comes after the noun.

(3.18) میں گھر کو جا رہا ہوں
    mein ghar ko ja raha hoon
    I home to go -ing am
    'I am going home'

There are similarities between Urdu and English regarding the semantics and use of prepositional adjuncts. For example, both languages make use of prepositions to indicate time, location, and direction. In Urdu, similar to English, سے *say* 'from' and تک *tak* 'until/to' are used to indicate the start and end of time periods, as illustrated in (3.19).

(3.19) میں 6 سے 9 تک اسکول میں ہوں۔
    mein 6 se 9 tak school mein hoon
    I 6 from 9 (till)to school in am
    'I am at school from 6 to 9.'

The tendency for prepositions to cover somewhat different semantic fields across languages is well known. This is a challenge for learners. For instance, English 'on' covers some meanings of Urdu پر *par* and کو *ko*, as illustrated in examples (3.20) – (3.21).

(3.20) میں صرف ویکنڈ پر کام کرتا ہوں۔
    Main sirf weekend par kaam karta hun
    I only weekend on work do am
    'I work on weekends only.'

(3.21) میں صرف اتوار کو کام کرتا ہوں۔
    Main sirf itwaar ko kaam karta hun
    I only Sunday on work do am
    'I work on Sunday only.'

Urdu prepositions can often be omitted when the meaning is clear without them. In example (3.22), the preposition کو *ko* can be omitted.

(3.22) میں اسکول (کو) جا رہا ہوں۔
    mein school (ko) jaa raha hoon
    I school (to) go -ing am
    'I am going to school'

3.5   Prepositional Verbs in Urdu

Urdu has prepositional verbs. The semantics of the verb is modified according to the preposition that combines with it. For example, the verb جانا *jana* 'to go' combined with the preposition سے *se* 'from' becomes جانا سے *se jana*, 'to leave [a place]', as illustrated in (3.23). The preposition must come directly between the object and the verb.

(3.23)   میں لاہور سے جا رہا ہوں۔
     mein Lahore se ja raha hoon
     I Lahore from go -ing am
     'I am leaving Lahore'

Some other examples of prepositional verbs in Urdu are کو مٹانا *ko mitana* 'to wipe out' and کو ڈھونڈنا *ko dhundna* 'to search for' as illustrated in (3.24) and (3.25).

(3.24)   طاعون نے شہر کو مٹا دیا
     Taoon ne shahar ko mita dia
     plague by city to wiped out
     'The plague wiped out the city.'

(3.25)   وہ کتاب کو ڈھونڈ رہا ہے
     Woh kitaab ko dhund raha hai
     he book to search -ing is
     'He is searching for the book.'

Evidently, the prepositions involved are quite specific to the verb and the target meaning. The choice of preposition cannot be inferred from its literal translational corresponding word in English. Consider the example in (3.26) in which کا 'ka' is used, which generally expresses a possessive meaning, roughly corresponding to English *of*.

(3.26)   وہ منشیات کا عادی ہے
     wo manshiaat ka aadi hai
     he drugs of addicted is
     'He is addicted to drugs.'

In contrast, English *to* generally has a directional sense, which would be expressed by differents prepositions in Urdu, or can be omitted, depending on context, as illustrated in (3.27) – (3.32).

(3.27)   ہم لاہور (کو) جا رہے ہیں
     Hum Lahore (ko) ja rahe hain
     we Lahore (to) travel -ing are
     'We are traveling to Lahore.'

(3.28)   میں بازار (کو) جا رہا ہوں
     Main bazaar (ko) ja raha hoon
     I market (to) go -ing am
     'I am going to the market.'

وہ موسیقی (کو) سن رہی ہے۔ (3.29)
Wo mosiki (ko) sun rahi hai
she music (to) listen -ing is
'She is listening to music.'

وہ پارک (کو) جا رہے ہیں۔ (3.30)
Wo park (ko) ja rahe hain
they park (to) go -ing are
'They are going to the park.'

وہ کل اسلام آباد کی طرف جا رہے ہیں۔ (3.31)
Woh kal Islamabad ki-taraf ja rahe hain
they tomorrow Islamabad to fly -ing are
'They are flying to Islamabad tomorrow.'

میں ہفتے کے لیے منتظر ہوں (3.32)
Main hafte ke-liye muntazir hoon
I weekend for look-forward am
'I am looking forward to the weekend.'

While these examples demonstrate the usage of verbs with prepositions to communicate various meanings in Urdu, the use of prepositional verbs in Urdu is dependent on specific phrases and idiomatic expressions which may or may not be similar to their English counterparts.

3.6 Challenges in learning prepositions and prepositional verbs

Prepositional verbs are an important aspect of Urdu grammar and are used frequently in daily conversation and written communication. Research has shown that the use of prepositional verbs can be challenging for second language learners, since there is usually little direct correspondence between prepositional verb patterns in source and target languages. Studies have found that learners often have difficulty acquiring the idiomatic meanings of prepositional verbs (Brinton and Celce-Murcia 2000; Goldberg 1995). Additionally, learners may struggle with the appropriate use of prepositions in general (Klein 1986).

There is limited research comparing the use of prepositional verbs in the Pakistani English language and Urdu language. However, some studies suggest that there are differences in the use of prepositional verbs between these two languages. Urdu has a complex system of postpositions, which are similar to prepositions in English (Ahmed 2015). However, unlike prepositional verbs in English, Urdu does not have a clear distinction between phrasal verbs and other types of verbs. This lack of distinction in Urdu may lead to difficulties for Pakistani learners of English, as they may have difficulty understanding and using prepositional verbs in English.

Pakistani learners of English often use Urdu syntax and grammar when speaking English, which can result in errors in the use of prepositions and prepositional verbs (Masood et al. 2020). For example, Pakistani learners may use prepositions inappropriately or omit them altogether, which can affect the clarity and accuracy of their English. In terms of the influence of Urdu L1 on English L2, research suggests that L1 transfer can play a role in the acquisition of prepositional verbs in English. For example, Kachru and Nelson (2006) found that Pakistani learners of English may transfer the postpositional structure of Urdu to English, resulting in errors in the use of prepositions and prepositional verbs.

ESL learners often unconsciously transfer Urdu grammar, syntax, and vocabulary while speaking or writing in English. This results in incorrect sentence construction, word choice,

and grammar mistakes. In addition, the teaching of English in Pakistan has traditionally emphasized rote memorization and grammar rules, rather than developing effective communication skills. This approach can lead to overemphasis on grammar rules and a lack of focus on spoken English, creating a language barrier. To overcome this interference, ESL learners in Pakistan need to develop a more practical and communicative approach to learning English. They should focus on listening, speaking, reading, and writing skills, and try to immerse themselves in an English-speaking environment as much as possible. Teachers should also use interactive and engaging teaching methods, such as role-play, discussion, and games, to help learners improve their communication skills. They should also provide feedback on errors and correct them immediately.

Regarding the influence of Urdu L1 on English L2, previous research suggests that the use of prepositions is one area where Pakistani learners of English may encounter difficulties. In particular, Pakistani learners of English tended to overuse prepositions such as *of* and *in*, and underuse prepositions such as *for* and *to*, compared to native English speakers (Khan and Tariq 2012). This suggests that the influence of Urdu L1 may lead to errors in the use of prepositions in English L2. The limited research suggests that the use of prepositional verbs in Pakistani English and Urdu may differ, and that Urdu L1 may influence the acquisition of prepositional verbs in English. Language teachers working with Pakistani learners of English may need to be aware of these differences and provide targeted instruction to address any difficulties or errors in the use of prepositions and prepositional verbs.

The present study continues along these lines and attempts to find further evidence for the challenges faced by Pakistani ESL learners regarding prepositional verbs. As English is a second language in Pakistan, students face difficulties in mastering its complex grammatical structures, including prepositions and prepositional verbs. Previous research has also shown that Pakistani learners tend to make more errors in using prepositions compared to other Asian ESL learners (Khan and Tariq 2012). Also other research has suggested that Pakistani ESL learners may face more difficulties with prepositions compared to other Asian ESL learners. A study has shown that Pakistani learners of English struggle with the correct use of prepositions more than Chinese or Korean learners (Ahmed 2015).

Therefore, investigating the hypothesis that Pakistani ESL learners may make more errors in prepositional verbs compared to other Asian ESL learners is relevant and important in understanding the unique challenges faced by this population. This is worth investigating because it addresses a specific and important issue related to language learning. The accurate use of prepositions and prepositional verbs is essential for effective communication in English, and any errors can cause misunderstandings and confusion. This study will analyze authentic examples in detail to test this hypothesis. The materials and method will be described in the following chapters.

Chapter 4

Data and method

4.1 Methodology

The present study is conducted using a corpus linguistics approach with a mixed method of data analysis. The mixed method is a research methodology that combines both quantitative and qualitative research methods to gain a more comprehensive understanding of a research problem. It can help researchers triangulate data from multiple sources and perspectives to gain a deeper understanding of a phenomenon.

Greene et al. (1989, p. 256) emphasized the mixing of methods and the disentanglement of methods and philosophy (i.e., paradigms) when they define "mixed method designs as those that include at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to any particular inquiry paradigm". Johnson et al. (2007, p. 123) describe mixed methods research as "the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the purposes of breadth and depth of understanding and corroboration".

Creswell and Plano Clark (2017) define mixed methods research as research designs with philosophical assumptions and investigative methods that combine qualitative and quantitative approaches in a single study or series of studies. They further explain that mixed method research involves the integration of different research approaches such as data collection, analysis, and interpretation to address research questions and problems. Similarly, Johnson and Onwuegbuzie (2004) describe mixed methods research as the practice of collecting, analyzing, and blending both quantitative and qualitative research and data in a single study or series of studies.

Overall, mixed methods research is characterized by combining both qualitative and quantitative research methods and integrating different approaches in order to comprehensively address the research question. The purpose of the current study is to identify the types and frequencies of prepositional verb errors in a corpus of writing by Asian ESL learners. Qualitative methods were used to identify and describe individual items as correct or incorrect with respect to standard English, while quantitative methods were employed to calculate the instances of correct and incorrect language use among groups, with the aim of suggesting differences between Pakistani learners compared to other Asian ESL learners.

More specifically, the method followed in the present work consists of the following methodological steps:

1. The qualitative identification of potential problem areas in ESL

2. Obtaining empirical data for the identified problems

3. Manual annotation of the data based on linguistic criteria

4. A quantitative analysis of the annotated observations

5. Discussion of the results

4.2 Data source

The data used in this study are from the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa 2023; Ishikawa 2011, 2013). ICNALE is an international learning corpus created by Dr. Makoto Ishikawa, Kobe University, Japan It has four modules: Spoken Monologue, Spoken Dialogue, Written Essays, and Edited Essays, as shown in Table 4.1. Only the Written Essays module is used in the current study, which focuses on writing skills. Information on the contents of the other modules can be found in the literature which documents the corpus (Ishikawa 2023; Ishikawa 2011, 2013).

The current study only used the written essays core module which contains 5,600 essays written by 2,800 writers, totaling about 1,300,000 words. The essays were written by students who were given 20-40 minutes to write one essay. Participants used a word processor and use of a spelling checker was allowed, but use of references was not allowed. For each essay, students were asked whether they agree or disagree with one of the following statements and to give reasons for their opinion.

1. It is important for college students to have a part-time job.

2. Smoking should be completely banned at all the restaurants in the country.

The participants were from ten South Asian and Southeast Asian countries/regions, listed in Table 4.2 with their codes. In addition, there is a section with essays by English native speakers (ENS), but this has not been used in the present work.

In ICNALE, all the learners have been classified into four kinds of CEFR-linked proficiency bands: A2 (Waystage), B1.1 (Threshold; B1 low), B1.2 (Threshold; B1 high), and B2+ (Vantage), based on their scores in the proficiency tests (cf. Chapter 1). The distribution of participants in levels is given in Table 4.3.

The corpus comes fully annotated with part of speech (PoS) tags. The tags are based on the English TreeTagger[1] part-of-speech tagset that was developed by Helmut Schmid in the TC project at the Institute for Computational Linguistics of the University of Stuttgart (Schmid 1994, 1995) and containing modifications developed by Sketch Engine.[2] An overview of this tagset is given in Appendix B.

4.3 Data selection

In order to study correct and incorrect use of prepositional verbs, samples containing prepositional verbs were obtained from the ICNALE corpus and annotated as described below. The resulting annotations from ICNALE were analyzed to examine the argument structure of these prepositional verbs by Pakistani ESL learners in comparison to other Asian ESL learners. More specifically, the goal was to reveal to what extent different groups of learners use prepositional verbs correctly, as in (4.1), or use them incorrectly, either with a wrong preposition, as in (4.2), or with a noun phrase, as in (4.2).

---

[1]https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/
[2]https://www.sketchengine.eu/english-treetagger-pipeline-2/

Table 4.1: ICNALE major modules (per November 2022)

Modules	Version	Updated	Participants	Samples	Words
Spoken Monologues (SM)	V2.0	2017/8	1,100	4,400	500,000
Spoken Dialogues (SD)	V1.2	2021/9	425	4,250	1,600,000
Written Essays (WE)	V2.4	2019/4	2,800	5,600	1,300,000
Edited Essays (EE)	V3.0	2022/4	328	1,312	150,000
Total	-	2022/11	4,653	15,562	3,550,000

Table 4.2: ICNALE country/region codes

Code	Country
CHN	China, mainland
HKG	Hong Kong
IDN	Indonesia
JPN	Japan
KOR	South Korea
PAK	Pakistan
PHL	Philippines
SIN	Singapore/Malaysia
THA	Thailand
TWN	Taiwan

Table 4.3: ICNALE participants and their distribution in CEFR levels

L1	A2	B1.1	B1.2	B2+	Sum
CHN	50	232	105	13	400
HKG	1	30	52	17	100
IDN	32	82	83	3	200
JPN	154	179	49	18	400
KOR	75	61	88	76	300
PAK	18	91	88	3	200
PHL	2	11	176	11	200
SIN	-	-	134	66	200
THA	119	179	100	2	400
TWN	29	87	61	23	200
ENS	-	-	-	-	200
Total	480	952	936	232	2800

(4.1) I agree with this argument.

(4.2) * I agree of this argument.

(4.3) * I agree this argument.

ICNALE has a specialized interface. However, all ICNALE data, i.e. features, textual data, participants, and tagsets were imported into the Corpuscle corpus management system (Meurer 2012), making it easier and faster to search the corpus as compared to the original interface. A frequency list of the tags in the Corpuscle version of the corpus is in Appendix C.

The first exploratory step consisted of identifying candidate verbs followed by prepositions by means of the search expression [pos = "V.*"] [pos = "IN"]. Their lemmas were obtained and a frequency list of their occurrences in the corpus was made using the Word List function in Corpuscle (cf. Appendix A). Many of these verbs are not prepositional verbs but phrasal verbs or verbs with preposition-initiated adjuncts. Candidates for further investigation as prepositional verbs were selected according to grammatical criteria (cf. Chapter 1). Furthermore, possible verbs which did not yield enough results, namely, *bear, come, need, pass, prohibit*, and *provide*, were eliminated. The following verbs with non-trivial frequencies were eventually selected:

1. addict
2. agree
3. apply
4. base
5. belong
6. benefit
7. care
8. cause
9. compare
10. concentrate
11. contribute
12. cope
13. deal
14. depend
15. disagree
16. focus
17. graduate
18. interfere
19. know
20. learn
21. look
22. prepare
23. relate

24. result

25. sit

26. suffer

27. think

28. worry

All instances of these prepositional verbs were searched in the corpus, excluding gerunds (*ing*-forms), because otherwise we would match many unhelpful adjectival uses. In order to classify these verb occurrences both with respect to argument structure and L1, four distinct queries were carried out for each verb. The queries were formulated in the Corpuscle query language (Meurer 2012, 2020). Examples of these four queries for *result* are given in (4.4) – (4.7). All of them search for all word forms of the lemma *result* used as a verb except the gerund form. Example (4.4) searches the verb followed by a preposition in texts by Pakistani learners, while (4.5) searches the same but in texts by non-Pakistanis. Example (4.6) searches the verb not followed by a preposition in texts by Pakistani learners, while (4.7) searches the same but in texts by non-Pakistanis. These searches were repeated for the other verbs.

(4.4) [lemma="result" & pos="V.*" & pos!="VVG"] [pos="IN"] :: country="PAK"

(4.5) [lemma="result" & pos="V.*" & pos!="VVG"] [pos="IN"] :: country!="PAK"

(4.6) [lemma="result" & pos="V.*" & pos!="VVG"] [pos!="IN"] :: country="PAK"

(4.7) [lemma="result" & pos="V.*" & pos!="VVG"] [pos!="IN"] :: country!="PAK"

The matches obtained by the query were obtained as KWIC lines ('keyword in context') and were stored in different files according to the query that was executed, while also the country was recorded. Table 4.4 presents an example of such a line.

Table 4.4: Example KWIC line

left context	match	right context	country
... Once a person is addicted to	smoking	it's difficult to stop him	PAK

4.4   Data annotation

The KWIC lines were assembled in tabular format with the following columns that were either based on the query results or to be manually annotated.

1. left-context (from KWIC)

2. match (from KWIC)

3. right-context (from KWIC)

4. lemma (as specified in query)

5. country (from query variable)

6. result (C, I or IR, manually annotated)

7. comment (manually annotated)

8. type (NP or PREP, from query result file)

Since the ICNALE annotation is restricted to PoS for each word, the instances of verbs followed by prepositions had to be manually assessed as to whether they are truly used as prepositional verbs or not. The instances of verbs not followed by prepositions had to be assessed as to what occurred after the verb, for instance, a noun phrase correctly used as direct object or otherwise.

Therefore, in the following step, which followed a qualitative method, each observation was manually annotated as correct (C) or incorrect (I) usage, or irrelevant (IR) to the present study. The annotation is based on distinguishing whether a potential prepositional verb is actually involved, cf. Chapter 3, and as already mentioned in that chapter, there are four cases, the criteria for which are as follows.

1. Correct use of PREP type: the verb is correctly used as a prepositional verb, i.e. it is followed by a PP functioning as a prepositional object, where the preposition is correctly selected by the verb.

2. Correct use of NP type: the verb is correctly used as a transitive verb, i.e. it does not require a preposition and it is followed by an NP functioning as a direct object. Since we are not studying transitive verbs, this type is excluded as irrelevant (IR) to the present purposes.

3. Incorrect use of PREP type: the verb used with preposition which is incorrect or superfluous.

4. Incorrect use of NP type: the verb should be used as a prepositional verb but lacks a selected preposition.

In addition, items could be marked as irrelevant (IR) in the following possible cases which led to their removal from the dataset.

1. Particle verbs (phrasal verbs).

2. Verbs with prepositions introducing adjuncts (adverbial complements).

3. Verbs with objects other than NP or PP, for instance, that-clauses.

4. Intransitive uses, i.e. verbs without any kind of object.

A total of 19027 observations were selected and annotated in this way. These were subsequently corrected and filtered. Because ICNALE has a few prepositions wrongly tagged with a part of speech other than a preposition, or the other way around, these cases were manually supplemented with comments indicating that items classified as occurring with prepositions were in fact NPs without prepositions or conversely, or that items given without prepositions in fact did have a specific preposition. These comments allowed for corrections by means of a script towards the final dataset. Furthermore, irrelevant items (result IR), were discarded, leaving 5106 items for use in the statistical analysis described in Chapter 5.

Chapter 5

Quantitative analysis

The quantification of data is one of the characteristics of corpus-based research; indeed, the quantitative findings may be used to make qualitative assessments (Mahmood 2009). This chapter will present a quantitative analysis of the results obtained by searching the corpus with the annotations as outlined in the Chapter 4.

5.1   Observation counts

As mentioned before, a total of 5106 relevant observations have been selected. These contain the following information that was used for further quantitative analysis:

1. lemma

2. country

3. result (C or I)

4. type (NP or PREP)

The data cover 28 lemmas in 10 countries/regions. The distribution of these observations over regions are shown in Table 5.1 and Figure 5.1. Unfortunately, there are substantial differences between observations per region, so that further analysis will focus more on relative rather than on absolute numbers.

The distribution of observations by lemma are shown in Table 5.2 and Figure 5.2. There are substantial differences between lemma observations. The overwhelming frequency of the verb *agree* is quite expected given the argumentative character of the essays, in which participants were asked to write whether they agree or disagree with one of two statements and to provide reasons for their opinion.

Table 5.1: Observations by region

	freq
JPN	1087
CHN	820
THA	700
KOR	574
IDN	396
PHL	364
SIN	357
TWN	351
PAK	260
HKG	197
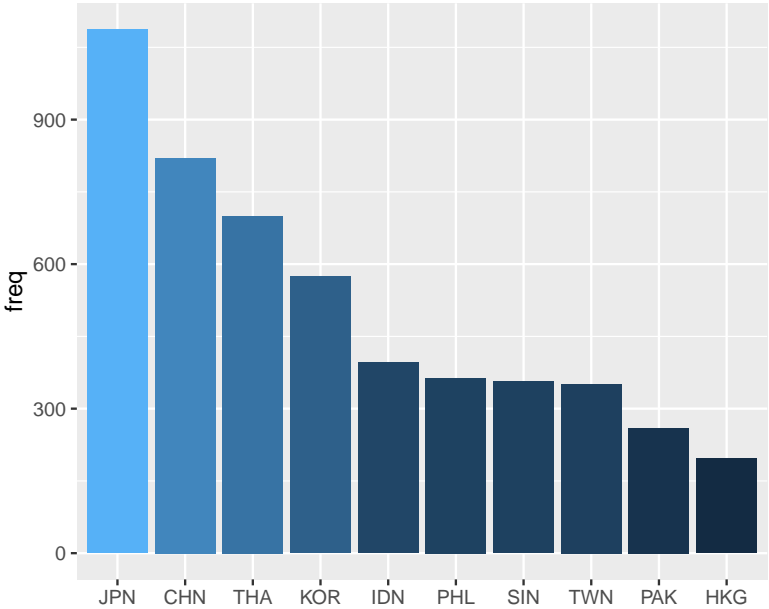
Figure 5.1: Observations per region

Table 5.2: Observations by lemma

	freq
agree	1325
think	381
depend	263
disagree	213
care	208
graduate	197
addict	192
focus	185
deal	181
look	164
suffer	154
concentrate	150
cause	148
apply	132
learn	130
relate	126
compare	100
prepare	100
belong	98
result	92
base	88
interfere	86
know	82
sit	80
worry	79
contribute	59
cope	52
benefit	41
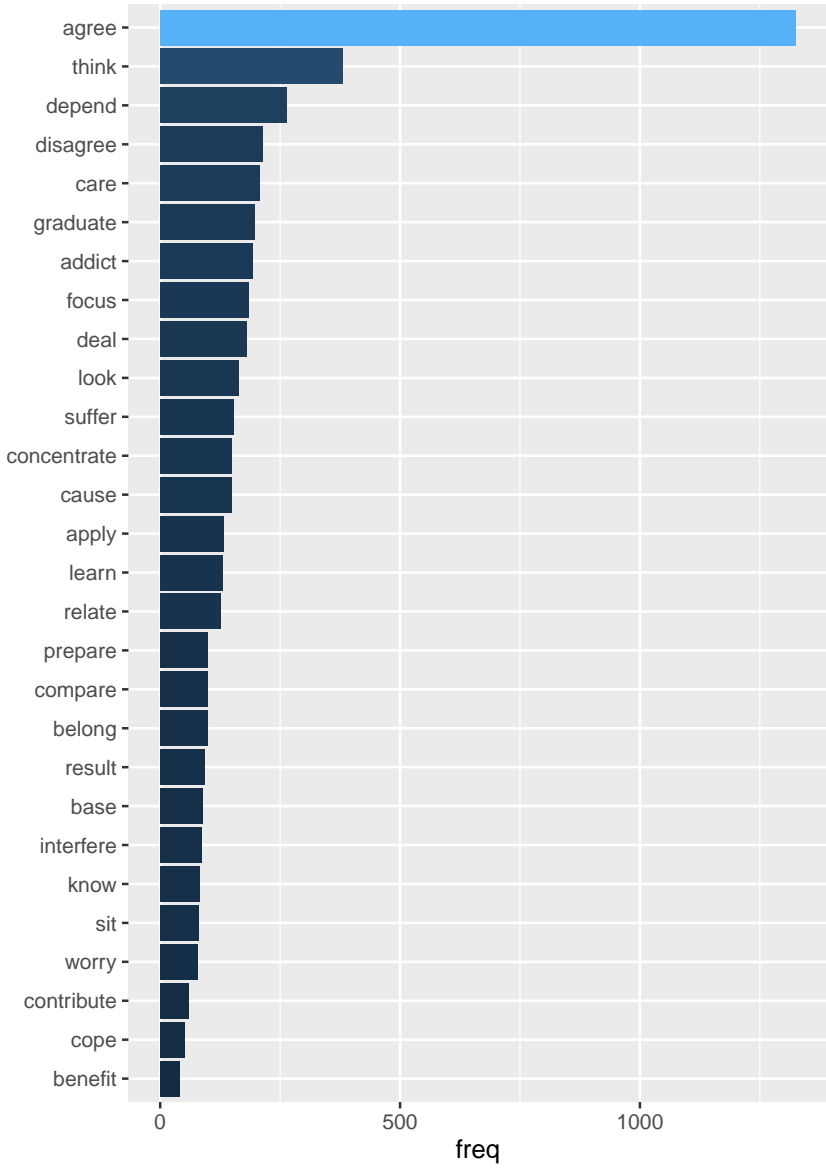
Figure 5.2: Observations per lemma

5.2   Correct and incorrect scores, grouped

An count was made of the correctness level, as percentages out of the totals. Correctness levels by lemma are shown in Figure 5.3; by region they are shown in Figure 5.4.
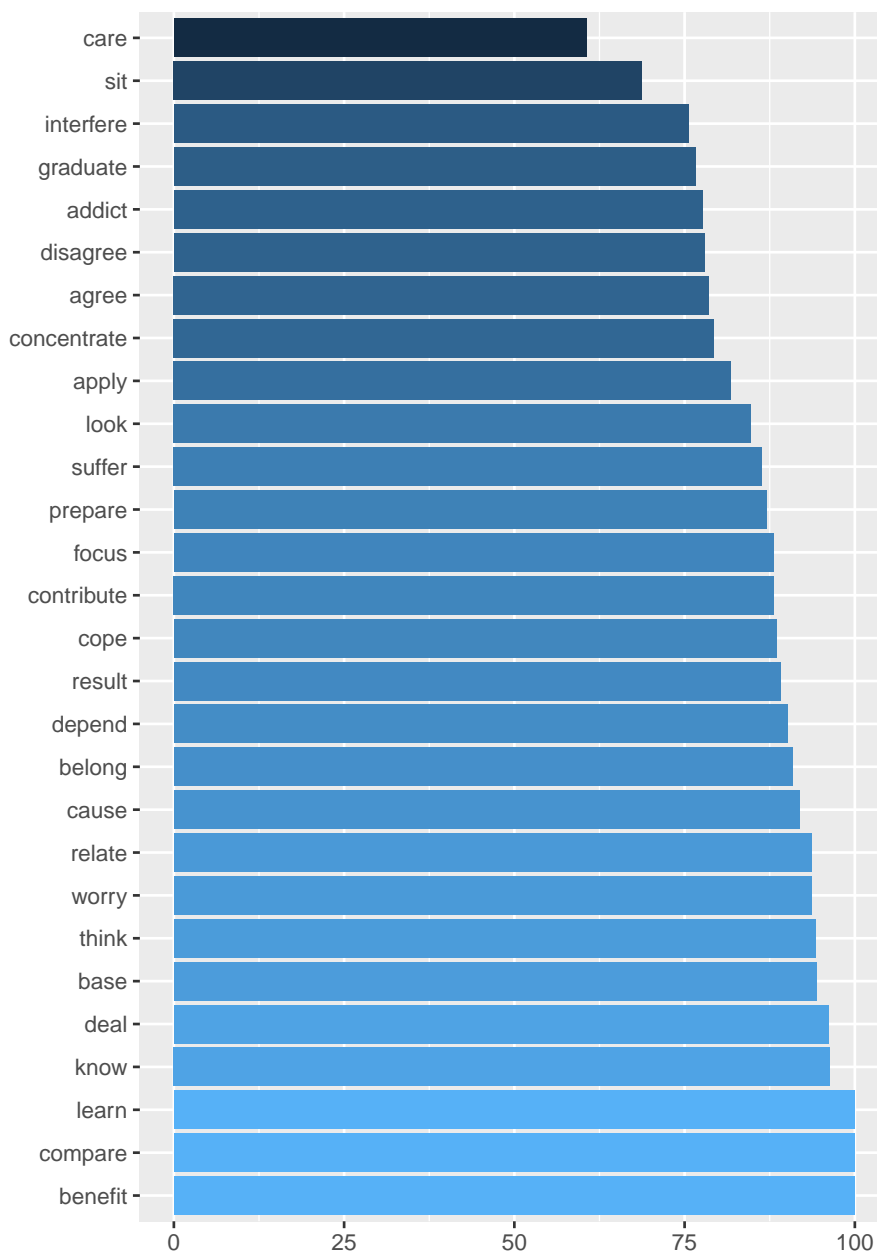
Figure 5.3: Total percentage correct by lemma

An overview of the correct and incorrect counts per country and type is given in Table 5.3, while the stacked barplot in Figure 5.5 gives percentages for the same data. For the NP condition, there are no correct items, since normal use of the transitive complement type is not relevant to the current study, so the counts for the NP condition are incorrect items only.

The data presented in Table 5.3 and Figure 5.5 suggests that PAK learners score the lowest of all regions. In order to test the hypothesis that the proportion of correct items to total number of items for all other regions together is significantly greater than that for PAK, a one-tailed
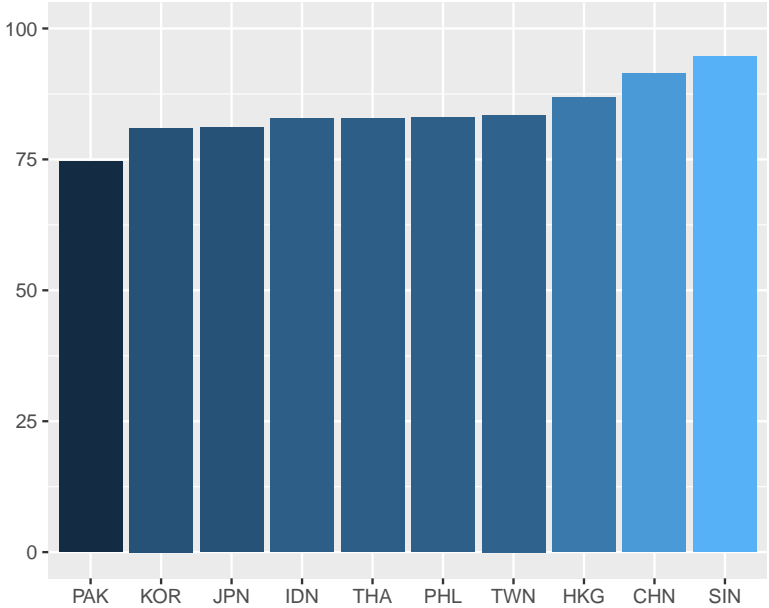
Figure 5.4: Total percentage correct by region
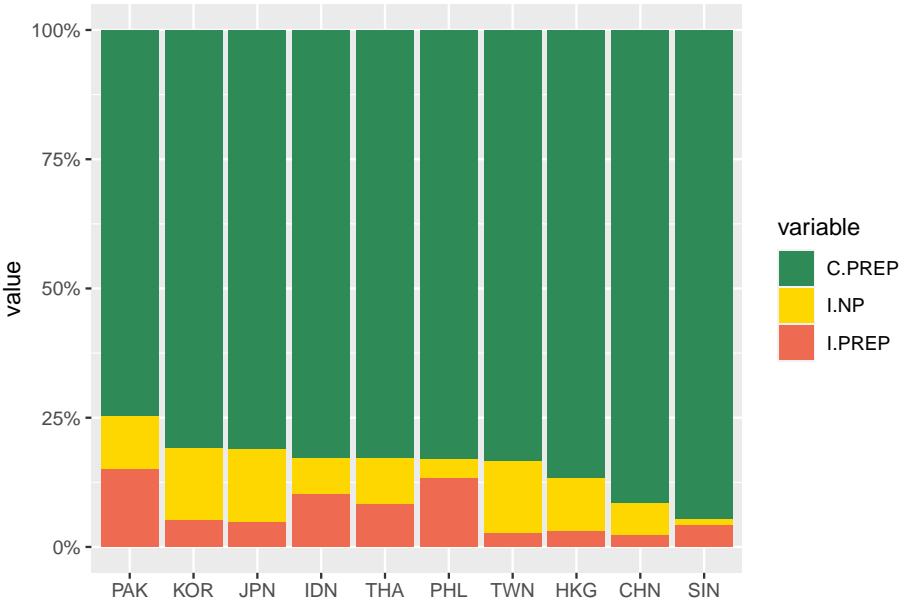
Figure 5.5: Breakdown of correct and incorrect percentages by region. C.PREP = correct use of preposition, I.PREP = incorrect use of preposition, I.NP = incorrect use of NP, i.e. missing preposition.

Table 5.3: Breakdown of counts by region. C.PREP = correct use of preposition, I.PREP = incorrect use of preposition, I.NP = incorrect use of NP, i.e. missing preposition, Correct = percentage of Total.

	country	C.PREP	I.NP	I.PREP	Total	Correct
1	PAK	194	27	39	260	74.62
2	KOR	465	80	29	574	81.01
3	JPN	881	154	52	1087	81.05
4	IDN	328	28	40	396	82.83
5	THA	580	63	57	700	82.86
6	PHL	302	14	48	364	82.97
7	TWN	293	49	9	351	83.48
8	HKG	171	20	6	197	86.80
9	CHN	750	52	18	820	91.46
10	SIN	338	4	15	357	94.68

two-proportions z-test was performed. The result of this test, as shown below, indicates that the difference is significant.

```
2-sample test for equality of proportions with continuity correction

data:  corrects out of totals
X-squared = 18.425, df = 1, p-value = 8.835e-06
alternative hypothesis: greater
95 percent confidence interval:
 0.05432928 1.00000000
sample estimates:
   prop 1    prop 2
0.8477095 0.7461538
```

5.3 Correctness by lemma and region

A more detailed analysis of correct and incorrect items by lemma and region is presented in Tables 5.4 to 5.31 and corresponding percentages as stacked barplots in Figures 5.6 to 5.33. Countries with zero observations for a lemma are not shown in the plots. Note that there is considerable variation in the counts by lemma. The next chapter will discuss some relevant findings.

Table 5.4: Counts for *addict*, 192

	C.PREP	I.NP	I.PREP
CHN	68	0	4
HKG	3	0	0
IDN	10	0	6
JPN	4	0	0
KOR	6	0	0
PAK	15	6	12
PHL	15	0	4
SIN	6	0	0
THA	9	9	2
TWN	13	0	0
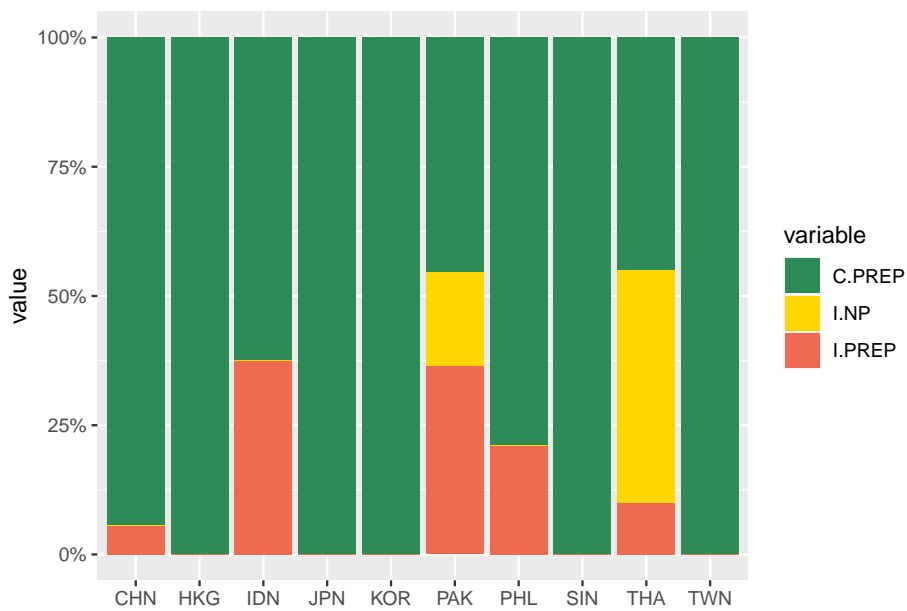
Figure 5.6: Percentages for *addict*

Table 5.5: Counts for *agree*, 1325

	C.PREP	I.NP	I.PREP
CHN	129	15	5
HKG	19	6	0
IDN	67	5	4
JPN	414	90	34
KOR	91	38	18
PAK	9	1	3
PHL	14	1	3
SIN	36	0	8
THA	207	11	15
TWN	55	24	3
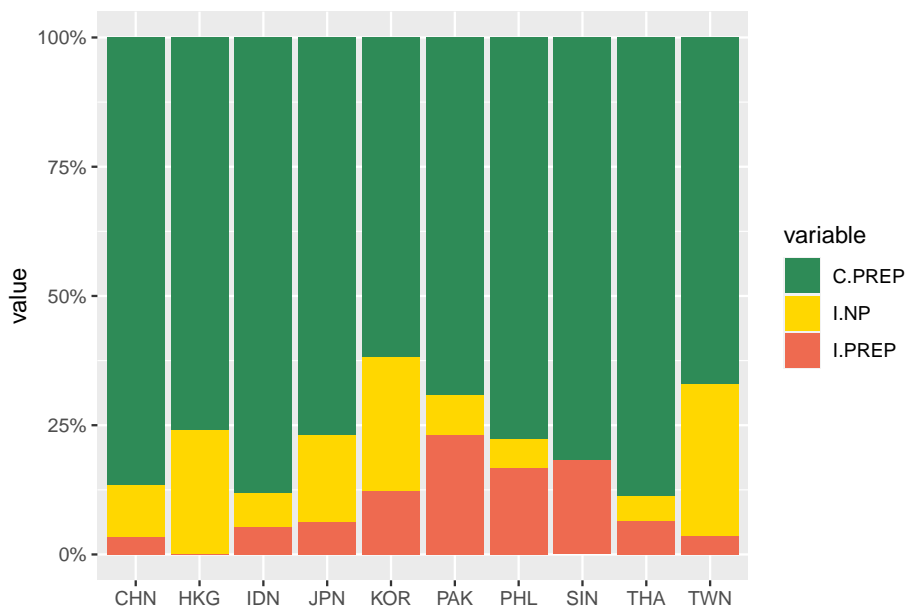
Figure 5.7: Percentages for *agree*

Table 5.6: Counts for *apply*, 132

	C.PREP	I.NP	I.PREP
CHN	9	0	0
HKG	9	4	0
IDN	7	2	0
JPN	9	0	0
KOR	6	1	0
PAK	1	0	2
PHL	37	2	4
SIN	8	0	1
THA	19	1	5
TWN	3	2	0
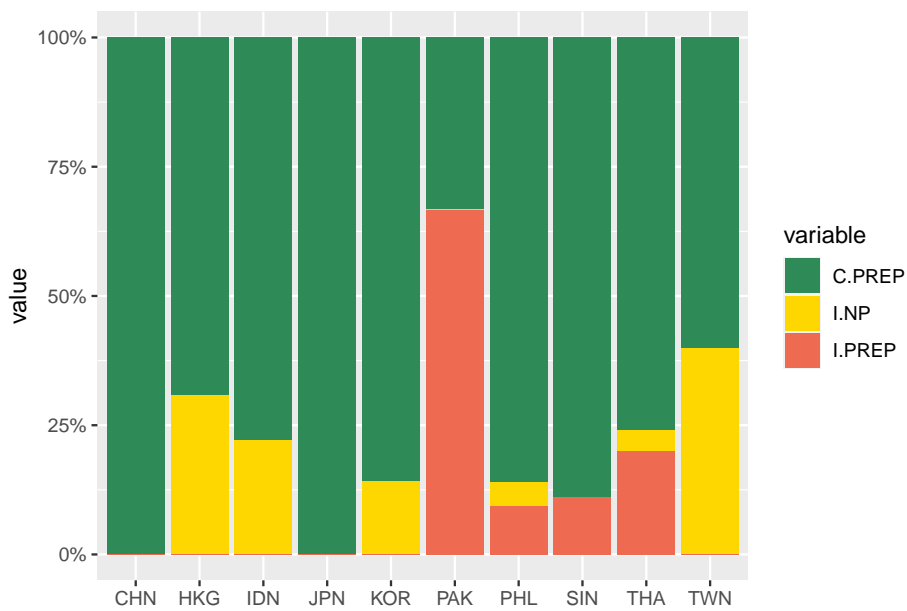
Figure 5.8: Percentages for *apply*

Table 5.7: Counts for *base*, 88

	C.PREP	I.NP	I.PREP
CHN	12	0	0
HKG	4	1	0
IDN	17	0	0
JPN	2	1	0
KOR	8	0	0
PAK	2	1	0
PHL	9	0	2
SIN	11	0	0
THA	4	0	0
TWN	14	0	0
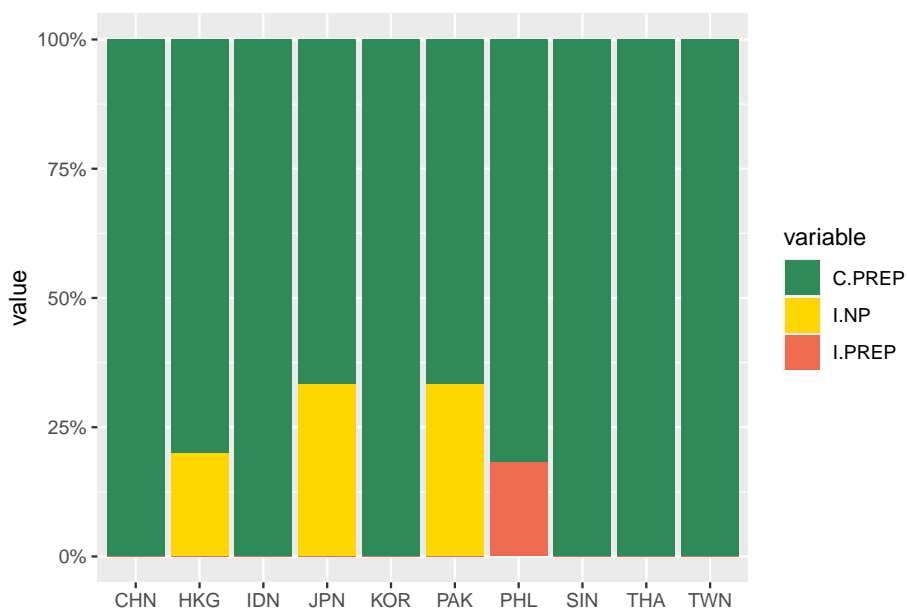
Figure 5.9: Percentages for *base*

Table 5.8: Counts for *belong*, 98

	C.PREP	I.NP	I.PREP
CHN	17	0	0
HKG	1	0	0
IDN	1	0	0
JPN	29	2	0
KOR	1	1	0
PAK	25	2	1
PHL	8	0	1
SIN	3	0	0
THA	3	1	1
TWN	1	0	0
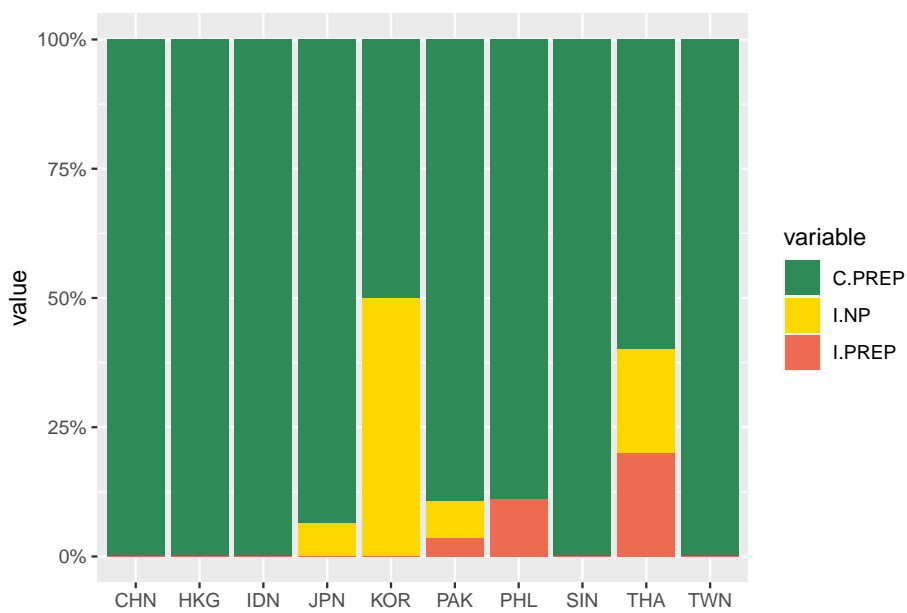
Figure 5.10: Percentages for *belong*

Table 5.9: Counts for *benefit*, 41

	C.PREP
CHN	18
HKG	5
KOR	1
PAK	1
PHL	3
SIN	6
THA	5
TWN	2
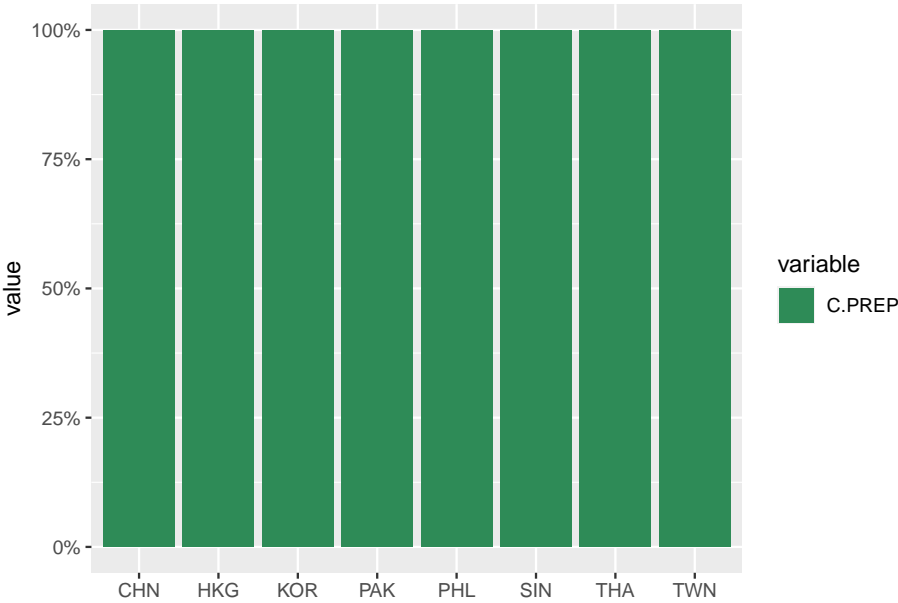
Figure 5.11: Percentages for *benefit*

Table 5.10: Counts for *care*, 208

	C.PREP	I.NP	I.PREP
CHN	27	12	0
HKG	3	2	0
IDN	6	4	3
JPN	17	3	1
KOR	14	9	1
PAK	9	6	2
PHL	10	0	0
SIN	3	0	0
THA	22	28	2
TWN	15	9	0
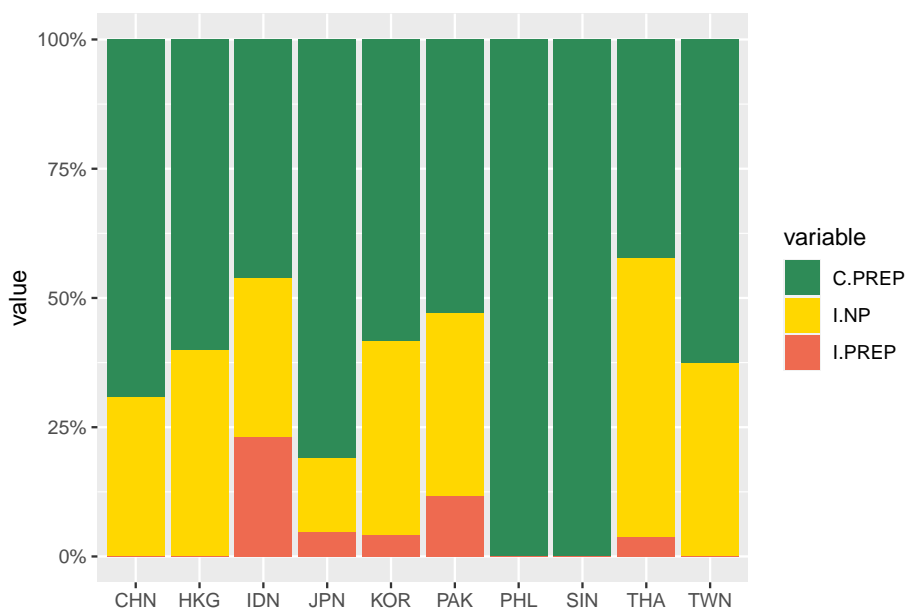
Figure 5.12: Percentages for *care*

Table 5.11: Counts for *cause*, 148

	C.PREP	I.PREP
CHN	32	1
HKG	11	0
IDN	19	0
JPN	9	0
KOR	16	0
PHL	9	5
SIN	5	0
THA	20	5
TWN	15	1
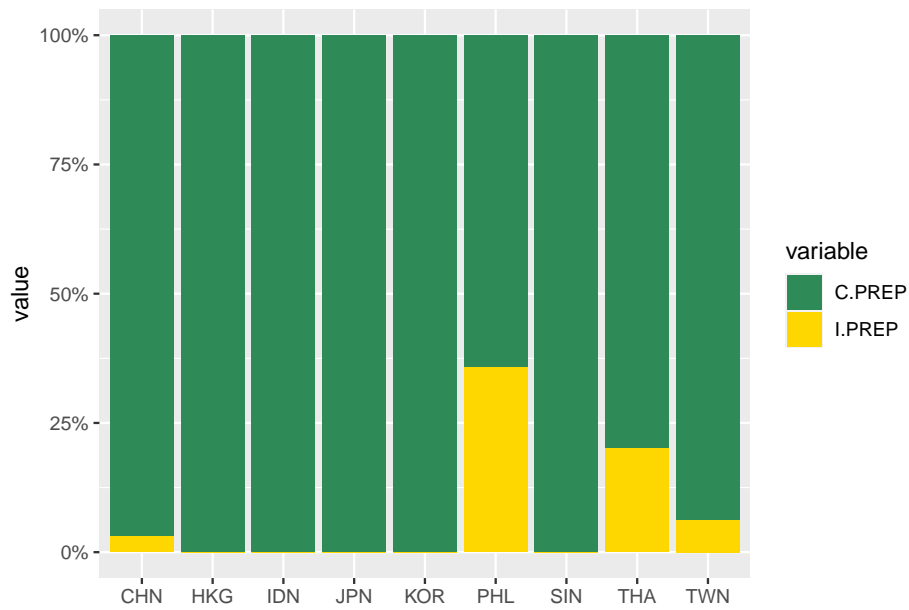
Figure 5.13: Percentages for *cause*

Table 5.12: Counts for *compare*, 100

	C.PREP
CHN	16
HKG	5
IDN	3
JPN	8
KOR	19
PAK	10
PHL	9
SIN	21
THA	5
TWN	4
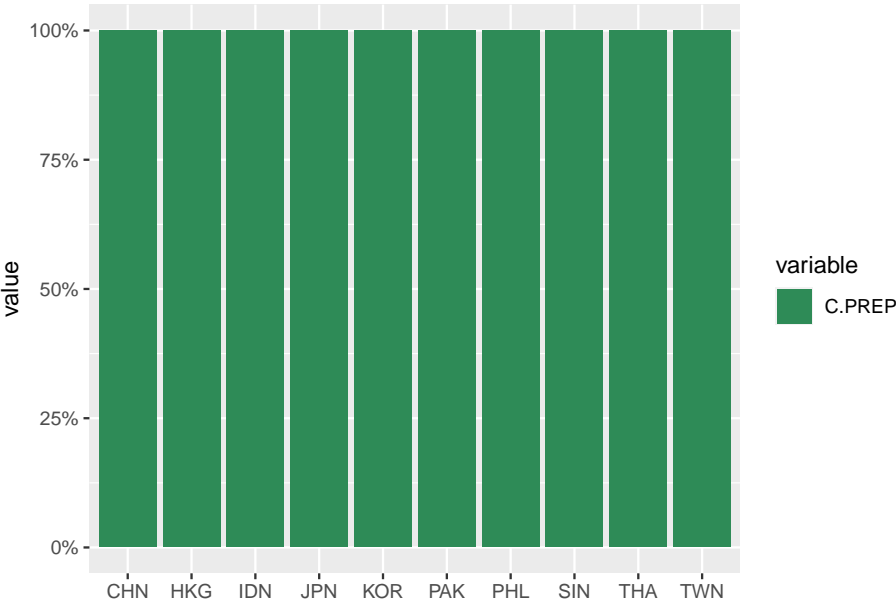
Figure 5.14: Percentages for *compare*

Table 5.13: Counts for *concentrate*, 150

	C.PREP	I.NP	I.PREP
CHN	14	0	0
HKG	12	1	2
IDN	3	3	2
JPN	11	1	4
KOR	30	5	3
PAK	15	0	2
PHL	6	0	2
SIN	12	0	0
THA	5	0	6
TWN	11	0	0
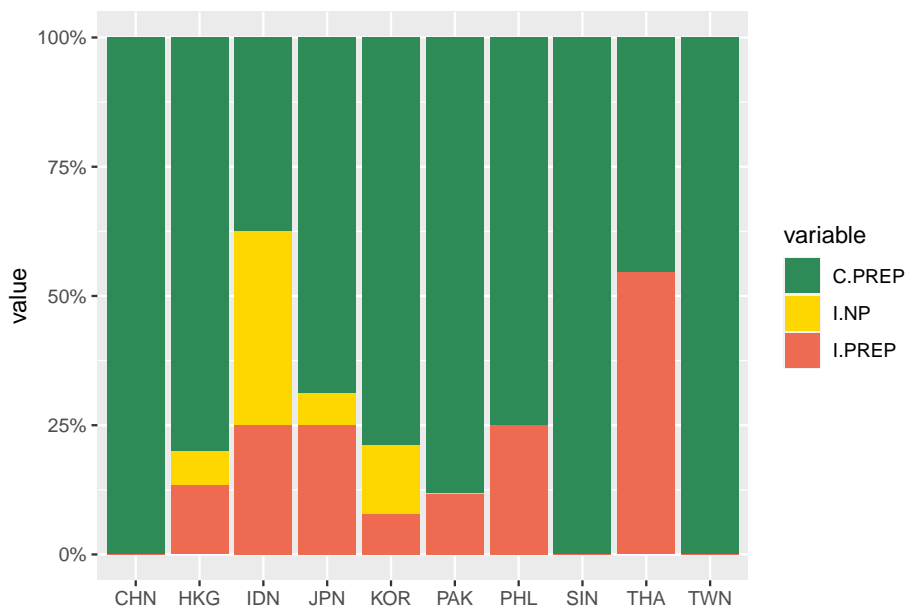
Figure 5.15: Percentages for *concentrate*

Table 5.14: Counts for *contribute*, 59

	C.PREP	I.NP	I.PREP
CHN	19	2	0
HKG	2	0	0
IDN	4	1	0
JPN	4	0	0
KOR	2	0	0
PAK	4	0	3
PHL	8	1	0
SIN	8	0	0
THA	1	0	0
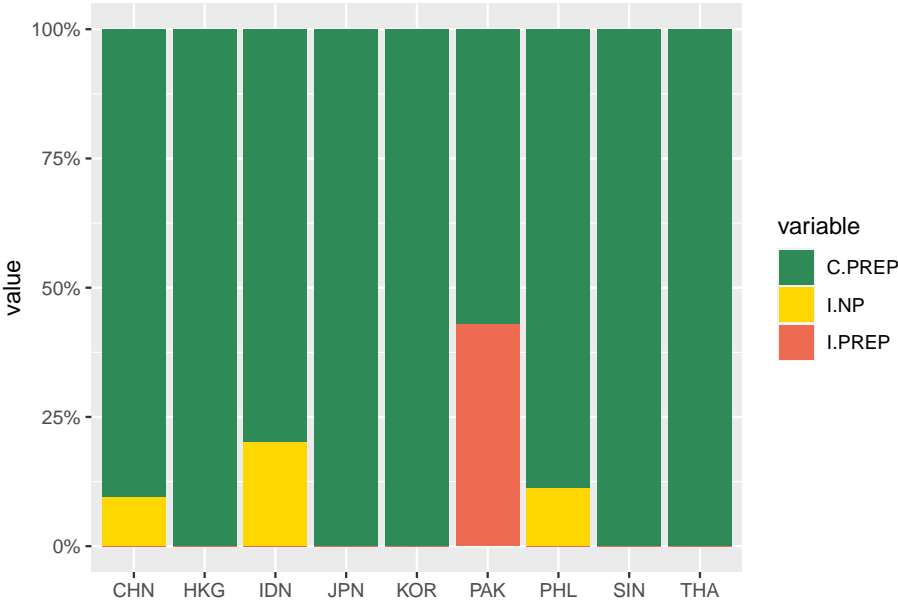
Figure 5.16: Percentages for *contribute*

Table 5.15: Counts for *cope*, 52

	C.PREP	I.NP	I.PREP
CHN	4	0	0
HKG	3	0	0
JPN	6	0	0
KOR	3	1	0
PAK	5	0	0
PHL	2	0	4
SIN	19	0	0
THA	2	1	0
TWN	2	0	0
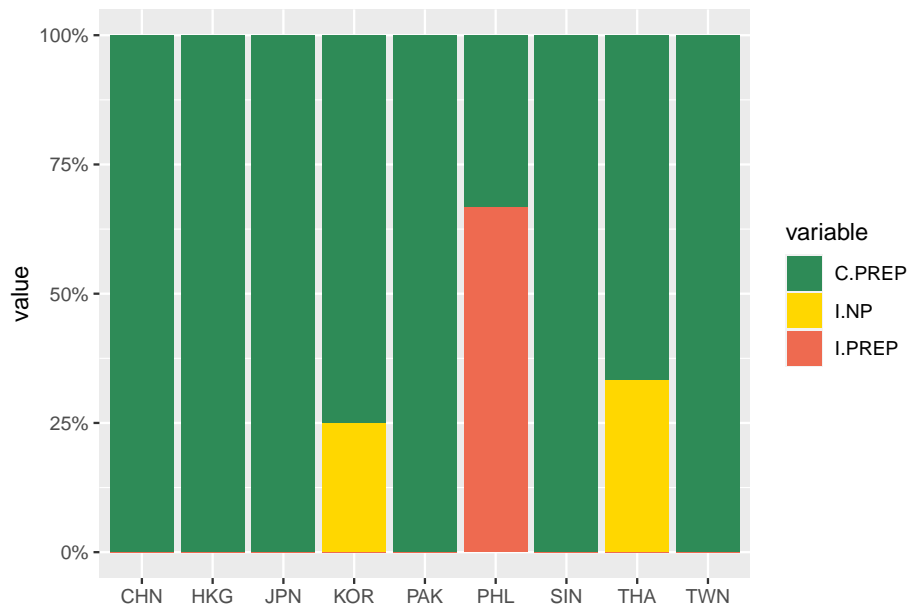
Figure 5.17: Percentages for *cope*

Table 5.16: Counts for *deal*, 181

	C.PREP	I.NP	I.PREP
CHN	61	1	1
HKG	7	0	0
IDN	4	1	0
JPN	15	0	0
KOR	16	1	0
PAK	7	0	0
PHL	16	0	1
SIN	8	1	0
THA	11	0	0
TWN	29	1	0
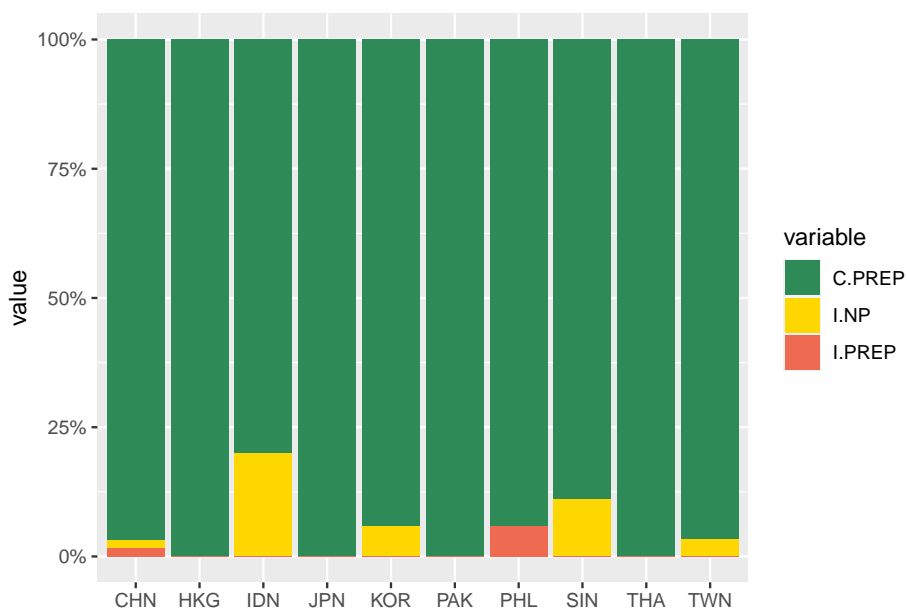
Figure 5.18: Percentages for *deal*

Table 5.17: Counts for *depend*, 263

	C.PREP	I.NP	I.PREP
CHN	45	0	2
HKG	7	0	0
IDN	20	0	11
JPN	47	0	0
KOR	13	0	2
PAK	18	1	0
PHL	26	0	9
SIN	26	0	0
THA	26	0	1
TWN	9	0	0
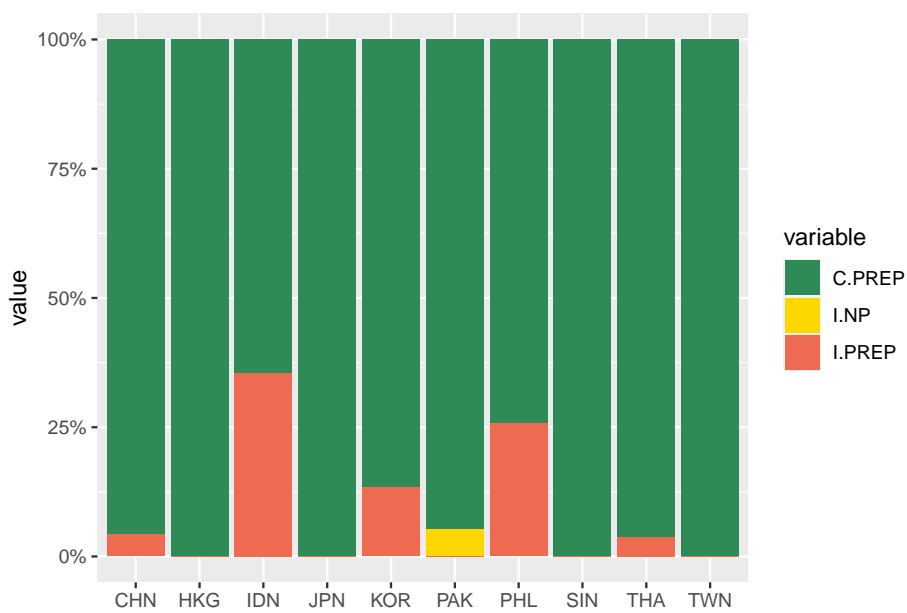
Figure 5.19: Percentages for *depend*

Table 5.18: Counts for *disagree*, 213

	C.PREP	I.NP	I.PREP
CHN	17	4	1
HKG	7	1	1
IDN	21	1	1
JPN	68	4	7
KOR	14	4	5
PAK	1	0	2
PHL	4	1	2
SIN	13	0	4
THA	10	0	3
TWN	11	4	2
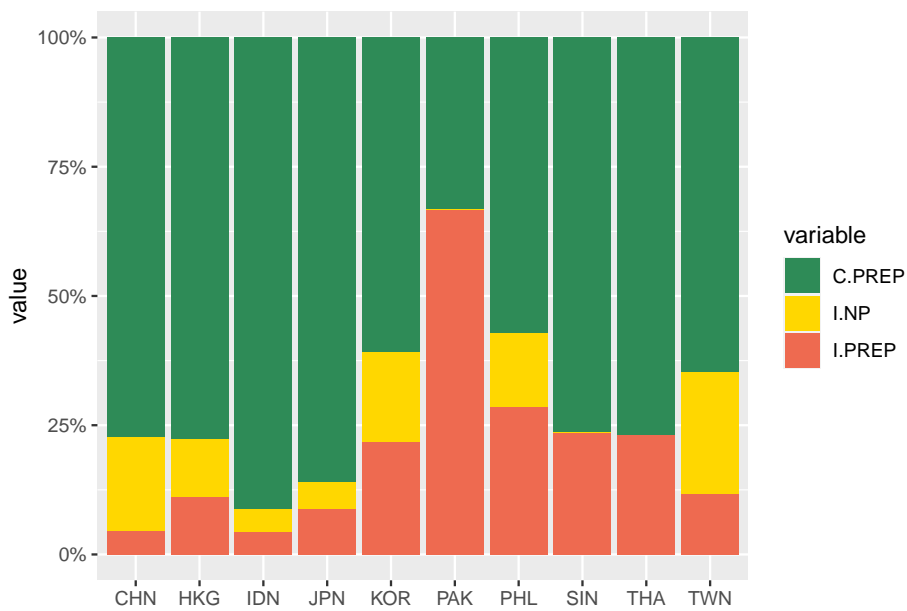
Figure 5.20: Percentages for *disagree*

Table 5.19: Counts for *focus*, 185

	C.PREP	I.NP	I.PREP
CHN	29	0	0
HKG	11	0	0
IDN	15	1	9
JPN	2	0	0
KOR	15	0	0
PAK	7	1	0
PHL	28	2	2
SIN	34	1	1
THA	15	2	2
TWN	7	1	0
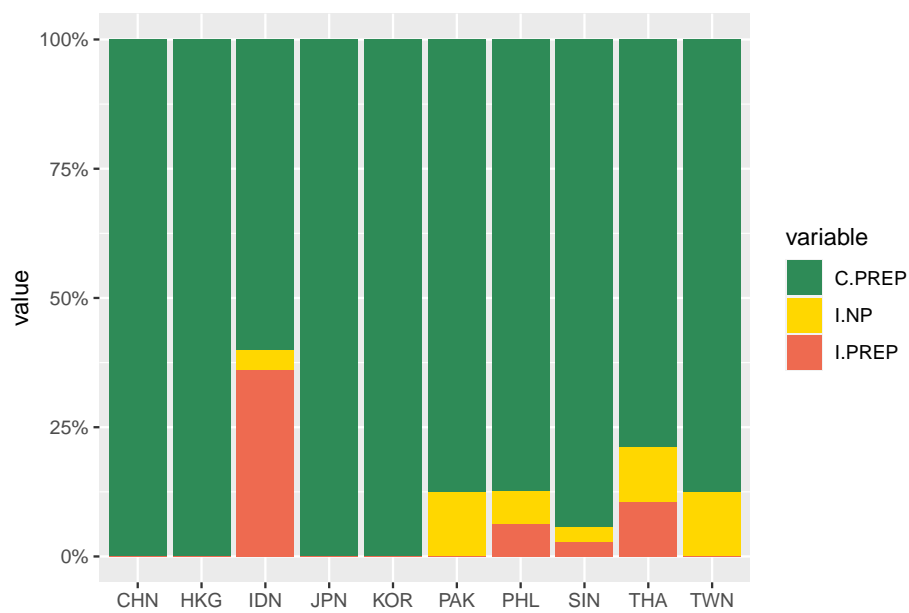
Figure 5.21: Percentages for *focus*

Table 5.20: Counts for *graduate*, 197

	C.PREP	I.NP	I.PREP
CHN	32	2	1
HKG	1	0	1
IDN	18	0	1
JPN	46	25	0
KOR	17	9	0
PAK	1	0	0
PHL	2	4	1
SIN	6	0	0
THA	11	1	1
TWN	17	0	0
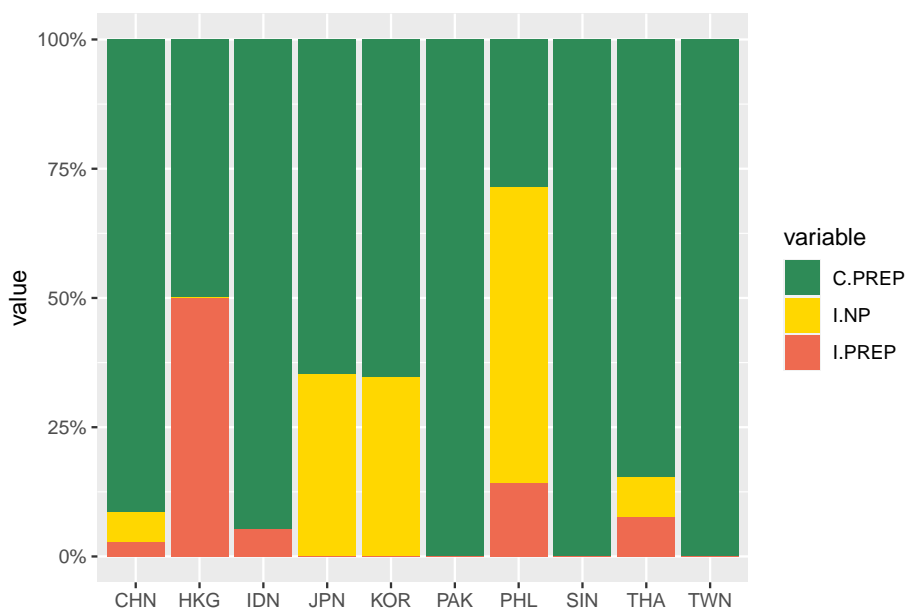
Figure 5.22: Percentages for *graduate*

Table 5.21: Counts for *interfere*, 86

	C.PREP	I.NP	I.PREP
CHN	5	3	0
HKG	0	1	0
IDN	18	1	0
JPN	5	1	0
KOR	8	2	0
PAK	4	6	0
PHL	3	1	0
SIN	7	0	0
THA	14	0	2
TWN	1	4	0
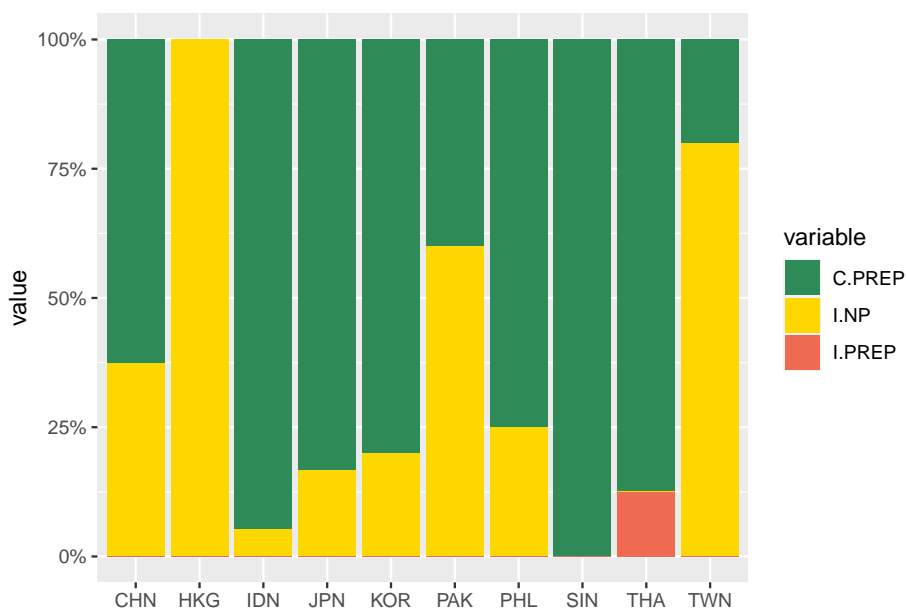
Figure 5.23: Percentages for *interfere*

Table 5.22: Counts for *know*, 82

	C.PREP	I.PREP
CHN	7	0
IDN	12	0
JPN	14	2
KOR	12	0
PAK	17	0
PHL	2	0
SIN	1	0
THA	11	1
TWN	3	0
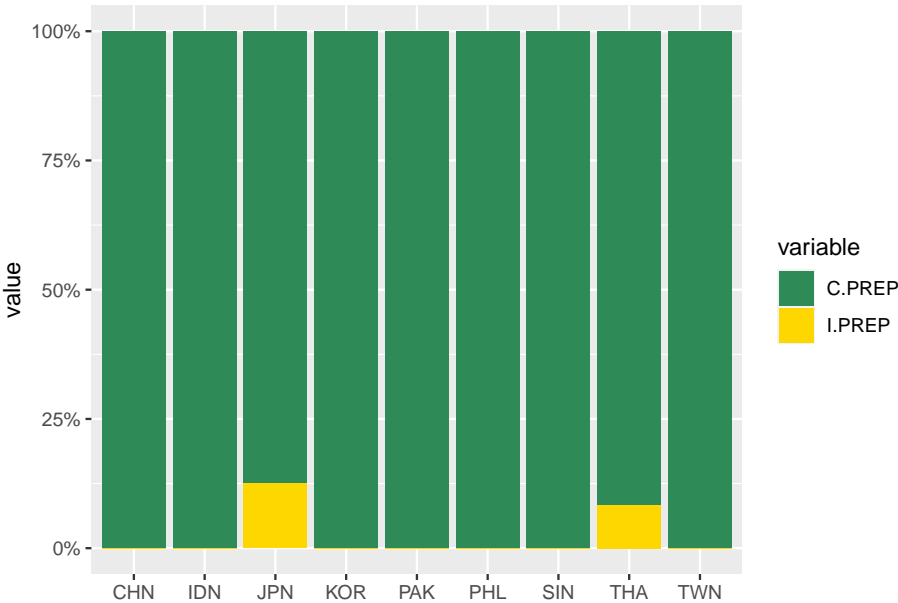
Figure 5.24: Percentages for *know*

Table 5.23: Counts for *learn*, 130

	C.PREP
CHN	26
HKG	11
IDN	9
JPN	16
KOR	14
PAK	3
PHL	5
SIN	11
THA	26
TWN	9
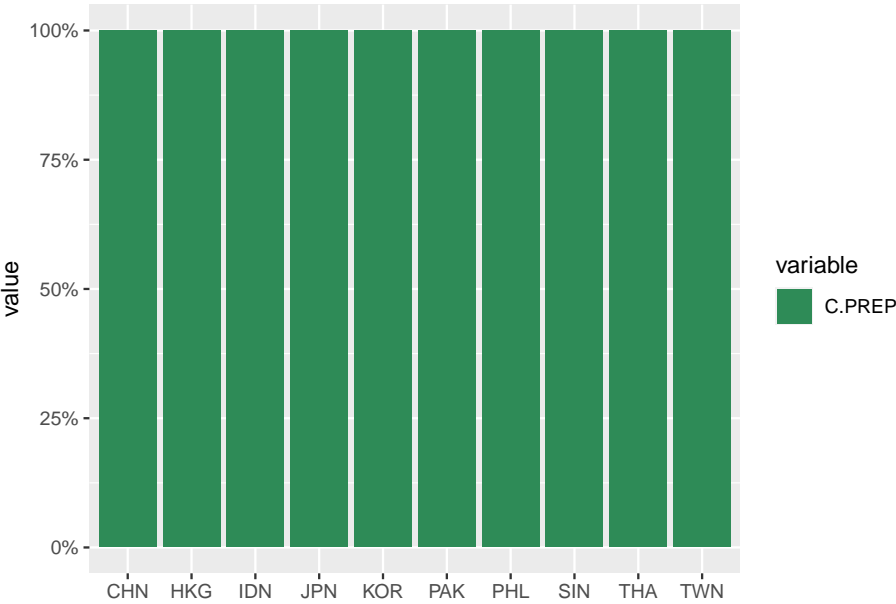
Figure 5.25: Percentages for *learn*

Table 5.24: Counts for *look*, 164

	C.PREP	I.NP	I.PREP
CHN	20	5	1
HKG	5	0	1
IDN	19	3	1
JPN	10	3	1
KOR	15	1	0
PAK	4	1	1
PHL	13	1	0
SIN	18	0	0
THA	31	5	0
TWN	4	1	0
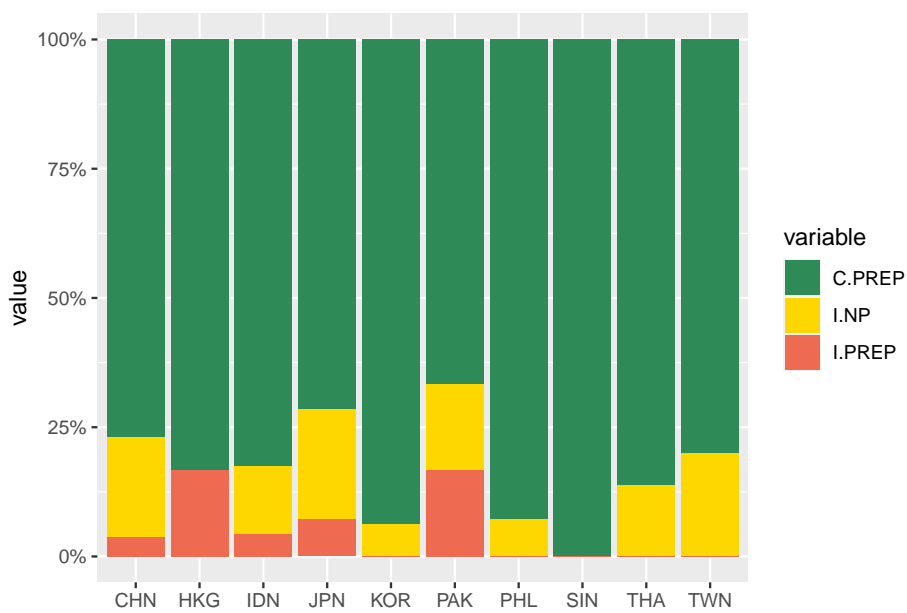
Figure 5.26: Percentages for *look*

Table 5.25: Counts for *prepare*, 100

	C.PREP	I.NP
CHN	19	1
HKG	5	2
IDN	0	3
JPN	18	2
KOR	17	1
PAK	1	1
PHL	3	1
SIN	12	0
THA	5	0
TWN	7	2
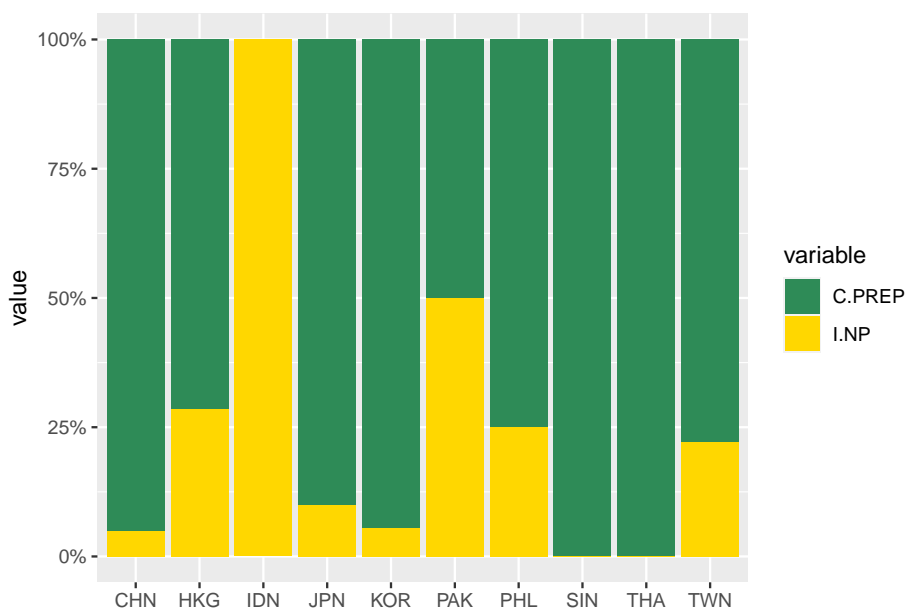
Figure 5.27: Percentages for *prepare*

Table 5.26: Counts for *relate*, 126

	C.PREP	I.NP	I.PREP
CHN	16	0	0
HKG	12	0	0
IDN	18	0	0
JPN	6	4	0
KOR	22	2	0
PAK	6	0	0
PHL	11	0	0
SIN	6	0	0
THA	13	1	1
TWN	8	0	0
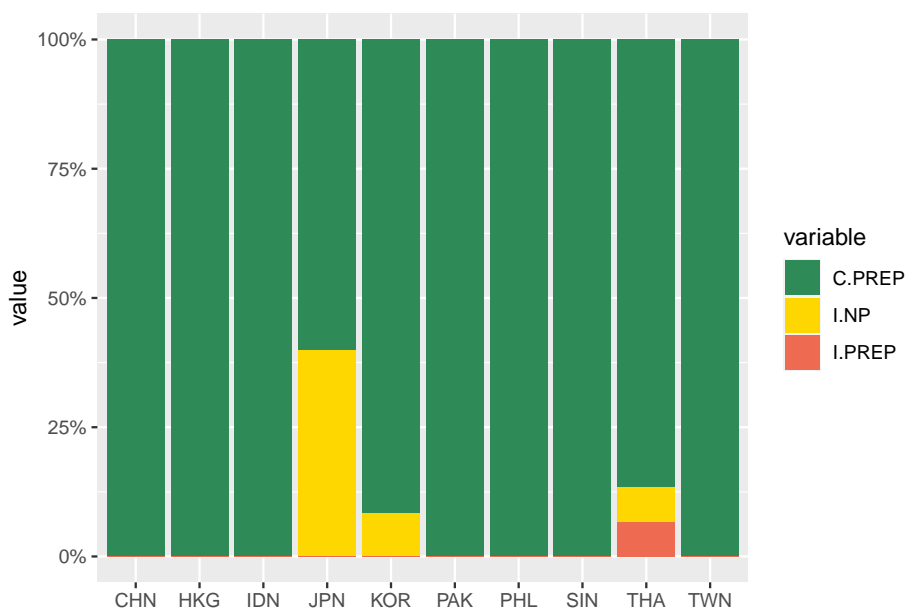
Figure 5.28: Percentages for *relate*

Table 5.27: Counts for *result*, 92

	C.PREP	I.NP	I.PREP
CHN	13	1	0
HKG	3	0	0
IDN	3	0	0
KOR	4	0	0
PAK	3	1	1
PHL	6	0	3
SIN	25	2	0
THA	16	2	0
TWN	9	0	0
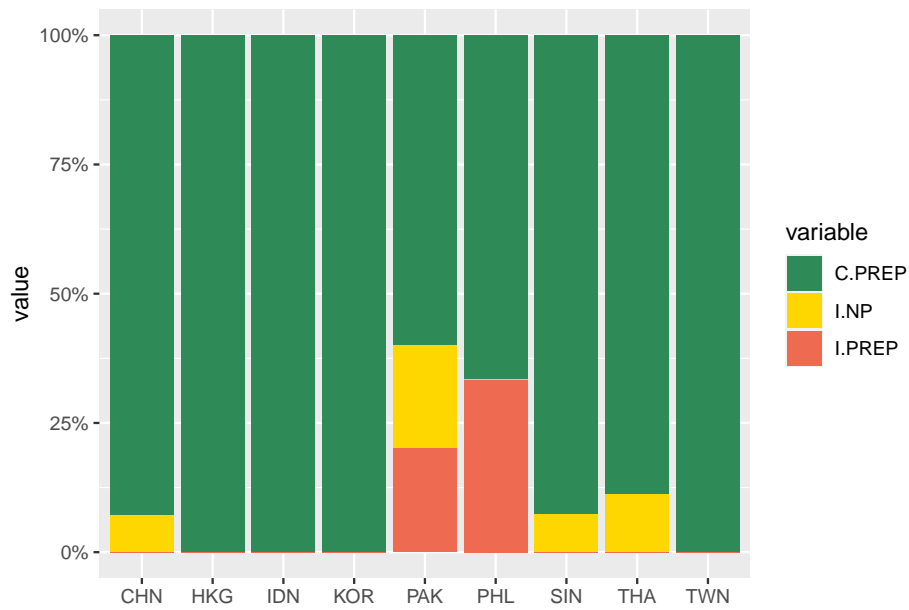
Figure 5.29: Percentages for *result*

Table 5.28: Counts for *sit*, 80

	C.PREP	I.NP	I.PREP
CHN	10	0	1
HKG	1	0	0
IDN	4	0	0
JPN	17	17	1
KOR	2	4	0
PAK	4	0	1
PHL	4	0	0
SIN	5	0	0
THA	7	1	0
TWN	1	0	0
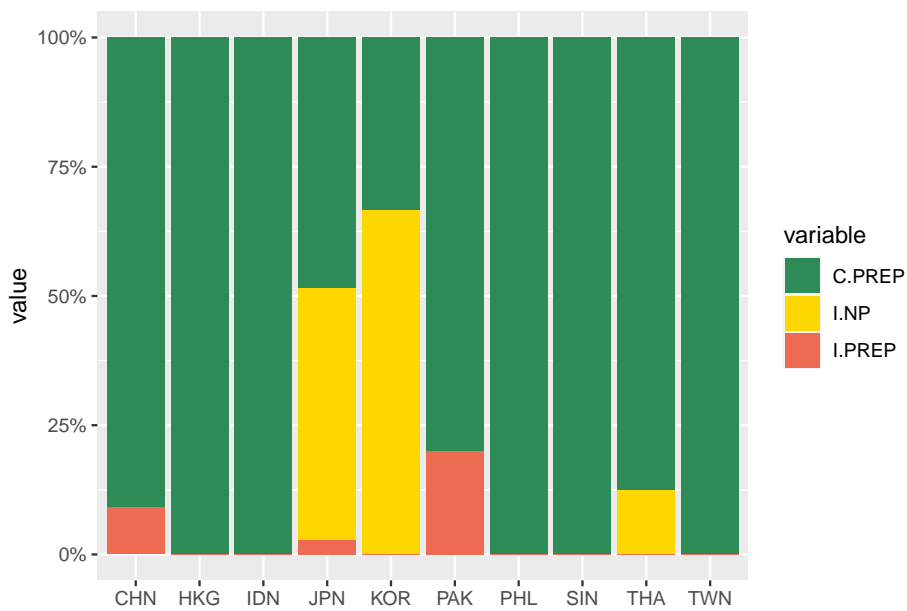
Figure 5.30: Percentages for *sit*

Table 5.29: Counts for *suffer*, 154

	C.PREP	I.NP	I.PREP
CHN	21	6	0
HKG	13	2	1
IDN	4	3	0
JPN	29	0	0
KOR	16	0	0
PAK	11	0	7
PHL	12	0	2
SIN	15	0	0
THA	6	0	0
TWN	6	0	0
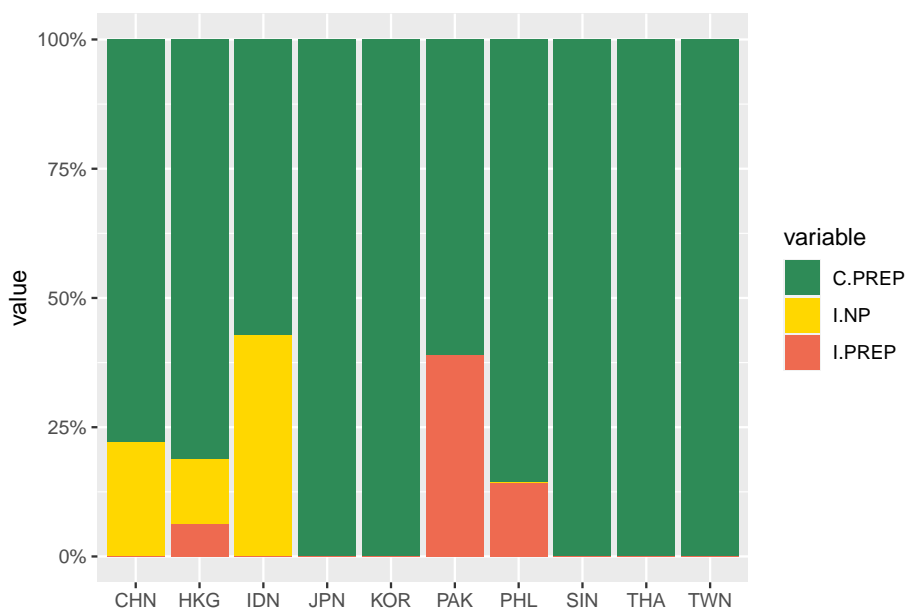
Figure 5.31: Percentages for *suffer*

Table 5.30: Counts for *think*, 381

	C.PREP	I.PREP
CHN	53	1
HKG	9	0
IDN	23	1
JPN	62	2
KOR	71	0
PAK	8	2
PHL	32	3
SIN	7	1
THA	70	9
TWN	24	3
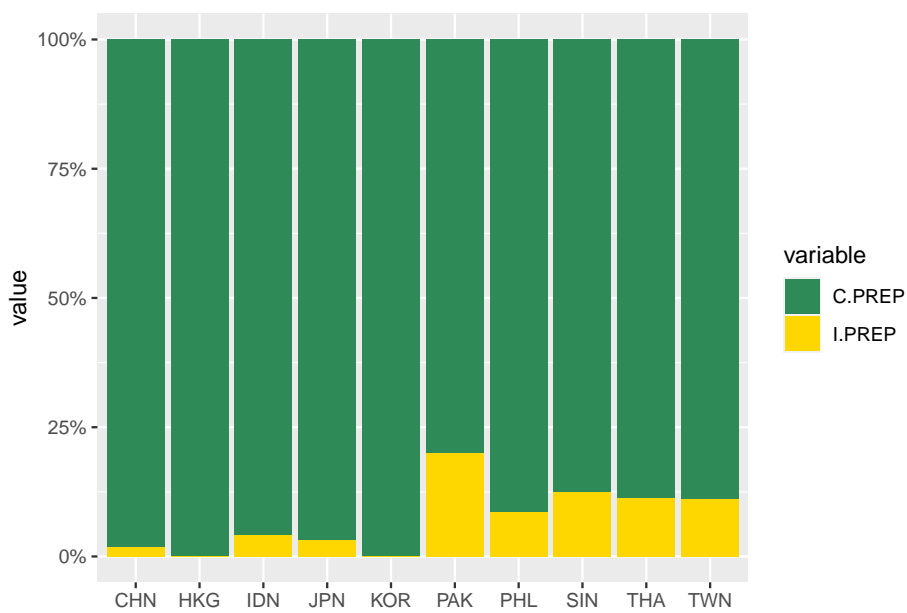
Figure 5.32: Percentages for *think*

Table 5.31: Counts for *worry*, 79

	C.PREP	I.NP	I.PREP
CHN	11	0	0
HKG	2	0	0
IDN	3	0	1
JPN	13	1	0
KOR	12	1	0
PAK	3	0	0
PHL	5	0	0
SIN	6	0	0
THA	6	0	1
TWN	13	1	0
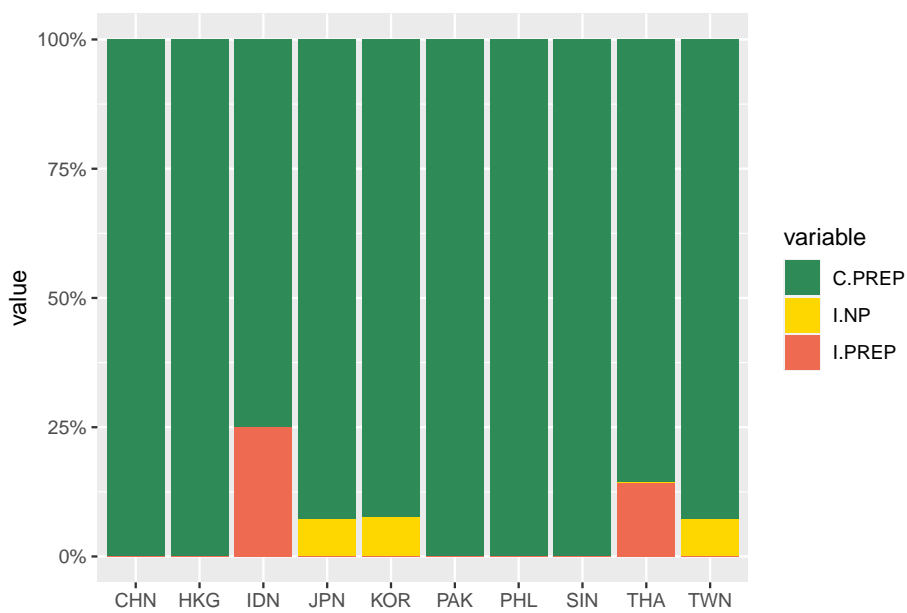
Figure 5.33: Percentages for *worry*

Chapter 6

Discussion

6.1   Possible causes for interlanguage errors

Interlanguage errors can result from learners' attempts to map the structures and rules of their native language onto the target language (in this case, English). Errors like inappropriate prepositional use are typical throughout the interlanguage period and L1 interference is a common cause of such errors (Dulay et al. 1982; Selinker 1972). Knowledge and use of the L1 may lead to interference at many levels, including morphology, structure, grammar, syntax, and prepositional and lexical interference (Dini 2020). L1 interference in SLA is also discussed by Iqbal et al. (2019), who argues that L1 leaves a lasting impression on a person's neurological and psychological structure; hence, learning further languages is influenced by this imprint and the learner becomes confused when certain first language components are present in the second language. Furthermore, Siyanova and Schmitt (2007) considered another factor responsible for errors in using multi word verbs. They found that it might take a while for learners, whose first language is not of Germanic origin, to get over their unease with unfamiliar multi-word verbs. Pakistani ESL learners have Urdu as their native language which is of Non-Germanic origin which may be a cause of challenges faced in using prepositional verbs. As rightly discussed by Islami (2015), Lorincz and Gordon (2012) and others, learning English prepositions may be difficult owing to their distinctive usage, frequency and contextual meaning, which vary greatly from language to language. This may lead to wrong preposition choices in learners' interlingua. Furthermore, learners' proficiency level, L1–L2 structural differences, and semantic complexity of the L2 structures, all interact to cause potential avoidance behavior in second language acquisition (Karim and Shahwar 2015). Previous research shows that Language competency, language use, and learning context all play a role in the transfer of L1 knowledge to the target language. Ringbom (1987) and Jarvis and Pavlenko (2008), for example, found that learners' proficiency in the target language influences the extent and manner of transfer from the L1. Similarly, Cook (2001) and Lyster and Ranta (1997) found that transfer varies depending on the context of language usage, such as whether the learner is communicating or studying in the target language. How these factors may play a role in the corpus observations in the present study will be further explored in this chapter.

6.2   Lemmas with a high error rate

The observations that were the subject of quantitative analysis in Chapter 5, in particular, selected analyses of lemma observations, cf. Section 5.3, will be further discussed here. Since the present study focuses on the prepositional verbs and interlanguage errors by PAK learners, the

emphasis will be on trying to provide explanations for PAK errors, often suggesting processes such as transfer and avoidance, although clearly not all errors will have an obvious explanation. Thus, the lemmas for which PAK errors are noteworthy will be discussed, in alphabetical order.

6.2.1 Addict

The verb *addict* has a higher percentage of errors by PAK than other learners in the PREP condition. Example (6.1) from the corpus shows incorrect use of the preposition *of* instead of *to*. Example (6.2) from the corpus shows an omitted preposition for the same verb, i.e. an error in the NP condition.

(6.1) The person who is addicted of smoking [...]

(6.2) [...] a person who addict smoking [...]

(6.3) وہ منشیات کا عادی ہے
Wo manshiat ka aadi hai
He drugs of addicted is
'He is addicted to drugs.'

The Urdu preposition with the corresponding verb would be کا *ka* which is associated with a possessive meaning, corresponding to English 'of", as in example (6.3). Thus, in (6.1) it is likely that some transfer has been at play here, in which the PAK learner chose the English preposition most closely corresponding to a possessive. As suggested by Saed and Yassin (2017), ESL learners use incorrect prepositions, but also delete necessary prepositions, and use additional (unnecessary) prepositions. Similarly, Siyanova and Schmitt (2007) argued that several studies have revealed that learners avoid multi-word verb constructions and struggle when they do try to use them, since L2 learners prefer to avoid those linguistic categories that do not exist in their L1. So, a possible explanation for omitting the preposition, as in (6.2), is that the learner is unsure of which preposition to choose, and therefore avoids using any preposition at all.

6.2.2 Agree

The verb *agree* has more errors by PAK than other learners in the PREP condition. Example (6.4) from the corpus shows incorrect use of the preposition *in*, while (6.5) shows incorrect use of *to*.

(6.4) I agree in that smoking should be completely banned [...]

(6.5) [...] we cannot agree to everyone to stop smoking [...]

(6.6) میں تم سے متفق ہوں
mein tum se muttafiq hun
I you from agree am
'I agree with you.'

As suggested by Siyanova and Schmitt (2007), polysemous multi-word verbs maybe the cause of errors. A possible explanation for using the wrong preposition, as in (6.4), is that the prepositions are ambiguous in nature (Saed and Yassin 2017) and the learner is confused about the difference in the use of 'with' and 'to' with the verb agree due to a lack of grammatical knowledge. L2 acquisition is influenced by the linguistic experience of the mother tongue. L2

speakers view the language at hand as a tool to assist them in resolving any issues that may come up throughout the learning process (Yadav 2014). In Urdu, the preposition in the corresponding sentence would be سے *se*, closely corresponding to English 'from', as in example (6.6). Keeping their L1 sentence form in mind, the PAK student most likely opted not to eliminate the English preposition. These types of errors are identified by Alhamadi (2015) as transfer errors, or mother-tongue interference, in which a student uses their first language skill to acquire a second language. This issue arises when familiar patterns are preferred over unfamiliar ones, leading to learning issues and errors. The influence of the first language can either positively or negatively affect the second language, but it leads to error in the above-mentioned scenario.

6.2.3   Apply

The verb *apply* has a higher percentage of errors by PAK than other learners in the PREP condition, but the counts are very low, so they need to be taken with a grain of salt. Example (6.7) from the corpus shows and unwanted use of the preposition *on* whereas the verb in this case seems to be an ordinary transitive verb expecting an NP as direct object. Example (6.8) shows incorrect use of the preposition *on* instead of *to*. This error could be explained through transfer and non-transfer influences. The latter may involve interference from the English constructions *levy/impose taxes on [something]*. But an effect of transfer cannot be excluded either, since Urdu uses پر *par* 'on' in similar contexts, as illustrated in (6.9). Pakistani learners of English struggle with the use of prepositional verbs and other construction types due to significant interference from L1 grammar, phonology, and vocabulary when utilizing the foreign structure of the second language (Masood et al. 2020).

(6.7) Police and other should make people to apply on this rule [...]

(6.8) Taxes should be applied on tobacco and its products [...]

(6.9) تمباکو پر ٹیکس لگانا چاہیے۔
       tambaacoo par tax lagana chahye
       Tobacco on tax apply should
       'Tax should be applied to tobacco.'

6.2.4   Care

The verb *care* has second highest percentage of errors by PAK after IDN learners in the PREP condition. Examples (6.10) – (6.11) from the corpus show incorrect use of the preposition *of* instead of *for*. Examples (6.12) – (6.13) from the corpus show an omitted preposition.

(6.10) We must care of our future [...]

(6.11) […] you doesn't care of yourself then can you please do care of others […]

(6.12) [...] He is not care his health [...]

(6.13) […] they don't care their health […]

(6.14) ہمیں اپنے مستقبل کا خیال رکھنا چاہیے۔
        humein apne mustaqbil ka khyal rakhna chahye
        We our future of care take must
        'We must care for our future.'

Using *of* instead of *for* may result from Urdu influence, where corresponding meanings are constructed differently. In Urdu, the preposition in the corresponding sentence would be کا *ka* which closely corresponds to English 'of', as in example (6.14). Thus, it is probable that some transfer occurred here, in which the PAK learner picked the English preposition most closely comparable to L1. It is also possible that readers confuse the use of the verb *care* with the multi-word expression *take care of.*

The preposition is omitted in (6.12), leading to an error in the NP condition. As discussed by Karim and Shahwar (2015), the concept of avoidance was initially introduced by Schechter in 1974. Because they lack the necessary language resources, L2 learners use this procedural method to replace obligatory L2 forms with other constructions. In addition to the structural similarities and differences between L1 and L2 as the reasons behind avoidance behavior, the various proficiency levels are another aspect of this phenomenon. This avoidance decreases as proficiency grows, which is associated with interlanguage development. Therefore, one possible reason for omitting or avoiding the preposition, is that the learner is confused of which preposition to use when influenced by L1, and hence avoids using any preposition at all. Prepositional use by learners is frequently influenced by their mother tongue. Prepositions are used differently in Urdu and English; certain English phrases have no direct equivalent in Urdu. Students can use prepositions or omit them altogether. Additionally, Urdu uses mostly postpositions rather than prepositions, which might add to the Pakistani learners' challenges in learning the use of prepositions in English. Their linguistic abilities may be impacted by this basic difference.

6.2.5 Contribute

The verb *contribute* has a higher percentage of errors by PAK than other learners in the PREP condition, but the counts are fairly low. The example (6.15) from the corpus is not in itself ungrammatical. However, from the context it appears that *to society* is more in line with the intended meaning than *in society*. The use of preposition *in* can be attributed to L1 transfer, as the corresponding Urdu preposition in this context would be میں *mein* 'in' as illustrated in example (6.16).

(6.15) […] how to deal with people or how to contribute in society.

(6.16) ہم سب کو معاشرے میں حصہ ڈالنا چاہیے۔
    hum sub ko muasherey mein hissa dalna chahye
    We all to society in contribute put should
    'We all should contribute to society.'

This can be attributed to factors like language transfer from the native language (Urdu), and interlanguage development. In this specific case, Urdu uses the preposition میں 'mein' with the equivalent of *contribute*, which could lead learners to choose *in* instead of *to* when forming English sentences. ESL learners, as they progress in learning English, go through different stages of interlanguage development (Dulay and Burt 1974). During these transitional stages, errors like incorrect prepositional usage are common. Similarly, Selinker (1972) has highlighted the influence of the native language on second language learning, which can manifest itself in errors like incorrect preposition use.

6.2.6 Disagree

Like its antonym *agree*, the verb *disagree* has a higher percentage of errors by PAK than other learners in the PREP condition, but the counts are very low, so they need to be taken with a

grain of salt. Example (6.17) from the corpus has the preposition *to* rather than *with*. This could be due to an analogy with English *agree to [an arrangement, a meeting, etc.]*, although in the context of example (6.17), the use of *with* is clearly preferable. Furthermore, as shown in (6.18), a corresponding Urdu preposition in this context would be سے, which has English equivalent 'for/to/from' depending on context. This difference and the semantic vagueness may lead to the choice of a less suitable preposition.

(6.17) […] But, I totally disagree to it.

(6.18) میں ان سے متفق نہیں ہوں۔
   mein in se muttafiq nahi hun
   I them for/to/from agree not am
   'I disagree with them.'

6.2.7   Look

The verb *look* has a higher percentage of errors by PAK than other learners; the counts are fairly low. Example (6.19) from the corpus shows incorrect use of the preposition *on* instead of *at*. Example (6.20) from the corpus shows an omitted preposition for the same verb in the NP condition.

(6.19) Today, if we are looking on the newspaper we shall notice […]

(6.20) They looked them with fully concentration […]

(6.21) وہ اخبار پر دیکھ رہے ہیں
   Wo akhbaar par dekh rahe hain
   They newspaper on look -ing are
   'They are looking at the newspaper.'

(6.22) میں نے اسے دیکھا
   mein ne isey dekha
   I have/by him looked
   'I looked at him.'

Example (6.21), illustrates that the preposition پر *par* which in most other contexts corresponds to 'on', would be used in the context of a newspaper. The learner's choice of preposition would thus reflect the influence of transfer, where the structures and prepositions used in the native language are transferred to the second language (English) even if they do not align with English grammar and usage. As discussed by Dagut and Laufer (1985), one reason for avoidance behaviour was the lack of "translation-equivalents" for the English structures which compelled the learners to choose less suitable but more L1-equivalent structures. Therefore, a possible explanation for omitting the preposition, as in (6.20), is interference from other Urdu constructions for looking, as in (6.22), where no preposition is required.

6.2.8   Result

PAK has the second highest number of errors in verb *result* in the PREP condition. Example (6.23) from the corpus shows incorrect use of the preposition *into* instead of *in*. These are easily confused since in some contexts, they are practically interchangeable in English. Both are covered by the Urdu preposition میں *mein* 'in' as illustrated in example (6.25) which can be

a translation of (6.23). Saed and Yassin (2017) discussed the same problem faced by Jordanian ESL learners and found that prepositions were used differently depending on the language. For instance, the same preposition in learners' native language might have completely different meanings and multiple prepositions in English in different contexts. They concluded that ESL learners should not rely solely on their first language's prepositional knowledge because of this.According to Islami (2015), this comparison to the learners' native language's prepositional system and the difference in number, meaning, and usage between L1 and English as a foreign language make English prepositions challenging for EFL learners. An example in the NP condition is (6.24). It is unclear why the preposition is omitted, but as suggested earlier, an avoidance strategy might be a factor. Lightbown and Spada (2013) also stated that EFL students occasionally omit words or word chunks that they find challenging to learn.

(6.23) […] it results into the destruction of states[…]

(6.24) […] which result a bad effect on their studies […]

(6.25) اس کے نتیجے میں ریاستوں کی تباہی ہوتی ہے۔
Is ke natijay mein riyaston ki tabahi hoti-hai
It of result in states of destruction is
'It results in the destruction of states.'

6.2.9 Sit

The verb *sit* has a higher percentage of errors by PAK than other learners in the PREP condition. The example (6.26) from the corpus shows incorrect use of the preposition *in* instead of *with*. In a similar way to what was discussed earlier in this section for the verb *contribute,* the use of *in* could be attributed to L1 transfer, as the corresponding Urdu preposition in this context would be میں *mein* 'in' as illustrated in example (6.27). Thus the error likely arises from interference from the Urdu language, where ساتھ *saath* 'with' is not commonly used in the context of 'sitting with friends'. This kind of prepositional error is a common occurrence in the interlanguage development of ESL learners (Dulay et al. 1982; Selinker 1972), particularly when their native language structures differ from those in English.

(6.26) […] To avoid smoking we should sit in good friends.

(6.27) سگریٹ نوشی سے بچنے کے لیے ہمیں اچھے دوستوں میں بیٹھنا چاہیے۔
cigarette-noshi se bachney ke-liye humen ache doston mein bethna chahye
Cigarette-smoking from to-avoid for we good friends in/among sit should
'To avoid smoking, we should sit with/among good friends.'

6.2.10 Suffer

The verb *suffer* has a higher percentage of errors by PAK than other learners in the PREP condition, but no errors in the NP condition. As previously discussed for verbs *sit, contribute* and *result,* incorrect use of the preposition *in* instead of *from* can be observed, as in example (6.28) from the corpus. This use could be attributed to L1 transfer as the corresponding Urdu preposition in this context would be میں *mein* 'in' as illustrated in example (6.30). The Urdu equivalent expressions for 'suffer from fever' would be بخار میں مبتلا , where the use of میں *mein* corresponding to English 'in' is common. The existence of this construction may lead to a transfer error. Similarly, example (6.29) from the corpus illustrates incorrect use of *to* instead of *from,*

which can also be due to the impact of transfer from L1, as argued by Masood et al. (2020), ESL learners tried to mimic the structure of their L1. They claimed that the influence of the first language was highly noticeable in the learning of English as a second language. The corresponding Urdu expression involves کو *ko* 'to' as in (6.31).

(6.28) With so much work you may suffer in fever.

(6.29) […] smokers are caught in suffer and also suffer to others […]

(6.30) وہ بخار میں مبتلا ہے
wo bukhar mein mubtala hai
He fever in suffer is
'He suffers from fever.'

(6.31) سگریٹ دوسروں کو تکلیف دیتا ہے
cigarette doosron ko taqleef deta hai
Cigarette others to hurt/suffer gives is
'Cigarettes hurt others.'

6.2.11 Think

The verb *think* has a higher percentage of errors by PAK than other learners in the PREP condition. Example (6.32) from the corpus shows an incorrect use of the preposition *on* instead of *about*.

(6.32) […] So please think on it.

(6.33) اسے اس مسئلے پر سوچنا چاہیے۔
usey is masley par sochna chahye
He/she it/this problem on think should/must
'He should think about this problem.'

Again, this error may result from L1 transfer, where similar expressions are constructed differently. In Urdu, the preposition in the corresponding sentence would be پر *par* which is closely corresponding to English 'on', as in example (6.33). This mechanism also seems to be at play in learners with other L1 backgrounds. In the context of Georgian L1, Gvarishvili (2013, p.1569) finds that, in a majority of the cases, the learner attempts to find a preposition that matches the native equivalent construction. Thus, it is probable that some transfer has occurred in which the PAK learner picked the English preposition which literally is most comparable to that in L1.

6.3 Correct use of prepositions

The findings indicate that PAK learners have a relatively low frequency of prepositional errors for the verbs *base, benefit, cause, compare, cope, deal, depend, focus, graduate, know, learn, prepare, relate,* and *worry*, for which several reasons can be suggested. These reasons may be linked to the complexity, transparency, and familiarity with these prepositions.

Early on in their language learning, ESL students are frequently exposed to basic prepositions. Prepositions like *in, on, under,* and *between* are some of the first ones that English language learners come upon. This early exposure enables students to get comfortable using

them. In spoken language, prepositional verbs are commonly employed. They are often used by learners in both written and oral communication, which causes them to stick in their memory. Also, certain prepositional verbs have more transparent usage and meanings. Consequently, it could be simpler for ESL students to comprehend and appropriately apply these prepositions.

Prepositions in Urdu may occasionally have meanings that are comparable to those of prepositions in English. Because of this resemblance, learning how to use them may be simpler for students. Prepositional verbs can sometimes have direct literal translations in both languages, which facilitates learners' comprehension and application of Urdu to English. Prepositions that present difficulties for learners, on the other hand, frequently have idiomatic, abstract, or context-dependent meanings. Prepositions that have no direct equivalents in the original language might also be challenging to learn.

English language teaching is frequently provided in differing degrees to learners at different skill levels. Higher achievers could have profited from more in-depth and advanced language training. Learners at these levels may have received specific teaching or emphasis on prepositional verbs, which has resulted in increased accuracy in their usage. Higher proficiency learners often possess a more sophisticated knowledge of linguistic structures, such as prepositional verbs. They could understand how to utilize prepositions more idiomatically and contextually.

It is also possible that students' ease of use and reduced error rates with certain prepositional verbs stem from the diversity of CEFR proficiency levels found in the ICNALE data. Data from English language learners of every proficiency level, from novices to fluent speakers, are included in the ICNALE corpus. This variety represents the real-world context in which English language learners have varied exposure to and experiences with the language due to their diverse linguistic origins. Higher proficiency students have usually encountered more English, either via immersion in English-speaking areas or through formal schooling. Prepositional verb usage and comprehension may improve as a result of this greater exposure. However, it is also possible that there are individual differences in language acquisition, and that some students may do well in some areas of grammar while struggling with others.

As they learn English, ESL students move through phases of interlanguage development. In order to build English sentences, learners with lower competence levels are more prone to rely on direct word-for-word translations from their home language — in this case, Urdu. Learners make fewer mistakes as they go to higher competence levels because they have a more sophisticated and context-dependent knowledge of English prepositions. For instructors and academics interested in comprehending the nuances of preposition learning in ESL, this diversity would offer valuable insights.

Chapter 7

Conclusion

Based on the data and the analysis presented in this study, the hypotheses that were put forward in Chapter 1 seem to be confirmed. Pakistani ESL learners make notable errors in their use of prepositional verb constructions. These errors include using wrong prepositions and omitting prepositions. The error rate in Pakistani learners is higher than in the other groups of Asian learners taken together. There are differences between specific lemmas and between error types, as discussed in some detail in the previous chapters. Plausible explanations for the Pakistani learners' choices may be found in transfer from L1, which is commonly Urdu. The usage of prepositions and how they are structured in Urdu differ on many points from usage in English.

The findings are in line with previous research that recognizes the impact of L1 on the process and outcome of language learning as the cause of partial L2 proficiency (Corder 1967, 1975, 1971). Language learners go through transitional periods when they learn a second language. Interlanguage is the developing linguistic system that learners construct as they learn a target language. Previous research on interlanguage errors often cites the influence of the native language as a significant factor. Errors occur as learners navigate this transition, often influenced by the grammatical structures and prepositions of their native language.

Multi-word verb constructions from a learner's L1 may conflict with the acquisition of L2 multi-word verb constructions (Arnon and Christiansen 2017). Incorrect prepositional use is typical in interlanguage and L1 interference is a common cause of such errors (Dulay et al. 1982; Selinker 1972). In their development of interlanguage, Pakistani learners may transfer the syntax and semantics of Urdu prepositions to English prepositional verb constructions or leave prepositions out entirely. These prepositional verb problems might result from Urdu interference and transfer since the two languages' prepositional use and structure are different. These distinctions are frequently difficult for ESL students to understand, which leads to the kinds of errors found in the corpus data.

In order to help students overcome these obstacles, teachers should be aware of typical interference problems and offer specialized education. It is therefore hoped that the insights gained in this work will enable the further improvement of teaching and teaching materials. Careful attention to challenging constructions involving prepositions, through targeted instruction and exercises, may result in more awareness and a higher level of competence among Pakistani ESL learners. In a Pakistani setting, ESL teachers usually make learners cram the prepositional verb lists without explaining their uses in context, which makes these verbs more problematic for Pakistani ESL learners. They are often used idiomatically and have multiple meanings, making it challenging to identify the correct preposition to use in a given context. Effectively teaching prepositional verbs to Pakistani ESL learners entails resolving L1 interference and encouraging a deeper comprehension. Some possible recommendations could therefore be given to English

teachers in Pakistan, although more research will be needed to assess the potential of specific measures along the lines of the following recommendations.

1. Understanding learners' interlanguage is crucial for language teachers as it helps them design language instruction and provide effective feedback tailored to the specific needs of the learner's interlanguage.

2. As mother tongue interference causes a substantial number of errors, considering learners' L1 while teaching prepositions should be one of the teacher's top objectives (Özışık 2014). Knowing where pupils are likely to make mistakes might help teachers address them in advance. Most prepositional verbs do not have equivalent expressions in the native language or they are non-existent at all in the mother tongue (Ella and Dita 2017). As stated by Masood et al. (2020), language teachers must be aware of the interference of L1 in L2 acquisition and devise suitable teaching techniques to meet the problems experienced by language learners.

3. Previous research has identified that Pakistani learners of English have difficulty with the correct use of prepositions in different contexts, and they often use prepositions inappropriately (Masood et al. 2020). Therefore, teachers should teach prepositional verbs in context and real-life situations and give examples of phrases and circumstances where these verbs are often utilized. They may use different teaching techniques, such as immersion programs, task-based learning, and grammar-focused instruction to promote understanding of prepositional verbs.

4. As discussed by Siyanova and Schmitt (2007), the mismatch between the idiomatic meaning and the individual verb and their polysemous nature might make multi-word verbs difficult to understand. So, ESL teachers should design activities to practice these areas of prepositional verb issues.

5. As Masood et al. (2020) stress the need to provide English language learners with adequate practice opportunities to improve their language skills, ESL teachers should include prepositional verbs in homework and classroom activities regularly and encourage students to use these verbs in their essays, tales, or dialogues.

6. L2 learners can benefit from specialized instruction in the usage of prepositions in English to address these issues. Teachers can explain the principles of preposition usage and give exercises that show the differences in argument structure between English and the learners' native language. Furthermore, reading and listening to authentic language input might assist L2 learners build a better comprehension of prepositions in context.

7. To reinforce learning, use prepositional verbs in frequent tests and examinations.

8. Assist students in identifying and correcting prepositional verb problems in their work. This type of self-correction can help with learning. Furthermore, provide constructive comments on assignments, emphasizing prepositional verb mistakes. Students should be encouraged to edit and resubmit their work.

9. Bilingual dictionaries can assist students in understanding the Urdu-to-English counterparts of prepositional verbs.

10. As suggested by Lin and Lin (2019), students would gain from exposure to corpus resources of real-world English use throughout their EFL studies if they were exposed to more realistic and varied uses of language. ESL teachers should give ICNALE examples that show how to utilize prepositional verbs correctly. Students can evaluate these sentences. Furthermore, prepositional verbs should be taught in conjunction with collocations, rather than in isolation (Islami 2015; Özışık 2014; Saed and Yassin 2017), utilizing concordance tools or online corpora as learning resources.

11. Pair students to work on tasks together and discuss prepositional verbs. Understanding complicated topics is frequently aided by peer learning.

12. Make the usage of prepositional verbs more approachable to pupils by relating them to parts of Pakistani culture.

13. To keep current on best practices in teaching prepositional verbs, teachers should engage in ongoing professional development.

14. Curriculum designers must pay greater attention to teaching prepositional verbs in textbooks and curricula (Abdalla 2021).

A few caveats and qualifiers regarding the current study must be noted, while these may also hint at possibilities for future studies. A general remark is that the whole distinction between a language and learner language tends to be somewhat unclear. Indeed, some authors claim that there is no such thing as 'English', an idea pushed further by (Makoni and Pennycook 2005, p. 138) who "start with the premise that languages — and the metalanguages used to describe them — are inventions". This idea is further pursued by May (2014) who points out that there are large differences between what is labeled 'English' across the globe and that people may have fluid ways of expression, including intensive code-switching such as Hinglish, the hybrid use of English and languages of the Indian subcontinent. Although such viewpoints are theoretical complications for the study of learner communities, they are not an obstacle for the present study, which merely wants to establish to what extent certain constructions are correctly used as required by standard English.

The criteria, then, as to what is standard English, rest with me as the only manual annotator of the corpus, with the advice of my thesis supervisor. In some cases, particularly when sentences had multiple grammatical problems, the annotation criteria were difficult to apply, and there is always the possibility that some errors and inconsistencies were introduced. Ideally, several people should have annotated every item and inter-annotator agreement should have been measured. No resources for such an effort were available at this time, but clearly this is still an interesting option, should the present study be replicated or extended in the future.

If the corpus had more detailed syntactic information (as a treebank), it might have been much easier to establish the specific complementation pattern of each verb use and whether the item is grammatically correct or not. A more fine-grained study of all grammatical aspects of interlanguage, both correct and incorrect use, is desirable, but a consistent description of interlanguage is ambitious and premature since it presupposes detectable systematicity, which may be difficult to establish, since interlanguage is individual and in flux (Rosén and De Smedt 2010; Tenfjord 1983). Because it does not seem possible to describe interlanguage without reference to the target L2 (Rosén and De Smedt 2010), it is more realistic to confine the annotation to error annotation, which must take the investigation of specific morphological, lexical or grammatical phenomena related to the intended target language into account.

Learner diversity, in terms of individual differences between the students who wrote the corpus texts, could not be taken into account in the present study. It is assumed that most Pakistani students either have Urdu as their L1 or have good knowledge of Urdu, which is not only an official language but also the lingua franca in the country. There is however no detailed information on participants' L1 or other L2 backgrounds in the ICNALE documentation. It is therefore possible that some individual differences in L1 and other language competence play a so far unknown role. Furthermore, the corpus is unfortunately not balanced with respect to levels, and the current study has not carried out separate analyses for each different CEFR level, since that would have faced a problem of data scarcity for each group. It is likely that students at higher proficiency levels demonstrate greater ease and accuracy in handling prepositional verbs. Thus, the finding that SIN learners have the highest overall correct percentage might be explained by the fact that their group does not include learners at A and B1.1 levels, but only at the B1.2 and B2+ levels. Future studies should attempt to find a larger corpus that is balanced between levels.

For now, the focus has been on the PAK observations in the corpus that is presently available. The gained knowledge may be extended through follow-up studies aimed specifically at various constructions and semantic contexts involving prepositions, and follow-up studies that use other empirical data, including speech. The present work may also be followed up with more studies focusing on other languages, since differences in the range and use of prepositions are challenging in the acquisition of many languages.

Bibliography

Abdalla, Ibrahim (2021). "Difficulties in Using Correct English Prepositions among EFL Students." In: *JEES (Journal of English Educators Society)* 6.2. DOI: 10.21070/jees.v6i2.1640.

Abdulaziz, Maimoona and Muhammad Asim Mahmood (2022). "Exploration of New Dimensions in Argumentative Essay Writing in a Corpus of World Englishes." In: *PalArch's Journal of Archaeology of Egypt / Egyptology* 19.1, pp. 1488–1502.

Abualzain, Osama Yousif Ibrahim (2017). "Writing Difficulties in Prepositions of Place Encountered by Albaha University Students' at Preparatory Year Program." In: *International Journal on Studies in English Language and Literature (IJSELL)* 5.5, pp. 1–8. DOI: https://www.arcjournals.org/pdfs/ijsell/early-view/1.pdf.

Ahmed, S (2015). "A Contrastive Analysis of Prepositions in English and Urdu." In: *Language in India* 15.7, pp. 150–162.

Akbarnezhad, Shima, Firooz Sadighi, and Mohammad Sadegh Bagheri (2020). "A Syntactic-Based Approach to the Perception and Production of English Verbs' Argument Structures by Iranian EFL Learners." In: *Cogent Arts & Humanities* 7.1, p. 1770956. DOI: 10.1080/23311983.2020.1770956.

Aldukhayel, Dukhayel M (2014). "The L2 Exposure Effect on Avoidance of Phrasal Verbs by Arab ESL Learners." MA thesis. Colorado: Colorado State University.

Alhamadi, Nesreen Saud (2015). "Language Transfer Errors in Speaking among Saudi Arabia Students; a Comparative Study between Students in Saudi Arabia and Students in the UK." PhD thesis. Leeds: University of Leeds. 299 pp.

Allen, Shanley E. M (2015). "Verb Argument Structure." In: *The Cambridge Handbook of Child Language*. Ed. by Edith L. Bavin and Letitia R. Naigles. 2nd ed. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, pp. 271–297. DOI: 10.1017/CBO9781316095829.013.

Alyas, Rafia (May 2019). "A Quantitative Approach Analysis on Subject-Verb Agreement Problem Among English as Second Language Learners: A Case Study of Pakistani Students." In: *Journal of Literature, Languages and Linguistics* 56, pp. 1–6. DOI: 10.7176/JLLL/56-01.

Anwar, Nazia, Moazzam Ali Malik, and Saima Jamshaid (2020). "Lexical Bundles in Academic Discourse: A Comparative Study of Functional Use of Lexical Bundles in Native and Non-Native Learner Corpora." In: *Epistemology* 7.3, pp. 28–43.

Arnon, Inbal and Morten H. Christiansen (2017). "The Role of Multiword Building Blocks in Explaining L1–L2 Differences." In: *Topics in Cognitive Science* 9.3, pp. 621–636. DOI: 10.1111/tops.12271.

Arslan, Muhammad Farukh, Muhammad Asim Mahmood, and Attia Rasool (2020). "A Corpus-Based Comparative Study of Derivational Morphemes Across ENL, ESL, EFL Learners Through ICNALE." In: *Linguistic Forum - A Journal of Linguistics* 2.4, pp. 1–12. DOI: 10.53057/linfo/2020.2.4.1.

Aslam, Aisha and Dr Shafaq Fayyaz (2021). "Investigating the Use of Nominalization: A Comparative Study of Pakistani Learners' and Native English Speakers' Argumentative Essays."

In: *Psychology and Education Journal* 58.2, pp. 3638–3648. DOI: 10.17762/pae.v58i2. 2625.

Baker, Susan C. and Peter D. MacIntyre (2000). "The Role of Gender and Immersion in Communication and Second Language Orientations." In: *Language Learning* 50, pp. 311–341. DOI: 10.1111/0023-8333.00119.

Bakken, Vivil Selliseth (2017). "The Acquisition of English Prepositions by Norwegian L2 Learners-a Corpus-Based Error Analysis." MA thesis. Bergen, Norway: Department of Foreign Languages, University of Bergen.

Barrot, Jessie and Mari Karen Gabinete (2021). "Complexity, Accuracy, and Fluency in the Argumentative Writing of ESL and EFL Learners." In: *International Review of Applied Linguistics in Language Teaching* 59.2, pp. 209–232. DOI: 10.1515/iral-2017-0012.

Block, David (2009). *Second Language Identities*. Bloomsbury Publishing.

Brinton, D.M. and M. Celce-Murcia (2000). "Teaching American Idioms to Overseas Students." In: *TESOL Quarterly* 9.3, pp. 14–18.

Brown, H. Douglas (2007). *Principles of Language Learning and Teaching*. 5th ed. White Plains, NY: Pearson Longman. 410 pp.

Cho Min, Hyun Soon (2020). "A Contrastive Corpus Analysis on the Use of Connectors in Students' Writing from 10 Asian Countries as Compared to Native Experts: Research from the ICNALE (The International Corpus Network of Asian Learners of English)." PhD thesis. United States – Tennessee: The University of Memphis. 80 pp.

Christiansen, Morten H. and Inbal Arnon (2017). "More Than Words: The Role of Multiword Sequences in Language Learning and Use." In: *Topics in Cognitive Science* 9.3, pp. 542–551. DOI: 10.1111/tops.12274.

COE (Council of Europe, Modern Languages Division) (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, U.K.: Cambridge University Press.

Cook, Vivian (2001). "Using the First Language in the Classroom." In: *The Canadian Modern Language Review* 57.3, pp. 402–423. DOI: 10.3138/cmlr.57.3.402.

— (2016). *Second Language Learning and Language Teaching*. 5th ed. Routledge. 334 pp. DOI: 10.4324/9781315883113.

Corder, S. P. (1967). "The Significance of Learner's Errors." In: *International Review of Applied Linguistics in Language Teaching* 5.1-4, pp. 161–170. DOI: 10.1515/iral.1967.5.1-4.161.

— (1975). "Error Analysis, Interlanguage and Second Language Acquisition." In: *Language Teaching & Linguistics: Abstracts* 8.4, pp. 201–218. DOI: 10.1017/S0261444800002822.

— (1971). "Idiosyncratic Dialects and Error Analysis." In: 9.2, pp. 147–160. DOI: 10.1515/iral.1971.9.2.147.

Creswell, John W. and Vicki L. Plano Clark (2017). *Designing and Conducting Mixed Methods Research*. 3rd ed. Thousand Oaks, Calif: SAGE Publications. 275 pp.

Dagut, Menachem and Batia Laufer (1985). "Avoidance of Phrasal Verbs—A Case for Contrastive Analysis." In: *Studies in Second Language Acquisition* 7.1, pp. 73–79. JSTOR: 44488547.

Dekeyser, Xavier, Betty Devriendt, Guy A. J. Tops, and Steven Geukens (1979). *Foundations of English Grammar: For University Students and Advanced Learners*. Vol. 3. Bibliotheca Anglistica. Antwerpen/Amsterdam: De Nederlandsche Boekhandel. 446 pp.

Demirel, Elif Tokdemir and Semin Kazazoğlu (2015). "The Comparison of Collocation Use by Turkish and Asian Learners of English: The Case of TCSE Corpus and Icnale Corpus." In: *Procedia - Social and Behavioral Sciences*. International Conference on New Horizons in

Education, INTE 2014, 25-27 June 2014, Paris, France 174, pp. 2278–2284. DOI: 10.1016/j.sbspro.2015.01.887.

Dini, Monara (2020). "An Analysis of Phrasal Verb Errors Used By Fifth Semester Students of English Study Program of IAIN Bengkulu (A Descriptive Qualitative Study on Students Phrasal Verb Errors at Fifth Semester Students of English Study Program of IAIN Bengkulu in Academic Year 2018/2019)." In: *STATE INSTITUTE OF ISLAMIC STUDIES (IAIN) BENGKULU*.

Dulay, Heidi, Marina Burt, and Stephen Krashen (1982). *Language Two*. Oxford: Oxford University Press. 315 pp.

Dulay, Heidi C. and Marina K. Burt (1974). "Errors and Strategies in Child Second Language Acquisition." In: *TESOL Quarterly* 8.2, pp. 129–136. DOI: 10.2307/3585536. JSTOR: 3585536.

EALTA (European Association for Language Testing and Assessment) (2016). *EALTA Guidelines for Good Practice in Language Testing and Assessment*.

Ella, Jennibelle and Shirley Dita (2017). "The Phrasal-Prepositional Verbs in Philippine English: A Corpus-based Analysis." In: *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation*. 31st Pacific Asia Conference on Language, Information and Computation (PACLIC 31). Philippines: National University (Philippines), pp. 34–41. DOI: https://aclanthology.org/Y17-1008.pdf.

Ellis, Rod (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press. 835 pp.

— (1997). *Second Language Acquisition*. Ed. by H.G. Widdowson. 1st ed. Oxford Introductions to Language Study. Oxford: OUP. 158 pp.

Farooq, Muhammad Shahid, Muhammad Uzair-Ul-Hassan, and S. Wahid (2012). "Opinion of Second Language Learners about Writing Difficulties in English Language." In: *South Asian Studies* 27.1, pp. 183–194.

Farukh, Ammara and Mila Vulchanova (2015). "L1, Quantity of Exposure to L2, and Reading Disability as Factors in L2 Literacy Skills." In: *Usage-Based Perspectives on Second Language Learning*. Ed. by Teresa Cadierno and Søren Wind Eskildsen. Germany: Walter de Gruyter GmbH & Co KG. Google Books: iKCnCgAAQBAJ.

Gardner, Dee and Mark Davies (2007). "Pointing out Frequent Phrasal Verbs: A Corpus-Based Analysis." In: *TESOL Quarterly* 41.2, pp. 339–359. DOI: 10.2307/40264356. JSTOR: 40264356.

Gass, Susan and Larry Selinker (1994). *Second Language Acquisition: An Introductory Course*. Vol. 71. DOI: 10.2307/416225.

Gass, Susan M. and Larry Selinker (2008). *Second Language Acquisition: An Introductory Course*. 3rd ed. New York: Routledge -Taylor & Francis. 616 pp.

Goad, Heather and Lydia White (2004). "Ultimate Attainment of L2 Inflection: Effects of L1 Prosodic Structure." In: *EUROSLA Yearbook* 4.1, pp. 119–145. DOI: 10.1075/eurosla.4.07goa.

Goldberg, Adele E.. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Greene, Jennifer C., Valerie J. Caracelli, and Wendy F. Graham (1989). "Toward a Conceptual Framework for Mixed-Method Evaluation Designs." In: *Educational Evaluation and Policy Analysis* 11.3, pp. 255–274. DOI: 10.3102/01623737011003255.

Gvarishvili, Zeinab (Jan. 2013). "Interference of L1 Prepositional Knowledge in Acquiring of Prepositional Usage in English." In: *Procedia - Social and Behavioral Sciences* 70, pp. 1565–1573. DOI: 10.1016/j.sbspro.2013.01.224.

Haugh, Sam and Osamu Takeuchi (2022). "Phrasal Verb Usage and L1 Typology : A Corpus-based Study on Chinese and Japanese EFL Learners." In: *The Journal of Asia TEFL* 19.4, pp. 1200–1218. DOI: 10.18823/asiatefl.2022.19.4.4.1200.

Heydari, Pooneh and Mohammad S. Bagheri (2012). "Error Analysis: Sources of L2 Learners' Errors." In: *Theory and Practice in Language Studies* 2.8, pp. 1583–1589. DOI: 10.4304/tpls.2.8.1583-1589.

Hunston, Susan (2002). *Corpora in Applied Linguistics*. Cambridge Applied Linguistics. Cambridge: Cambridge University Press. 241 pp. DOI: 10.1017/CBO9781139524773.

Ionin, Tania and Silvina Montrul (2010). "The Role of L1 Transfer in the Interpretation of Articles with Definite Plurals in L2 English." In: *Language Learning* 60.4, pp. 877–925. DOI: 10.1111/j.1467-9922.2010.00577.x.

Iqbal, Liaqat, Irfan Ullah, and Ayaz Ahmad (2019). "Error Analysis: A Study of Prepositional Errors at Advanced and Intermediate Level in Pakistan." In: *Global Regional Review* IV.1, pp. 112–119. DOI: 10.31703/grr.2019(IV-I).13.

Ishaq, Qudsia and Nadeem Haider Bukhari (2016). "Persistent Errors in English Writing and Speech Made by Urdu Undergraduates at UAJK: A Case of Mother Tongue Interference." In: *Kashmir Journal of Language Research* 19.1, pp. 249–264.

Ishikawa, Shin'ichiro (2023). *The ICNALE Guide: An Introduction to a Learner Corpus Study on Asian Learners' L2 English*. 1st ed. London: Routledge.

Ishikawa, Shin'ichiro (2011). "A New Horizon in Learner Corpus Studies: The Aim of the IC-NALE Project." In: *Corpora and Language Technologies in Teaching, Learning and Research*. Ed. by George Weir, Shin'ichiro Ishikawa, and Kornwipa Poonpon. Vol. 404. Glasgow, UK: University of Strathclyde Press, pp. 3–11.

— (2013). "The ICNALE and Sophisticated Contrastive Interlanguage Analysis of Asian Learners of English." In: *Learner corpus studies in Asia and the world* 1. Ed. by Shin'ichiro Ishikawa, pp. 91–118. DOI: https://hdl.handle.net/20.500.14094/001002007.

— (2019). "The ICNALE Spoken Dialogue: A New Dataset for the Study of Asian Learners' Performance in L2 English Interviews." In: *KoreaScholar* 74.4, pp. 153–177. DOI: 10.15858/engtea.74.4.201912.153.

Ishikawa, Yuka (2015). "Gender Differences in Vocabulary Use in Essay Writing by University Students." In: *Procedia - Social and Behavioral Sciences* 192, pp. 593–600. DOI: 10.1016/j.sbspro.2015.06.078.

Islami, Avni (2015). "The Main Difficulties in Mastering the English Language Prepositions." In: *THESIS* 4.3, pp. 17–28. DOI: https://rep.thesis-journal.net/handle/123456789/91.

Jabeen, Farhat, Asim Rai, and Sara Arif (2011). "A Corpus Based Study of Discourse Markers in British and Pakistani Speech." In: *International Journal of Language Studies* 5.4, pp. 69–86.

Jackendoff, Ray and Ray S Jackendoff (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. USA: Oxford University Press.

Jarvis, S. and A. Pavlenko (2008). *Crosslinguistic Influence in Language and Cognition*. Taylor & Francis.

Jiang, Nan and Kenneth I. Forster (2011). "Word Order and Grammatical Gender Processing in Sentence Reading: A Cross-Linguistic ERP Study." In: *Journal of Cognitive Neuroscience* 23.10, pp. 2822–2838.

Johnson, Gerry, Ann Langley, Leif Melin, and Richard Whittington (2007). *Strategy as Practice. Research Directions and Resources*. Cambridge University Press.

Johnson, R. Burke and Anthony J. Onwuegbuzie (2004). "Mixed Methods Research: A Research Paradigm Whose Time Has Come." In: *Educational Researcher* 33.7, pp. 14–26. DOI: 10.3102/0013189x033007014.

Kachru, Yamuna and Cecil L Nelson (2006). *World Englishes in Asian Contexts*. Vol. 1. Hong Kong: Hong Kong University Press. 436 pp.

Kanno, Y., N. Hikima, and O. Inbar (2017). "English as a Second Language Education." In: *Handbook of Educational Linguistics*. Ed. by B. Spolsky. John Wiley and Sons, Inc., pp. 381–395.

Karim, Shahzad and Dur-e- Shahwar (2015). "Avoidance of English Phrasal Verbs: A Study Based on the Pakistani ESL Learners." In: *ELF Annual Research Journal* 17, pp. 125–144.

Khan, M.A and R. Tariq (2012). "Preposition Errors in the Written English of Pakistani University Students." In: *Language Testing in Asia* 2.2, pp. 99–116.

Klein, W. (1986). *Second Language Acquisition*. Cambridge University Press.

Komba, Sotco and Stephen Bosco (2015). "Do Students' Backgrounds in the Language of Instruction Influence Secondary School Academic Performance?" In: *Journal of Education and Practice* 6.30, pp. 148–156.

Krashen, Stephen (1982). *Principles and Practice in Second Language Acquisition*. 1st ed. Language Teaching Methodology Series. Oxford: Pergamon Press.

Krashen, Stephen D. (1981). *Second Language Acquisition and Second Language Learning*. 1st ed. Language Teaching Methodology Series. Oxford: Pergamon Press.

Leroux, William and Tyler Kendall (Aug. 2018). "English Article Acquisition by Chinese Learners of English: An Analysis of Two Corpora." In: *System* 76, pp. 13–24. DOI: 10.1016/j.system.2018.04.011.

Li, Ping and Yasuhiro Shirai (2000). *The Acquisition of Lexical and Grammatical Aspect*. Vol. 16. Studies on Language Acquisition (SOLA). Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110800715.

Lightbown, Patsy M. and Nina Spada (2013). *How Languages Are Learned*. 4th ed. Oxford Handbooks for Language Teachers. United Kingdom: Oxford University Press. 252 pp. Google Books: 5PadBgAAQBAJ.

Lin, Ching-Hsi and Yen-Liang Lin (2019). "Grammatical and Lexical Patterning of Make in Asian Learner Writing: A Corpus-Based Study of ICNALE." In: *3L; Language,Linguistics and Literature,The Southeast Asian Journal of English Language Studies*. 25.3, pp. 1–15. DOI: 10.17576/3L-2019-2503-01.

Lorincz, Kristen and Rebekah Gordon (2012). "Difficulties in Learning Prepositions and Possible Solutions." In: *Linguistic Portfolios* 1.1. DOI: https://repository.stcloudstate.edu/stcloud_ling/vol1/iss1/14.

Lyster, Roy and Leila Ranta (1997). "Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms." In: *Studies in Second Language Acquisition* 19.1, pp. 37–66. JSTOR: 44488666.

Mahmood, Muhammad Asim (2009). "A Corpus Based Analysis of Pakistani English." Bahauddin Zakariya University Multan.

Makoni, Sinfree and Alastair Pennycook (2005). "Disinventing and (Re)Constituting Languages." In: *Critical inquiry in language studies* 2.3, pp. 137–156. DOI: 10.1207/s15427595cils0203_1.

Masood, Muhammad Hamzah, Shahzeb Shafi, Muhammad Yousaf Rahim, and Maqsood Ali Darwesh (2020). "Interference of L1 (Urdu) in L2 (English) in Pakistan: Teaching English as a Second Language." In: *International Journal of Applied Linguistics and English Literature* 9.5, pp. 110–118. DOI: http://dx.doi.org/10.7575/aiac.ijalel.v.9n.5p.110.

May, Stephen (2014). *The Multilingual Turn: Implications for SLA, TESOL and Bilingual Education*. New York: Routledge. 240 pp.

McEnery, Tony and Andrew Wilson (2001). *Corpus Linguistics An Introduction*. Edinburgh University Press. JSTOR: 10.3366/j.ctvxcrjmp.

Meurer, Paul (2012). "Corpuscle – a New Corpus Management Platform for Annotated Corpora." In: *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian*. Ed. by Gisle Andersen. Studies in Corpus Linguistics 49. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 31–49.

— (2020). "Designing Efficient Algorithms for Querying Large Corpora." In: *Oslo Studies in Language* 11.2, pp. 283–302. DOI: 10.5617/osla.8504.

Mohamed, Kabashi Alagab, A. O. Mohammed, H. M. Abdellatif, Anwer Abde-Elrhman, and Mohamed Mohammed Osman (2014). "Difficulties Facing First Year Students in Using English Prepositions." In.

Norton, Bonny and Kelleen Toohey (2011). "Identity, Language Learning, and Social Change." In: *Language Teaching* 44. DOI: 10.1017/S0261444811000309.

Odlin, T., S.M. Gass, and L. Selinker (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press.

Özışık, Cem (2014). "Identifying Preposition Errors of Turkish EFL Students." In: *ELT Research Journal* 3.2, pp. 59–69.

Park, Sunyoung (2023). "Corpus Analysis of L2 English Article Usage Patterns & Pedagogical Implications." In: *Cogent Education* 10.1. DOI: 10.1080/2331186X.2023.2197662.

Ragnhildstveit, Silje (2009). "Genustildeling og morsmålstransfer i norsk mellomspråk: En korpusbasert studie [Gender assignment and first language transfer i Norwegian interlanguage: A corpus-based study]." University of Bergen.

Richards, Jack C., ed. (1974). *Error Analysis: Perspectives on Second Language Acquisition*. 1st ed. Applied Linguistics and Language Study. London: Longman. 228 pp.

Ringbom, H. (1987). *The Role of the First Language in Foreign Language Learning*. Multilingual Matters Limited.

Rosén, Victoria and Koenraad De Smedt (2010). "Syntactic Annotation of Learner Corpora." In: *Systematisk, variert, men ikke tilfeldig*. Ed. by Hilde Johansen, Anne Golden, Jon Erik Hagen, and Ann-Kristin Helland. Novus forlag, pp. 120–132.

Saed, Hadeel Ali and Baderaddin Yassin (2017). "Jordanian Undergraduate Students' Use of English Prepositional Verbs: Analytical Study." In: *Canadian Social Science* 13.4, pp. 68–75. DOI: 10.3968/9379.

Saeed, Abdul, Sajid Hussain, and Surhan Fatima (2015). "Prepositional Errors in the Writings of Pakistani ESL Students." In: *International Journal of English Linguistics* 5.33. DOI: 10.5539/ijel.v5n3p120.

Schenck, Andrew (Dec. 2020). "Examining the Influence of Native and Non-Native English-speaking Teachers on Korean EFL Writing." In: *Asian-Pacific Journal of Second and Foreign Language Education* 5.1, p. 2. DOI: 10.1186/s40862-020-00081-3.

Schmid, Helmut (1994). "Probabilistic Part-of-Speech Tagging Using Decision Trees." In: *Proceedings of International Conference on New Methods in Language Processing, Manchester, UK*.

— (1995). "Improvements in Part-of-Speech Tagging with an Application to German." In: *Proceedings of the ACL SIGDAT-Workshop. Dublin, Ireland*.

Selinker, Larry (1972). "Interlanguage." In: *International Review of Applied Linguistics in Language Teaching* 10.1-4, pp. 209–232. DOI: 10.1515/iral.1972.10.1-4.209.

Shah, Muhammad Imran, Hafiz Muhammad Qasim, and Masroor Sibtain (2021). "Locus of Conjunctions in Academic Writing: A Corpus-Driven Approach to Developing Writing among EFL Learners." In: *PAKISTAN LANGUAGES AND HUMANITIES REVIEW* 5.II, pp. 432–443. DOI: 10.47205/plhr.2021(5-II)1.35.

Short, Deborah J and Shannon Fitzsimmons (2007). *Double the Work: Challenges and Solutions to Acquiring Language and Academic Literacy for Adolescent English Language Learners– A Report to Carnegie Corporation of New York*. New York: Carnegie Corporation of NewYork.

Siyanova, Anna and Norbert Schmitt (2007). "Native and Nonnative Use of Multi-Word vs. One-Word Verbs." In: 45.2, pp. 119–139. DOI: 10.1515/IRAL.2007.005.

Sorace, Antonella and Francesca Filiaci (2006). "Anaphora Resolution in Near-Native Speakers of Italian." In: *Second Language Research* 22.3, pp. 339–368. DOI: 10.1191/0267658306sr271oa.

Spada, Nina and Patsy M. Lightbown (2019). "Second Language Acquisition." In: *An Introduction to Applied Linguistics*. 3rd ed. London: Routledge, pp. 111–127.

Sultana, Kishwar (2018). "Common Syntactic Errors in English Composition Committed By Pakistani Learners Of Various Linguistic Backgrounds." MA thesis. Islamabad: NUML.

Suzuki, Yoko (2015). "The Uses of Get in Japanese Learner and Native Speaker Writing : A Corpus-based Analysis." In: *Komaba Journal of English Education, Department of English Language, The University of Tokyo, Komaba* 6, pp. 3–18.

Tarone, Elaine (1979). "Interlanguage as Chameleon." In: *Language Learning* 29.1, pp. 181–191. DOI: 10.1111/j.1467-1770.1979.tb01058.x.

— (2006). "Variability in Interlanguage Use: A Study of Style-Shifting in Morphology and Syntax." In: *Language Learning* 35, pp. 373–403. DOI: 10.1111/j.1467-1770.1985.tb01083.x.

Tenfjord, Kari (1983). "Systematisk, variert, tilfeldig: Ein kontrastivt basert analyse av mellomspråket til ei gruppe vietnamesiske ungdommar [Systematic, varied, arbitrary: A contrastively based analysis of the interlanguage of a group of Vietnamese youths]." University of Bergen.

Tenfjord, Kari, Paul Meurer, and Knut Hofland (2006). "The ASK Corpus - a Language Learner Corpus of Norwegian as a Second Language." In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*, pp. 1821–1824.

Tsimpli, Ianthi Maria and Maria Dimitrakopoulou (2007). "The Interpretability Hypothesis: Evidence from Wh-Interrogatives in Second Language Acquisition." In: *Second Language Research* 23.2, pp. 215–242. DOI: 10.1177/0267658307076546.

Wang, Lulu and Shaolong Liu (2022). "A Corpus-Based Study: A Comparison between Native Speakers and Chinese EFL Learners in Using Adjectives in Argumentative Writings." In: *Open Journal of Modern Linguistics* 12.02, pp. 158–178. DOI: 10.4236/ojml.2022.122013.

Yadav, Manoj (2014). "Role of Mother Tongue in Second Language Learning." In: *International Journal of Research* 1.11, pp. 572–582.

Yazdi, SS Hejazian and MJ Rezai (2015). "Language Learnability of the Argument Structures of English Transitivity Constructions by EFL Learners." In: *International Journal of English and Education.* 4.3, pp. 506–528.

Zafar, Yasin (2018). "OBSTRUCTION OF MOTHER'S TONGUE (URDU) IN THE SUBSEQUENT LANGUAGE LEARNING (ENGLISH) AMONG THE UNDERSTUDIES." In: *JOURNAL OF APPLIED LINGUISTICS AND TESOL* 1.4, pp. 39–48. DOI: https://www.jalt.com.pk/index.php/JALT/article/view/636/558.

Appendices

A   Frequency list of verbs with prepositions (no hapax)

With counts and percentages.

957 3.22 agree with
734 2.47 go to
266 0.89 accord to
264 0.89 depend on
210 0.71 lead to
209 0.70 think about
205 0.69 work for
203 0.68 deal with
203 0.68 pay for
201 0.68 come to
162 0.55 suffer from
158 0.53 disagree with
149 0.50 cause by
149 0.50 focus on
141 0.47 come from
136 0.46 look for
133 0.45 addict to
117 0.39 concentrate on
111 0.37 relate to
106 0.36 prepare for
106 0.36 think of
94 0.32 get from
93 0.31 care about
92 0.31 affect by
92 0.31 apply for
91 0.31 learn about
85 0.29 belong to
82 0.28 result in
81 0.27 worry about
80 0.27 know about
79 0.27 base on
79 0.27 sit in
76 0.26 think if
73 0.25 compare to
70 0.24 agree to
70 0.24 look at
68 0.23 take on
68 0.23 use for
66 0.22 rely on
64 0.22 come in
61 0.21 interfere with
59 0.20 work on
58 0.20 use to
54 0.18 contribute to
53 0.18 get in
51 0.17 ask for
51 0.17 go into
49 0.16 die from
48 0.16 cope with
45 0.15 prohibit in
40 0.13 adapt to
40 0.13 go for
38 0.13 apply to

38 0.13 compare with
37 0.12 affect to
36 0.12 die because
36 0.12 die of
35 0.12 go through
33 0.11 consist of
33 0.11 give by
33 0.11 go on
32 0.11 need for
31 0.10 cooperate with
31 0.10 take into
30 0.10 agree because
30 0.10 think for
28 0.09 benefit from
27 0.09 go in
27 0.09 refer to
26 0.09 escape from
26 0.09 influence by
26 0.09 prohibit at
25 0.08 happen to
24 0.08 adjust to
24 0.08 listen to
24 0.08 put in
23 0.08 bring about
23 0.08 look like
23 0.08 provide for
22 0.07 agree on
22 0.07 fail in
22 0.07 succeed in
21 0.07 search for
20 0.07 allow at
20 0.07 associate with
20 0.07 keep on
19 0.06 die in
19 0.06 forbid at
19 0.06 lack of
18 0.06 apply in
18 0.06 benefit to
18 0.06 refrain from
17 0.06 come for
17 0.06 get by
17 0.06 live on
17 0.06 vary from
16 0.05 agree about
16 0.05 appeal to
16 0.05 care for
16 0.05 effect on
16 0.05 make for
16 0.05 stick to
15 0.05 account for
15 0.05 complain about
15 0.05 concentrate in
15 0.05 disagree to

15 0.05 relate with
15 0.05 sit around
15 0.05 sit on
15 0.05 socialize with
14 0.05 cause of
14 0.05 deprive of
14 0.05 indulge in
14 0.05 owe to
14 0.05 seek for
13 0.04 believe in
13 0.04 cater to
13 0.04 connect with
13 0.04 dine at
13 0.04 equip with
13 0.04 result from
12 0.04 accustom to
12 0.04 avoid from
12 0.04 concern with
12 0.04 put on
12 0.04 take by
11 0.04 addict in
11 0.04 affect on
11 0.04 agree for
11 0.04 bring by
11 0.04 come across
11 0.04 fall in
11 0.04 focus in
11 0.04 get through
11 0.04 learn of
11 0.04 pass by
11 0.04 sit at
11 0.04 turn into
10 0.03 affect in
10 0.03 allow for
10 0.03 annoy to
10 0.03 bear in
10 0.03 call for
10 0.03 compose of
10 0.03 decide for
10 0.03 disagree about
10 0.03 excel in
10 0.03 follow by
10 0.03 get on
10 0.03 make of
10 0.03 oppose to
10 0.03 provide by
10 0.03 put into
10 0.03 speak of
10 0.03 suffer in
10 0.03 take for
9 0.03 bear with
9 0.03 beg for
9 0.03 bother by

9 0.03 concern about
9 0.03 depend to
9 0.03 die by
9 0.03 differ from
9 0.03 effect by
9 0.03 effect to
9 0.03 fascinate by
9 0.03 fit in
9 0.03 give in
9 0.03 go without
9 0.03 hear of
9 0.03 pass through
9 0.03 set by
9 0.03 subject to
8 0.03 acquire in
8 0.03 addict of
8 0.03 approve of
8 0.03 contain of
8 0.03 dedicate to
8 0.03 impose on
8 0.03 infringe on
8 0.03 require for
8 0.03 sit by
8 0.03 stand in
8 0.03 use of
7 0.02 abstain from
7 0.02 affect with
7 0.02 annoy for
7 0.02 arrive at
7 0.02 ask from
7 0.02 benefit for
7 0.02 carry on
7 0.02 come on
7 0.02 commit to
7 0.02 comply with
7 0.02 disagree for
7 0.02 get over
7 0.02 go beyond
7 0.02 go by
7 0.02 hope for
7 0.02 impact on
7 0.02 live for
7 0.02 look after
7 0.02 prohibit for
7 0.02 prohibit from
7 0.02 suffer for
6 0.02 act like
6 0.02 adhere to
6 0.02 agree in
6 0.02 aim at
6 0.02 blame for
6 0.02 care of
6 0.02 combine with

6 0.02 come at
6 0.02 concentrate to
6 0.02 consider for
6 0.02 depend upon
6 0.02 distract by
6 0.02 fight for
6 0.02 fill in
6 0.02 get at
6 0.02 interfere in
6 0.02 linger in
6 0.02 partake in
6 0.02 practice for
6 0.02 recommend for
6 0.02 stay around
6 0.02 strive for
6 0.02 take from
6 0.02 tend to
5 0.02 act in
5 0.02 addict by
5 0.02 advocate for
5 0.02 bring in
5 0.02 call on
5 0.02 compensate for
5 0.02 confine to
5 0.02 confront with
5 0.02 contaminate with
5 0.02 contribute in
5 0.02 disagree on
5 0.02 discriminate against
5 0.02 distract from
5 0.02 fit for
5 0.02 forbid for
5 0.02 forbid to
5 0.02 give on
5 0.02 hunt for
5 0.02 implement in
5 0.02 inhale into
5 0.02 insist on
5 0.02 negotiate with
5 0.02 object to
5 0.02 opt for
5 0.02 plan for
5 0.02 prohibit because
5 0.02 prohibit by
5 0.02 rob of
5 0.02 set in
5 0.02 suffocate by
4 0.01 abide by
4 0.01 accumulate in
4 0.01 acquire through
4 0.01 affect around
4 0.01 argue against
4 0.01 argue for
4 0.01 argue whether
4 0.01 assign to
4 0.01 bring into
4 0.01 bring on
4 0.01 bring up
4 0.01 bring with
4 0.01 cause from
4 0.01 come after
4 0.01 compete with
4 0.01 concentrate with
4 0.01 decrease in
4 0.01 decrease of
4 0.01 depend in
4 0.01 designate as
4 0.01 die for
4 0.01 die if
4 0.01 forbid by
4 0.01 frown upon
4 0.01 get after
4 0.01 overwhelm with

4 0.01 pertain to
4 0.01 provide with
4 0.01 realize about
4 0.01 require at
4 0.01 resort to
4 0.01 restrict for
4 0.01 result to
4 0.01 serve at
4 0.01 set for
4 0.01 stand for
4 0.01 suggest to
4 0.01 take at
4 0.01 take to
4 0.01 utilize by
4 0.01 walk along
4 0.01 work through
3 0.01 acquaint with
3 0.01 acquire by
3 0.01 affect for
3 0.01 allot for
3 0.01 annoy from
3 0.01 argue on
3 0.01 argue with
3 0.01 arise from
3 0.01 attribute to
3 0.01 bear from
3 0.01 benefit by
3 0.01 blame by
3 0.01 bombard with
3 0.01 care with
3 0.01 cling to
3 0.01 comprise of
3 0.01 condone in
3 0.01 depend whether
3 0.01 depend with
3 0.01 designate for
3 0.01 die at
3 0.01 die with
3 0.01 disapprove of
3 0.01 discuss for
3 0.01 disturb from
3 0.01 divert from
3 0.01 dream of
3 0.01 effect for
3 0.01 effect through
3 0.01 embark on
3 0.01 encourage among
3 0.01 encourage by
3 0.01 encourage for
3 0.01 encourage in
3 0.01 engage to
3 0.01 entitle to
3 0.01 exhale from
3 0.01 expect for
3 0.01 find out
3 0.01 full of
3 0.01 gain for
3 0.01 give up
3 0.01 go around
3 0.01 hinder to
3 0.01 hold in
3 0.01 impress from
3 0.01 improve by
3 0.01 improve on
3 0.01 include of
3 0.01 include with
3 0.01 inculcate in
3 0.01 irritate to
3 0.01 irritate with
3 0.01 kill at
3 0.01 lack in
3 0.01 leave by
3 0.01 leave with

3 0.01 live of
3 0.01 long for
3 0.01 look in
3 0.01 look into
3 0.01 look on
3 0.01 make at
3 0.01 make on
3 0.01 mean by
3 0.01 mean in
3 0.01 mean to
3 0.01 obsess with
3 0.01 occupy by
3 0.01 prepare in
3 0.01 prevail in
3 0.01 provide in
3 0.01 provide to
3 0.01 punish by
3 0.01 push for
3 0.01 qualify for
3 0.01 quit for
3 0.01 recommend by
3 0.01 recommend to
3 0.01 reflect on
3 0.01 regret for
3 0.01 require by
3 0.01 restrict to
3 0.01 sacrifice for
3 0.01 separate by
3 0.01 separate for
3 0.01 separate in
3 0.01 set up
3 0.01 spread on
3 0.01 spread over
3 0.01 succumb to
3 0.01 support on
3 0.01 support with
3 0.01 surround with
3 0.01 take of
3 0.01 take over
3 0.01 teach for
3 0.01 teach with
3 0.01 think as
3 0.01 think because
3 0.01 think by
3 0.01 think on
3 0.01 think without
3 0.01 turn on
3 0.01 understand about
3 0.01 understand by
3 0.01 understand in
3 0.01 understand with
3 0.01 walk behind
3 0.01 walk by
3 0.01 walk to
3 0.01 want after
3 0.01 want for
3 0.01 wrap in
3 0.01 write about
3 0.01 write to
2 0.01 absorb into
2 0.01 accept for
2 0.01 accept on
2 0.01 accept with
2 0.01 accommodate to
2 0.01 accompany by
2 0.01 accuse with
2 0.01 achieve by
2 0.01 achieve through
2 0.01 acknowledge of
2 0.01 acquire before
2 0.01 act for
2 0.01 addict on
2 0.01 addict with

2 0.01 admit by
2 0.01 admit in
2 0.01 adopt by
2 0.01 advertise on
2 0.01 affect because
2 0.01 affect of
2 0.01 affect than
2 0.01 afford for
2 0.01 aid in
2 0.01 align with
2 0.01 allocate for
2 0.01 allocate to
2 0.01 amount to
2 0.01 annoy of
2 0.01 answer for
2 0.01 appear on
2 0.01 applaud for
2 0.01 apply as
2 0.01 argue in
2 0.01 assemble of
2 0.01 attend about
2 0.01 avoid because
2 0.01 avoid if
2 0.01 avoid of
2 0.01 avoid to
2 0.01 bear for
2 0.01 bother with
2 0.01 breathe for
2 0.01 breathe on
2 0.01 breathe without
2 0.01 bring from
2 0.01 bring upon
2 0.01 burden by
2 0.01 burden with
2 0.01 carry by
2 0.01 carry out
2 0.01 categorize as
2 0.01 cause at
2 0.01 cause for
2 0.01 cause within
2 0.01 challenge for
2 0.01 challenge to
2 0.01 change by
2 0.01 change for
2 0.01 change from
2 0.01 change into
2 0.01 check in
2 0.01 come of
2 0.01 come outside
2 0.01 come through
2 0.01 come towards
2 0.01 comfort with
2 0.01 complain of
2 0.01 complain to
2 0.01 comply to
2 0.01 compromise for
2 0.01 compromise on
2 0.01 compromise with
2 0.01 conclude with
2 0.01 condemn in
2 0.01 consider by
2 0.01 consider from
2 0.01 consider in
2 0.01 consider of
2 0.01 consume by
2 0.01 consume in
2 0.01 contain about
2 0.01 contend with
2 0.01 control about
2 0.01 control in
2 0.01 count on
2 0.01 criticize by
2 0.01 deal to

2 0.01 decline because	2 0.01 emphasise on	2 0.01 find at	2 0.01 irritate without
2 0.01 decrease by	2 0.01 emphasize on	2 0.01 find by	2 0.01 like after
2 0.01 decrease if	2 0.01 engulf in	2 0.01 find from	2 0.01 like at
2 0.01 decrease with	2 0.01 enjoy at	2 0.01 forbid from	2 0.01 like during
2 0.01 depart from	2 0.01 enjoy for	2 0.01 forbid of	2 0.01 like with
2 0.01 deter from	2 0.01 enjoy of	2 0.01 forbid without	2 0.01 look around
2 0.01 develop by	2 0.01 enter by	2 0.01 gather at	2 0.01 look to
2 0.01 devote to	2 0.01 enter for	2 0.01 gather for	2 0.01 offend by
2 0.01 die after	2 0.01 enter through	2 0.01 give as	2 0.01 patronise by
2 0.01 die as	2 0.01 enter with	2 0.01 give from	2 0.01 pay after
2 0.01 die before	2 0.01 entice to	2 0.01 go about	2 0.01 pay like
2 0.01 die once	2 0.01 establish in	2 0.01 go against	2 0.01 pay on
2 0.01 differ in	2 0.01 establish on	2 0.01 go along	2 0.01 pay through
2 0.01 differ to	2 0.01 evade from	2 0.01 go because	2 0.01 prevent about
2 0.01 diffuse to	2 0.01 excite about	2 0.01 go of	2 0.01 prevent by
2 0.01 discourage at	2 0.01 exhaust after	2 0.01 go out	2 0.01 put out
2 0.01 discourage in	2 0.01 exhaust by	2 0.01 help at	2 0.01 recall of
2 0.01 discuss at	2 0.01 exhaust in	2 0.01 influence for	2 0.01 suffer at
2 0.01 discuss by	2 0.01 expand of	2 0.01 influence of	2 0.01 suffer because
2 0.01 discuss in	2 0.01 expect as	2 0.01 influence than	2 0.01 suffer by
2 0.01 disgust with	2 0.01 expect of	2 0.01 inform of	2 0.01 suffer if
2 0.01 earn after	2 0.01 experience about	2 0.01 infringe upon	2 0.01 suffer on
2 0.01 earn on	2 0.01 experience at	2 0.01 inhale as	2 0.01 suffer to
2 0.01 effect with	2 0.01 experience of	2 0.01 inhale for	2 0.01 suffocate with
2 0.01 eliminate from	2 0.01 experience outside	2 0.01 inhale of	2 0.01 think after
2 0.01 embarrass in	2 0.01 experience through	2 0.01 injure of	2 0.01 think around
2 0.01 emerge from	2 0.01 fall behind	2 0.01 irritate because	2 0.01 think at
2 0.01 emit through	2 0.01 fall on	2 0.01 irritate for	2 0.01 think with

B PoS tags (Treetagger set)

POS Tag	Description	Example
(empty tag)	HTML and other entities enclosed in angle brackets	
” “	single or double quotation marks	” ‘
(	left brackets	([{
)	right brackets	)] }
,	comma	,
$	currency symbols	$ £ €
#	hash (number sign)	#
:	dashes, ellipsis, underscore, (semi)colon	– … .. _ ; :
CC	coordinating conjunction	and
CD	cardinal number	1, one
CDZ	possessive numeral	one's
DT	determiner	the
EX	existential there	there is
FW	foreign word	d'hoevre
IN	preposition, subordinating conjunction	in, of, like
IN/that	that as subordinator	that
JJ	adjective	green
JJR	adjective, comparative	greener
JJS	adjective, superlative	greenest
LS	list marker	1)
MD	modal (verbs)	could, will, should, would
NN	noun, singular or mass	table
NNS	noun plural	tables
NNSZ	possessive noun plural	people's, women's
NNZ	possessive noun, singular or mass	year's, world's
NP	proper noun, singular	John
NPS	proper noun, plural	Vikings
NPSZ	possessive proper noun, plural	Boys', Workers'
NPZ	possessive noun, singular	Britain's, God's
PDT	predeterminer	both the boys
PP	personal pronoun	I, he, it
PPZ	possessive pronoun	my, his
RB	adverb	however, usually, naturally, here, good
RBR	adverb, comparative	better
RBS	adverb, superlative	best
RP	particle	give up
SENT	Sentence-break punctuation	. ! ?
SYM	Symbols (except for those listed above)	/ = *
TO	infinitive 'to'	togo
UH	interjection	uhhuhhuhh
VB	verb be, base form	be
VBD	verb be, past tense	was, were
VBG	verb be, gerund/present participle	being
VBN	verb be, past participle	been
VBP	verb be, present, non-3d person	am, are
VBZ	verb be, 3rd person sing. present	is
VH	verb have, base form	have
VHD	verb have, past tense	had
VHG	verb have, gerund/present participle	having
VHN	verb have, past participle	had
VHP	verb have, sing. present, non-3d	have
VHZ	verb have, 3rd person sing. present	has
VV	verb, base form	take
VVD	verb, past tense	took

VVG	verb, gerund/present participle	taking
VVN	verb, past participle	taken
VVP	verb, present, not 3rd person	take
VVZ	verb, 3rd person sing. present	takes
WDT	wh-determiner	which
WP	wh-pronoun	who, what
WPZ	possessive wh-pronoun	whose
WRB	wh-abverb	where, when
Z	possessive ending	's

C   PoS tag frequencies in Corpuscle

Freq	POS Tag
232389	NN
153548	
139347	IN
107529	DT
83712	NNS
81604	JJ
80380	RB
78526	PP
76573	SENT
65938	VV
52317	,
39568	CC
38605	MD
35014	VVP
30154	TO
29291	VBZ
21659	PP$
19712	VVG
16499	VVN
12430	VBP
11823	VB
10836	IN/that
10126	VVZ
8064	WRB
7295	VHP
7204	WP
6167	VH
5603	WDT
5543	JJR
5456	CD
5378	NP
5249	VVD
3749	EX
3522	RBR
3225	POS
2735	RP
2505	PDT
2454	VHZ
2347	JJS
1990	VHG
1470	VBD
1056	:
808	RBS
687	VBG
633	VBN
572	VHD
511	”
441	PPZ
418	“
408	FW
403	SYM
336	)
328	(
206	UH

96	NPS
63	NNSZ
63	WP$
59	LS
58	VHN
26	$
24	NNZ
3	CDZ
3	WPZ
2	NPZ