Corpus Assisted Critical Discourse Analysis of Pakistan’s Language Education Policy Documents: What are the Existing Language Ideologies?

Muhammad Asim Khan1 and Sajida Zaki1

Abstract
Language education policies contain discourses that have language ideologies embedded within them. This study explores the language ideologies in official language education policy documents of Pakistan from 2000 to 2020. Using Corpus Assisted Critical Discourse Analysis techniques, a 1.28 million-word-corpus was generated from 32 policy documents collected from official websites of UNESCO and the Ministry of Federal Education and Professional Training, Pakistan. The corpus analysis of frequency, collocation, and MI significance values were utilized to find out the typicality of topics and semantic preferences, while the concordance analysis revealed discourse prosody around the languages mentioned in the corpus. Moreover, seven documents were downsampling through dispersion plot. Discourse Historical Analysis was carried out on these documents based on discursive strategies of argumentation, predication, and perspectivization with reference to intertextuality and interdiscursivity. The findings show that English and to some extent Urdu occupy the ideological space related to education. The local languages were discussed in positive light, but they were marginalized in educational ideological space. The analysis also showed that the policies had underlying monoglossic ideology that can be seen having tolerance-oriented and expediency-oriented policy ideologies for local languages. This implies that the official policies of Pakistan marginalized the local languages in the discourse of language education. These findings invite policymakers to bring the local languages in the ideological space of education and introduce high level of ideological uniformity, consistency and goal setting for English, Urdu, and local/indigenous languages in the language-in-education policies of Pakistan that matches, captures and promotes the linguistic diversity in Pakistan.

Keywords
language ideology, language policy, critical discourse analysis, corpus, discourse, discourse historical studies

Introduction
Language policy and planning is one of the most critical domains due to the power of the language policy that makes some ideologies official and prevalent (Canese, 2018). For the subfield of language-in-education policy, Tollefson (2002, p.6) identifies two important arenas for language ideology research: (1) how language ideologies shape policies and how these are naturalized through discursive practices, and (2) the interplay of contesting language ideologies in language-in-education policies. Schools are the primary sites for promoting language ideologies through language-in-education policies (Canese, 2018; Cushing, 2021). Cushing (2021) calls the language policy-related texts, including curriculum documents, tests, linguistic landscapes, and rules and regulations, as a “mechanism” to implement those ideologies into practice. Through these mechanisms, the language ideologies become de facto language practices in centralized education systems (Shohamy, 2006).

The language-in-education in Pakistan has witnessed the power of ideologies implemented through 10 education or language-in-education policies over the last seven decades since the country’s independence in 1947, shown in Table 1. The table has been adapted from Siddiqui (2016) by adding 2017 policy as the education policy (2017) was not captured in the book. Besides, two columns have been added to show the ideological trends related to language-in-education policy in the government education policies. Siddiqui (2016) believes that most policies were led by the ideology of the ruler of the time instead of a democratic consultative process. He argued that it is essential to look at these policies...
from the ideological standpoint of the ruler (Siddiqui, 2016, p. xxi). de Jong et al. (2016) compared the language policies of Taiwan, Pakistan and China. They argue that historically different languages have been treated from different ideological stances. In pre-partition India, Urdu was considered a resource as the language of intellectuals and literature. During British rule, English did not immediately replace Urdu as the language of formal communication. However, it ripped Urdu off its heritage of being a resource for intellectuals, administration, economy, political and social mobility. After the independence, the Bengali language emerged from the language-as-right orientation in response to Urdu that was declared the national language. Bengalis in East Pakistan felt it was challenging to communicate in English or Urdu as they had limited proficiency in both languages. However, their study is based on the selected review of literature which traces its history in a concise space and lacks any systematic reference and comparison to the language policies. Shamim (2008) traced the ideological trends in Pakistan’s language-in-education policy up to National Education Policy (2009). She broadly divides them into four phases. The phase from 1947 to 1977 was identified as a trend toward one-nation-one-language ideology, that is, Urdu as the symbol of national identity and unity. 1978 to 1988 was the phase of the revitalization of Urdu under the Islamization efforts by General Zia ul Haq. The democratization of English ideology marked the phase from 1989 to 1999. Lastly, in the period from 1999 to 2009, the ideology of English as the language of development was dominant. Mahboob (2002) discussed the connection between political shift and language policy till 1999, as indicated in Table 1. It may be noted that these studies show varying ideologies during those political eras. Shamim (2008) believes that the one-nation-one-language ideology remained dominantly consistent over a long period, whereas, Mahboob (2002) shows a fluctuation or shift. Both agree that from 1979 until 1992, the policy reflected pro-Urdu ideology, which changed into pro-English post-1992.

The above studies show the investigation of language ideologies underlying in language-in-education policies till 2009, primarily based on randomly selected policies and their analysis. Moreover, mostly researchers selected national-level policy documents only. It suggests a selective presentation of ideologies in official documents related to language-in-education without including the provincial governments’ documents. We argue that after the 18th Constitutional Amendment of the 1973 Constitution of Pakistan, the provincial governments have autonomous educational decision-making, including language-in-education. Consequently, the national and provincial language-in-education policy documents would reflect a broader ideological representation of different languages. Therefore, the purpose of this study is to explore the language ideologies embedded in the discourse of language-in-education policies not only in the selected national education language policies but across all national and provincial government education policies. The study is significant because it addresses the limitation of contextual studies discussed earlier and adopts a structured linguistic analysis of policy documents from 2000 to 2020 guided by Corpus Assisted Critical Discourse Analysis (Baker et al., 2008). The research question for the study is:

- What are the ideologies related to the languages present in the education policy texts of Pakistan?

| Year | Policy Document | Type of government | Shamim (2008) | Mahboob (2002) |
|------|----------------|--------------------|---------------|---------------|
| 1947 | Pakistan Educational Conference | Civil | One nation, one language (1947–1977) | One nation, one language (1947–1958) |
| 1959 | Commission on National Education (Urdu to replace English in 15 years) | Military | ------ | Openly Pro-English |
| 1969 | Proposal for a New Education Policy | Military | ------ | Urdu in West Pakistan/ Bengali in East Pakistan |
| 1970 | The New Education Policy | Military | ------ | Urdu (but English legalized) |
| 1972 | National Education Policy | Civil | ------ | ------ |
| 1973 | Constitution of Pakistan (Urdu to replace English in 15 years) | Civil | ------ | ------ |
| 1979 | National Education Policy | Military | Revitalization of Urdu as the language of Muslim identity (1978–1988) | Urdu or Provincial language + Arabic (Compulsory) |
| 1992 | Education Policy | Civil | Democratisation of English (1989–1999) | English permitted from class I |
| 1998 | National Education Policy | Civil | ------ | Seemed pro-English |
| 2009 | National Education Policy | Civil | English as the language of development (1999 to date) | ------ |
| 2017 | National Education Policy | Civil | ------ | ------ |

Source: Compiled from Mahboob (2002), Shamim (2008), Siddiqui (2016).
Literature Review and Theoretical Framework

Language Policy and Language Ideologies

Ideology, discourse and language are interconnected phenomena. It is problematic to capture ideology with a single agreed-upon definition of the term (Tollefson, 2011). For Blommaert (1999), language is the leading site of ideological debate. The language beyond sentence level is discourse (Lin, 2014); this “Discourse” reflects ideologies or the way of thinking and represents certain phenomena under discussion (Gee, 1996). Krzyżanowski and Wodak (2001) show the inter-relatedness between language policy and language ideologies. Language ideologies are “cultural ideas, presuppositions and presuppositions with which different social groups name, frame and evaluate linguistic practices” (Gal, 2006 cited in Krzyżanowski & Wodak, 2011). These ideologies are reconstructed and perpetuated in discussions where language is the main topic of debate in public domains. Language-in-education policy is one such example of public sphere discourse that reflects the ideologies of institutional social actors who define how languages should co-exist and what roles they should perform in social, political and economic contexts (Krzyżanowski & Wodak, 2011). The process of language and planning takes place at several levels. Ricento and Hornberger (1996) identified those levels of language policy and planning with three layers of an onion. They consider that the outer-most layer or macro level policy is regulated by the state that decides which language receives national status, which receives educational support, which stay repressed, and which stay ignored. The second inner layer as meso-level policy is concerned with institutions such as schools, religious organization, media, civic and other private and public organizations, and business communities involved in language policy and planning development. The inner-most layer or micro level policy is where teachers are the main people, but it is debatable that whether they can influence or have direct contribution to language policy and planning process.

The powerful groups such as authors and politicians who control significant public discourses influence and help reproduce dominant ideologies (Baker et al., 2008). Tollefson (2000) maintains that language ideology present in the language policy is so strong that it makes school systems translate the language ideology as something necessary that is in favor of the society. He argues that language ideology has a direct impact on language education policies. In countries like the United States, England, and Australia, the standard language ideology is enacted through language policies of those countries. Spolsky (2004) believes there is no difference between language ideology and language policy as he puts down “language policy is language ideology with the manager left out, what people think should be done. Language practices, on the other hand, are what people actually do” (p.14). For the “legitimization” (Canese, 2018) and “concretization” (Cushing, 2021) of language ideologies symbolized in language-in-education policies, they need to be practiced in educational institutions and in the communities outside those institutions. The four categories of actors with different roles for implementation of language policy and planning are people-with-power, people-with-expertise, people-with-influence and people-with-interest (Zhao, 2011; Zhao & Baldauf, 2012). The people-with-power are the policymakers at the macro level with official role of deciding the status of languages. The other three groups are described as the professionals such as linguists, social elites who influence the masses with their work such as scholars and celebrities, and the communities and language users at the micro level. According to Zhao (2011), the postmodernist perspective emphasizes the roles of agency and actors from micro levels, and on individual agency from the four categories on the effectiveness of language policy and planning efforts. These micro level actors lack the power that the macro level policymakers possess as they have the privilege to take part in policy formation and often these policies benefit themselves (Shao & Gao, 2019, pp. 257–258). Although the process of successful language-in-education policy and planning takes place at macro to micro levels with a variety of policy actors, the textual analysis of language policy documents is an important arena. A rigorous policy textual analysis can reveal the underlying hegemonic ideologies that perpetuate inequalities in and through language-in-education policies without being questioned. It is aptly stated by Shao and Gao (2019), p. 261), “A critical-discursive perspective is able to deconstruct ideology laden statements and taken-for-granted positions.” The critical discourse analysis of policy texts has the power to make the implicit ideologies explicitly visible for a variety of actors at different levels for them to address the language ideologies in alignment with the needs of the rapidly changing society and the world.

Language Ideology as a Theoretical Framework

Tollefson (2006) presents Fairclough’s (1989) definitions of hegemony and ideology together within the realm of Critical Theory in language policy as both are interconnected. Hegemony is a process through which institutions help reside and maintain power within the control of a few people, supported by the ideology that is “naturalized” unconscious belief and assumptions in society. The focus of critical theory in language policy remains on uncovering the implicit and explicit policies. These policies act as tools for creating and sustaining inequalities through hegemony portrayed as “natural” via social institutions (Tollefson, 2006, pp. 47–48). More specifically, certain language ideologies are presented as natural by language-in-education policies. For example, Ricento (2013, p. 530) refers to the ideology of English monolingualism prevalent in English dominant countries, initially introduced by Wiley and Lukes (1996) and
Silverstein (1996). Wiley and Lukes (1996) believe that monolingual ideology considers that immigrants introduced linguistic diversity in those originally English dominant societies. Silverstein (1996) terms it as “monoglot” ideology which disregards the sociolinguistic reality replete with practices of multilingualism and diversity. The monoglot ideology influences language-in-education practices, social life, identities, and academic life (Blommaert 2006, cited in Ricento, 2013). Garcia (2009) separates monoglossic ideology from heteroglossic ideology. Monoglossic ideology embedded in bilingual programs considers command on two languages independent of each other, resulting in “double monolingualism.” Heteroglossic ideology maintains that languages are inseparable and uncountable, and reflects a fluid use of multiple languages by multilingual speakers known as translanguaging (Flores & Baetens Beardsmore, 2015, p. 213). Similarly, Wiley and Lukes (1996) identify standard language ideology. A specific variety of a language of a powerful social group is given a higher status while other varieties are attached with a lower status. The political and social influential groups make a particular variety standard through codification against which all other varieties are judged and made as a norm through schooling gradually (Wiley & Lukes, 1996). Busch (2011) notes that schools in Europe have actively repressed other languages and language varieties by implementing standard languages as a nation-building process under monolingual and homoglossic ideologies. Despite prevalent multilingualism, schools in Belgium and Switzerland still have several monolingually oriented systems. Learners are assessed in the standard academic language against a standard based on an ideal native speaker, creating barriers for equal access and opportunities for other language speakers (Busch, 2011, pp. 545–546).

Ideologies are embedded in language policies. Kloss (1998) studied official language policies from an ideology perspective and developed a schema that helps categorize official language policies’ orientations or ideologies. He identified four types of language policies: (1) promotion-oriented, (2) expediency-oriented, (3) tolerance-oriented, and (4) restrictive-oriented. Wiley and Tollefson (2013) elaborates that promotion-oriented policies are those in which the governmental resources are utilized to support a language or languages in official domains. Expediency-oriented policies are a “weaker version” of promotion-oriented policies. They do not target the enhanced use of minority languages. Tolerance-oriented policies leave the responsibility of maintaining minority languages on language communities with no governmental resources and support. Restriction-oriented policies “make social, political, and economic benefits, rights, and opportunities conditional on knowing or using the dominant language. Language restrictions usually target communication in work-related or official domains” (p. 326). Wiley and Tollefson (2013) expanded Kloss’s (1998) framework by adding two more types: null- and repression-oriented policies. Repression-oriented policies are those that deliberately try to eliminate minority languages. Null policies are those where there is an apparent absence of policy recognizing minority languages or language varieties. Historically, language policies of the US can be categorized across all of these types (Wiley & García, 2016). Similarly, Ruiz (1984, p.16) introduced three orientations or ideologies to language policy: language-as-resource, language-as-right, and language-as-problem. According to him, these orientations indicate “a complex disposition towards languages and their roles,” and language attitudes are formed within these ideologies. Johnson (2013, p.36) further elaborated that language-as-problem ideology views minority languages as a problem in acquiring the dominant language. Therefore, transitional bilingual policies hold this ideology. Language-as-right ideology is a unidirectional bilingual program in which the minority language students are required to learn the dominant languages and maintain their language simultaneously. Language-as-resource ideology values linguistic diversity through bidirectional programs aiming at additive bilingualism. Students from both the dominant and minority language learn in both languages. The above literature shows that language ideology is an important phenomenon that is inseparable from and strongly influences language policies and has been studied from different perspectives. This study contributes to the existing literature on language ideology by expanding the application to language-in-education policy text from Pakistan. It also contributes to the methodologically growing field of corpus assisted critical discourse analysis studies (CACDA) by applying it to the corpus of language-in-education documents. Other studies have applied CACDA mainly in language ideologies presented in mainstream, social media, and language-in-education policy (see Doolan, 2011; Vessey, 2017).

**Language Ideologies, CDA and Corpus Linguistics**

Critical Discourse Analysis (CDA) has three dominant approaches with different foci: discourse-historical, cognitive, and dialectical-relational-dialectic (see Fairclough, 2013; Van Dijk, 2009; Wodak, 2001). Although all three dominant approaches of CDA have been widely used in language ideology research, CDA is often criticized for its random data collection (Cheng, 2013). Similarly, frequency, collocation and concordance techniques under Corpus Linguistic (CL) have been applied to ideology research (e.g., Cheng, 2013; Vessey, 2017). Nevertheless, the focus on textual analysis of data in corpus linguistics is considered a disadvantage in language ideology research since useful “contextual” information beyond the texts is usually absent (Vessey, 2017). Texts are significant instances for witnessing the embedded language ideology (Vessey, 2017); nonetheless, the socio-political context and discourses from which those ideologies emerge are equally important (Milani & Johnson, 2008 cited in Vessey, 2017). Within texts or documents, the explicit discursive representation or metalanguage shows the explicit
language ideologies, whereas the implied assumptions about the role of languages show implicit language ideologies (Vessey, 2017). Combining corpus tools with other methods can enhance research on language ideology (Cheng, 2013; Vessey, 2017) as demonstrated by CACDA (Baker, 2006; Haider et al., 2021; Partington, 2010). A valuable combination of CDA and CL has been utilized by Baker et al. (2008), combining corpus linguistics approaches with Wodak’s (2001) discourse-historical studies (DHS). This combination helps address the disadvantages associated with both CDA and CL approaches noted above and enable triangulation of findings across methods (Baker et al., 2008, p. 295).

Wodak’s DHS provides a robust systematic methodological framework enabling researchers to carry out systematic data collection around a macro topic over a defined period to investigate changes in discourse practices (Reisigl & Wodak, 2009). Catalano and Waugh (2020) argue that DHS is one of the approaches of CDA with the greatest emphasis on wide-ranging historical data. DHS targets three types of critiques: (1) exploration of inconsistencies and contradictions within text or discourse; (2) exposition of manipulative discursive practices; and (3) improvement in communication for the future. The analytical strategies include intertextuality and interdiscursivity across a range of texts, genres, and discourses around a macro topic that helps identify a set of ideologies and inequalities that are perpetuated and reinforced by different texts and sites over time (Reisigl & Wodak, 2009). Five discursive strategies or intentional plans that authors use to attain specific goals through a text are investigated (see Reisigl & Wodak, 2009). These strategies are nomination, predication, argumentation, perspectivization and mitigation. The analysis of relevant strategies is guided by the nature and focus of a study (Mulderrig, 2014).

Method

This study incorporated Baker et al. (2008) methodology of combining CL with DHS. Incorporating Baker et al. (2008) and Vessey’s (2017) strategies, the word frequency, collocation and concordance analysis was carried out, followed by a detailed qualitative analysis of downsampled policy documents.

Data collection

For collection of language-in-education policy documents, provincial and federal government websites, offline government archives through government personnel, and officials in Higher Education Commission of Pakistan were approached. However, very few documents were available. Therefore, “Portal of Education Plans and Policies” (UNESCO, 2021.) website was selected because it provided the most comprehensive list of documents on language education in a single place. It also reduced the researcher’s bias for the selection of the documents. A few drafts of the recent Single National Curriculum (SNC) were gathered from the Federal government website (Ministry of Federal Education and Professional Training, n.d.) as these were not updated on the UNESCO’s portal yet.

Compilation of corpus. Based on the selected portal, a “purpose-built corpus” was developed from policy texts/documents on “particular topic” of language ideologies in language-in-education policy documents as suggested by Mautner (2016, p. 165). Total 71 documents were collected (54 from UNESCO and 17 from Ministry of Federal Education and Professional Training website). Using Vessey’s (2013) strategy of core query term (CQT), documents were manually searched for CQT “language”. This step ensured that only the documents that potentially had a discourse on language-in-education were selected for the corpus, while documents that did not contain the word “language” were excluded. This process yielded 26 documents from the UNESCO website, whereas 06 documents from the Ministry of Federal Education and Professional Training website. Other CQTs frequently used in the literature were also searched through those documents to ensure that no potential document was ignored. These terms included “tongue” (for mother tongue), “medium,” “instruction” (for medium of instruction), “linguistic” “Urdu,” “English,” “local” (for local language), “regional” (for regional language), minority (language). Based on this process, 32 PDF document files, shown in Table 2, were selected for analysis and converted to (.txt).

Downsampling. Downsampling is a common technique adopted in CACDA as a strategy to counter the challenge that it is too “labour-intensive” to perform CDA on a number of texts (see Baker et al., 2008; Vessey, 2017). In CACDA, samples of texts are selected based on high frequency of the CQTs. For detailed CDA analysis of 07 downsampled documents shown in Figure 1, a combination of strategies used by Baker et al. (2008), and Vessey (2017) including high frequency, dispersion plot (see appendix) and data relevance were applied.

Data analysis. The research question is answered by combining the corpus and DHS analysis of data. The micro analysis of corpus data is carried out through frequency, collocation, and concordance analysis. These three analyses respectively reveal the typicality of topics/CQTs, strength of relationship with other words found in close environment, and the collocational environment of the words. In other words, the ideologies of the policymakers are revealed through what they discuss most often in the policy documents, what other words are most often used in discussing those topics, and whether negative or positive terms are used by the policymakers while discussing those topics/CQTs. The DHS analysis reveals the ideologies of the policymakers in macro discussion of the languages.

Corpus software packages such as WordSmith Tools (Scott, 2013) allows researchers to generate frequency lists,
and explore measures of statistical significance such as Mutual Information (MI value/scores) and t-scores through collocation analysis (Mautner, 2016). Frequency analysis is the first most frequently employed and essential step in corpus linguistics (Archer, 2009). The frequency analysis is meaningful when interpreted as a reflection of the users’ knowledge of discourse norms (Stubbs, 2001). In ideology research, word and phrase frequency may indicate the typicality of topics and how they are discussed (Vessey, 2013). Vessey (2017) argues that high, low, and statistically significant frequency can assist in identifying and exploring ideology in a corpus. Not only high-frequency words but also low-frequency words carry importance when compared to each other (Stubbs, 2001; Vessey, 2017). The frequency analysis of the languages identified in Census 2017 report were explored to find out the which languages were typically discussed in the corpus indicating the level of significance attached to different languages.

Similarly, Mutual Information (MI) test scores show co-occurrence of words in a corpus that indicates the strength of collocation relationship between the words (Baker, 2006). Researchers have used various MI scores. For example, Vessey (2017) suggests MI score higher than 3.0 is considered to have statistically significant collocation, whereas Baker (2006) used an MI score of 4.0 and above for a significant relationship between words. Following Vessey’s work on language ideology, an MI score of 3.0 and higher was considered significant. WordSmith also allows researchers to examine the “collocational environment” of search words through patterns of concordance lines that show “evaluative polarity” (Mautner, 2016). In other words, whether the search word is found repeatedly in negative or positive collocational

| Corpus                                         | Total tokens (running words) in text | Types (distinct words) | Type/token ratio (TTR) | % of total words in corpus |
|------------------------------------------------|--------------------------------------|------------------------|------------------------|---------------------------|
| Balochistan Education Sector Plan 2013–2018    | 55,564                               | 6,016                  | 11.76                  | 4.34                       |
| Draft National Education Policy 2017          | 65,278                               | 5,323                  | 11.66                  | 5.10                       |
| Draft Single National Curriculum English (6–8) 2020 | 28,681                               | 3,283                  | 11.66                  | 2.24                       |
| Draft Single National Curriculum History (6–8) 2020 | 8,708                                | 1,822                  | 21.56                  | 0.68                       |
| Early Childhood Care and Education Policy Sindh 2015 | 16,493                               | 2,661                  | 17.21                  | 1.29                       |
| Education in Pakistan_ A whitepaper 2007      | 47,392                               | 4,967                  | 10.87                  | 3.70                       |
| Education Sector Plan 2010–2015 Khyberpakhtunkhwa | 54,436                               | 4,921                  | 10.15                  | 4.25                       |
| Education Sector Reform_ Action Plan 2001–2006 | 38,839                               | 3,818                  | 10.46                  | 3.03                       |
| Gilgit-Baltistan Education Strategy 2015–30    | 69,972                               | 4,667                  | 7.36                   | 5.46                       |
| Green Papers_ National Education Policy Review Process 2006 | 19,772                               | 3,096                  | 16.21                  | 1.54                       |
| Higher Education Medium Term Development Framework II 2011–15 | 21,584                               | 2,717                  | 13.41                  | 1.69                       |
| Minimum Standards for Quality Education in Pakistan 2016 | 26,210                               | 3,638                  | 14.30                  | 2.05                       |
| National Curriculum Framework Pakistan         | 33,496                               | 3,765                  | 11.52                  | 2.62                       |
| National Education Policy 2008 Draft          | 25,861                               | 3,298                  | 13.39                  | 2.02                       |
| National Education Policy 2009                | 28,784                               | 3,477                  | 12.72                  | 2.25                       |
| National Education Policy Framework 2018      | 4,873                                | 1,140                  | 24.08                  | 0.38                       |
| National Plan of Action to Accelerate Education-Related MGDs 2013–2016 | 53,963                               | 2,752                  | 6.22                   | 4.21                       |
| Pakistan 2025_ One Nation One Vision          | 37,634                               | 4,792                  | 13.15                  | 2.94                       |
| Pakistan Education For All Review Report 2015 | 18,254                               | 2,617                  | 15.62                  | 1.43                       |
| Pakistan National Education Policy 2017–2025  | 65,278                               | 5,323                  | 8.58                   | 5.10                       |
| Pakistan National Education Response Resilience Plan for COVID-19 | 7,495                                | 1,398                  | 19.06                  | 0.59                       |
| Pakistan National Human Development Report 2003 | 89,262                               | 5,757                  | 7.08                   | 6.97                       |
| Pakistan National Human Development Report 2017 | 118,366                              | 9,995                  | 9.63                   | 9.24                       |
| Punjab Education Sector Plan 2019–2024        | 41,536                               | 3,655                  | 9.46                   | 3.24                       |
| Punjab School Education Sector Plan 2013–2017 | 51,053                               | 4,195                  | 8.86                   | 3.99                       |
| School Education Sector Plan and Roadmap for Sindh 2019–2024 | 23,953                               | 2,711                  | 12.17                  | 1.87                       |
| Sindh Education Sector Plan 2014–2018         | 110,706                              | 6,356                  | 6.26                   | 8.64                       |
| Single National Curriculum ECCE 2020–Grade Pre I | 39,028                               | 4,852                  | 12.66                  | 3.05                       |
| Single National Curriculum English 1–5 2020   | 28,560                               | 3,326                  | 11.90                  | 2.23                       |
| Single National Curriculum General Knowledge 1–3 2,020 | 21,891                               | 3,095                  | 14.34                  | 1.71                       |
| Single National Curriculum Social Studies 4–5 2020 | 14,676                               | 2,654                  | 18.52                  | 1.15                       |
| The National Skills Strategy 2009–2013        | 13,168                               | 2,290                  | 17.84                  | 1.03                       |
| **Total Primary Corpus**                       | 1,280,766                            | 28,672                 | 2.41                   | 100%                       |
environment. This evaluative polarity is known as semantic prosody or discourse prosody. Similarly, “semantic preference” reveals with which category of other words a search word frequently occurs (Mautner, 2016; Vessey, 2013). For example, Baker (2006) concluded that the term refugees had frequent occurrences with numbers and phrases such as more and more; therefore, she concluded that the word had a semantic preference of quantification. Semantic preference and discourse prosody together show “what kind of social issues a particular lexical item is bound up in” and the ideologies surrounding those issues (Mautner, 2016, p. 161). The collocational environment MI values showed the collocational environment of the languages as they were discussed.

The DHS techniques guided the macro and detailed textual analysis. The focal strategies from Reisigl and Wodak (2009) used for this study were predication, argumentation, and perspectivization along with intertextual and interdiscursive analysis. The authors of texts use different discursive strategies to persuade the reader or audience. Through predication strategy, they assign positive or negative evaluative traits through adjectives, phrases, and clauses. Therefore, the predication techniques revealed what other semantic terms the authors attributed with the languages. Argumentation strategy involves the claims and how they are justified through culture-specific topoi or loci that are the underlying assumptions for concluding the argument. The arguments with the underlying topos revealed the policy authors’ use of different arguments to support their views. Perspectivization strategy, to some extent, relies on intertextuality to other texts and interdiscursivity by drawing upon references to other discourses for legitimization and support for the author’s perspective (Reisigl & Wodak, 2009). The perspectivization techniques found in the text showed the use of intertextual and interdiscursive references by the policymakers to support their perspectives.

The corpus techniques and analysis of DHS techniques together revealed the dominant ideologies through analysis of typicality, significance, attribution and discursive strategies in the language-in-education policy documents of Pakistan.

**Findings and Discussion**

This section answers the research question revealing the ideologies related to the languages of Pakistan through highlighting the typicality of topics/CQTs, significant words around the CQTs, and the collocational environment of the CQTs through expanded concordance lines. These techniques capture the ideologies of the policymakers through corpus analysis.

**Frequency Analysis: The Typicality of Topics/CQTs**

One of the key findings that surfaced through the distribution plot was that the CQT LANGUAGE had low frequency in National Education Policy (2009) and (2017) documents, while it occurred more frequently in other recent policy-related documents. Nevertheless, it occurred in 100% of the documents (32) as it was the main CQT for corpus development. This reflects the validity of the corpus data as the study aims to explore discourse and ideology in language-in-education policy documents. The wordlist showed that LANGUAGE occurred 1,009 times in the corpus. In contrast, the CQT EDUCATION appeared as the first most frequent content word ranked 8 with a frequency of 17,131, contributing 1.34% in the entire corpus across 32 texts. This is because the issue of language is a
subject discussed in small portions in language-in-education policy documents, whereas the issue of education dominates the discourse. The next word of interest ENGLISH occurred 617 times in 26 texts. Compared to English, the CQT URDU had a frequency of 243 in 23 texts of the corpus. It reflects that ENGLISH has been discussed roughly three times more frequently than URDU. CQT MOTHER occurred 118 times across 17 texts and the phrase MOTHER TONGUE 75 times in 11 documents. This is the term commonly used for local languages the learners speak at home and has been heavily discussed in the domain literature. Other adjectives as CQTs such as “local” and “indigenous” showed even lower frequency. However, they are not reported here due to limited space. Overall, the low frequency of adjectives commonly used to name the local languages implies that the discussion of local languages is very limited in the education document corpus. Besides the significant CQT from literature in the domain of language-in-education, it was essential to find out which languages are frequently discussed in the corpus and the discourse around them. For this reason, the latest language data from census report of Pakistan 2017 (Pakistan Bureau of Statistics, 2017) served as the starting point. It showed 09 local languages, including six major languages Urdu, Punjabi, Sindhi, Pushto, Balochi, Saraiki, and 03 minor languages, including Kashmiri, Hindko, and Brahvi. Table 3 shows frequency analysis from the corpus for 09 languages identified in Pakistan Census (2017) (Figure 2).

Urdu and Sindhi appeared to be the most frequently occurring local languages in the corpus, besides Balochi and Punjabi 9 and 8 occurrences, respectively. English stands as the most frequently discussed language in the entire corpus. It is pertinent to compare the representation of languages in the discourse in language-in-education documents and the number of speakers. In Figure 1 above, the most prominent Pakistani language is Punjabi with around 80 million speakers, followed by Pushto roughly (40 million), Sindhi (30 million), Saraiki (25 million), and Urdu (15 million). The representation of languages in the corpus shows almost the reverse trend. Punjabi and Pushto are the two major languages according to the latest Pakistan Population Census (2017). The reference to these languages occurred only eight and three times respectively in five and three documents each. It means that the discourse of major local languages has been largely absent in language-in-education policies of Pakistan. Given this situation, it is hard to expect that other minor or smaller indigenous languages would receive any representation in language-in-education policy discourse.

The frequency analysis shows that English dominates as the typical topic among the languages discussed in the corpus. Urdu (the national language of Pakistan) is the second most discussed topic while the regional and local languages of Pakistan are least discussed topics. Nevertheless, frequency analysis has its limitation as sometimes frequency analysis in isolation could be misleading (Archer, 2009). These frequencies do not show the context of discourse. Other tools within CL such as collocation and concordance (Baker et al., 2008; Vessey, 2017) show the contexts of the topics or CQTs.

Table 3. Major Language Frequencies in Primary Corpus.

| Word     | Frequency | % of words in corpus | Occurrence in no. of texts in corpus | % of texts in corpus |
|----------|-----------|----------------------|--------------------------------------|----------------------|
| Urdu     | 243       | 0.02                 | 23                                   | 71.88                |
| Punjabi  | 8         | 0.00                 | 5                                    | 15.63                |
| Sindhi   | 41        | 0.00                 | 8                                    | 25.00                |
| Pushto   | 3         | 0.00                 | 3                                    | 9.38                 |
| Balochi  | 9         | 0.00                 | 5                                    | 15.63                |
| Kashmiri | 1         | 0.00                 | 1                                    | 3.13                 |
| Saraiki  | 1         | 0.00                 | 1                                    | 3.13                 |
| Hindko   | 1         | 0.00                 | 1                                    | 3.13                 |
| Brahvi   | 3         | 0.00                 | 2                                    | 6.25                 |

*English presented for comparison of languages present in the corpus.

Figure 2. Mother tongues and number of speakers Pakistan. Graph generated from Pakistan Bureau of Statistics (2017) data.
The analysis of statistically significant collocates started with the search of the CQT LANGUAGE showing the strength of the relationship with other words. LANGUAGE and EDUCATION showed a low MI value (0.5); therefore, not even a weak relationship was indicated between the words. This implies that discourse on the specific term “language-in-education” in the educational documents is minimal.

The selected significant collocates of CQT LANGUAGE/ES in Table 4 shows that it mostly co-occurred with words associated with education or school (learning, school, teaching, instructions, medium), besides three languages (Urdu, English, and Sindhi) and a few hinting at the context of discourse about language issues (need, mother (tongue), policies, and problem). Thus collocation analysis reveals that CQT LANGUAGE has semantic preferences with international, national, and local languages and policy issues in the discourse of language-in-education policy community. These results further helped identify CQT MOTHER TONGUE analyzed later in this section.

URDU collocates with a range of words with various perspectives related to education with MI values between 5 and 8 (see Table 5). URDU collocates with ENGLISH 8 times (MI value 8–11). MOTHER occurs with TONGUE, TONGUES, and LOCAL 77, 14, and 6 times and with education-related terms MEDIUM and INSTRUCTION 7 and 5 times respectively (MI value 6–13.41). This means unlike English and Urdu, mother tongues were largely absent in proximity of educational context showing the policymakers sidelined the local languages in educational domain. There was no significant collocation relationship between the CQT PUNJABI/PANJABI, PUSHTO/PASHTO, BALOCHI, KASHMIRI, SARAIKI, HINDKO, and BRAHVI with any other word. However, they were found in the word frequency analysis. During the analysis, the corpus showed that the use of the term mother tongue(s) for local language was more common, whereas very few local languages were discussed by their names. Therefore, the concordance analysis below has used mother tongue rather than the specific local language names. Urdu and English have been explored by names. The following section presents the concordance analysis of selected lines from the corpus related to English, Urdu, and mother tongues.

### Table 4. Education Related Significant Collocates of LANGUAGE.

| Word 1    | Freq. | Word 2   | Freq. | Texts | Gap | Joint | MI  |
|-----------|-------|----------|-------|-------|-----|-------|-----|
| LANGUAGES | 183   | LEARNING | 3,837 | 7     | 3   | 9     | 4.04|
| LANGUAGES | 183   | SCHOOL   | 4,768 | 4     | 2   | 7     | 3.36|
| LANGUAGES | 183   | URDU     | 243   | 4     | 4   | 7     | 7.66|
| LANGUAGE  | 1,009 | LEARNING | 3,837 | 11    | 1   | 35    | 3.53|
| LANGUAGE  | 1,009 | ENGLISH  | 617   | 11    | 2   | 29    | 5.90|
| LANGUAGE  | 1,009 | INSTRUCTION | 303 | 10    | 2   | 34    | 7.15|
| LANGUAGE  | 1,009 | MEDIUM   | 632   | 10    | 2   | 25    | 5.65|
| LANGUAGE  | 1,009 | TEACHING | 1,582 | 8     | 1   | 24    | 4.27|
| LANGUAGE  | 1,009 | APPROPRIATE | 472 | 8     | 3   | 13    | 5.13|
| LANGUAGE  | 1,009 | NEEDS    | 1,517 | 8     | 4   | 12    | 3.33|
| LANGUAGE  | 1,009 | MOTHER   | 118   | 5     | 2   | 8     | 6.43|
| LANGUAGE  | 1,009 | SINDHI   | 41    | 5     | 1   | 6     | 7.54|
| LANGUAGE  | 1,009 | POLICIES | 554   | 4     | 1   | 5     | 3.52|
| LANGUAGE  | 1,009 | PROBLEM  | 315   | 5     | 5   | 5     | 4.33|

**Concordance Analysis: The Immediate Context of CQTs**

The frequency and collocational significance are valuable techniques for finding out the topics’ typicality and the
The contextual use of the words can be derived from the frequency of concordance lines showing the examples of the use of a word in context and functions of those words (Baker, 2006).

The pattern of concordance lines in Table 7 reveals that words around English explicitly suggest that English is necessary for “economic advancement” in the “globalised and competitive world” therefore “not just the elite,” but all schools should teach it or use it as a medium of instruction (MoI) as “children with better English language skills tend to have more opportunities.” The repeated use of the adjective “important” positively evaluates it as a language necessary for “advancement,” in the “modern world” that “continues to globalize.” This is a clear example of language-as-resource. The expanded concordance lines in the examples below show the importance of English.

**Example 1**

Private schools predominantly use English as medium of instruction, with a very strong focus on active use of English language by children, while public sector schools mostly use Urdu and the regional mother tongue as a language of instruction. Although, English is taught in many public-sector schools but because of the low capacity of teachers to teach English as a second language, children are unable to achieve even basic competency levels. The disparity in acquisition of language skills is one of the most divisive factors when it comes to application of education skill for jobs. Children with better English language skills tend to have more opportunities because of its glorified status in society and its use as a de facto official language. Generally, students in high-cost private schools are more advantaged than children studying in public sector schools, low-cost private schools and madrassas.

(National Education Policy Framework 2018)

**Example 2**

Furthermore, the predominant medium of instruction in rural areas, Urdu, makes the transition to the English medium higher education system difficult for students. The English language proficiency requirement poses a severe challenge to education.
Table 7. Selected Concordance Lines of ENGLISH.

| English necessary | everyone economic advancement. After 1947, English has continued to enjoy privilege of not just the elite. 2. English language teaching in schools should skill for jobs. Children with better English language skills tend to have more opportunity. Globalized and competitive world. 21. English is an international language, and it
| Table (02) (book/Lesson) (01) c.2) English Reading Assessment Sheet (Sentence) citizens, and not just the elite. 2. English language teaching in schools should skill for jobs. Children with better English language skills tend to have more opportunity. Globalized and competitive world. 21. English is an international language, and it
| Importance of English | If a century or so, the influence of English language has broadened way outside possible sources. The importance of English language for advancement in the comprehension of increased importance of English language. It has been felt that the 08. Recognition of the importance of English in the modern world has been manifested by 2009 recognises the importance of English in a world that continues to global

Table 8. Selected Concordance Lines for ENGLISH.

ta prefer to enroll their children in English medium schools to seek better employment. Low socio-economic strata to learn English language. 9. A comprehensive school advantage. Everyone now needs to learn English, not because of any colonial trimmission across the board demand for learning English language in the country. 75. Differ

Table 9. Selected Concordance Lines of ENGLISH.

rd. 9.60.1 Policy Recommendations: 1. English should be made a compulsory subject. Medium of instruction will be English from Class 1. Urdu to be taught as a subject from Class III 2008 Education, medium of instruction should be English for physical and natural sciences. For higher education, Mol may be English Ed in Pak 2006 Curriculum Uniform c

equity even in urban areas for students coming from Urdu medium and disadvantaged backgrounds. This makes them more vulnerable amidst fierce competitions in universities’ admission tests. This increasing institutional heterogeneity creates classes of education where the advantaged retain their right to privileges that define their upward social mobility and worsens the development divides.

(United Nations Development Program, Pakistan, 2017)

Examples 1 and 2 exhibit that while English is considered the language of employment and opportunities, the students taught in Urdu medium schools are at a disadvantage. The negative words “disadvantaged,” “vulnerable,” “worsen” show the negative evaluation associated with Urdu.

Concordance trend in Table 8 indicates that the policy documents establish that English is not only considered necessary by the policymakers; instead, it is also the “demand,” “need,” and “prefer(ence)” of the parents and communities for their children to acquire English language proficiency to “seek better employment” opportunities.

The concordance trend in Table 9 manifest mixed opinions on the introduction of English in the curriculum discussed as “compulsory subject,” and as a “medium of instruction as early as Class 1” or “Class III” and at “higher education.” These and other examples indicate that the documents argue that children should have good English proficiency for the highly competitive world.

The concordance lines in Table 10 reveal a trend of Urdu being discussed as a national language and a regional lingua franca.

Table 11 reveals multiple opinions about the introduction of Urdu in various roles and stages such as in “upper classes,” “as a subject from class 1,” “as medium of instruction at the primary level” and according to the “local decision” and “wishes.”

The concordance trend on mother tongue in Table 12 reflects it being perceived as a resource that should be capitalized on by using more and more mother tongue-based education. The benefits of mother tongue are positively predicated with the help of several positive lexical items “additional,” “better expressions,” “superior achievement,” “perform better,” “international research supports.” All these words in the discourse around mother tongue are intentionally used as a persuasive technique. In context of the Urdu language, the discourse was positive or neutral. It was positive as emphasized by higher frequency (243 occurrences) in the entire corpus; moreover, it was predicated as a “national”
language or “a lingua franca.” MOTHER TONGUE with 75 occurrences, on the other hand, showed lower frequency, however, several positive attributes are carefully used.

Table 13 reveals the trend that mother tongue is encouraged at “early” stage of education for students to have a strong foundation in their own language as they “need confidence” to “value” their own language. Neither distance to schools nor mother-tongue as the language of instruction were considered a barrier in education. This triangulates the finding from wordlist and collocation analysis that local languages are greatly ignored in language-in-education policy discourse.

The policymakers discussed English the most compared to other languages of Pakistan. The typicality of topics appeared in a hierarchy where English occupied the top place, followed by Urdu and mother tongues. English occurred in discourse of education, followed by Urdu while mother tongues only very marginally. English, Urdu, and mother tongues are positively attributed shown through concordance analysis. Together these findings reveal that English occupies the center place as being the most important language sharing some grounds with Urdu language in the discourse of the policymakers evident in the policies documents. The mother tongues have been marginalized in the educational domain. They are valued as a resource for the society, but no space is indicated for them in the education. English remains the central of policymakers’ discourses of education in the policy documents.
Discourse Historical Analysis

In this section, discourse historical analysis of the downsampled language-in-education policy documents is presented, highlighting ideologies of policymakers and their discursive strategies of argumentation, predication, and perspectivevisation under DHS principles in order to answer the research question through detailed document analysis. First, four Federal government documents are discussed, followed by three documents from provincial governments.

One of the downsampled policy documents Education in Pakistan: A Whitepaper (2007) reveals that the topoi underlying the arguments are equality, problem, solution, and balance. The sections related to language-in-education show discourse of problem that builds up before presenting the innovative language policy recommendations. The document refers to both the historical and contemporary major discourses around language-in-education and especially swaying medium of instruction of English, Urdu, and mother tongue and then proposes its new proposal. The topos of problem related with language-in-education policy is reflected in the use of negative evaluative lexis “debate,” “friction,” “conundrum,” “blunted,” “tension,” “confusion at the policy level,” and “half-baked measures.” All these semantically negative terms are used to highlight the issue of language-in-education that Pakistan has been facing since its independence. The document uses predication strategy positively referring to the “rich regional languages,” and “natural multilingualism of Pakistani people.” It positively acknowledges that “mother tongue allows better expressions and conceptual understanding.” However, there is negative attribution for the mother tongue which “does not prepare a child for secondary education in another language.” Detailed inter-discursive references are made to the main positions or discourses surrounding the topic of language-in-education. These positions are related to mother tongue, national language (Urdu), and international language (English). First, mother tongue is considered effective for better expression and cognition, but it does not prepare the child for later stages of education in the dominant language, nor is it “perceived to be the language of economic gain.” Moreover, teachers may not know the mother tongue or may not be able use all mother tongues spoken in a linguistically diverse area. There is a reference to the “international research that supports mother tongue as the best medium of instruction, but in the context of a complex population mix in Pakistan, Urdu and English are both essential.” Second, Urdu performs as the “regional lingua franca”; however, it is considered “imperialist language” in some areas. Third, English “helps people to move up economically”; nevertheless, it is not “considered a national language” nor is there enough readiness to teach it properly due to teachers and teaching resources. The topos of solution is closely linked with the topos of balance since no language is sufficient for all the domains of social life. Again such a balance is not easy as it notes, “a difficult balance must be struck to ensure language proficiency, conceptual learning, effective administration, religious learning and economic and cultural equity.” The author tries to convincingly present the solution to the “difficult” problem as “Federal Ministry of Education has found another answer to the issue by declaring Urdu as the medium-of-instruction at the middle and secondary level for social sciences and English as the medium of instruction for mathematics and natural sciences.” This means that new policy would allow teaching of different subjects in Urdu and English. However, the “rich regional languages” are not mentioned and marginalized in the statement. Later the document presents the policy recommendations as 1. First 5 years left to provincial governments to decide the language-in-education policy matters, but ensuring mother tongue as MoI in first 3 years in linguistically less diverse areas and any other language [most probably Urdu] in urban multilingual settings; and Urdu as a subject where it is not MoI. 2. Class VI onward Urdu for Social Sciences and English for Mathematics and Natural sciences. 3. Regional languages should be taught as a compulsory subject up to class VIII if the provincial government desire. 4. English for physical and natural sciences in higher education including Colleges and English or Urdu for Social Sciences in Colleges. The topos of equality is supported by the strategy of perspectivisation through intertextual reference to Convention Against Discrimination in Education (UNESCO, 1960) to accentuate the effect of the argument that social inequality is due to the widespread social stratifications and notes that “It is our purpose to address issues of gender, economic and geographic disparities, parallel systems of education and medium of instruction, especially teaching of the English language” (p.28) [emphasis added]. The underlying assumption related to language is that medium-of-instruction and English language teaching and resulting English language proficiency are the contributing factors to the discrimination and divide among people. This indicates that perhaps acquiring English language proficiency could help reduce discrimination in society. Later in the document, a comprehensive section on policy recommendations for “Teaching of English Language” is detailed. Several positive evaluative terms are attributed to English as “vehicle for social and economic advancement, the language of international communication, cosmopolitan life, and transnational trade and commerce,” “basic necessity,” “international necessity,” “dynamic,” “language of advancement,” “important[t] for pursuit and participation in international research, trade, commerce and other avenues of international communications.”

National Education Policy Framework (2018) document is comparatively not very detailed, so is the discussion on language-in-education. The discourse mainly centers around the English language with eight occurrences, while Urdu and “regional mother tongue” occur once each. The
central argument is to provide equal opportunities to all students. The topos used to support the argument is uniformity. Perspectivation strategy to signify the importance of uniformity across the provinces for language-in-education policy is created through intertextuality. One reference is made to the Article 25 A and the other to the inaugural speech of Prime Minister Imran Khan, in which he committed to bringing uniformity in education. The lexeme “standard” occurred nine times in the context of language education as the document highlights there are “huge differences in standard” in “various education systems across the country” that results in “children with better English language skills have more opportunities.” Another positive word associated with English is “advantaged.” The public school where Urdu and regional languages are used as MoI are considered at a disadvantage. Further predication strategy for the topos of uniformity is achieved through repeated use of lexis “common” “agreed” “agreement” on “multilingual policy” and “proficiency” ensuring “proficiency” in “English as a second language.”

National Curriculum Framework Pakistan (2018) makes several intertextual references to all the education policies issued over the history of Pakistan from 1947 to 2009 to situate the “framework” in the historical and political context. Reference to the 1973 Constitution of Pakistan is made to show that educational matters were initially a “joint concern” of the Federal and Provincial governments. Nonetheless, the importance and need of the policy document is highlighted by referring to the 18th Constitutional Amendment that allowed autonomy to the provinces resulting in a “number of issues and challenges. They need to be addressed urgently” as they are of “great concern” related to uniformity in educational matters. The negative predication strategy is used for local language speaking families through a series of negative lexis “low income”, “illiterate or less educated,” “having proficiency in local language only.” Local languages are given the role of “supporting” language for the acquisition of the official language. A detailed intertextual reference is made to the document “Minimum standards for quality education in Pakistan 2016” that briefly on the teaching of English and different strategies and goals related to “learning in English” and “learning of English as a language.” The argumentation scheme of uniformity draws on topos of responsibility, suggested by Wodak (2001). The topos of responsibility assumes that the “governments are responsible for the emergence of a problem, it must act in order to find solutions to these problems” (Wodak, 2001, p. 73). The document repeatedly urges provincial governments and stakeholders to draw consensus for uniformity of MoI as the solution to the problem of MoI. While it does not explicitly urge for English to be the medium of instruction, several references favoring the “official language” imply it when seen together with the phrase “official language (English, at present).” The document hints that Sindhi has been progressing as the regional language, but no other provincial language have been mentioned. The document lays recommendations for provincial governments to use local mother tongues until grade 5 as MoI and English onward.

Single National Curriculum English 6 to 8 (2020) reflected the greatest number of CQT LANGUAGE hits (251 occurrences). The main argument scheme of the document is effectiveness. The overall argument revolves around making English language teaching “holistic” emphasizing speaking and listening skills through “interactive teaching and learning practices in the classroom” to equip learners to achieve success in different domains of life. The topoi of usefulness and comprehensiveness are used to support the argument in favor of the effectiveness of English language learning. The topos of usefulness in favor of the English language goes as “if one wants to achieve high in academic and professional life, English language proficiency is useful to avail it.” Several positive phrases are used as predication strategy using “important skills when it comes to education at all levels, personal development, global communication, and making better professional choices” (p.2). The positive evaluation is approached with phrases “to excel in any field,” “further education,” and “becoming productive members” is only possible through comprehensive English languages skills. Intertextual references to United Nation’s Education for Sustainable Development and Global Citizenship Education are employed to strengthen the argument as “the rationale is to empower learners of all ages to become proactive contributors to a fair, peaceful, tolerant, inclusive and sustainable world” (p.3). English language skills are supposed to be the vehicle to gain those characteristics envisioned by the UN.

The downscaled Balochistan Education Sector Plan [BESP] (2013) document briefly presents the linguistic profile of the province dominated by Brahvi and Balochi speakers, besides Punjabi and Urdu ethnic groups in Quetta, Hazara, and “small pockets of other ethnic groups.” The document proposes evaluating language teaching, Balochi, Brahvi, and Pashto as MoI and other mother tongues, and a system to use these languages in education alongside “second and third languages.” The main arguments are autonomy, problem, and promotion. It states the vision “BESP considers teaching in mother tongue at the primary level as critical to the cognitive development of the child. . . [and] problems of higher classes where Urdu and English courses have been designed on unrealistic assumptions about children’s capacity.” The employed topos for the argument of local language promotion can be related to the usefulness and uselessness suggested by Wodak (2001) that manifests that if something is useful under certain circumstances, it should be performed. In this case, the implementation and promotion of the mother tongue will result in a range of “useful” outcomes.
The downsampled *Gilgit-Baltistan Education Strategy (GBES)* (2015) document argues in favor of mother tongue as MoI and “calls for mother-tongue instruction in early grades.” Predication strategy to highlight the importance of mother tongue use in education is carried out through multiple positive terms “important,” “essential,” “culturally,” “linguistically,” “early,” “threshold” in support of mother tongue use at an early age. Whereas the negative lexis “very limited” resulting in “barrier” and “loss of learning” shows the result of its absence. The perspective strategy through interdiscursivity of the “international research” favoring mother tongue at an early age is achieved coupled with intensification strategy through repeated use of “should.” The author discusses the rich linguistic diversity of the region with the term “abundance of mother languages” and mentions the main languages Shina, Barushiski, Balti, Wakhi, and Khowar, and migrant languages including Pashto, Punjabi, Gujari, Ughur, and Kashmiri.

The *Early Childhood Care and Education Policy Sindh (ECCEPS)* (2015) document draws intertextual reference to the then existing National Education Policy (2009). It pledges to continue the policy guidelines for Sindh early year education. In a paragraph-long discussion on the language, culture, and diversity, the document shows a positive stance toward language diversity as it uses phrases showing intensification through “must,” “ensure” along with the words “recognition” of “language” and “identity” and “respect for language” and “each family unit” to be “acknowledged” and “represented” in schools and communities. Urdu is stated as the “national language, Sindhi as the provincial, and English as a foreign language.” Sindhi is prescribed as MoI as the first language, Urdu as the second language, and English as the third language. Without naming any local languages, the document states that “Sindh must provide support to address those students whose spoken language /mother-tongue is another language than the language stated above in an early childhood classroom in any particular locality (urban/rural).” The local languages are created as outgroup referred to as “another language” or mother tongues while the only locally recognized mother tongue remains Sindhi.

The Federal government’s *Education in Pakistan: A Whitepaper* (2007) document reveals a complex and even contradictory interplay of ideologies. For instance, the mother tongue-based education in the initial 3 years for some learners in specific geographical regions alongside Urdu indicates language-as-right ideology and the ideology of monolingualism aimed at command in two languages separate from each other resulting in “double monolingualism” (Garcia, 2009). For people in the Urban linguistically more complex settings, the document reflects monoglot ideology disregarding the linguistic diversity (Silverstein, 1996) in that part of the society. Also, it shows the more robust version of double monolingualism (Garcia, 2009) from secondary school or education in grade 6 onward in two dominant languages, English and Urdu, simultaneously. In fact, the author highlighted this solution of using two languages simultaneously for teaching different subjects to address the problem of language-in-education from secondary to higher education. The balance is further desired by leaving the decision of teaching regional or local languages as compulsory subjects up to grade eight to the discretion of provincial governments. Two of the downsampled Federal government policy documents can be considered tolerance-oriented policies, leaving the responsibility of maintaining minority languages on the language communities without providing governmental resources and support (Kloss, 1998; Wiley, 2013). The policies’ stated implementation plan also reflects its proximity with the transitional version of a dual language or bilingual program (Palmer et al., 2015). It incorporates the learners’ mother tongues only for a few years and shifts to English and Urdu later. Contrary to it, the developmental program focuses on developing literacy in English and the mother tongue of the child. The child can learn a minority language for an extended period (Palmer et al., 2015).

Two of the downsampled provincial government policy documents appear expediency-oriented policies (Wiley, 2013). *BESP* document seems to be expediency-oriented or a “weaker version” of promotion-oriented policies as it does not intend for enhanced use of minority language beyond middle and secondary school level and not in official domains. The discourse suggests that the provincial government does not leave local language promotion to the communities and is ready to promote those languages, unlike tolerance-oriented policies. However, there is no discussion of the government resources allocated for local language promotion. From its implementation plan, the policy does not specify any “transition” of MoI. However, it considers the role of English, Urdu, and mother tongues in education, mother tongue at primary while second and third languages at later stages. The overall emphasis remains on language use concerning students’ cognition and teaching other languages with a more communicative focus. Since it does not clearly state the “transition” of MoI at later stages, it is difficult to pin down the ideological orientation of the document. Nevertheless, it seems to be falling under the language-as-right orientation, a unidirectional bilingual program in which the minority language students must learn the dominant languages and maintain their language simultaneously (Kloss, 1998; Wiley, 2013).

Similarly, the *GBES* document can be considered an expediency-oriented language policy (Wiley & Tollefson, 2013) as it favors the local languages more than the dominant languages; however, the extent of its usage is limited to 3 years with the unclear application of mother tongues beyond 3 years. Language programs based on the transition model aim to prepare the learners for the dominant language, unlike developmental programs where students develop their
language and the dominant language for a very long period (Helot & Garcia, 2019). Therefore, the discourse within the document relates to language-as-right and language-as-resource ideologies. The implementation plan indicates that it may be considered a monoglossic ideology (Garcia, 2009). The aim is to shift to English and Urdu after the first 3 years if the local authorities within the province deem it desirable by adopting a transition to major languages after a few years.

It is doubtless that the ideologies present in the policy documents have significant impact on the stakeholders. It is the actors and stakeholders that make the implementation of any policy successful. Therefore, all the actors identified by Zhao and Baldauf (2012) working at the macro, meso and micro levels (Ricento & Hornberger, 1996) contribute to the implementation of the ideologies embedded in the language-in-education policies.

**Conclusion and Implications**

This study attempted to answer the research question focusing on the ideologies related to languages through the discourse language-in-education policy documents of Pakistan. The findings revealed that English dominated the ideological space of discourse existing in the language-in-education policy texts explored through typicality, significance and attribution techniques of corpus analysis. The discourse around Urdu and mother tongues had positive prosody generally. Whereas Urdu was associated with educational discourse, mother tongues were completely marginalized from language-in-education related discourse resulting in complete absence from ideological orientation of considering them important for education.

The detailed DHS discursive analysis showed that English, Urdu and mother tongues were discussed in positive light. English and Urdu had their own importance in education. It also indicated agreement on the use of mother tongue or local language in education by the provinces, however, there is variation in the approach of different provincial governments. The analysis confirmed that over all the policies had underlying monoglossic ideology evident from their early-exit transition implementation plans for local languages. The policies were ideologically tolerance-oriented or expediency-oriented for local languages. This means that either no government resources were allocated for development of languages to be used in education or the role of local languages is constrained for first few years of education. These ideologies do not seem to promote the linguistic diversity of the country.

These findings are significant for language policy actors or “people” at macro, meso and micro levels. The people-with-power need to balance their ideological orientations for international, national and local languages as all these languages perform significant roles in lives of people. It has been established the ideologies enacted through policies are given huge importance by schools and those ideologies are translated through schools. People-with-expertise should work for corpus planning of several local languages so that they can be ready for educational domain. People-with-influence such as different NGOs working for local languages should understand that local languages have been largely ignored in the educational space, therefore, they should coordinate with the policymakers and inform provide guidelines for effectively including local languages in education. People-with-interest including communities can benefit from these findings and create pressure from the bottom through building confidence, pride and realization among the local language speakers. More specifically, these findings have implications for the policymakers both in the Federal government and the regional or provincial governments of Pakistan. First, since the Federal government is consistently urging the provinces for uniformity, there should be similarities of discourses and ideologies about the languages across the government documents from these layers of government. The provincial governments’ discourses need some consistent efforts to present the linguistic profile of each province and seek support from all corners given their autonomous powers after 18th Constitutional Amendment. Furthermore, the Federal government should recognize and allow discourse space to more major and minor local languages and provide guidelines to the provinces for the development and projection of the local languages. This study was limited to the analysis of official documents only. Future research may focus on the discourse in the media, both electronic and print, to explore the discourse and ideologies around languages in Pakistan.
### Appendix

Dispersion Plot CQT LANGUAGE.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Muhammad Asim Khan https://orcid.org/0000-0001-7261-5467

References

Archer, D. (2009). Does frequency really matter? In D. Archer (Ed.), What's in a word list? (pp. 1–16). Ashgate.

Baker, P. (2006). Using corpora in discourse analysis. Continuum.

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McNeney, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. Discourse & Society, 19(3), 273–306.

Blommaert, J. (1999). The debate is open. In J. Blommaert (Ed.), Language ideological debates (pp. 1–38). Mouton de Gruyter.

Busch, B. (2011). Trends and innovative practices in multilingual education in Europe: An overview. International Review of Education, 57, 541–549. https://doi.org/10.1007/s11159-011-9257-1

Canese, V. (2018). Language ideology as a conceptual framework to analyze issues related to language policy and language education. Revista Científica de la Facultad de Filosofía, 6(1), 20–42.

Catalano, T., & Waugh, L. R. (2020). Critical discourse analysis, critical discourse studies and beyond. Springer International Publishing.

Cheng, W. (2013). Corpus-based linguistic approaches to critical discourse analysis. In C. A. Chapelle (Ed.), The encyclopedia of applied linguistics (pp. 1353–1360). Blackwell Publishing.

Cushing, I. (2021). Policy mechanisms of the standard language ship, and/or publication of this article. Routledge.

Flores, N., & Baetens Beardsmore, H. (2015). Programs and structures in bilingual and multilingual education. In W. Wright, S. Boun, & O. Garcia (Eds.), The handbook of bilingual and multilingual education (pp. 205–222). Wiley.

Gal, S. (2006). Migration, minorities and multilingualism. In C. Mar-Molinero & P. Stevenson (Eds.), Language ideologies, policies and practices: Language and the future of Europe (pp. 13–28). Palgrave Macmillan.

Garcia, O. (2009). Bilingual education in the 21st century: A global perspective. Wiley-Blackwell.

Gee, J. P. (1996). Social linguistics and literacies: Ideology in discourses. Taylor & Francis Group.

Haider, A. S., Olimy, S. M., & Al-Abbas, L. S. (2021). Media coverage of Syrian female refugees in Jordan and Lebanon. Sage Open, 11(1), 2158244021994811.

Helot, C., & Garcia, O. (2019). Bilingual education and policy. In J. Schweiter & A. Benati (Eds.), The Cambridge handbook of language learning (pp. 649–727). Cambridge University Press.

Johnson, D. C. (2013). What is language policy? Language policy (pp. 3–25). Palgrave Macmillan.

Kloss, H. (1998). The American bilingual tradition. Language in education: Theory and practice No. 88 (p. 1400). Delta Systems Co., Inc.

Krzyżanowski, M., & Wodak, R. (2011). Political strategies and language policies: The European Union Lisbon strategy and its implications for the EU’s language and multilingualism policy. Language Policy, 10, 115–136.

Lin, A. (2014). Critical discourse analysis in applied linguistics: A methodological review. Annual Review of Applied Linguistics, 34, 213–232.

Mahboob, A. (2002). “No English, no future!” Language policy in Pakistan. In Gyasi, O. Samuel & H. Beverly (Eds.), Political independence with linguistic servitude: The politics about Languages in the developing world (pp. 15–39). Nova Science Publishers.

Mautner, G. (2016). Checks and balances: How corpus linguistics can contribute to CDA. In R. Wodak & M. Meyer (Eds.), Methods of critical discourse studies (3rd ed., pp. 154–180). Sage.

Ministry of Federal Education and Professional Training. (n.d.). Single National Curriculum. Government of Pakistan. Retrieved July 3, 2021, from https://snc.gov.pk/

Mulder, J. M. (2014). “Enabling” participatory governance in Education: A Corpus-based critical analysis of policy in the United Kingdom. In P. Smyers, D. Bridges, N. Burbules, & M. Griffiths (Eds.), International handbook of Interpretation in educational research (pp. 441–470). Springer.

National Curriculum Framework Pakistan. (2018). Ministry of federal education & professional training. Islamabad. https://planipolis.iiep.unesco.org/en/2019/national-curriculum-framework-ncf-6581

National Education Policy. (2009). Ministry of education. Government of Pakistan. http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/pakistan_national_education_policy_2009.pdf

National Education Policy Framework. (2018). Ministry of federal education & professional training. Islamabad. https://planipolis.iiep.unesco.org/en/2018/national-education-policy-framework-2018-6524
Pakistan Bureau of Statistics. (2017). Salient features of final results census-2017. Retrieved June 20, 2021, from https://www.pbs.gov.pk/sites/default/files/population/2017/sailent_feature_census_2017.pdf

Palmer, D. K., Zuñiga, C. E., & Henderson, K. (2015). A Dual language revolution in the United States?: On the bumpy road from compensatory to enrichment education for bilingual children in Texas. In W. E. Wright, S. Boun, & O. Garcia. (Eds.), The handbook of bilingual and multilingual education (pp. 447–460). Wiley Blackwell.

Partington, A. (2010). Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: An overview of the project. Corpora, 5(2), 83–108.

Reisigl, M., & Wodak, R. (2009). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), Methods of critical discourse analysis (pp. 87–121). Sage.

Ricento, T. (2013). Language policy, ideology, and attitudes in English-dominant countries. In R. Bayley, R. Cameron, & C. Lucas (Eds.), The Oxford handbook of sociolinguistics (pp. 525–544). Oxford University Press.

Ricento, T. K., & Hornberger, N. H. (1996). Unpeeling the onion: Language planning and policy and the ELT professional. TESOL Quarterly, 30(3), 401–427.

Ruiz, R. (1984). Orientations in language planning. ABE Journal, 8(2), 15–34.

Scott, M. (2013). WordSmith tools version 6: Lexical Analysis Software.

Shamim, F. (2008). Trends, issues and challenges in English language education in Pakistan. Asia Pacific Journal of Education, 28(3), 235–249. https://doi.org/10.1080/02188790802267324

Shao, Q., & Gao, X. (2019). Methods and approaches in language policy research. In J. McKinley & H. Rose (Eds.), The Routledge handbook of research methods in applied linguistics (pp. 253–263). Routledge.

Shohamy, E. (2006). Language policy: Hidden agendas and new approaches. Routledge.

Siddiqui, S. (2016). Education policies in Pakistan: Politics, perceptions, and practices (1st ed.). Oxford University Press.

Silverstein, M. (1996). Monoglot “Standard” in America: Standardization and metaphors of linguistic hegemony. In D. Brenneis & R. K. S. Macaulay (Eds.), The matrix of language: Contemporary linguistic anthropology (pp. 284–306). Routledge.

Spolsky, B. (2004). Language policy: Key topics in sociolinguistics. Cambridge University Press.

Stubbs, M. (2001). Words and phrases: Corpus studies of lexical semantics. Wiley-Blackwell.

Tollefson, J. (2006). Critical theory in language policy. In T. Ricento (Ed.), An introduction to language policy: Theory and method (pp. 42–59). Blackwell Publishing.

Tollefson, J. W. (2000). Language ideology and language education. In J. Shaw, D. Lubelska, & M. Noullet (Eds.), Partnership and interaction: Proceedings of the Fourth International Conference on Language and Development (pp. 43–52). Asian Institute of Technology.

Tollefson, J. W. (2002). Language policies in education: Critical issues. Lawrence Erlbaum.

Tollefson, J. W. (2011). Ideology in second language education. In E. Hinkel (Ed.), Handbook of research in second language teaching and learning (Vol. 2, pp. 801–816). Routledge.

UNESCO. (1960, December 14). Convention against discrimination in education 1960. http://portal.unesco.org/en/ev.php-URL_ID=12949&URL_DO=DO_TOPIC&URL_SECTION=201.html

UNESCO. (2021, July). Portal of education plans and policies. https://planipolis.iiep.unesco.org/#:~:text=Planipolis%20is%20a%20portal%20of,policy%20makers%2C%20donors%20and%20partners.

United Nations Development Program, Pakistan. (2017). Pakistan national human development report 2017: Unleashing the potential of a young Pakistan. https://planipolis.iiep.unesco.org/en/2017/pakistan-national-human-development-report-2017-unleashing-potential-young-pakistan-6460

Van Dijk, T. A. (2009). Critical discourse studies: A sociocognitive approach. In R. Wodak & M. Meyer (Eds.), Methods of critical discourse analysis (pp. 62–86). Sage Publications Ltd.

Vessey, R. (2013). Language ideologies and discourses of national identity in Canadian newspapers: A cross-linguistic corpus-assisted discourse study [Unpublished doctoral dissertation]. Queen Mary, University of London.

Vessey, R. (2017). Corpus approaches to language ideology. Applied Linguistics, 38(3), 277–296.

Wiley, T. G., & Garcia, O. (2016). Language policy and planning in language education: Legacies, consequences, and possibilities. Modern Language Journal, 100(S1), 48–63. https://doi.org/10.1111/modl.12303

Wiley, T. G., & Lukes, M. (1996). English-only and standard English ideologies in the U.S. TESOL Quarterly, 30(3), 511–535.

Wiley, T. G. (2013). A brief history and assessment of language rights in the United States. In J. W. Tollefson (Ed.), Language policies in education: Critical issues (2nd ed., pp. 61–90). Routledge.

Wodak, R. (2001). The discourse historical approach. In R. Wodak & M. Meyer (Eds.), Methods of critical discourse analysis (pp. 63–94). Sage.

Zhao, S. (2011). Actors in language planning. In E. Hinkel (Ed.), Language rights in the United States. In J. W. Tollefson (Ed.), Language planning: Chinese script reform as a case study. Language Problems & Language Planning, 36(1), 1–24.