NUML Journal of Critical Inquiry
Vol. 23 (I) June 2025

e ISSN 2789-4665 p ISSN 2222-570
10.52015/numljci.v23iI.291

The Development of Nominal Synsets for the Saraiki Language: A Corpus-based Analysis

Madiya Asgher[1] & Musarrat Azher[2]

ABSTRACT

Article History:
Received:
August 9, 2024
Accepted:
June 23, 2025

Funding:
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest:
The authors have declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

This paper focuses on developing nominal synsets for the Saraiki language (SL), a lesser-studied language spoken in Pakistan. Nominal synsets are groups of nouns that share semantic characteristics and are crucial for natural language processing tasks such as information retrieval, machine translation, and text classification. The research aims to create Saraiki Nominal Synsets (SNS) using the Gurumukhi Punjabi WordNet. The study employs a hybrid approach, combining merge and expansion techniques for analysis and gathers data from PDF textbooks, online sources, and the Saraiki Wikimedia incubator. The collected data is limited to texts published between 2000 and 2019, and manually tagged using Antconc 3.4.4.0 wordlist due to the unavailability of a tagger for the Saraiki Language. The study builds a 2.2 million Saraiki word corpus and a list of 750 nouns, then categorizes and semantically organizes the Saraiki Nominal Synsets based on the list of Saraiki nouns. To identify and classify nouns in SL based on their semantic properties, a corpus-based approach is utilized, and nominal synsets are constructed using a combination of manual and automatic methods. Evaluating the quality of the synsets involves comparing them to existing lexical resources and conducting a semantic similarity analysis. The results demonstrate the effectiveness of the approach in capturing semantic relations among nouns in SL and producing synsets useful for various NLP applications. Overall, this study contributes to the development of linguistic resources for lesser-studied languages and provides valuable support for researchers and developers working on natural language processing tasks involving SL.

Keywords: *Saraiki language, Saraiki Nominal Synsets, Antconc, NLP, Corpus, WordNet*

[1] Lecturer in English at the University of Management and Technology, Sialkot Campus, Pakistan. She may be accessed at madya.asghar@skt.umt.edu.pk

[2] A Fulbright Alumna, currently working as a professor at Government Sadiq College for Women University, Bahawalpur, Pakistan. She may be accessed at musarratazher@gmail.com

Introduction

This study aims to develop nominal synsets for the Saraiki language (SL) using the Gurumukhi Punjabi WordNet. While the Shahmukhi Punjabi WordNet remains under development, the Gurumukhi WordNet provides accessible resources for research. Due to the linguistic and cultural proximity of Urdu, Punjabi, and Saraiki, this study leverages the Gurumukhi Punjabi WordNet for the creation of SNS (Khaled et. al., 2020). A 2.2 million-word Saraiki corpus was constructed from literary books, newspapers, and textbooks, meticulously tagged and analyzed. To ensure authenticity and usability, native speakers and Saraiki dictionaries were consulted. This research is a pivotal step toward an online lexical database for SL, aiming to facilitate language learners and advance Saraiki NLP applications.

Saraiki, spoken by over 20 million people, has received limited linguistic attention. This study develops a Saraiki corpus exceeding 2 million words, encompassing data from Multan, Bahawalpur, and Muzaffargarh. The research seeks to provide a robust foundation for linguistic advancements in SL by addressing this gap. Furthermore, this work aligns with global efforts to preserve minority languages and cultural heritage through digital documentation, as seen in projects like the Endangered Languages Project (ELP) and the World Atlas of Language Structures (WALS) (Austin & Sallabank, 2011; Dryer & Haspelmath, 2013). This study focuses on developing Saraiki Nominal Synsets (SNS) using a hybrid approach. The corpus is limited to written script and constrained to 740 high-frequency nouns due to time and funding limitations. Manual tagging was necessary owing to the absence of automated tools for SL.

Review of Related Literature

The Saraiki language, an Indo-Aryan tongue with significant regional and historical importance, remains understudied compared to other Pakistani languages like Urdu and Punjabi. Recent corpus-based research endeavors aim to bridge this gap, focusing on Saraiki's lexico-semantic relationships and resource development for computational linguistics. Awais et al. (2023) explored Saraiki verbs' lexical semantics, developing a corpus of three million words from diverse sources, including literary texts, newspapers, and online archives. The study utilized Fellbaum's (1993) semantic categorization to create verb synsets, including glosses, example sentences, and semantic relations such as troponymy and entailment. This work advances the creation of a WordNet for Saraiki, providing foundational resources for machine translation and semantic analysis. Similarly, Nazeer et al. (2024) focused on the lexico-semantic properties of Saraiki nouns. Using a similar corpus size and a combination of manual and semi-automated techniques, the research identified 173 synsets for 39 high-frequency nouns. The study highlighted hierarchical relationships like hyponymy, hypernymy, and meronymy, contributing to Saraiki's lexical database development.

Both studies adopted a hybrid approach, leveraging existing lexical frameworks and consulting native speakers for cultural and contextual accuracy. For instance, Nazeer et al. (2024) implemented the expansion approach for borrowing synsets from related languages like Punjabi while maintaining Saraiki's linguistic independence. Similarly, Awais et al. (2023) combined corpus analysis with dictionary consultations to validate verb senses. These methodological innovations underscore the challenges of limited linguistic resources for regional languages. They also highlight the potential applications of Saraiki WordNet in natural language processing (NLP), including semantic search, machine learning algorithms, and language preservation. This aligns with global trends in computational linguistics, contributing to multilingual and cross-lingual resource integration. Additionally, both studies contextualize their work within Saraiki's rich linguistic heritage, emphasizing its unique blend of Indo-Aryan and regional linguistic traits. These efforts are seen as pivotal in acknowledging Saraiki's status as a distinct language while enhancing its digital and academic presence.

Another study conducted by Gull et al. (2021) focuses on the development of a Saraiki WordNet by mapping Urdu word senses to Saraiki word senses. Saraiki, a regional language spoken in Pakistan, has similarities with Punjabi and Sindhi. The researchers used the existing Urdu WordNet as a basis and mapped Urdu word senses to Saraiki word senses using dictionaries, literary sources, and corpus-based approaches. The development of a Saraiki WordNet is significant for natural language processing applications and can aid in the creation of bilingual dictionaries in the future. The researchers employed the expansion approach, a widely used method in WordNet development, to build the Saraiki WordNet. They utilized various dictionaries, both monolingual and bilingual, to map the Urdu and Saraiki word senses. The researchers also compiled a diverse corpus from various sources, including newspapers, stories, essays, and poetry, to provide necessary examples and elaborate on the concepts. The use of corpus technology enabled the researchers to create a resource that adequately reflected the distribution of Saraiki words and their lexical-semantic variants in real contextual environments. The corpus was analyzed using the AntConc software, which provided information on the frequency of words and helped in finding the correct and reliable senses of Saraiki words.

Overall, these studies contribute significantly to the field of natural language processing and language resource development. They provide foundational frameworks for the creation of bilingual dictionaries, semantic analysis tools, and applications in language preservation. The advancement of a Saraiki WordNet using corpus-based approaches is a pivotal step toward enhancing the digital and linguistic representation of Saraiki, ensuring its relevance and integration into modern computational systems.

Methodology

The process of developing Saraiki nominal synsets (SNS) involves three major steps. Firstly, a corpus of 2.2 million words is created, followed by manual tagging of the corpus using a POS tagging pattern. Secondly, the tagged data is used for creating Saraiki nominal synsets. The production of SWN involves the use of

merging and expansion techniques. In the merge approach, the senses of words are recorded first, followed by recording the words in which the senses are used. In the expanded model, the senses of the source language are translated into the target language.

Development of Corpus

Different sources were utilized for the creation of the corpus. These sources included newspapers, fiction, essays, and columns, and the corpus developed through these sources comprises 2.2 million words, now available at the University of Sargodha library. For development, the 2.2 million-word corpus Sample Text (ST), passed through certain stages:

1) Data collected from online available sources and books published in Siraiki, but available in hard form

2) Hard-form books scanned and converted into PDF form

3) PDF form changed into the form of images manually

4) Image files uploaded into Google Docs that were converted into text

5) Online available text and converted text combined according to their genre

After these steps, the data was processed in Antconc 3.4.4.0 to create a word list. During this process of Saraiki nominal synsets development, the Gurumukhi Punjabi WordNet is used.

Saraiki's word list is translated into PL, and its equivalents are found manually. After finding equivalents, the concepts of words are extracted for the best results. Then the untagged corpus is tagged with the help of Antconc 3.4.4.0 wordlist manually, as no tagger is available for the Siraiki Language. Some dictionaries and Saraiki speakers were also consulted for correct POS tagging. These dictionaries include Punjabi and Siraiki dictionaries.

Table 1

Dictionaries used in the study and their publishers

Sr. No.	Source	Name of the Dictionary	Publishers of Dictionaries
1		*Dictionary of the Jhatki or Western Punjabi Languagealso available online at https://archive.org/details/204912920SaraikiDictionary/page/n5/mode/2up*	*Religious Books and Tract Society Lahore*

2		*Glossary of the Multani Language by E.O' Brian also available online at https://skr.m.wiktionary.org/*	*Saraiki Adabi Board, Multan*
3	*Books*	*Siraiki English Dictionary by Andrew Jukes also available online at https://skr.m.wiktionary.org/*	*Siraiki Adabi Board, Multan*
4		*Pehli Wadi Siraiki Lughat by Saad Ullah Khatran also available online at https://skr.m.wiktionary.org/*	*Siraiki Area Study Centre, BZU, Multan*
5	*Online available at https://www.shabdkosh.com/dictionary/english-punjabi/*	*Shabakdosh a English-Punjabi Dictionary*	
6	*Online available at http://dic.learnpunjabi.org/default.aspx*	*Akhar (2016) a Punjabi-English Dictionary*	*Punjabi University, Patiala, India*
	Online available at https://skr.m.wiktionary.org/	*Ijunoon a English Siraiki Dictionary*	

Data Conversion into Machine-Readable Form

All data was collected from various sources and in various forms. All the data needed to be converted into machine-readable form for further applications. To achieve this aim, various tools and methods were applied by the researcher, which took tremendous effort and time. The process of these conversions is

described in Figure 3.1.

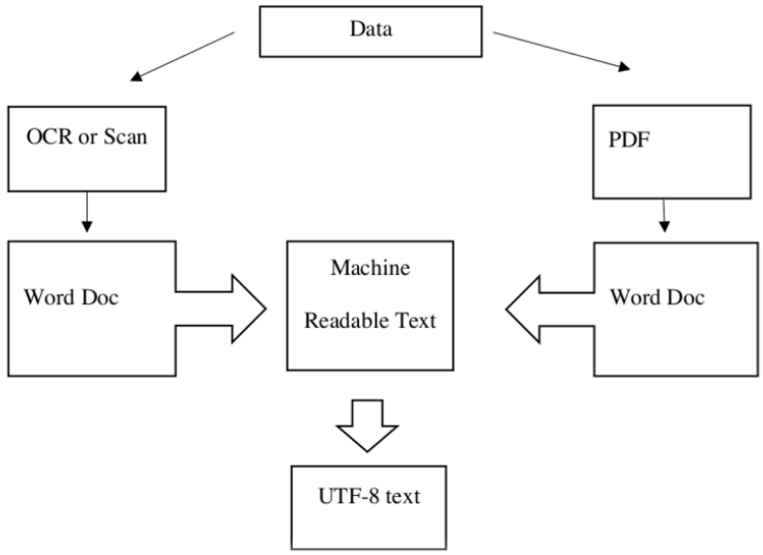

Figure 3.1: Process of converting Data into machine-readable form

At first, all books were scanned using the HP DeskJet All-in-One Printer and then converted into PDF form using the iLovePDF site. While some of the data was not readable for the machine then OCR was done using Google Lens. It changed the data into image files. After making image files, the data was processed into Google Docs, which read the image and converted it into text form. After this process, data was available for the machine-readable form, which was later combined with online data (directly). Then all the data was saved into Word 2010 for the researcher's convenience. After going through all these stages, the researcher saved all the data in UTF-8 format using Notepad++ which was processed in Antconc 3.4.4.0 and tagged to develop SNS.

- **Coding Corpus**

All data were collected from various parts, and giving codes to these parts was necessary to avoid ambiguity. The corpus of Newspapers was assigned the code of NP. The fiction corpus was assigned a unique code FT, while the essay corpus was coded with ES. The translated corpus was given with TR. These unique codes were mentioned properly during corpus compilation, which also assisted in the identification of the source of the corpus.

- **Process of POS Tagging Saraiki Corpus**

POS tagging is also known as grammatical tagging, used to tag data for further applications based on its context and definition. In this study, the process of tagging is also used, which includes certain steps. First, the data is converted

from Word Doc to Notepad++ and coded properly. Second, after encoding, the data is processed into AntConc 3.4.4.0, which provides a wordlist of the Saraiki corpus, which tells the frequency of a word in the corpus (2.2 million words Saraiki corpus) as in Figure 3.2.

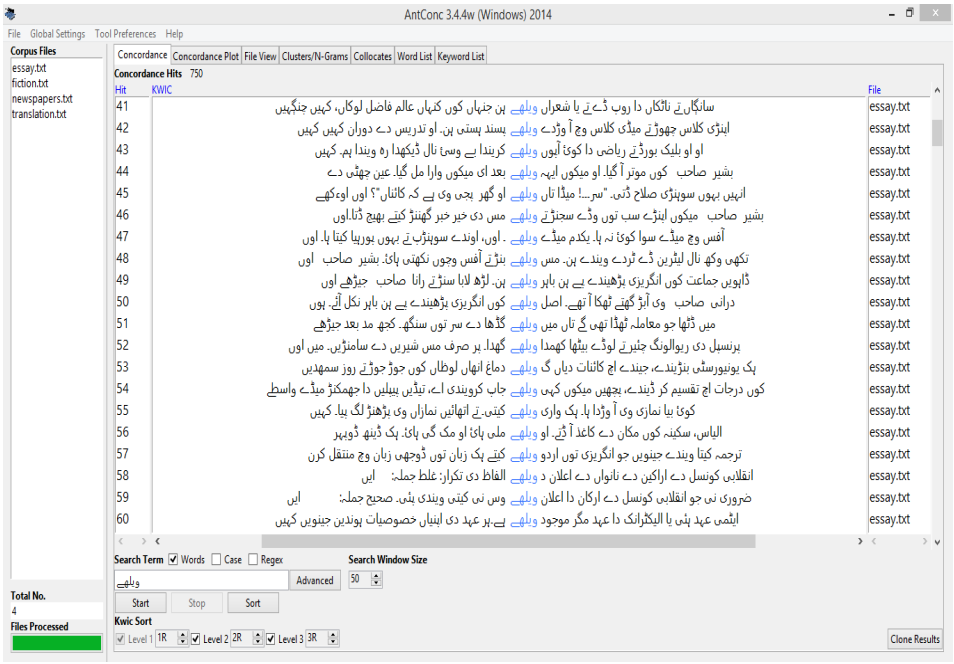

Figure 2: Most Frequent Nouns' Concordance in AntConc 3.4.4.0

Third, the words from the wordlist are copied one by one and found in a Word document for tagging manually as in Figure 3. Fourth, the Lexical technique is kept in view while tagging the data.

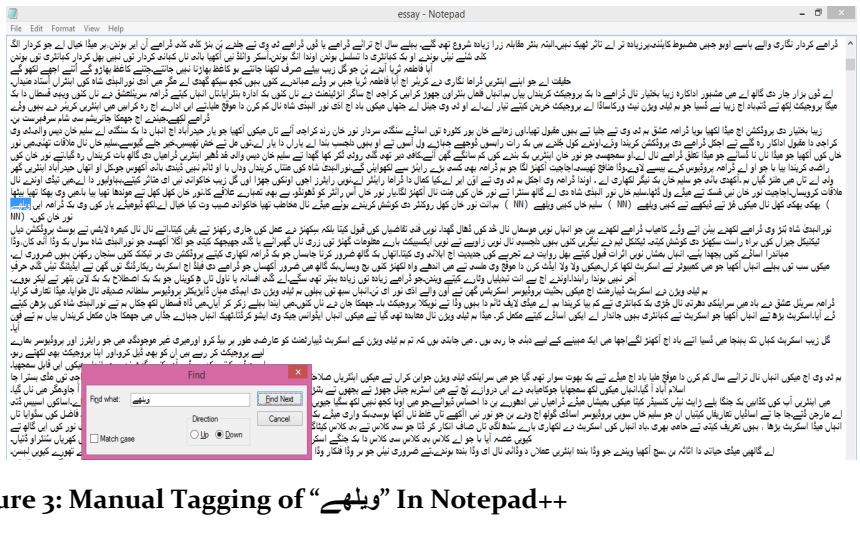

Figure 3: Manual Tagging of "ویلھے" In Notepad++

Manual tagging is done because Tagger for Saraiki Language is not available. This manual tagging provides accurate results because the context of every word is checked, and then the word is tagged. This also helped in extracting examples for *Saraiki's* noun synsets.

The universal POS Tagset defined by Bird et al. (2009) is used specifically for nouns because the focal point of this study is to develop Saraiki Noun Synsets.

Semantically Classification and Finalization of Saraiki Nouns

The nouns that were highly frequent in the wordlist were finalized and classified semantically. It comprises a list of 750 noun words accessed from fiction, essays, newspapers, and translations. The details of these nouns have been given in the Appendix based on their classification.

Development of Nominal Synsets of Saraiki Language

The purpose of this study was to develop nominal Synsets of Saraiki Language. To develop Nominal Synsets, the following components were devised in the form of entry number, nouns, senses' number, synsets of noun words, gloss of synset, and example sentences (extracted from the developed Saraiki corpus). *Synsets* are the sense developed from a word while gloss is what a word is.

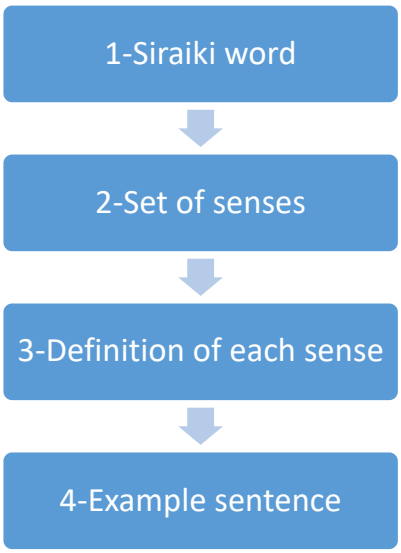

Figure 4: Basic Steps Involved in Synset Creation

Results and Discussions

The finalized noun words, based on comparisons with the developed corpus, are detailed in the following tables. As a result, we have compiled a list of

3,000 Saraiki synsets derived from 750 noun words from the Saraiki language, as illustrated in Table 3.

Table 2

A list of Saraiki Nouns along with their Semantic types

Sr. No	Semantic Type	Roman Urdu	Saraiki
1	Fasal	Kanark, Sitta, Wataun, Alloo, Gungloo, Sawani, Kapah, Jawainr, Bajrah, Kamad, Rayi, Jantar, Gunwaar, Turi, Bhun, Rerh, Manjhi/Sariyan, Chanrhy, Mungherian	کنڑک، سٹا، وتاؤں، آلو، گونگلوں، سوانی، کپاہ، جوائنڑ، باجرہ، کماند، رائی، جنتر، گنوار، ٹُری، بُھوں، رڑھ، مُنجی/ سریاں، چنھڑے، مونگیریاں
2	Khurak	Phaal, Sitta, Makhanrh, Routi, Basta, Gosat, Sabzi, Daal, Keema,Gheu, Ghurh, Chelrha, Atta, Bhaji, Bhorh	پھل، سٹا، مکھنڑ، روٹی، بستہ، گوست، سبزی، دال، قیمہ، گھیو، گُڑ، چیلڑا، اٹا، بھاجی، بوڑ
3	Phal	Amb, Peelun, Toot, Baer, Khajoor, Tar, Naakh, Saeb, Amrood, Akhroot, Jammun, Mateera, Khoprha, BidaamBoor, Ambian, Anghoori, Hadwana, Kharmarian, Ghiri, Darakh, Dokky	امب، پیلھوں، توت، بیر، پنڈہ، تر، ناکھ، سیب، امرود، اخروٹ، جموں، متیرا، کھوپڑا، بدام، بُور، امبیاں، انگوری، ہدوانہ، خرمایڑیاں، گری، دراخ، ڈوکے
4	Phul Ty Ondy Hissy	Phul, Patti, Kandy/ Kanjy, Jarh, Akh, Moundh, Ghulab, Chamaeli	پُھل، پتی، کنڈے/ کنجے، اکھ، مونڈھ، گُلاب، چمیلی
5	Ghaa Boty	Bhakaat, Baala, Bota, Ghaa, Phoog, Kakh, Kunwaar, Aak, Bael, Jantar, Turhi	بھکاٹ، بالا، بوٹا، گھا، پھوگ، ککھ، کُنوار، اَک، بیل، جنتر، توڑی
6	Khaeti Barhi	Kanark, Zameen, Sawani, Aallo, Thal, Mitti, Killa, Khali, Khaal, Khaad, Hal, Zegal, Khoo, Gara	کنڑک، زمین، سوانی، آلو، تھل، مٹی، کلہ، کھالی، کھال، کھاد، ہل، زیگلی، کھوہ، گارا
7	Zaar	Saam, Oog, Matheera, Aar, Takrhi, Sotti, Sotta, Datari, Chamoota, Talwar, Halya, Chorhi, Hal, Phalla, Dasta, Ranba, Kurh, Phana, Sharat, Danna, Churri, Chaqu, Bandook	سام، اوگ، مُٹھیڑا، آر، تکڑی، سوٹی، سوٹا، ڈاتری، چموٹا، تلوار، ہلیہ، چوڑی، ہل، پَھلّا، دستہ، رَنبا، کُڑ، پھانہ، شارت، ڈنہ، چُھری، چاقو، بندوق

8	Jism Dy Zaa	Nooh, Choti, Khuthe, Ghetty, Book, Matha, Cham, Cheechi, Mondha, Dhaidh, Talli, Maas, Kandh, Gheechi, Lahoo, Zaban, Damagh, Poorh, Nak, Mukh, Dhandh, Kuch, Ungal, Dheela, Ghoda, Baanh, Aakh, Mounh, Lat, Nasan, Hoth, Darhi, Moch, Irak, Waal, Gal, Bothi, Bheja, Kan, Wakhi, Chelah, Zoban, Nain	َونہ، چُوٹی، کُھتھی، گِٹّے، بُک، متھا، چم، چیچی، مونڈھا، ڈھیڈ، تلی، ماس، کنڈھ، گیچی، لہو، زبان، دماغ، پوڑ، نک، مُکھ، ڈنڈ، کُچھ، اُنگل، ڈیلھا، گوڈا، بانہہ، اکھ، مُنہہ، لت، ناساں، ہوٹھ، ڈاڑھی، مُچھ، ارک، وال، گل، بُوتھی، بھیجا، کن، وکھی، چیلھ، زوبان، نین
9	Sabziyan	Saagh, Allo, Wataun, Ghoonglu, Kachalu, Paalak, Ghobhi, Maethi, Thoom, Wasal, Bey, Gunwaar	ساگ، آلو، وتاؤں، گونگلوں، کچالو، پالک، گوبھی، میتھی، تھوم، وسل، بے، گنوار
10	Kapry	Suthan, Leer, Patti, Ghaghri, Paag, Khaisa, Bukaal, Buchka, Ghandhre, Choola, Tamboo, Patka, Cholli, Chunni, Buchanrh, Jaeeb , Pandh, Rassi, Romaal, Neefa, Kameez, Banain, Aghat, Jorha, Sata, Ghut, Pallu, Sata, Jorha, Sanwherha	سُتھن، لیر، پٹی، گھگھری، پگ، کھیسہ، بُکل، بُچکا، گنڈھری، چولا، تمبو، پٹکا، چولی، چُنی، بوچھنڑ، جیب، پنڈھ، رسی، رومال، نیفہ، قمیض، بنین، آگھٹ، جوڑا، ساٹا، گھت، پلو، جوڑا، سنویڑہ
11	Zewar	Mundari, Wangan, Tikka, Mala, Haar, Waliya, Jhumar, Chanjar, Koka	مُندری، ونگاں، ٹکا، مالا، ہار، والیاں، جھمر، چانجر، کوکا
12	Mosam	Jharhi, Seet, Meenh, Saun, Andhari, Baddu, Chandra, Sayala, Hunala, Patar Kaer, Waan Phaphurh	جھڑی، سیت، مینہہ، ساؤن، اندھاری، بادو، چندرا، سیالا، ہُنالا، پتر کیر، وٹ پھپھوڑ
13	Rang	Kaala, Chitta, Surkhi, Sawa, Peela, Neela, Bagha, Laal, Khaki, Kesar, Sanwala, Sunehra, Ratti, Badami, Ghandmi, Nuswari, Narangi, Ratta,Anghori,Saleiti, Jamhun, Kaleiji	کالا، چٹا، سُرخی، ساوا، پیلا، نیلا، بگھا، لال، خاکی، کیسری، سانولا، سُنہرا، رتی، بدامی، گندمی، نسواری، نارنگی، رتّا، انگوری، سلیٹی، جمھوں، کلیجی
14	Bimari	Tonda, Rat, Borhy, Kanna, Khangh, Bakhar/Kosa, Tabheer-E-Maeda, Thand, Korh, Langhra, Botha, Ghanja, Kanrha, Thakerha, Phaat, Sora, Mally, Matan	ُنڈا، رت، بوڑے، کانا، کھنگ، بخار/ کوسا، تبخیر معدہ، ٹھنڈ، کوڑھ، لنگڑا، بوتھا، گنجا، کانڑھا، تھکیڑا، پھٹ، سورا، مللے، ماتاں

15	Paandy	Changheair, Thaali, Ghadwi, Thal, Karhchi, Parhopi, Handi, Prhoopa, Chabbi, Doyi, Katori, Kunni, Degarhi, Degharha, Doye,Daigar, Kanjheer, Koop, Katori, Payali	چنگیر، تھالی، گڈوی، تھال، کڑچھی، پڑھوپی، پڑوپا، چھبی، ڈوئی، کٹوری، کُنی، دیگڑی، دیگڑا، ڈوئی، دیگر، کنجیر، کوپ، کٹوری، پیالی
16	Bayen Layi Cheezan	Kathrha, Manjhi, Parhchi, Peerhi, Peerha, Moorha, Kursi, Peengha	کٹھڑا، منجھی، پڑچھی، پیڑھی، پیڑھا، موڑھا، کُرسی، پینگھا
17	Lakri Tun Bani Cheezan	Taakrhi, Lakarh, Kaath, Pawa, Peengh, Peerhi, Bal, Berhi, Tahat, Teer, Shateerh, Chokat, Dar, Darri	تاکڑی، لکڑ، کاٹھ، پاوا، پینگھ، پیڑھی، بل، بیڑھی، تخت، تیر، شہتیر، چوکھٹ، در، درّی
18	Khaed	Aason Panjun, Luk Chuparh, Banrhi Qitaar, Taash, Kanga Maari, Douda, Kushti, Gheeti Danna, Telkanrh, Dhi Urhi Dhi, Kukry Chaek Jummaraat Aayi Ha, Ghaind Balla, Barf Paani, Laatu, Chibian, Stappu, Luddi	اسوں پنجوں، لُک چھپڑ، بنڑی قطار، تاش، کانگا ماری، دُودا، کُشتی، گیٹی ڈنڈاں، تلکنڑ، دھی اُو ڑی دھی، کُوکڑے چھپک جمعرات آئ ہے، گیند بلا، برف پانی، لاٹو، چبیاں، سٹاپو، لوڈی
19	Khaun Layi Cheezan	Ghorh, Khalwa, Kheer, Methaye, Loon, Khandh, Chogh, Thoom, Kaaj, Tikka, Pokorha, Rotti, Zahar, Ghandherian, Atta, Tukar, Salad, Bhorry, Mirchan, Duda	گُڑ، حلوہ، کھیر، مٹھائی، لون، کھنڈ، چوگ، تھوم، کاج، تکہ، پکوڑا، روٹی، زہر، گنڈھیریاں، آٹا، ٹُکر، سلاد، بھورے، مرچاں، ڈوڈا
20	Look	Kaaj, Waeri, Rani, Tarimat, Baba, Pakhi, Saein, Jhanjh, Kath, Waseeb, Porhiya, Look, Saein, Tabar, Banda, Pleas, Awam, Fakeer, Sodagar, Kotarein, Maela,Chandra, Lucy, Naist, Meesna, Chatar, Chabal, Bebt	کاج، وَیری، راٹی، تریمت،بابا، پکھی، سَیڑ، جنج، کٹھ، وسیب، پورھیا، لوک، سئیں، ٹبر، ، بندہ، پلیس، عوام، فقیر، سوداگر، کوٹاریں، میلا، چندرا، لوسی، نیست، میسنا، چتر، چبل، بیبت
21	Marat Ty Ondy Hissy	Alhanrha, Watta, Maseet, Ghar, Salh, Kotha, Werha, Porhi, Jhok, Boha, Rasoye, Bagh, Parhcha, Mahal, Darbaar, Madrissah, Askool, Nukar, Kachari, Chabara, Chaat, Aent, Baaly, Kamra, Batti, Kandh, Kundi, Jhumar, Bharti, Sil, Rorhy, Chapra, Kothi, Makaan, Khuddi, Bhanan, Bandur, Pakha, Palli	آلہنڑا، وٹا، مسیت، سالھ، کوٹھا، ویڑھا، پوڑی، جھوک، بوہا، رسوئی، باغ، پاڑچھا، محل، دربار، مدرسہ، اسکول، نُکڑ، کچیری، چبارہ، چھت، بالے، کمرہ، بتی، کندھ، کُنڈی، جُھمر، بھرتی، سل، روڑھے، چھپرہ، کوٹھی، مکاڑ، کُھڈی، بھنان، بندور، پکھا، پلی

22	Waela	Raat, Dainh, Pooh, Bangh, Fajar, Saman, Karhi, Dhup, Chaan, Sawael, Waela, Dupahar	رات، ڈینہہ، پوہ، بانگ، فجر، سماں،کڑی، دھپ، چھاں، سویل، ویلا، ڈوپاہر
23	Jhah	Aroorhi, Hatti, Barz, Ranarh, Choki, Bhuk. Goth, Khoo, Chulah, Tanoor, Khud, Bazar, Cheerya-Ghar, Wasti, Shaher, Ghalli, Mohallah, Chotti, Jungle, Darya, Khal, Karbala, Wanrha	اُروڑی، ہٹّی، برز، رنڑ، چوکی، بھک، گوٹھ، کھوہ، چُلھ، تنور، کُھڈ، بزار، چڑیا گھر، وستی، شہر، گلّی، محلہ، چوٹی، جنگل، دریا، کھل، کربلا، وانڑہ، دیرہ
24	Rishty	Junwaye, Bhen, Putra, Budha, Pahaj, Piyo, Zaal, Budhi, Balrhi, Bhara, Baal, Mitar, Chohar, Miyan, Putar, Chokri, Babu, Amaan, Saas, Sorha, Tabar, Dhadhi, Nani, Kasoli, Malook, Rishta, Maa, Mama, Mami, Chachi, Chacha, Baeli, Saenghi, Phoopharh, Malear, Masaar, Sabala, Kanwar, Zanani, Juwan	جَنّوائی، بھیڑ، پوترا، بُڈھا، پہاج، پیو، ذال، بُڈھی، بالڑی، بھرا، بال، متر، چھوہر، میاں، پُتر، چھوکری، بابو، اماں، ساس، سورھا، ٹبر، ڈاڈی، نانی، کسولی، ملوک، رشتہ، مأ، ماما، مامی، چاچی، چاچا، بیلی، سینگی، پھوپھڑ، ملیر، مسیر، سبالا، کُنوار، زنانی، جُوان
25	Pakhi	Tateerh, Terkala, Lali, Bhagla, Badak, Kaan, Ghij, Chirhi, Chirha, Talur, Chanjhur , Chapak, Koyal, Ghorakh, Chandur, Kanwrihi, Batera, Ratha, Tooba, Jal-Kukarh, Mamola, Mamhala, Haal, Tatuhan, Tetar, Ghera, Toota, Dodar-Kaan, Bagh, Tillar, Baaz, Marghabi, Krainh, Waah, Chakori	ٹٹیہر، ترکلا، لالی، بگلا، بدک، کاں، گیجھ، چڑی، چڑا، تلور، چنجھور، چپک، کوئل، گورکھ، چندور، کانوڑی، بٹیرا، رٹھا، ٹوبا، جل-کُکڑ، مامولا، ممہالا، ہل، ٹٹوہاں، تتر، گھیرا، طوطا، ڈوڈر-کاں، باگھ، تلر، باز، مرغابی، کرینہہ، واہ، چکوری
26	Waan	Neem, Lasoorha, Harnoli, Kareer, Shareenh, Sohanjrhan, Saar, Jind, Jammun, Peelun, Taali, Toot, Pepal, Berhi, Bouharh, Kaanh/Tolha, Kikar, Kath, Layi, Phoog, Rukh, Waan, Khajji, Safaida, Kachnar, Jhaal, Jhatar, Baans	نم، لَسُوڑا، ہرنولی، کَریر، شَرینہہ، سُہانجنڑاں، سَر، جنڈ ، جَمّوں، پیلھُوں، ٹالھی، توت، پِیپَل، بیری ، بُوہڑ، کانہہ/ٹُولہا، ککر، کاٹھ، لئی، پھوگ، رُکھ، ون، کھجی، سفیدا، کچنار، جھال، جھتر، بانس
27	Zanwar	Uth, Arghalli, Khotti, Shenh, Nang, Danghrh, Khota, Wachi, Kukrhi, Khattun, Bhaedh, Poongh, Dachi, Kirhi, Manjh, Saeharh, Ghalarh, Lyla, Nyola, Dhedhar, Cham-	اُٹھ، آرگالی، کھوتی، شِینہہ، نانگ، ڈنگر، کھوتا، وچھی، کُکڑی، کھٹوں، بھیڑ، پونگ، ڈاچی، کرڑی، منجھ، سیہڑ، گاہلڑ، لیلا، نیولا، ڈیڈر، چم-چڑھ، بلی، چوہا، شیر، گدڑ،

		Chicharh, Billi, Choha, Shaer, Ghedarh, Bandari, Ghorha, Bakri, Cheeta, Shairni, Kutta, Ghular, Ghadan, Lumarh, Rech, Dhand, Mainh, Ghau, Bloongrha, Ghaba, Jhoota, Phandar, Jaaha, Machi	باندری، گھوڑا، بکری، چیتا، شیرنی، کُتا، گُلہر، گڈاں، لومڑ، رچھ، ڈھانڈ، مینہہ، گاؤ، بلونگڑا، گابا، جھوٹا، پھنڈر، جاہا، مچھی
28	Ehsaas	Roosna, Saek, Chaa, Man, Khaab, Wachorha, Sawad, Rahmat, Mounjh, Rees, Sanrap, Bhuk, Hanju, Muhabbat, Dosti, Makholl, Mahangh, Ghilla, Nafrat, Hussan, Payaar, Khabas, Hawas, Ruthi, Bhoog, Khuwari, Kanbarhi, Wasal, Dukh, Tap, Kawarh, Neer	رُسنا، سیک، چاہ، من، کھاب، وچھوڑا، سواد، رحمت، مونجھ، ریس، سنڑھپ، بُکھ، ہنجو، محبت، دوستی، مخول، مہانگ، گلہ، نفرت، حُسن، پیار، خبس، ہوس، رُٹھی، بھوگ، خواری، کنمبڑی، وصل، ڈکھ، تپ، کاوڑ، نیر
29	Dhatan	Sona, Chandi, Loya, Kola, Heera, Tanba, Sang-E-Marmar,	سونا، چاندی، لویا، کولا، ہیرا، تانبا، سنگ مرمر
30	Chezan	Purhi, Jhandra, Lafafa, Watta, Basta, Waag, Lota, Ghandh, Sheesha, Moundh, Tohfa, Dabba, Kitab, Kawaz, Kapi, Sawarhi, Radhi, Bhan-Bhosrha, Dhool, Tallian, Taar, Jutti, Subbi, Buhaari, Mandi, Chata	پوڑی، جندرا، لفافہ، وٹہ، بستہ، واگ، لوٹا، گنڈھ، شیشہ، گڈاں، مونڈھ، تحفہ، ڈبہ، کتاب، کاوز، کاپی، سواری، ردھی، بھن-بھوسڑھا، ڈھول، ٹلیاں، تار، جُتی، سُبی، بوہاری، مینڈی، چھاتا
31	Pakhi Dy Zaa	Chunj, Poochal, Khanmb, Chamby, Gheechi, Narghat, Sirri,	چُنج، پوچھل، کھنب، چمبے، گیچی، نرگھٹ، سری
32	Keerhy	Makhi, Pissun, Tooka, Sondha, Joon, Machar, Titli, Jaaz, Wathuhan, Makrha, Makhi,Kaweli, Seewi, Bhondh	ماکھی، پسوں، ٹوکا، سونڈھا، جوں، مچھر، تتلی، جاز، وٹھوہاں، مکڑھا، مکھی، کویلی، سیوی، بھونڈ
33	Bank	Maal, Raqam, Karza, Udhar, Jaib, Paisa, Rishwat/Dallali, Kisat, Sood, Manafah, Sarmaya/Dhan, Khata, Bill, Hatti, Khatti, Chatti, Bha	مال، رقم، قرضہ، اُدھار، جیب، پیسہ، رشوت، قسط، سود، منافع، سرمایہ، کھاتہ، بل، ہٹی، کھٹی، چٹی، بھا
34	Ghaer Insani Cheezan	Dain, Parri, Balan, Jin, Farishty, Dewta, Rooh, Churail, Deu	ڈین، پری، بلاں، جن، فرشتے، دیوتا، روح، چڑیل، دیو
35	Kudrati Cheezan	Hawa, Paani, Ag, Chan /Chandar, Taary, Dhoop, Chanan, Andhara, Bhaa, Dharti, Mitti, Phal, Sabzian	ہوا، پانی، اگ، چن، تارے، دھپ، چانن، اندھارا، بھاہ، دھرتی، مٹی، پھل، سبزیاں
36	Aoun Jaan Layi Cheezan	Ghaddi, Weghan, Sawari, Gadhan,Pandh,Sarak, Tracktor, Larri, Tanga, Jaaz, Saikal, Real-	گڈی، ویگن، سواری، گڈھاں، پندھ، سڑک، ٹریکٹر، لاری، ٹنگہ، جاز، سیکل، ریل گڈی،

		Ghaddi, Dhala, Raksha, Chakrha, Rarhi, Tralli, Tracktor	ڈالا، رکشہ، چھیکڑا، ریڑھی، ٹرالی، ٹریکٹر
37	Paishy	Arhti, Nokarhati, Dayi, Mistari, Marasi,Mashara/ Bhand, Dakhdar, Mouzeera, Darkhaan	آڑھتی، نوکڑاتی،دائی، مِستری، میراثی، مسخرہ/ بھانڈ، ڈاکھدار، موزیرا، درکھاٹ

These Saraiki nouns have been considered for analysis. Moreover, these semantic categories of Saraiki noun words have also been considered for data analysis. Some of these noun words, along with their Saraiki Synsets, have been discussed below.

1. **Semantic Type: فصل (Fasal)**

Table 3

Saraiki Noun کنڑک (Kanark)'s Synsets

Semantic Type	EN	Words	Sense No.	Grammatical Type	Senses	Glosses	Examples
	1	کنڑک	Sense 1	Noun	پکی کنڑک	پکی ہووی کنڑک	"سال کھن کنڑک کپ تے نال توں کنڑک کپڑ والیاں اِکڑ دُکڑ مشیناں تاں آیاں ودیاں ہن۔"
			Sense 2	Noun	کنڑک دی فصل	کنڑک ہک پیداوار	"زرعی پیکج دے مثبت اثرات کنڑک اتے مکئی دی مثالی پیداوار دی شکل اچ ظاہر تھئے ہن ۔"
			Sense 3	Noun	کنڑک دی بنی چیزاں	کنڑک دی بنی روٹی تے ویسن وغیرہ	"حکومت نے 5 لکھ ٹن کنڑک اتے کنڑک دیاں مصنوعات برآمدکرنڑ دی منظوری ڈتی اے۔"

			Sense 4	Noun	کنڑک ویلا	سخت گرمی	"زمینداریں کنے لئو سانگے کنڑک ویلے او ہتھ منہ دھویندا پروتھی بھاچی نال ہک ٹُو پٹھے سدھے گرانہہ مریندا اللہ دی آس تے گئی۔"

The mentioned word in Table 3, کنڑک categorized under the semantic Type of fasal. This word shows polysemic relation as all senses of *"kanark"* sound the same but have four different related meanings: *"pakki kanark, kanark di fasal, kanark di bani cheezan, ty sahat garmi".* Three senses (*kanark di fasal, pakki kanark, ty kanark di bani cheezan*) are directly acquired from the Punjabi WordNet. These are also part of *Shabdkosh* as these are found in *Shabdkosh,* and *Akhar* (2016). But the fourth sense is generated from the developed Saraiki corpus manually because it is not present in Punjabi dictionaries, but in Saraiki. It is extracted by using the merge approach that is also used for the construction of gloss. Furthermore, all the examples are taken from the Saraiki language corpus.

2. Semantic Type: خوراک (Khurak)

Table 4

Saraiki Noun کھیر (Kheer)'s Synsets

Semantic Type	EN	Words	Sense No.	Grammatical Type	Senses	Glosses	Examples
	2	کھیر	Sense 1	Noun	کھیر	دودھ دی بنی کھیر	"کھیر تے ھیر ہئی اوندا مزہ سب توں وکھری ہئی۔"
			Sense 2	Noun	مٹھاس	کھیر دی طرحاں مٹھا	"نال ماء دے کھل الیناں میں اوندے لباں توں کھیر آندے ہن۔"
			Sense 3	Noun	خالصدو دھ	ملاوٹ توں پاک گاڑھی	"جیویں پاݨیاں ہووے کھیر جدا جیویں ہال جدا ما اپݨی توں"

			Sense 4	Noun	دودھ	دودھ	"ہک ڈاچی روزانہ ۱۰ کِنوں لا تے کلو کھیر ڈیندی ہِ ۔"

In Table 4, the root word "kheer" has been taken from the same Type: *khurak*. It shares four various but related senses and shows polysemic relations. Three of these senses *'dodh de kheer, dodh, khalas dodh'* in Punjabi WordNet and dictionaries: Akhar (2016), but *'mithas'* is a pure Saraiki sense used in Saraiki literature that is extracted by applying the merge approach.

3. Semantic Type: گھا تے بوٹی (Ghaa ty Booti)

Table 5

Saraiki Noun بوٹا (Bota)'s Synsets

Semantic Type	EN	Words	Sense No.	Grammatical Type	Senses	Glosses	Examples
	27	بوٹا	Sense 1	Noun	بوٹا	پھل دا بوٹا	"مقروض نے اپنے گھر دے نیڑے گلاب دا ہک بوٹا لاو ناں ہئی۔"
			Sense 2	Noun	اولاد	نشانی	"خاتون اول بیگم محمودہ ممنون نوں اللہ نے اولاد دے تے اس دا بوٹا لایا۔"
			Sense 3	Noun	پُھل بوٹا	کپڑے تے بنیا تصویری پھل بوٹا	"لال کُرتی تے پیلے رنگ دے بوٹے سوہنے تھیندے ہن۔"
			Sense 4	Noun	ایٹ یاں نیہ	کسے کم دی نیہ گھننا	"اساڈی حکومت جیر ھا بوٹا 1997اچ لاتا ہئی او اج پروان چڑھ تے ہک پھل آلے بوٹے

							دی حیثیت اختیار کر گیا ہے۔"
			Sense 5	Noun	پیار دا بوٹا	پیار	"سانول پیار دا بوٹااوکھا پلدے ہن۔"

In Table 5, the word بوٹا belongs to the semantic Type *ghaa ty botti.* This specific word has been used in five different senses that make it polysemous. All these senses are taken from the Punjabi WordNet under expansion approach. These are used similarly in the Saraiki Language and culture. One sense of *bota* as a *'phal ala bota'* is also described in online Punjabi dictionaries: Akhar (2016) and Shabdkosh. Moreover, the gloss of the Saraiki synset is constructed through the merge approach.

4. Semantic Type: پھل (Phal)

Table 6

Saraiki Noun امب (Amb)'s Synsets

Semantic Type	EN	Words	Sense No.	Grammatical Type	Senses	Glosses	Examples
	14	امب	Sense 1	Noun	امب	کھاؤن آلا پھل، امب	"لنگڑا امب کتنے دا تھیندا ہوسی۔"
			Sense 2	Noun	امب	امب دا بوٹا	"میں کھٹڑے امب دی چھاں تھلے ہاں۔"
			Sense 3	Noun	امب رس	امب دا جوس	"امب رس بلیں دا پسندیدہ مشروب ہئے۔"
			Sense 4	Noun	بور	امب دے پھل جو بعد اچ امب بنیدا	"اندھیاریاں دی وجہ توں بور گھٹ گیا ہے۔"
			Sense 5	Noun	امبی	کچا امب	"امبیاں دا چار بہوں سواد اے۔"

In Table 6, امب comes under the semantic Type*phal*. It shares four senses in the source corpus that are *'amb, amb da wan, amb-ras, and boor'*. It is also considered as polysemous. These extracted senses of *amb* have been used in Punjabi WordNet, Akahr (2016), and Shabdkosh but *ambi* is created manually through a merge approach from Saraiki.

5. Semantic Type: کھیتی باڑی (Khaeti Barhi)

Table 7

Saraiki Noun زمین (Zameen)'s Synsets

Semantic Type	EN	Words	Sense No.	Grammatical Type	Senses	Glosses	Examples
	39	زمین	Sense 1	Noun	زمین	سیارے دا ناں	"زمین ہک بہوں چھوٹا سیارہ ہا۔"
			Sense 2	Noun	سر زمین	قوم دی رہن آلی تھاں	"دہشتگردی کیتے پاکستان دی سرزمین استعمال تھیونڑ دا سوال ہی پیدا نئیں تھیندا۔"
			Sense 3	Noun	احاطہ یاں پلاٹ	گھر بناون کیتے زمین	"راولپنڈی اسلام آباد اچ زمین دی قیمت آسمان نال گالھیں کریندی پئی اے۔"
			Sense 4	Noun	کرۂ ارض	دنیا	"آبادی اچ ودھارے پوری دنیا دی زمین تے پانڑیں دے ذخیرے تے بوجھ پاتا ہے۔"
			Sense 5	Noun	سُکی زمین	جاہ	"دریاویں دے کناریں یاں وچلی سُکی جاہ تے آبادکار یاں ول کاشتکاریکرڈ والے واسی بہوں ہن۔"
			Sense 6	Noun	جائیداد	ملکیت	"زن، زر، تے زمین فساد دی جَڑھ ہن۔"
			Sense 7	Noun	کاشت لئی زمین	زرعی رقبہ	"ہک سر دا مُل ڈاہ توں ویہ روپے نقد ہک مربع زمین ہا۔"
			Sense 8	Noun	ملک	زمینی حدود	"ترکی دی زمین یونان نال گھندی اے۔"

The word زمین, in Table 7, uses the above-mentioned same Type *KhaetiBarhi*. It is categorized as polysemous because of its multiple senses. These senses are acquired from the Punjabi WordNet by using the expansion approach. All these senses are also mentioned in online dictionaries, Akhar (2010) and Shabdkosh.

Conclusion

The research is focused on two main areas: the development of nouns in the Saraiki language and the challenges encountered in the data analysis process. The first part of the research involved the development of a Saraiki language corpus, comprising 2.2 million words. From this corpus, a list of 750 Saraiki nouns was

finalized and divided into different categories. To develop Saraiki's nominal synsets, a hybrid approach was adopted, which involved both the merge and expansion approaches. The merge approach was used to create glosses, example sentences, and some synsets because some of the senses were not mentioned in the Punjabi WordNet due to the cultural gap. The expansion approach was used to develop synsets of Saraiki nouns.

The research methodology involved the conversion of data into machine-readable form, coding of data, and POS tagging to develop identification numbers for nouns, a list of noun words, and a synset of Saraiki nouns. The second part of the research focused on the challenges encountered during the development of Saraiki Nominal Synsets. Since this was the first-ever research on WordNet development for SL, POS tagging was done manually due to the unavailability of the Saraiki tagger. The data was not in machine-readable form, so it had to be converted and tagged manually. The creation of a noun list was time-consuming, as the entire corpus had to be cross-checked, and synsets had to be developed. Each word in the list was checked in the Gurumukhi Punjabi WordNet, Punjabi dictionaries, and Saraiki dictionaries. Glosses and example sentences that were not part of the corpus were constructed by the researcher. Native speakers of Saraiki were consulted to ensure accurate and appropriate results. Finally, the research has opened new avenues for future research in this area.

The present study offers valuable insights into the development of noun synsets in Saraiki, which can be extended to other Pakistani languages such as Sindhi and Pashto. The study provides a sturdy foundation for the development of Saraiki Adjectives, verbs, and adverbial synsets. Furthermore, the study can facilitate the creation of multilingual and bilingual dictionaries for Saraiki language learners, as well as contribute to the development of lexico-semantic relations for other WordNet components. The research also offers a list of nouns, which can be increased to a thousand nouns, and the developed corpus can be expanded to 5 or 10 million, making it an ideal source for the development of online thesauri and dictionaries for the Saraiki language. The study is a significant step towards the creation of the Saraiki Language WordNet, as it provides a comprehensive understanding of contextual meanings of nouns, which can help comprehend words and their proper usage.

References

Akhter, N., Mahmood, M. A., & Nadeem, M. (2019). Desarrollo de nombres en Punjabi y relaciones léxico-semánticas. *Dilemas Contemporáneos: Educación, Política Y Valores*, 1-31. https://doi.org/10.46377/dilemas.v27i1.1529

Awais, M., Azher, M., & Arslan, F. (2023). Developing lexical resources of Saraiki verbs: A corpus-based study. *Linguistic Forum-A Journal of Linguistics*, 5(3), 136-158.

Austin, P. K., & Sallabank, J. (Eds.). (2011). *The Cambridge handbook of endangered languages*. Cambridge University Press. https://doi.org/10.1017/CBO9780511975981

Bhattacharya, P. (2010). IndoWordNet.

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python. *O'Reilly*.

Dryer, M. S., & Haspelmath, M. (Eds.). (2013). *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info

Garcia, M. (2016). Saraiki: Language or dialect? *Eurasian Journal of Humanities,* 1(2), 40-53.

Gul, S., Azher, M., & Nawaz, S. (2021). Development of Saraiki WordNet by mapping of word senses: A corpus-based approach. *Linguistics and Literature Review, 7*(2), 46-66.

Hasan, E., Iqbal, M., Azeemi, Q. and Javeed, A., (2015). An online Punjabi Shahmukhi lexical resource. *Science International,* 25(3), 2529-2535.

Hashmi, R. S., & Majeed, G. (2014). Saraiki ethnic identity: Genesis of conflict with state. *Journal of Political Studies,* 21(1), 79-101.

Kaur, R., Sharma, R. K., Preet, S., & Bhatia, P. (2010). Punjabi WordNet relations and categorization of synsets. In *third national workshop on IndoWordNet under the aegis of the 8th international conference on natural language processing, Kharagpur, India.*

Khaled, S., Noor, U., Imran, M., & Younas, M. (2020). The study of orthographical difference between Punjabi language and Siraiki dialect in Punjab province. *Hamdard Islamicus: Quarterly Journal of the Hamdard National Foundation, Pakistan*, 43, 175-193.

Mladenovic, Miljana & Mitrović, Jelena & Krstev, Cvetana. (2014). Developing and maintaining a WordNet: Procedures and Tools. In *Global WordNet Conference, Tartu, Estonia.*

Moldovan, D. & Novischi, A., (2004). Word sense disambiguation of WordNet glosses. *Computer Speech & Language*, 18(3), 301-317.

MultiWordNet. Multiwordnet.fbk.eu. https://multiwordnet.fbk.eu/english/home.php.

Mushtaq, M. & Shaheen, M., (2017). The Siraiki province movement in Punjab, Pakistan: Prospects and challenges. *Journal of the Punjab University Historical Society*, 30(2), 139-150.

Nazeer, M., Azher, M., Pervaiz, A., & Yasmeen, I. (2024). Developing lexico-semantic relations of Saraiki nouns: A corpus-based study. *University of Chitral Journal of Linguistics and Literature*, *8*(I), 162-182. https://doi.org/10.33195/

Online Dictionary: https://archive.org/details/204912920SaraikiDictionary/page/n5/mode/2up

Online Dictionary: https://skr.m.wiktionary.org/

Online available at https://www.shabdkosh.com/dictionary/english-punjabi/

Online available at http://dic.learnpunjabi.org/default.aspx

Pham, B. (2020). Parts of speech tagging: Rule-based. https://digitalcommons.harrisburgu.edu/cisc_student-coursework/2

Prabhu, V., Desai, S., Redkar, H., Prabhugaonkar, N., Nagvenkar, A., & Karmali, R. (2012). An efficient database design for IndoWordNet development using hybrid approach. *COLING,* 229-236

Shackle, C. (1977). Siraiki: A language movement in Pakistan. *Modern Asian Studies,* 11(3), 379-403.http://www.jstor.org/stable/311504

Shackle, C. (2015). *Siraiki language*. *Encyclopedia Britannica*. https://www.britannica.com/topic/Siraiki-language

Appendix

Semantic Type List and Saraiki Nouns

Serial No.	Semantic Type	Saraiki Nouns	Serial No.	Semantic Type	Saraiki Nouns
1	*Fasal*	*20*	*20*	*Look*	*35*
2	*Khurak*	*15*	*21*	*Amarat ty ondy Hissy*	*40*
3	*Phal*	*23*	*22*	*Waela*	*13*
4	*Phul ty ondy hissy*	*8*	*23*	*Jhah*	*22*
5	*Ghaa Booty*	*9*	*24*	*Rishty*	*40*
6	*Khaeti Barhi*	*14*	*25*	*Pakhi*	*37*

7	*Zaar*	*23*	*26*	*Waan*	*29*
8	*Jism Dy Hissy*	*40*	*27*	*Zanwar*	*47*
9	*Sabziyan*	*12*	*28*	*Ehsaas*	*33*
10	*Kapry*	*33*	*29*	*Dhatan*	*7*
11	*Zewar*	*9*	*30*	*Chezan*	*24*
12	*Mosam*	*13*	*31*	*Pakhi dy Zaa*	*7*
13	*Rang*	*22*	*32*	*Keerhy*	*17*
14	*Bimarian*	*19*	*33*	*Bank*	*16*
15	*Paandy*	*25*	*34*	*Ghaer Insani Cheezan*	*9*
16	*Bayen Layi Cheezan*	*10*	*35*	*Kudrati Cheezan*	*15*
17	*Lakri Tun Bani Cheezan*	*15*	*36*	*Aoun Jaan Layi Cheezan*	*17*
18	*Khaedan*	*16*	*37*	*Paishy*	*8*
19	*Khawan Layi Cheezan*	*20*			