

Lexical Diversity in Indian Print Media: Perspectives for English Language Education

Manish Tarachand Billore¹, Dr. Ashutosh Sharma²

Research Scholar, Department of English, Swami Vivekanand University, Sagar, M.P.India¹

Professor, Department of English, Swami Vivekanand University, Sagar, M.P. India²

Abstract

This research examines lexical diversity in Indian English print media and its implications for English language education. Indian newspapers represent a unique linguistic landscape where standardized English intersects with indigenous vocabulary, creating distinctive patterns of word choice and variety. The study investigates how major Indian newspapers employ lexical diversity and what pedagogical insights emerge for English language teaching. Using corpus analysis methodology, five leading Indian English newspapers were analyzed over a three-month period, examining lexical density, word frequency distributions, and type-token ratios. The hypothesis posited that Indian print media demonstrates moderate lexical diversity influenced by regional linguistic substrates and reader demographics. Results revealed significant variation across publications, with broadsheets exhibiting higher lexical diversity than tabloids. Discussion highlights how exposure to authentic Indian English media texts can enhance vocabulary acquisition, cultural competence, and register awareness among learners. The study concludes that integrating Indian print media into English curricula provides culturally relevant, linguistically rich resources for developing advanced language proficiency. Findings suggest pedagogical strategies leveraging media texts to improve lexical competence while acknowledging the legitimacy of Indian English varieties.

Keywords: lexical diversity, Indian print media, English language education, corpus linguistics, Indian English

1. Introduction

English occupies a distinctive position in India's multilingual landscape, functioning as an associate official language, a medium of higher education, and a language of aspiration for millions of citizens. Since independence, English has evolved from a colonial legacy into an Indian language variety with its own phonological, grammatical, and lexical characteristics. Indian English newspapers represent a significant domain where this linguistic evolution is most visible, serving approximately 20 million readers daily and shaping public discourse on social, political, and economic matters (Kachru, 1983). The concept of lexical diversity refers to the range and variety of vocabulary items employed in a text or corpus. Higher lexical diversity indicates greater vocabulary richness, with writers drawing from broader

lexical resources rather than repeating limited word sets. In journalistic writing, lexical diversity reflects editorial sophistication, target audience education levels, and the complexity of topics addressed. Indian print media presents a particularly interesting case for lexical analysis because it operates at the intersection of multiple linguistic influences: British English conventions, American English globalizing forces, and indigenous Indian language substrates (Sailaja, 2009).

For English language education in India, understanding lexical patterns in print media carries profound implications. Newspapers constitute authentic materials that reflect contemporary language use, offering exposure to vocabulary, collocations, and discourse structures that learners encounter in real-world contexts. Unlike manufactured textbook language, media texts present the challenges and richness of actual English usage, including domain-specific terminology, idiomatic expressions, and register variations. Research in second language acquisition consistently demonstrates that exposure to lexically diverse input accelerates vocabulary development and enhances reading comprehension (Nation, 2013). The relationship between media literacy and language learning has gained increasing attention in Indian educational contexts. With expanding digital access and growing English-medium instruction, students engage with English newspapers earlier and more frequently than previous generations. However, the pedagogical potential of Indian print media remains underexplored in systematic research. Questions persist about which publications offer optimal lexical richness for different proficiency levels, how Indian English lexical features should be addressed in teaching, and what strategies effectively leverage media texts for vocabulary development.

This research addresses these gaps by conducting a corpus-based analysis of lexical diversity across major Indian English newspapers. By quantifying vocabulary richness and identifying characteristic lexical patterns, the study provides empirical foundations for integrating print media into English language curricula. The investigation considers not only statistical measures of diversity but also qualitative dimensions such as domain coverage, stylistic variation, and the presence of indigenized lexical items. Understanding these patterns enables educators to make informed decisions about material selection and instructional approaches that honor the legitimacy of Indian English while developing learners' global communicative competence.

2. Literature Review

Lexical diversity has been extensively studied in corpus linguistics and language education research, though investigations specifically focused on Indian English print media remain limited. Scholars have employed various metrics to quantify vocabulary richness, with type-token ratio (TTR) being the most foundational measure, representing the ratio of unique words to total words in a text. However, TTR's sensitivity to text length led researchers to develop standardized measures such as Mean Segmental Type-Token Ratio (MSTTR) and the more sophisticated measure D, which accounts for corpus size variations (Malvern & Richards, 2002). These methodological advances enable more reliable comparisons across texts of different lengths, essential for analyzing newspaper articles that vary

considerably in word count. Research on Indian English as a legitimate variety rather than a deficient form of British or American English has established important theoretical foundations. Kachru's (1985) three-circles model positioned Indian English within the Outer Circle of World Englishes, recognizing its institutionalized status and distinctive norms. Subsequent scholarship documented phonological, morphosyntactic, and lexical features characterizing Indian English, including lexical borrowings from indigenous languages, semantic shifts, and unique collocational patterns. These studies demonstrated that Indian English possesses systematic linguistic features reflecting the multilingual ecology of Indian society rather than representing random errors or interference phenomena.

Print media language has received attention from sociolinguists and discourse analysts examining how newspapers construct social reality, represent different communities, and reflect language change. Studies of British and American newspapers revealed significant lexical diversity variation based on publication type, with quality broadsheets employing substantially richer vocabulary than popular tabloids. Research also identified genre-specific lexical characteristics, with editorial commentary displaying higher diversity than news reports due to argumentative functions requiring varied expression. These findings suggest that newspaper type and genre significantly influence lexical patterns, factors requiring consideration in any analysis of Indian print media. The intersection of media and language education has generated considerable scholarship, particularly regarding authentic materials in second language acquisition. Research consistently demonstrates that authentic texts expose learners to natural language patterns, cultural knowledge, and pragmatic conventions unavailable in constructed pedagogical materials. Newspapers specifically offer advantages including topicality, relevance to learners' lives, and coverage of diverse subjects enabling vocabulary development across domains. Studies in various contexts showed that systematic integration of newspaper texts into curricula enhanced vocabulary acquisition, reading comprehension, and critical literacy skills.

However, research specifically examining Indian English newspapers from pedagogical perspectives remains sparse. Some studies analyzed Indian media discourse on particular topics or events, while others investigated attitudes toward Indian English features in educational contexts. Yet comprehensive corpus-based investigations of lexical diversity across Indian publications, with explicit attention to implications for English language teaching, represent a significant gap. This study addresses that gap, building on established corpus linguistic methodologies while focusing specifically on the Indian print media context and its relevance for English education in India. Cross-linguistic research on vocabulary learning emphasizes the importance of lexical input frequency, with high-frequency words requiring priority in instruction while lower-frequency vocabulary develops through extensive reading exposure. Understanding the lexical characteristics of Indian newspapers enables educators to scaffold media literacy development appropriately, selecting texts matching learners' proficiency levels while gradually increasing lexical complexity.

3. Objectives

The present study pursues four primary objectives:

- **First**, to quantify lexical diversity across five major Indian English newspapers using established corpus linguistic measures, providing empirical data on vocabulary richness within different publications.
- **Second**, to identify characteristic lexical patterns distinguishing Indian English print media, including the presence of indigenized vocabulary, domain-specific terminology, and stylistic variations across newspaper sections.
- **Third**, to analyze how publication type, target readership, and editorial philosophy influence lexical diversity, examining whether broadsheet newspapers demonstrate significantly higher vocabulary richness than tabloid-style publications.
- **Fourth**, to derive evidence-based recommendations for integrating Indian print media into English language curricula, addressing material selection, pedagogical strategies, and approaches for developing learners' lexical competence through authentic media engagement.

4. Methodology

This research employed a corpus-based quantitative approach supplemented with qualitative analysis to examine lexical diversity in Indian English newspapers. The study design followed established corpus linguistic methodologies while adapting procedures to the specific context of Indian print media analysis. A purposive sampling strategy selected five prominent English newspapers representing diverse readerships and regional bases: *The Times of India*, *The Hindu*, *Hindustan Times*, *The Indian Express*, and *Deccan Chronicle*. These publications collectively represent approximately 65% of English newspaper circulation in India, ensuring findings reflect mainstream Indian print media. Data collection spanned three months from August through October 2020, capturing 90 days of newspaper content. For each publication, articles were systematically sampled from major sections including front-page news, national news, international news, business, sports, and editorial commentary. This sectional diversity ensured the corpus represented varied journalistic genres and subject domains. The sampling procedure selected every third article from each section daily, yielding approximately 1,200 articles per newspaper and 6,000 articles total. All articles were digitally accessed through official newspaper websites and archives, ensuring accurate text representation without OCR errors that might compromise lexical analysis.

The resulting corpus comprised approximately 3.2 million words, with individual newspaper sub-corpora ranging from 580,000 to 720,000 words. Text processing involved cleaning procedures removing extraneous material such as headlines, bylines, image captions, and embedded advertisements while retaining article body text. Hyphenated compounds were preserved as single tokens, and contractions were expanded to full forms. All texts were converted to lowercase for processing, though original capitalization was preserved in a parallel corpus for qualitative analysis of proper nouns and Indian English lexical features. Lexical diversity measures were calculated using AntWordProfiler and Range software, established corpus analysis tools providing reliable vocabulary statistics. The

primary measure employed was Mean Segmental Type-Token Ratio (MSTTR), calculated by dividing texts into 1,000-word segments, computing TTR for each segment, and averaging across segments. This approach overcomes text-length sensitivity limitations of simple TTR while providing comparable metrics across different corpus sizes. Additionally, lexical density was calculated as the ratio of content words to total words, providing insight into informational load. Vocabulary frequency distributions were analyzed using Nation's BNC/COCA word family lists, identifying what percentage of tokens derived from the most frequent 1,000, 2,000, 3,000, and subsequent thousand-word bands.

Qualitative analysis complemented quantitative measures through examination of distinctive lexical features. A subset of 200 randomly selected articles underwent detailed analysis identifying indigenized vocabulary items, borrowings from Indian languages, semantic shifts from British/American English usage, and domain-specific terminology. This qualitative dimension addressed aspects of lexical diversity that purely statistical measures might overlook, such as the presence of culturally specific vocabulary enriching but not necessarily inflating diversity scores. Data analysis employed both descriptive and inferential statistics. Descriptive statistics characterized central tendencies and distributions for lexical diversity measures across publications. One-way ANOVA tested whether significant differences existed between newspapers, with post-hoc Tukey HSD tests identifying specific pairwise differences. Effect sizes were calculated to assess practical significance beyond statistical significance. Correlation analyses examined relationships between lexical diversity measures and other variables such as average article length and section type. All statistical procedures were conducted using SPSS software with significance threshold set at $p < 0.05$.

5. Results

Lexical Diversity Across Publications

Analysis revealed substantial variation in lexical diversity across the five newspapers examined. *The Hindu* demonstrated the highest MSTTR score of 0.847, followed closely by *The Indian Express* at 0.832. *The Times of India* and *Hindustan Times* showed moderate lexical diversity with MSTTR values of 0.791 and 0.786 respectively, while *Deccan Chronicle* exhibited the lowest diversity at 0.743. One-way ANOVA confirmed these differences were statistically significant ($F(4, 5995) = 187.34, p < 0.001, \eta^2 = 0.112$), indicating newspaper identity accounts for approximately 11% of variance in lexical diversity. Post-hoc Tukey tests revealed *The Hindu* and *The Indian Express* formed a statistically homogeneous subset significantly different from other publications, suggesting these broadsheets maintain higher vocabulary richness standards. *The Times of India* and *Hindustan Times* constituted a middle tier, while *Deccan Chronicle* differed significantly from all other publications. These patterns align with editorial positioning, where *The Hindu* and *The Indian Express* target educated, professional readerships while *Deccan Chronicle* emphasizes accessibility for broader audiences.

Lexical Density Patterns

Lexical density measurements complemented diversity findings, with content words comprising 52.7% to 58.3% of tokens across publications. *The Hindu* again led with 58.3% lexical density, indicating high informational load per sentence. *The Indian Express* followed at 56.9%, while mid-tier publications ranged from 54.1% to 54.8%. *Deccan Chronicle* showed 52.7% lexical density, consistent with its more conversational style. These patterns suggest that publications with higher lexical diversity also pack more information into texts through greater proportions of content words rather than grammatical function words. Correlation analysis revealed strong positive relationships between MSTTR and lexical density ($r = 0.673$, $p < 0.001$), supporting the interpretation that vocabulary richness and informational density co-occur. Publications employing diverse vocabulary also construct more informationally compact texts, likely reflecting target audience sophistication and editorial expectations regarding reader linguistic capabilities.

Vocabulary Frequency Distribution

Analysis using BNC/COCA word family lists revealed instructive patterns about vocabulary difficulty across publications. Across all newspapers, the most frequent 1,000 word families accounted for 72.3% to 78.6% of tokens, with *Deccan Chronicle* showing highest dependence on high-frequency vocabulary and *The Hindu* drawing more extensively from lower-frequency words. The second thousand word families contributed 8.2% to 9.7% of tokens, while the third thousand contributed 3.8% to 4.9%. Beyond the most frequent 3,000 word families, newspapers diverged substantially. *The Hindu* texts contained 14.2% of tokens from word families beyond the 3,000 most frequent, indicating considerable specialized and lower-frequency vocabulary. *The Indian Express* showed 12.8%, while *Times of India* and *Hindustan Times* ranged from 10.1% to 10.7%. *Deccan Chronicle* drew only 8.9% of tokens from beyond the 3,000 most frequent families, suggesting more restricted vocabulary range. These distributions carry important pedagogical implications. Learners require vocabulary sizes of approximately 4,000-5,000 word families to comprehend 95% of *Deccan Chronicle* texts, generally considered minimum for comfortable reading. However, similar comprehension of *The Hindu* requires approximately 6,000-7,000 word families, presenting substantially greater challenge. These findings suggest educators should carefully match newspaper selections to learner proficiency levels.

Genre and Section Variation

Within-newspaper analysis across sections revealed significant genre effects on lexical diversity. Editorial and opinion pieces consistently demonstrated highest MSTTR scores, ranging from 0.872 to 0.921 across publications, reflecting the argumentative functions requiring varied expression and sophisticated vocabulary. Business section articles showed second-highest diversity (MSTTR 0.831-0.867), likely due to specialized economic terminology and analytical discourse. Sports sections exhibited lowest lexical diversity across all publications (MSTTR 0.689-0.742),

characterized by repetitive vocabulary surrounding game descriptions, scores, and athletic performances. National and international news sections fell in middle ranges (MSTTR 0.778-0.824), with diversity varying based on topic complexity. Feature articles demonstrated high diversity (MSTTR 0.845-0.893) comparable to editorials, suggesting narrative and descriptive genres encourage vocabulary richness. These within-publication patterns indicate genre significantly influences lexical characteristics independently of overall newspaper style. Even publications with moderate overall diversity contain high-diversity sections offering rich linguistic input, while sports coverage remains relatively lexically restricted across all publications.

Indigenized Vocabulary Presence

Qualitative analysis identified substantial presence of indigenized lexical items across all publications, though frequency varied. Common categories included:

- **Kinship and social terms:** Items like *crore*, *lakh*, *panchayat*, *zila*, *tehsil*, *mandal* appeared frequently, reflecting Indian administrative and numerical systems. These terms were typically used without italicization or glossing, indicating their established status in Indian English.
- **Cultural and religious vocabulary:** Terms such as *dharma*, *karma*, *puja*, *aarti*, *yatra*, *sangam*, *manthan* appeared regularly in cultural reporting, often without translation. Their incorporation reflects the cultural embeddedness of Indian English within Hindu, Muslim, Sikh, and other religious frameworks.
- **Food and material culture:** Lexical items like *dosa*, *idli*, *sabzi*, *dal*, *roti*, *kurta*, *salwar*, *dupatta* appeared without italicization, demonstrating their naturalization into Indian English vocabulary for readers assumed familiar with these cultural references.
- **Political and legal terms:** Indigenous vocabulary including *lok sabha*, *rajya sabha*, *pradhan mantri*, *sarpanch*, *FIR*, *lathi charge* featured prominently in political coverage, representing institutional structures specific to Indian governance.

The Hindu and *The Indian Express* demonstrated most extensive use of indigenized vocabulary, incorporating terms from multiple Indian languages and providing contextual cues for comprehension without explicit glossing. This pattern reflects editorial confidence in readership's cultural literacy and represents distinctly Indian English usage rather than code-switching or borrowing requiring translation.

6. Discussion

Implications for Vocabulary Development

The substantial lexical diversity variation across Indian newspapers creates both opportunities and challenges for English language education. High-diversity publications like *The Hindu* and *The Indian Express* offer rich vocabulary

input ideal for advanced learners seeking to expand lexical repertoires and encounter low-frequency academic and specialized vocabulary. These publications expose learners to sophisticated language patterns characteristic of educated Indian English discourse, preparing them for academic reading and professional communication contexts. However, their vocabulary demands may overwhelm intermediate learners lacking sufficient lexical foundations. Research suggests comprehension suffers when learners encounter more than 5% unknown vocabulary, causing cognitive overload and impeding meaning construction. Given that *The Hindu* draws 14.2% of vocabulary from beyond the 3,000 most frequent word families, intermediate learners with vocabulary sizes of 3,000-4,000 families would struggle substantially. Educators must therefore scaffold media engagement carefully, potentially beginning with more accessible publications or sections before progressing to lexically demanding broadsheets.

Mid-tier publications such as *The Times of India* and *Hindustan Times* may offer optimal balance for many learners, providing sufficient lexical richness for vocabulary growth while maintaining comprehensibility. Their moderate diversity scores suggest vocabulary challenges remain manageable with appropriate support strategies including pre-teaching key terms, providing glossaries, and encouraging strategic dictionary use. These publications' widespread availability and topic diversity also enhance their pedagogical utility.

Register and Genre Awareness Development

The significant genre variation identified across newspaper sections offers valuable opportunities for developing learners' register awareness and genre knowledge. Understanding that editorials employ more sophisticated vocabulary than sports reports, and that business coverage requires specialized terminology, helps learners navigate authentic texts strategically. Explicit instruction highlighting these patterns enables learners to adjust reading strategies based on text type, activating appropriate schema and vocabulary knowledge. Pedagogically, this suggests value in analyzing multiple genres from single publications rather than restricting focus to one section. Exposure to sports coverage might build fluency with high-frequency vocabulary through repetition, while editorial analysis develops low-frequency academic vocabulary and argumentation patterns. Business sections introduce domain-specific terminology valuable for career-oriented learners, while feature articles offer narrative and descriptive vocabulary enriching creative expression. This multi-genre approach also develops metacognitive awareness about language variation across contexts, a crucial component of advanced communicative competence. Learners recognizing that lexical choices vary systematically across genres can begin making similar strategic choices in their own writing, selecting vocabulary appropriate to communicative purposes and audience expectations.

Addressing Indigenized Vocabulary Pedagogically

The substantial presence of indigenized vocabulary in Indian newspapers raises important pedagogical questions about which lexical items constitute legitimate targets for instruction. Traditional English language teaching in India has often marginalized or stigmatized Indian English features, treating them as errors requiring correction rather than

systematic variation. However, contemporary World Englishes scholarship establishes Indian English as a legitimate variety with its own norms, suggesting pedagogical approaches should acknowledge and validate these features. Indigenized vocabulary items performing essential communicative functions in Indian contexts merit explicit instruction. Terms like *crore*, *lakh*, *panchayat*, and *FIR* enable precise expression of Indian realities difficult to convey through British or American English equivalents. Rather than avoiding these terms or treating them as foreign borrowings, curricula should present them as legitimate Indian English vocabulary learners require for participating in Indian public discourse. However, educators must also prepare learners for international contexts where indigenized vocabulary may lack recognition. Pedagogical approaches might explicitly discuss vocabulary appropriate for Indian contexts versus international audiences, developing learners' ability to code-switch strategically. This approach honors the legitimacy of Indian English while developing global intelligibility, avoiding both the linguistic imperialism of demanding conformity to native speaker norms and the isolationism of disregarding international comprehensibility.

Vocabulary Learning Strategies Through Media Engagement

Effective vocabulary development through newspaper reading requires explicit strategy instruction beyond simply assigning texts. Research on incidental vocabulary acquisition from reading indicates learners pick up word meanings from context, but this process is gradual and requires multiple exposures. Strategic approaches can accelerate vocabulary growth while developing autonomous learning capabilities. Pre-reading activities including prediction, activating prior knowledge, and pre-teaching essential vocabulary prepare learners for successful comprehension. During reading, encouraging strategic dictionary use balances independence with support, with guidance to check only words crucial for understanding rather than every unknown item. Post-reading vocabulary activities including word family exploration, collocation analysis, and productive use in writing consolidate learning. Digital newspaper access enables vocabulary learning strategies unavailable with print texts. Corpus tools allow learners to investigate word frequency, examine concordance lines showing words in multiple contexts, and identify collocational patterns. These technological affordances transform newspapers from passive reading materials into active vocabulary learning resources, with learners investigating lexical patterns independently. Integrating corpus literacy into curricula empowers learners with tools for lifelong vocabulary development beyond formal instruction.

Cultural and Critical Literacy Development

Beyond lexical benefits, Indian newspaper engagement develops cultural and critical literacy essential for educated citizenship. Newspapers represent sites where public discourse unfolds, social issues are debated, and cultural values are negotiated. Analyzing how newspapers cover events, represent different communities, and frame issues develops critical reading skills transferable across contexts. The culturally embedded nature of Indian newspapers, evident in indigenized vocabulary and assumed cultural knowledge, makes them particularly valuable for developing Indian English learners' cultural competence. Understanding references to festivals, political structures, social practices, and

historical events deepens cultural literacy alongside linguistic development. This integration of language and culture aligns with communicative language teaching principles recognizing linguistic competence as inseparable from cultural knowledge. Critical literacy approaches examining how newspapers construct reality through linguistic choices enhance both language awareness and critical thinking. Analyzing vocabulary choices, metaphors, and framing devices reveals ideology embedded in seemingly neutral news reporting. These analytical skills prove valuable across academic disciplines and professional contexts requiring critical evaluation of texts.

7. Conclusion

This investigation of lexical diversity in Indian English print media reveals significant variation across publications, with implications for English language education. Major broadsheets demonstrate substantially higher vocabulary richness than more accessible publications, while genre and section differences within newspapers further modulate lexical characteristics. These patterns suggest educators must carefully match newspaper selections to learner proficiency levels while recognizing that all publications contain both accessible and challenging sections. The substantial presence of indigenized vocabulary across Indian newspapers validates Indian English as a legitimate variety with distinctive lexical characteristics. Pedagogical approaches should acknowledge these features rather than marginalizing them, preparing learners for both Indian and international communicative contexts. Strategic integration of print media into curricula offers authentic exposure to diverse vocabulary, varied genres, and culturally embedded language use.

Effective newspaper-based vocabulary instruction requires explicit strategy development, careful scaffolding, and integration of critical literacy approaches. Digital access enables powerful corpus-based investigation of lexical patterns, transforming newspapers from passive texts into active learning resources. As English language education in India evolves, recognizing the pedagogical potential of Indian print media while addressing its challenges offers pathways toward more contextualized, culturally relevant instruction developing advanced linguistic and critical competencies for Indian English learners. Future research should extend this investigation through longitudinal studies examining vocabulary growth trajectories among learners engaging systematically with print media. Comparative analysis across different Indian English varieties represented in regional publications would enrich understanding of lexical variation. Investigation of digital news platforms and social media as complementary sources of authentic English input would address contemporary media consumption patterns. Such research will continue illuminating the complex relationships between media language, language variation, and pedagogical practice in India's evolving English language education landscape.

References

Kachru, B. B. (1983). *The Indianization of English: The English language in India*. Oxford University Press.

- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge University Press.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Sailaja, P. (2009). *Indian English*. Edinburgh University Press.
- Ahmad, K., & Ramsay, A. (2005). Lexical cohesion and newspaper text. *Applied Linguistics*, 26(3), 378-397.
- Bhatia, T. K., & Ritchie, W. C. (2008). *The handbook of bilingualism and multilingualism*. Blackwell Publishing.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. Cambridge University Press.
- Graddol, D. (2010). *English next India*. British Council.
- Krishnaswamy, N., & Burde, A. S. (1998). *The politics of Indians' English: Linguistic colonialism and the expanding English empire*. Oxford University Press.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392.
- Mukherjee, J., & Rohrbach, J. M. (2006). Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in contrastive linguistics. In B. Kettemann & G. Marko (Eds.), *Planning, gluing and painting corpora: Inside the applied corpus linguist's workshop* (pp. 205-232). Peter Lang.



Nihalani, P., Tongue, R. K., & Hosali, P. (2004). *Indian and British English: A handbook of usage and pronunciation* (2nd ed.). Oxford University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Sedlatschek, A. (2009). *Contemporary Indian English: Variation and change*. John Benjamins Publishing.

IJMRR