

## CLEFT CONSTRUCTIONS IN BRITISH AND PAKISTANI ENGLISH NEWSPAPERS: A COMPARATIVE ANALYSIS

**Uswa Shahid**

PhD Scholar, Department of English, University of Gujrat, Gujrat

**Behzad Anwar**

Associate Professor, Department of English, University of Gujrat, Gujrat

**Muhammad Waqar Amin**

Lecturer in English, Govt. Graduate Murray College, Sialkot

### **Abstract**

*This article provides a detailed analysis of cleft constructions (i.e., it-clefts, wh-clefts, and all-clefts) in health-related reports of Pakistani and British journalistic English. Based on a comparable corpus of fifty texts (25 Pakistani, 25 British) taken from newspaper articles (total 178096 words), a mixed-methods approach has been adopted with the aim of unearthing cross-cultural patterns related to syntax and their rhetorical purposes. Data were extracted over 12-week time period (March–May 2025) by automated Python scripts, independently validated (Cohen's  $\kappa = 0.87$ ). Texts were preprocessed using NLTK, and segmented for pattern extraction; the quantitative frequencies were normalised per 1000 words and analysed by chi-square tests using SPSS 27, and the qualitative discourse functions were coded in NVivo 14. Quantitatively, British newspapers use clefts at a frequency approximately double as compared to Pakistani newspapers (2.05 vs. 1.09 per 1000 words;  $\chi^2(1)=8.23$ ,  $p<.01$ ), as well as for it-clefts ( $p=.006$ ) and wh-clefts ( $p=.024$ ). In qualitative terms, Pakistani clefts are mostly explanatory (48.4%), in accordance with the didactic focus on clarity, while British journalists resort to clefts more often for contrast (20.0%) and emotional appeal (12.1%), emphasising narrative and (linguistic) persuasive purposes. Both corpora, however, exploit emphasis clefts at similar frequencies ( $\approx 27-30\%$ ) highlighting their cross linguistic nature in emphasizing focal items. These findings reveal the interplay between editorial policy and audience expectation as reflected in syntactic preferences in global English journalism. The Pakistani health reporting also privileges a pedagogical framing, whereas the British outlets deploy the rhetorical versatility of the cleft as a persuasive vehicle to affect emotion and comparison for their readers. Methodologically, by combining automated extraction, manual validation and mixed-methods coding, our approach to building the corpus provides a strong corpus-based framework for future research in syntactic variation. This research both adds to the theory of World Englishes and discourse analysis and points the way to further work on digital and broadcast media in order to monitor emergent editorial practices.*

**Keywords:** Cleft constructions, Journalistic discourse, Newspapers, British English, World Englishes. Pakistani English

### **1. Introduction**

The cleft construction helps us bring attention to particular parts of our sentences through direct focus (Standing et al., 2021). These sentence patterns push information forward by selecting and placing new content with a clear focus or opposition in mind for direct communication. Cleft sentences work in three main types in English texts: It was the prime minister who announced the policy (it-cleft), What needs focus is transparency (wh-cleft) and All she needed was peace (all-cleft). People use these word patterns in everyday speaking and writing to organize the flow of their messages in any discourse.

Cleft constructions serve as important strategic tools for media reporters since they format information in newspaper content. Through cleft sentences journalists organize intricate details and lead readers to particular parts of their content. The style of cleft structures works well in opinion writing or human-interest writing to build force and relate emotions between various points of view. Examining how journalists use cleft constructions in their work helps us

understand how they structure their reports as well as how readers react to the presented information (Kwon et al., 2018).

This research analyses the number and application of cleft structures in British and Pakistani English newspapers. Our research targets these areas due to their special social and cultural environment and press industry characteristics. The UK, as a native English-speaking country, offers a well-established tradition of English journalistic writing. Pakistan shows the characteristics of a post-colonial society which uses English as a secondary language yet gives it high-status positions in institutions (Rahman, 2020). Pakistani newspapers written in English target their content to educated urban readers through a language mix of native and local linguistic practices. Examining cleft constructions in these two settings brings us valuable insights into their different usage.

Studies on cleft structures remain scarce when looking at their use in news publications and between different national settings. This research study creates a special collection of news articles to investigate differences in cleft construction use between Pakistani and British news outlets. All pieces only cover health topics and cleft palate care making their content similar between both sets of documents. The analysis uses authoritative sources that contain news reporting with health information plus reports on humanitarian efforts.

The research project combines machine-based regular expression finding with human review to find cleft construction examples throughout its target news articles. Researchers divide identified cleft sentences into *it*-clefts, *wh*-clefts, and *all*-clefts before studying their number of occurrences and contexts in each newspaper group (Yousefi & Esfandiari, 2024). The research examines how cleft sentences deliver emphasis, present focus areas, create emotional appeal and help identify differences.

Cleft constructions gained selection because they have specific syntax that makes them useful for multiple purposes (Seitanidi et al., 2024). The normal verbalization pattern of English sentences changes through cleft structure to direct readers toward important information in relaxed and professional writing situations. Journalists use these specific sentence structures in their accredited pieces because they make information easier to read while increasing its influence on readers. Clefts in Pakistani journalism may help narrate stories using techniques specific to the nation's tradition due to differences in British and Pakistani reporting standards. Through this study World Englishes research becomes apparent as authors from separate regions utilize English syntax which matches their unique cultural backgrounds (León et al., 2018). Our work checks if Pakistani and British news outlets have different cleft use patterns which English speakers usually consider informal or optional. The research assesses if Pakistani newspaper clefts change because of second language impact from Urdu or regional languages or because of global media writing standards.

### **1.1 Research Questions**

This study aims to answer the following research questions:

1. How frequently do cleft constructions occur in British and Pakistani English newspapers?
2. What discourse functions do these clefts serve in British and Pakistani English newspapers?
3. How can syntactic variations in journalistic English help to understand different English varieties in terms of syntactic choices?

### **2. Literature Review**

Syntacticians together with discourse analysts have continuously studied cleft constructions because these structures bring distinctive features that direct how information receives focus (Dorgeloh et al., 2022). The syntactic arrangement of cleft constructions serves essential

purposes because they enable users to restructure sentences actively for altering focus on specific elements. This review of literature shows how theoretical viewpoints relate to published research along with the conceptual use of cleft constructions within journalistic texts and comparison studies.

### **2.1 Cleft Constructions and Typology**

The conventional definition of cleft constructions indicates syntactic patterns which divide unified propositions into two clauses to highlight specific elements (Shipova, 2024). English uses cleft constructions as a syntactical tool that applies structural marking to focus one sentence element. English language has three main classification groups of clefts which include it-clefts and wh-clefts (also called pseudo-clefts) alongside all-clefts. The grammatical construction of it-clefts consists of "It + be + focused element + clause" pattern that appears in statements like "The minister who made the announcement" was "It." Through this structure the construction creates a highlighted subject by removing it from its natural sentence placement. The structure of a wh-cleft also known as a pseudo-cleft begins with "Wh-" followed by a clause and a linker connecting to the final focused element ("What she needs is help"). The latter part of these constructions serves two functions which include explanation and contrastive comparison in discourse. The use of "all" in all-clefts represents one form of wh-cleft which places selected clauses under emphasis ("All he wanted was peace"). Clefts function within discourse to create specific information management structures. From his perspective clefts achieve dual functions by emphasizing subject matter while assisting in topic flow and creating contrast effects and reference ambiguity solution. According to his theoretical work clefts act as structural units that demonstrate the interplay between syntax and discourse pragmatics since they frequently signal new and contrasting information elements in textual contexts.

### **2.2 Theoretical Perspectives on Clefting**

Modern research about clefts exists across various syntactic approaches (Wilson & Stevens, 2016) Transformational-generative grammar in its initial phase analyzed cleft constructions as outputs of transformational rules which modified standard word order (Hussain et al., 2023). During the subsequent stage of information structure theories which adopted Halliday's view clefts functioned as discourse focus instruments that establish thematic prominent elements. According to Halliday (1994), clefts enable maximum effectiveness in Theme-Rheme organization because they place topics or themes before new important information or rhemes. Writers can use clefts to enhance their message effectiveness through strategic placement of discourse components. Corpus-based and usage-based models have started analyzing clefts using empirical data to demonstrate their methods of functioning within different registers and genres.

### **2.3 Corpus-Based Studies on Clefts**

Empirical research of cleft constructions has shifted in a significant way because of corpus linguistics. The analysis represents a foundational research on English clefts where he tracked their distribution patterns and functional discourse roles in language use. Through his research he demonstrated that cleft constructions appear infrequently yet demonstrate rich pragmatic qualities that change based on register. Studies show it-clefts occur most often in spoken and informal writing because they serve as an effective focus-marking mechanism in such contexts. Cleft structures emerge in specific ways across different registers when they studied the grammar of English through corpus analysis (Biber et al. 1999). It-clefts frequently appeared in conversations and fiction yet academic writing showed infrequent use of these structures according to their research (Bourgoin et al., 2022). Standard written arguments demonstrated increased usage of wh-clefts because their explanatory function makes them appropriate for

formal writing. Research demonstrates that cleft constructions function as both linguistic decisions related to syntax and discourse choices which stem from genre-particular communication needs. Cleft structures by analyzing native and learner corpora through their investigations of cleft structures. The authors discovered that non-native speakers showed minimal usage of cleft constructions especially *wh*-clefts which they attributed to possible difficulties with syntax and unfamiliarity with the forms. Research findings concerning cleft constructions remain applicable to the current investigation because Pakistani English shows L2 status and its writers possibly demonstrate syntax distinct from native speakers.

#### **2.4 Cleft Constructions in Media and Journalism**

The scholarly literature contains many studies about cleft constructions within spoken and literary discourse yet the same literature lacks extensive research on their usage in journalistic papers. The hybrid role of newspaper language makes it an optimal research site for cleft investigation because it both delivers information facts and guides reader understanding. In his critical discourse analysis of media concluded that newspaper syntactic choices always carry some form of bias. Cleft constructions have distinct uses in creating both attribution assignments and responsibility reinforcement as well as achieving emphasis which media writers employ to mold public perception (Purcell 2020). The sentence "It was the government that delayed the response" uses clefts to show blame even though a neutral expression of this fact exists. Throughout his evaluation of media articles, Fairclough (1995) identified how clefts along with passives and nominalizations served to guide reader attention yet conceal actual agents. Rhetorical tools such as clefts exist within news reporting since journalists intentionally use them to modify reader perception. The work presented by Fairclough (1995) concentrated on ideological interpretations through qualitative methods although it did not provide a thorough syntactic analysis of clefts. Corpus-based methods are used to study news discourse because syntactic patterns create specific journalistic stances and tones (Sánchez-Ramírez & León, 2018). Their analysis backs the adoption of corpus methods to study linguistic constructions called clefts which influence narrative delivery in stories. Clefts express minimal interest in news corpora and especially so when examined in non-native English contexts.

#### **2.5 Clefts in World Englishes and South Asian English**

World Englishes as a scholarly field led to increased research about syntactic variations between different English language varieties. The linguistic influence of Urdu together with regional languages shapes the post-colonial Pakistani English which shows distinct syntactic and lexical patterns (Halo et al., 2024). Numerous studies analyze Pakistani English phonological and lexical variations but research on its syntactic patterns particularly cleft constructions remains limited. The evolution of second-language English varieties results in the formation of separate norms (Pérez-Llantada, 2012). The conventional language used by speakers often shows itself through syntax that integrates native discourse customs, educational learning methods or translated language patterns. Pakistani journalists writing in English tend to use British standard cleft construction norms by editorial convention or they develop native language structure-based preferences for clefts. An unresolved question drives the research aim of this study. Cleft structures remain infrequent in academic writing from Pakistani English researchers because of natural L2 reduction patterns. This paper addresses the missing research on the use of clefts in Pakistani news discourse by focusing on their analytical usage. Research on cleft appearances between Pakistani and British news articles expands existing scientific materials about World Englishes while investigating the syntactic patterns between different language varieties and publishing genres.

Research on cleft constructions extensively details their theoretical and spoken discourse applications but insufficient research exists regarding their actual usage in journalistic writing.



The study of cleft structures shows limited research regarding native British English against post-colonial Pakistani English. Research evidence indicates cleft constructions respond to genre and sociolinguistic setting which makes a corpus-based investigation of Pakistani and British newspaper use both critical and beneficial now.

The research intends to establish a connection through syntactic and functional analysis between cleft constructions in two distinct English versions within the same journalistic genre. Through theoretical and corpus-based frameworks this research seeks to identify how linguistic choices reveal patterns of language behavior alongside communicative methods and language-based identity expressions.

### **3. Methodology**

#### **3.1 Corpus Compilation**

A comparable corpus consisting of fifty health newspaper articles was compiled for the purpose of carrying out cross-cultural comparison of use of cleft construction. Twenty-five articles from major Pakistani newspapers (The Dawn, The News International, The Express Tribune) and twenty-five from British newspapers (The Guardian, BBC News, The Telegraph) were selected for the compilation of a corpus. All articles were published from March to May 2025. The Pakistani sub-corpus consists of 85312 words, and the British sub-corpus consists of 92784 words. Eligibility criteria dictated combined consideration of only those written works that presented in-depth information on health issues and were available digitally in full text.

#### **3.2 Data Collection Process**

Collection of data continued for 12 weeks (1 March–24 May 2025). Automated harvesting was performed with Python 3.9 scripts, which used a controlled vocabulary of search terms—“cleft lip”, “cleft palate”, “craniofacial anomaly”—to compile article metadata (title, author, publication date, URL). Data Imported URLs and PDF. URLs were imported and logged in a central database. Each article was then downloaded in either html or pdf format, converted to plain utf-8 text, and sorted into country- and outlet-specific directories. A comprehensive log was made on the retrieval of studies with any access difficulties (e.g. paywalls, broken links), making the dataset entirely reproducible.

#### **3.3 Preprocessing and Preparation of the Text Data**

All raw text files were cleaned of boilerplate—headlines, captions, ads—programmatically and whitespace-normalized. Sentences were tokenized using NLTK’s Punkt tokenizer (retrained on a development subset of ten articles to improve accuracy for journalistic text). Tokenization and part-of-speech tagging were performed using NLTK. Words were counted with custom scripts and corpus sizes were verified before starting analysis (Pakistani: 85312 words; British: 92784 words).

#### **3.4 Automated retrieval and manual confirmation**

Automatic extraction of cleft constructions (it-clefts, wh-clefts, all-clefts) was conducted with custom Python scripts employing NLTK for tokenization and regular expression pattern matching (e.g., `\bIt\s+was\s++\s+that` for it-clefts) to extract the desired constructions. Candidate instances were exported to spreadsheets in Excel 2016 for manual annotation by two individual annotators. Differences were settled through joint review sessions with a Cohen’s  $\kappa$  inter-annotator agreement of 0.87. Cases with equivocal findings were not included in the quantitative analyses but recorded for qualitative assessment.

#### **3.5 Quantitative Analysis**

To enable cross-corpus comparisons, validated cleft instances were normalized to per 1000 words for each corpus, taking into account that the two corpora were of different sizes. Descriptive statistics (average frequency, variability) were calculated in SPSS 27, and chi-square tests were used to verify the significance of the differences between the two groups

( $\chi^2(1) = 8.23$ ,  $p < .01$ ). Practical significance was measured by effect sizes ( $\phi$ ). All statistical analyses met the standard of an  $\alpha$  level of .05).

### 3.6 Qualitative Thematic Coding

To identify these functions, each confirmed cleft construction was brought in NVivo 14 and hand coded into either one of four function groups: emphasis, contrast, explanation or emotional appeal. A coding framework was derived inductively during pilot analysis and through repeated comparison. Thematic frequency differences between the corpora were contrasted and distributional differences between the functions were tested with nonparametric Mann–Whitney U tests.

### 3.7 Quality control and references check

Retrieval logs, preprocessing scripts, and annotation files were version controlled at all stages of the experiment for the sake of transparency (Jurafsky & Martin, 2009). All cited references mentioned in the methodology and subsequent chapters were looked up with regard to the list of references; if there was something missing, it was included, and if there were any orphaned bibliography items, they were deleted to ensure a one-to-one correspondence.

## 4. Results

### 4.1 Quantitative Findings

#### 1. Raw Counts of Cleft Constructions

Cleft Type	Pakistani Corpus (n = 85312 words)	British Corpus (n = 92784 words)
It-clefts	53	107
Wh-clefts	28	62
All-clefts	12	21
<b>Total</b>	<b>93</b>	<b>190</b>

British newspapers exhibit roughly double the raw number of cleft constructions (190) compared to Pakistani outlets (93). The largest disparity is in it-clefts (107 vs. 53), suggesting a stronger editorial preference for this structure in UK journalism.

#### 2. Normalized Frequencies (per 1000 words)

Cleft Type	Pakistani (per 1000 wds)	British (per 1000 wds)
It-clefts	0.62	1.15
Wh-clefts	0.33	0.67
All-clefts	0.14	0.23
<b>Total</b>	<b>1.09</b>	<b>2.05</b>

After normalizing for corpus size, British articles still show nearly twice the density of cleft usages (2.05 vs. 1.09 per 1 000 words). This confirms that the raw-count differences are not driven solely by text length.

#### 3. Chi-Square Tests by Cleft Type

Cleft Type	$\chi^2(1)$	p-value	$\phi$ (effect size)	Significance
It-clefts	7.56	.006	0.16	<b><math>p &lt; .01</math></b>
Wh-clefts	5.12	.024	0.14	$p < .05$
All-clefts	2.34	.126	0.09	n.s.
<b>Overall</b>	<b>8.23</b>	<b>.004</b>	<b>0.17</b>	<b><math>p &lt; .01</math></b>

- **It-clefts:** Significant difference ( $\chi^2=7.56$ ,  $p=.006$ ), moderate effect ( $\phi=0.16$ ).
- **Wh-clefts:** Also significant ( $\chi^2=5.12$ ,  $p=.024$ ), small-to-moderate effect ( $\phi=0.14$ ).
- **All-clefts:** Difference not statistically significant ( $p=.126$ ).
- **Overall usage:** Highly significant ( $\chi^2=8.23$ ,  $p=.004$ ),  $\phi=0.17$ .

- These results indicate that the British corpus uses *it*- and *wh*-clefts significantly more often than the Pakistani corpus, while the rarer all-clefts show no clear cross-cultural difference.

#### 4. Summary of Descriptive Statistics

Statistic	Pakistani Corpus	British Corpus
Total words	85 312	92 784
Total clefts	93	190
Mean clefts per article	3.72	7.60
SD of clefts per article	1.45	2.12
Mean normalized freq (‰)	1.09	2.05
SD normalized freq (‰)	0.38	0.47

- On average, each Pakistani article contains 3.7 clefts versus 7.6 in British articles.
- Variability is greater in the British set (SD 2.12 vs. 1.45), reflecting more diverse stylistic practices across outlets.
- Consistent with earlier tables, the mean normalized frequency underscores a substantially higher reliance on clefts in UK journalism.

### 4.2 Qualitative Findings

#### 1. Discourse Function Distribution

Each validated cleft instance was coded for one of four primary discourse functions—**emphasis**, **contrast**, **explanation**, or **emotional appeal**—using NVivo 14. Table 1 summarizes the frequency and relative proportion of each function in the two corpora.

**Table 1. Distribution of Cleft Functions**

Function	Pakistani (n = 93)	% of Total	British (n = 190)	% of Total
Emphasis	25	26.9%	57	30.0%
Contrast	15	16.1%	38	20.0%
Explanation	45	48.4%	72	37.9%
Emotional Appeal	8	8.6%	23	12.1%
<b>Total</b>	<b>93</b>	<b>100%</b>	<b>190</b>	<b>100%</b>

- Explanation clefts dominate in both corpora but are proportionally more prevalent in the Pakistani articles (48.4% vs. 37.9%), suggesting a stronger tendency to use clefts for clarifying health information.
- Emphasis constructions are prominent in both sets ( $\approx 27\%$  vs. 30%), reflecting a shared stylistic role in foregrounding key facts.
- Contrast and emotional appeal are relatively more frequent in the British corpus, indicating a greater use of clefts for persuasive or reader-engagement purposes in UK outlets.

A nonparametric Mann–Whitney U test confirmed that the distribution of **explanatory** clefts differs significantly between the two corpora ( $U = 652.5$ ,  $p = .032$ ), whereas **emphasis**, **contrast**, and **emotional appeal** distributions did not reach statistical significance at the  $\alpha = .05$  level.

#### 2. Illustrative Examples

To deepen understanding of how functions manifest, Table 2 provides representative excerpts for each category.

**Table 2. Sample Cleft Excerpts by Function**

Function	Corpus	Excerpt
Emphasis	Pakistani	<i>“It was the lack of early intervention, not the surgery itself, that delayed the child’s recovery,”</i>
Contrast	British	<i>“What sets this new procedure apart is its minimally invasive nature, not its cost — that remains comparable to traditional methods,”</i>
Explanation	Pakistani	<i>“It was the irregular feeding schedule, rather than genetic factors, that most contributors identified as the root cause of postoperative complications,”</i>
Emotional Appeal	British	<i>“All that stood between the family and hope was a single, transformative operation,”</i>

### 3. Cross-Cultural Patterns and Insights

- **Explanatory Focus in Pakistani Journalism:** Nearly half of Pakistani clefts serve to unpack or clarify technical or procedural details. This aligns with an informational ethos in local health reporting, where writers prioritize ensuring reader comprehension of complex medical topics.
- **Persuasive and Engaging Use in British Outlets:** The higher proportions of contrast (20%) and emotional appeal (12.1%) clefts in the British corpus reflect a journalistic style that blends factual reporting with narrative engagement and reader persuasion—perhaps due to more competitive media markets and diverse readership expectations.
- **Shared Emphasis Strategies:** Both corpora deploy emphasis clefts at similar rates. This suggests that, regardless of cultural context, writers recognize the rhetorical power of foregrounding key information through the cleft construction.
- **Implications for World Englishes:** These patterns underscore how reporting conventions and audience expectations shape syntactic choices. Pakistani health journalists lean on clefts for explanatory clarity, while British counterparts use them more variably to engage and persuade.

### 5. Discussion

The convergence of quantitative and qualitative results thus unambiguously suggests a clear-cut syntactic contrast in the deployment of cleft constructions between Pakistani and British journalism, implying that law-governed editorial norms and reader expectations crucially determine syntactic selections. The much higher general rate of cleft usage – and especially of it-clefts and wh-clefts – shown in the British newspapers, suggests that British journalists are more likely to use the construction partially as a device to grab attention and improve readability, as not just a tool to convey information. This is not new and fits the pattern documented in World Englishes of written communication in more disputed, competitive media environments being characterized by complex syntactic structures, as writers try to exploit foregrounded syntax in the name of attracting readers' eye.

In contrast, the preponderance of explanatory clefts in the Pakistani corpus — almost 1 of 2 instances — reflects a basic communicative orientation towards explicitness and pedagogy. Pakistan health journalists seem to favour cleft construction as an instrument for decoding technical information hence minimising risk of non-understanding among an audience with diverse levels of bio-medical knowledge. This trend mirrors those previously noted, after whom non-native English press appears to be more concerned with explicit information framing than stylistic application. It reminds us of an effective strategy of localizing it for



audience preference and of making complex casual or procedural information about the most critical health messages available and memorable by clefting them.

The British preference for contrast and emotion-appeal clefts—constructions which are less common in Pakistani news—indicates a more narrative style of health reporting. Contrast clefts (“What makes X different is”) enable UK journalists to contextualize any new development relative to a known base couple of lines, while proving to be attractively framed innovation stories. Emotional appeal: clefts (“All that stood between the family and hope was ...”) that access human-interest angles, strengthening reader empathy and buy-in. This rhetorical flexibility might be indicative of the commercial constraints that are common to British news and in which audience stickiness has a strong editorial force. These results lend empirical weight to claim that the frequency of syntactic complexity in journalistic English is related mainly to persuasive or affective purposes, something that is reflected in our higher quantities of contrast and emotional-appeal clefts in the British corpus.

Notwithstanding these cross-cultural variations, the identical proportions of emphasis clefts in both corpora suggest a common recognition of their rhetorical power in focusing on salient information in the preceding discourse. Clefts of emphasis (“It was not the surgery, not the failure to act early ...”) shed light on salient causes or corrective statements—a rhetorical device with universal value in health journalism, where precision and impact are everything. The commonality here indicates that, despite regional norms, journalists are using emphasis clefts instinctively to match the competing requirements of precision and involvement.

Methodologically, our hybrid computational and manual system—automatic extraction using Python (2020) and subsequent annotator validation—was successful in identifying both quantitative patterns and functional nuance. A high inter-annotator agreement ( $\kappa = .87$ ), which points to the validity of our coding system, and the software packages SPSS (Field, 2024) and NVivo (2021) helped with a multilayered analysis. Further studies could potentially expand this framework by incorporating more genres (e.g., social media health blogs, broadcast transcripts) or investigate diachronic change over narrower time scales with a view to identify changing syntactic preferences.

There are some limitations, however, that should be mentioned. We may not recognize the more proximate changes in journalism, especially as news media increasingly embrace the digital age. In addition, while our statement selection, including the front pages of the popular newspapers, makes our sample visible, it might fail to cover the heterogeneous nature of regional or tabloid press styles. Finally, manual validation was laborious and manual, although sufficient to ensure validity; future work may scrutinize semi-automated classification methods including machine learning algorithms, in order to scale analysis.

## 6. Conclusion

In this study, we have shown that clefts are a powerful stylistic means to write about health in the media, but that the realization of cleft construction is conditioned by different editorial construals and audience exigencies in the Pakistani and British contexts. Quantitatively, British newspapers employ clefts—especially it-clefts and wh-clefts—at almost double the rate observed in Pakistani counterparts, a difference which remains statistically significant after text length is controlled for. From a qualitative point of view, Pakistani health-related articles make extensive use of clefts to explain and to simplify complicated medical information, and this seems to be justified by a didactic orientation oriented at achieving maximum reader understanding. While, British newspapers show more rhetorical flexibility in the usages of the

cleft, not limiting to a role of emphasis and explanation, in a news text they also used to opposition and creating an affective link between text and reader accommodating the imperatives of the information and narrative status (discourse participant roles).

These trends highlight how a cleft syntax can be two-faced: it can be a pedagogical device, used to present information clearly, or a rhetorical device, used to argue or persuade. Our contribution to the general theories of syntactic variation in World Englishes is to demonstrate how local communicative needs and commercial pressures interact in the determination of grammatical choice. Methodologically, the amalgamation of automated extraction, careful manual verification, and mixed-method analysis is a model that can be replicated in future studies of syntactic phenomena across genres and languages.

But the narrowness of this study's focus—just six major newspapers—may not reflect patterns on the rise in smaller digital and local newsrooms. Future research should be expanded to social media, broadcast transcripts, and newer digital technologies, and should include more detailed temporal comparisons that can track changing editorial routines. By helping to bridge the gap in our knowledge of clefting use in different media, studies of this kind could contribute knowledge of the relationship between syntax, generic expectations, and reader involvement in global English journalism.

### References

- Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.
- Bourgoin, C. (2022). *A corpus-based study of the prosody and information structure of English it-clefts and French c'est-cleft* (Doctoral dissertation, Cardiff University and University of Leuven).
- Dorgeloh, H., & Wanner, A. (2022). *Discourse syntax*. Cambridge University Press.
- Fairclough, N. (1995). *Media Discourse*. London: Edward Arnold.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Halo, L. A., Rustamani, S., & Muhammad, N. (2024). Study of linguistic components of Pakistani English: An indigenized English variety. *Pakistan Journal of Humanities and Social Sciences*, 12(1), 287-293.
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Edward Arnold.
- Hussain, B., & Mehdi, M. (2023). A Comparative Study of Punjabi and English Open-Ended Syntactic Constructions Through the Revised Extended Standard Transformational Generative Theory. *Journal of English Language, Literature and Education*, 5(3), 63-88.
- Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing* (2nd ed.). Prentice Hall.
- Kwon, M. H., Staples, S., & Partridge, R. S. (2018). Source work in the first-year L2 writing classroom: Undergraduate L2 writers' use of reporting verbs. *Journal of English for Academic Purposes*, 34, 86-96.
- León, M. (2018). Standard language ideologies, world Englishes, and English language teaching: An overview. *Journal of Second Language Acquisition and Teaching*, 25.
- Pérez-Llantada, C. (2012). Authorial voice in English scientific research articles published by Spanish scholars: A corpus-based study. *English for Specific Purposes*, 31(1), 18-29. <https://doi.org/10.1016/j.esp.2011.06.002>
- Purcell, V. (2020). *Understanding Visible Differences: Working Therapeutically With Individuals Who Look Different*. Springer Nature.

- Python Software Foundation. (2020). Python Language Reference (Version 3.9) [Computer software]. <https://www.python.org>
- NVivo (2021). [Computer software]. (Version 14). QSR International.
- Rahman, T. (2020). English in Pakistan: Past, present and future. *Functional variations in English: Theoretical considerations and practical challenges*, 127-148.
- Sánchez-Ramírez, D., & León, M. (2018). Cleft constructions in news discourse: A comparative study of English and Spanish corpora. *Journal of Pragmatics*, (134), 112–125. <https://doi.org/10.1016/j.pragma.2018.08.004>
- Seitanidi, E., Pöldvere, N., & Paradis, C. (2024). Everything-cleft constructions in spoken British English: A neglected construction. *Functions of Language*, 31(3), 289-326.
- Shipova, E. (2024). *Formal analysis of èto-clefts in Russian: syntax and semantics* (Doctoral dissertation, Universität Potsdam).
- Standing, W. (2021). *It-cleft change across the lifespan in early modern authors: variation, rhetoric and the individual in constructional change* (Doctoral dissertation, University of Antwerp).
- Wilson, F., & Strevens, P. (2016). Syntactic strategies for emphasis and engagement in modern journalism. *Journalism Studies*, 17(4), 455–472. <https://doi.org/10.1080/1461670X.2015.1088492>
- Yousefi, A. M., & Esfandiari, R. (2024). To Cleave or Not to Cleave: Distributional Frequencies of Cleft Structures in Research Articles, Textbooks, and PhD Dissertations. *Applied Research on English Language*, 13(1), 111-136.