

MORPHOLOGICAL REPRESENTATION OF GENDER IN ARABIC, ENGLISH, GERMAN, AND PUNJABI: A COMPARATIVE LINGUISTIC ANALYSIS

Nimra Noor 1, Muhammad Farukh Arslan 2, Huda Noor 3

1. M.Phil. Linguistics Scholar, English Graduate Studies, National University of Modern Languages (Faisalabad Campus) nimranoor1105@gmail.com
2. Lecturer, English Graduate Studies, National University of Modern Languages (Faisalabad Campus) farukh.arslan@numl.edu.pk
3. M.Phil. Linguistics Scholar, English Graduate Studies, National University of Modern Languages (Faisalabad Campus) hudanoor2112@gmail.com

Abstract

This research provides insights into the morphological representation of gender across Arabic, English, German, and Punjabi, where the last three belong to same language family (Indo-European) while Arabic belongs to an entirely different family (Afro-Asiatic). Yet, Arabic shows more similarities with Punjabi and German as opposed to English, in terms of gender representation. The Distributed Morphology Theory piloted with Language Typology is used to analyze the selected dataset of short stories for language learners in the four mentioned languages. The findings revealed more gender-neutral approach within the English language while highly rich gender-systems in Arabic, German and Punjabi as a result of morphological analysis of all the grammatical categories (parts of speech). English and German differ significantly on the basis of linguistic gender system despite belonging to same branch of same language family, while both have gender-based differences from Punjabi language belonging to another branch of the same family. On the other hand, Arabic, a member of an entirely different family, shows similarities with German and Punjabi on the basis of grammatical gender. It questions on the classification system of languages in various language families. Such differences appear challenging for the language learners while navigating through the rules of second language. The investigation of interplay between language and gender presents significant contributions in the field of education and linguistic typology. It will also assist the language learners to understand the rules of these languages based on gender and how these rules shape the sociocultural narratives respectively.

Keywords: morphological representation of gender, distributed morphology, linguistic typology, short stories for language learners, gender-based linguistic differences.

1. INTRODUCTION

The most significant social tool to express ideas, emotions, and build social relationships and identities is language. It not only performs communicative functions, but also shapes our perceptions of reality and society. The most debated aspect of language across decades is its role in representation of gender. Some languages are gender-neutral such as Mandarin, while others show rich gender-systems such as French.

This study discusses the similarities and differences across languages of same or different families on the basis of morphological representation of gender. English and German both belong to the **Germanic branch** of the *Indo-European language family*, but their treatment of linguistic gender differs significantly. English represents gender only through nouns based on biological gender or the personal pronouns. However, the modern English is becoming more and more gender-neutral, replacing gendered words with suitable neutral terms (Papadopoulos, 2021). German, on the other hand, retains a three-gender system (masculine, feminine, neuter), at various lexical and functional morphemes (Mills, 2012).

In contrast, Punjabi language belongs to the **Indo-Aryan branch** of *Indo-European language family*, having binary system of gender, revealed by distinct gender markings (Kokab, 2020). It is different from both the above mentioned languages not only on the basis of gender but also due to difference in script and structure. However, it shows similarities with Arabic, a

language belonging to the **Semitic branch** of *Afro-Asiatic language family*, written in the script similar to Shahmukhi script of Punjabi. On the basis of gender, not only Punjabi but also German can be classified along with Arabic rather than English, in a particular classification system based on gender contrary to the present one. It also shows the common evolutionary history of Arabic and other Indo-European languages (Jassem, 2013).

1.1. Background of the Study

English language being lingua franca has been spread worldwide and is developing rapidly with the change of time. It is becoming more and more gender-neutral (Papadopoulos, 2021) by replacing gendered words with gender-neutral words, for example, policeman/policewoman has been changed into police officer; fireman/firewoman into firefighter; mankind into humanity; boy/girl into kid; mother/father into parent etc. however, there are certain nouns and pronouns mentioning the biological gender only. On the other hand, many languages highly encode gender in their basic structures and agreements such as Arabic, German, and Punjabi (Kramer, 2020), which are also analyzed in this research. In such languages, not only the nouns and pronouns but other parts of speech are also gendered according to their agreement. Moreover, unlike English, they also assign linguistic gender to the non-living entities too.

Punjabi language is also highly gendered language with two genders: masculine and feminine (Shafiq & Iqbal, 2023), reflecting through various parts of speech. This language also differs in script from English and German. There are two script systems of Punjabi i.e. Gurmukhi script and Shahmukhi script (Arslan, Mehmood & Kanwal, 2024). As the Shahmukhi script is derived from Perso-Arabic script, so this script is analyzed in this study due to its similarities with Arabic language. Arabic also has two genders, masculine and feminine, where the endings of feminine with ة (-ah) represent the linguistic gender assigned to every noun and adjectives (Kramer & Winchester, 2018). Since all these four languages are widespread round the globe, this study intends to explore the gender systems of Arabic, English, German and Punjabi at morphological level employing the theoretical framework grounded in Distributed Morphology theory and Language Typology theory.

1.2. Statement of the Problem

This study addresses the similarities and differences among languages (specifically Arabic, English, German, and Punjabi) on the basis of gender at morphological level. Though English, German, and Punjabi belong to same language family, yet they show differences in gender representation. On the other hand, Arabic being member of an entirely different typological category resembles Punjabi a lot and German to some extent on the basis of gender systems. This study aims to reveal these similarities and differences and to mention the limitations of current language family classification systems.

1.3. Research Objectives

The main objectives of this research are:

- To analyze how grammatical categories reveal the gender system in Arabic, English, German, and Punjabi.
- To compare the morphological structures of gender marking in these languages.
- To evaluate the similarities and differences among languages belonging to same family (English, German and Punjabi) and those from different families (Arabic) on the bases of linguistic gender.

1.4. Research Questions

The above mentioned objectives raise the following questions:

- 1.4.1. What are the basic morphological differences in gender representation between English (a minimally gendered language) and Arabic, German, and Punjabi (languages with rich gender systems)?
- 1.4.2. How are gender-based morphological patterns represented in Arabic, English, German, and Punjabi?
- 1.4.3. How can the cross-linguistic similarities and differences in gender marking be explained by Distributed Morphology and Language Typology?

1.5. Theoretical Framework

The major theoretical framework used in this study is grounded in Distributive Morphology theory piloting with Language Typology. *Distributed Morphology (DM)* first introduced by Halle and Marantz (1993), states that morphological structure of words are in correspondence with the syntactic features, having no division in between. It also recommends that the syntax along with the late insertion of features derived morphological structures. In this study, this framework is used to analyze the gender-based morphemes in various grammatical categories i.e., nouns, pronouns, adjectives, articles, verbs etc. across Arabic, English, German, and Punjabi languages. It also expressed various agreement patterns on the basis of gender.

For the study of languages belonging to different families, the *Language Typology* framework is piloted with Distributed Morphology. This framework helps to compare various languages on the basis of their similarities and differences with respect to the linguistic universals as well as typological features. In this research, the impact of gender markings on grammatical categories in the four languages belonging to different typological families is assessed. Together with the Distributed Morphology, this framework provided a comprehensive insight into how gender is assigned to morphological features in Arabic, English, German, and Punjabi.

2. LITERATURE REVIEW

2.1. Overview of the Literature Review

Morphological representation of gender across languages has been an interesting topic of debate in the domain of linguistics since long ago, in order to investigate how gender is encoded and represented in various languages. By understanding the gender representations in languages, the study of underlying linguistic structures can be assisted. Many researches have established the foundations for the current study by exploring the morphological representations and shifts in various languages:

2.2. Language and Gender:

Gender in languages reinforces, or challenges the social gender identities and influences the language structures in terms of the grammatical gender-based agreements. Gender is polysemous in nature that means it not only represents sex-animacy, but is also associated with individuality, collectiveness, abstractness, number, size, evaluation, and perceptivation (Fehri, 2016). According to Yule (2016), it is reflected in three ways: natural gender, grammatical gender, and the social gender. Natural or biological gender is the distinction in sex; grammatical gender is the masculine and feminine classifications of nouns in languages; while the social gender is the classification of individuals due to their social roles. Veeman, Berdicevskis, Allasonnière-Tang, and Basirat (2020) argued that the rules for the assignments of grammatical

gender are lineage-specific as well as universal. However, the distinctions across languages are the results of phylogenetic distance. Prewitt-Freilino, Caswell, and Laakso (2011) found link between grammatical and social gender, i.e. more gender in language results in less gender equality and vice versa.

2.3. Arabic Gender: Binary Morphological System

Arabic has an extensive system of grammatical gender assigning masculine or feminine gender to every noun, its pronoun, adjective and also the verbs in agreement with the noun. According to Steriopolo (2021), the grammatical gender reversals in Arabic are due to its sociocultural influences. Arabic has gendered morphological markings through affixes in the similar way as the markings of Punjabi gendered morphemes, highlighted by Arslan, Kanwal, Mehmood, and Haroon (2024). Jassem (2013) claimed that Arabic phonological structures (especially the consonantal features) have provided historical foundations for linguistic gender in Indo-European languages. This also represents the gender based historical similarities among languages of different families.

2.4. Gender Representation in English

Being an analytic language, English has a minimal system of linguistic gender. Kramer (2020) stated that in English linguistic gender highly restricted to nouns or pronouns indicating natural gender (i.e., pronouns like “he”, “she”) and lexical gender in certain nouns (e.g., "actor" for male and "actress" for female). This demonstrates the limited morphological gender representation in English as compared to German and Arabic, where gender is grammatical. Papadopoulos (2021) challenges traditional notions of gender in English, claiming that language evolve with the evolutions in social norms as the recent evolution in English has declared it a non-binary and gender-inclusive language (e.g., the use of "they" as a singular pronoun using where the biological gender is not mentioned). However, such use of “they”, despite of its widespread, is still considered non-grammatical (Konnelly & Cowper, 2020).

2.5. Gender in German: Grammatical Complexity

Comparatively, German has a robust grammatical gender system showing masculine, feminine and neuter gender in nouns, pronouns, articles, and adjectives. As demonstrated by Jassem (2013), the gender in German language along with many other European languages has been evolved from Arabic cognates indicating a chronological evolution. It is related to Kramer’s (2020) typological analysis according to which grammatical gender in German is linked with the semantic and morphological influences. This intricate relationship between language and gender in German makes its morphological structure more complex as compared to English (Mills, 2012).

2.6. Punjabi Gender: Rich Morphological System

Punjabi also has a binary gender system like Arabic revealing gender markings through agreements at various grammatical categories. Arslan, Mahmood, Shoaib, Idrees, and Tariq (2023) investigated inflectional patterns for number and gender in Punjabi nouns. Also, Arslan, Kanwal, Mehmood, and Haroon (2023) studied case markers of Punjabi nouns to have deeper understanding of complexity of Punjabi nouns. Similarly, Arslan, Kanwal, Mehmood, and Haroon (2024) explored Punjabi adjectives and found the distinct morphemes that inflect for gender and number through morphological analysis of adjectives. Shafiq and Iqbal (2023) conducted a comparative analysis of English and Punjabi, while Kokab (2020) compared Punjabi

with Urdu on the basis of gender. Findings of both these studies represented a diverse gender system of Punjabi close to Urdu and entirely different from English.

2.7. Comparative Analysis: Arabic, English, German, and Punjabi

The framework for a comparative analysis of grammatical gender in Arabic, English, German, and Punjabi is provided by the previous researches. Kramer (2020) and Papadopoulos (2021) highlighted the minimal morphological gender system in English mostly associated with the semantic gender distinctions. On the other hand, Arabic, German and Punjabi show extensive and complex systems of grammatical gender, as evident in the researches of Jassem (2013), Steriopo (2021) and Kokab (2020). Such gender-based similarities in Arabic, German and Punjabi emphasizes the link between Indo-European and Afro-Asiatic languages on the basis of gender (Veeman, Berdicevskis, Allasonnière-Tang & Basirat, 2020).

Still there is a question mark on the divergence of these languages on the basis of their treatment of gender at morphological level. The answer is that in comparison to Arabic, German and Punjabi, the shift of English towards a gender-neutral language is due to its sociocultural influences on its linguistic development, as Jassem (2013) found common phonological origin between Arabic and Indo-European languages in his historical analysis of cognates, leading to the morphological developments. He also suggested the sociocultural factors as the reason of divergence over time.

2.8. Gaps in Literature and Rationale for the Current Study

Despite a comprehensive existing literature on the linguistic gender systems, there are certain noticeable gaps, for example, some studies investigated gender markings in various languages but targeting of language families on the basis of similarities and differences has not been focused. Arabic language has minimally analyzed with English and German at morphological level before due to its entirely different script and morphology. Also, Punjabi being a regional language has not been compared to these international languages especially on the basis of gender. Furthermore, Distributed Morphology has not been piloted with Language Typology in any gender-based language study before. Therefore, the differences among languages of same or different families have not being focused comprehensively.

On the basis of these gaps, this study aims to analyze the gender-based morphological differences among Arabic, English, German, and Punjabi languages comparatively using Distributed Morphology and Language Typology frameworks. It will provide deeper insights into the treatment of gender by these languages at morphological level, contributing to the practical and pedagogical implications of this study.

3. RESEARCH METHODOLOGY

This research employed a comprehensive mixed method approach utilizing Distributed Morphology framework piloted with Language Typology for the comparative analysis of gender representations in Arabic, English, German, and Punjabi languages at morphological level.

3.1. Research Design

In this study, a *mixed method approach* is adopted combining both quantitative and qualitative methods. The *corpus based quantitative research design* helped to provide statistical evidences for the theoretical descriptions. While a *comparative qualitative research design* helped to describe the key differences among languages theoretically by applying Distributed

Morphology and Language Typology. For this purpose, short stories designed for language learners having simple language and standard grammar are analyzed comparatively at morphological level in order to evaluate the gender markings in various grammatical categories.

3.2.Data Collection and Sampling

The data collected for this research includes short stories designed for language learners in the four languages under discussion, i.e. Arabic, English, German, and Punjabi. A corpus of 15,420 words is collected by 20 short stories (5 in each language) using purposive sampling technique. Each short story contained 500-1000 words. The stories are gathered from books designed for language learners, having standard and simple language. Besides, the everyday language and situations were discussed in the stories, preventing them to become part of some complex register. This kind of sampling ensured the appropriate data selection for the gender-based comparative analysis of languages.

The data is then converted into notepad files by using AntFileConverter software, separately for each language. Then, AntConc software tagged the data grammatically from which the gender encoded morphemes were extracted and managed in the excel sheet to analyze along with frequencies in order to provide authentic statistical data. This tool is also used by Arslan, Mahmood, and Rasool (2020) for corpus based analysis of morphemes.

3.3.Data Analysis

The data is analyzed at morphological level using Distributed Morphology piloted with Language Typology both quantitatively and qualitatively.

3.3.1. Quantitative Analysis

The corpus for each language is tagged separately by AntConc and frequencies are recorded, as done by Arslan, Mahmood, and Rasool (2020). Then the gendered words at each grammatical level are pinpointed i.e. the gender agreements and gender-based inflectional changes and represented in the form of a table. Finally, the results of all the four languages are compared to interpret the differences.

3.3.2. Qualitative Analysis

After the identification of the gendered markings in each language, the descriptive interpretation of how gender is morphologically encoded, and their comparative analysis is given to identify the extent of use of linguistic gender in each language. This qualitative analysis is conducted through the lens of Distributed Morphology piloted with Language Typology for comparative analysis across different languages.

3.4.Significance of the Study

The significance of this research lies in its contribution to the gendered language study with a special focus on the comparative analysis of grammatical gender. Moreover, it highlights similarities among languages belonging to different language families, and differences among languages belonging to same language family on the basis of linguistic gender. It also provides significant contribution for the language learners to understand gendered nuances among various languages, by addressing the most common challenge learners face during learning a new language related to gender markings. The dataset collected for this research (i.e., short stories for language learners) also reveals its pedagogical implications. Furthermore, this study also

challenges the current language classification system by cross-linguistic comparisons of gendered languages.

3.5. Delimitations

This study is focused on the morphological analysis of Arabic, English, German, and Punjabi languages on the basis of gender markings in each grammatical category. The selection of these four languages Arabic, English, German, and Punjabi (last three belonging to same family and first one to a different family) for analysis is intended to investigate the typological differences in gender systems across different language families. The short stories for language learners are selected to analyze the standard use of grammar in simple language. The framework of Distributed Morphology along with Language Typology is selected for structural comparative analysis of gendered languages instead of semantics or pragmatics of gender.

As for the ethical concerns, the researcher has chosen the publically available data for which there were no consent issues. The data is used honestly without editing for manipulation of results in order to justify the research hypothesis. Furthermore, proper citations are given for each previous study used to validate this research. The limitations and delimitations of study are also mentioned explicitly.

4. DATA ANALYSIS

The gender is encoded in Arabic, English, German, and Punjabi differently, i.e. minimal representation in English through only the natural genders, while three genders in German, and two genders in Arabic and Punjabi distributed extensively across these languages. The comparative analysis of gender morphology across these languages provides deeper insights into different extents of gender marking. The findings of the research are discussed separately for quantitative and qualitative analysis of the data.

4.1. Quantitative Analysis

The data collected for this research included a corpus of 15,420 words from the short stories in Arabic, English, German, and Punjabi designed for the language learners. The data is then tagged along with frequencies and converted to the form of word list through AntConc and analyzed through gender insights and then recorded in tables form given below separately for each language:

4.1.1. Arabic

Arabic has a diverse system of gender due to its complex structure in which a word conveys the meaning of a complete sentence. The verbs, for example, have prefix related to the subject, while a suffix in agreement with the object. Such a complex system is problematic to handle. So, in the Table 1, the features are tagged grammatically in relation to the subject:

Table 1
Grammatical gender in Arabic

Grammatical Feature	Masculine	Frequency	Feminine	Frequency
Noun + Noun Phrase	الشاطر	22	يد	11
	محمد	13	الحبوب	7
	السلطان	11	القافلة	5
	الخروف	10	أمه	3
	الحمار	9	العصافير	3
	صديقي	8	يدان	3
	الوزير	7	أكوام	2
	الكلب	6	أم	2
	جال	5	ابنة	2
	رجال	5	الأم	2
	*others	322	*others	101
total	418	total	141	
Pronoun	هذا	15	ها	24
	الذي	4	هذه	5
	عنه	2	التي	3
	هم	2	بهذه	3
	أنه	1	فيها	2
	منهم	1	خلالها	1
	ناك	1	لها	1
	هو	1	منها	1
	total	27	هي	1
			total	41
Adjective	جميلا	4	كبيرة	3
	آخر	3	أخرى	2
	التالي	3	إحدى	2
	طويلا	2	الأخرى	2
	عال	2	الطويلة	2
	نائما	2	الكبيرة	2
	أحق	1	جميلة	2
	أشداء	1	كثيرة	2
	أعلى	1	أخرى	1
	الآخر	1	الخصبة	1
	*others	49	*others	23

	total	69	total	42
Verb	راق	2	كانت	3
	وعاد	2	ينسجها	2
	وقف	2	أخذها	1
	آمن	1	اعتنت	1
	أثار	1	القاها	1
	أحضره	1	تبيع	1
	أخفى	1	ترك	1
	أذكره	1	تسير	1
	أرجوك	1	تصح	1
	*others	110	*others	24
	total	122	total	36

(* only top 10 most frequent words are added in the table, while frequencies of the rest of words are added)

4.1.2. English

In English, the gender is minimally encoded in the language demonstrating only the natural gender in the language. Only the nouns and pronouns referring to the male and female sex are represented through gendered language in English. In the Table 2, the gendered words in English language from the selected corpus are listed in the form of table along with their frequencies:

Table 2

Grammatical gender in English

Grammatical Feature	Masculine	Frequency	Feminine	Frequency
Noun	Armando	37	Sylvia	40
	George	29	Julia	35
	man	29	sister	20
	dad	18	mum	15
	Daniel	14	mother	2
	father	8	Sarah	1
	son	7	woman	1
	Antonio	3	total	114
	waiter	2		
	Arthur	1		
	brother	1		
	men	1		
total	150			
Pronoun	he	56	she	42

	his	11	her	14
	him	6	total	56
	total	73		

4.1.3. German

In case of German, there are three genders: masculine, feminine, and neuter assigned to every noun and also demonstrated by the related articles, adjective and pronouns. Following Table 3 shows the words of German language that have assigned gender as well as their frequencies:

Table 3

Grammatical gender in German

Grammatical Feature	Masculine	Frequency	Feminine	Frequency	Neuter	Frequency
Noun	Michael	28	Julia	31	Haus	13
	Mann	26	Silvia	22	Restaurant	8
	Jochen	16	Currywurst	14	Handy	7
	Lastwagen	13	Schwester	12	Geld	5
	Vater	13	Wohnung	10	Wesen	5
	Bus	12	Mutter	9	Boot	4
	Daniel	12	Kisten	5	Zimmer	4
	Fahrer	8	Frage	4	Currywurstrestaurant	3
	Sohn	6	Geschichte	4	Flugzeug	3
	Tag	6	Mama	4	Holz	3
	*others	121	*others	109	*others	64
	total	261	total	224	total	119
Pronoun	er	20	meine	18	es	52
	mein	18	ihre	5	mein	18
	jeden	1	jede	1	alles	4
	total	39	total	24	nichts	3
					jedes	1
					total	78
Article	der	42	sie	46	das	56
	den	38	die	44	dem	27
	dem	27	eine	25	ein	23
	ein	23	keine	9	einem	10
	einem	10	einer	4	des	7
	des	7	total	128	diesem	4
	diesem	4			total	127

	einen	4				
	keinen	3				
	total	158				
Adjective	kölner	1	kleine	1	längeres	1
	langer	1	leckere	1	schlimmes	1
	total	2	letzte	1	total	2
			total	3		

(* only top 10 most frequent words are added in the table, while frequencies of the rest of words are added)

4.1.4. Punjabi

Comparatively, Punjabi has robust gender system represented through almost all the parts of speech i.e. nouns, pronouns, verbs, adverbs, adjectives, postpositions etc., where postpositions are prominent in Punjabi as opposed to the other three languages in comparison. These are demonstrated in the Table 4, along with frequencies of the gendered words:

Table 4

Grammatical gender in Punjabi

Grammatical Feature	Masculine	Frequency	Feminine	Frequency
Noun	اکرم	36	چا	19
	عمران	32	صحت	12
	پنڈ	18	گل	11
	افضل	14	دھی	10
	گھر	14	اکھاں	9
	امجد	13	دولت	8
	پیو	11	سمجھ	8
	پنر	10	سوئی	7
	دن	9	گلاں	7
	کم	9	حفاظت	6
	*others	493	*others	290
total	659	total	387	
Pronoun	اپنے	17	اپنی	18
	اوبدے	16	اوبدی	9
	اوبدا	6	اوبدیاں	1
	اپنا	5	اپہدی	1
	اتنے	3	اپہوائی	1
	اپہدے	3	جیہڑی	1
	ایسے	2	ساڈی	1
	ایڈا	2	میری	1

	تیرے	2	total	33
	جیہدا	2		
	*others	12		
	total	70		
Adjective				
	بڑا	12	ساریاں	9
	دوچے	7	بڑی	8
	وڈا	7	وڈی	4
	سارے	6	پکی	4
	سارا	5	ساری	3
	وڈے	5	یوڈ	3
	پہلے	3	پوری	3
	پیلے	3	بولی	3
	کھلے	3	بوہنی	2
	بند	2	تھوڑی	2
	*others	65	*others	36
	total	118	total	77
Verb				
	گیا	25	سی	20
	دتا	16	گئی	10
	گا	15	سوئی	7
	بوندا	15	چاپیدی	5
	چاپیدا	10	آئی	4
	لگا	7	گی	4
	پیا	7	آگئی	3
	کیتا	7	جاندی	3
	ٹے گ	7	چاہی	3
	آیا	6	دیا	3
	*others	233	*others	64
	total	348	total	126
Adverb				
	آج	4	چھیتی	4
	سویرے	4	کدی	4
	پہلے	3	پہلاں	3
	ایڈے	1	total	11
	بچھے	1		
	کٹھے	1		
	کل	1		
	واپس	1		
	total	16		

Postposition	کے	70	دی	64
	دے	66	سی	20
	دا	43	دیاں	16
	کا	7	کی	7
	والے	4	total	107
	والا	3		
	جیسے	2		
	سا	1		
	total	196		

(* only top 10 most frequent words are added in the table, while frequencies of the rest of words are added)

4.1.5. Comparative Analysis

The comparative analysis of these grammatical genders and their frequencies in all the three languages under discussion is shown in the table given below:

Table 5

Comparative grammatical gender analysis across Arabic, English, German, and Punjabi on the basis of frequencies

Grammatical Feature	Arabic		English		German			Punjabi	
	Masculine	Feminine	Masculine	Feminine	Masculine	Feminine	Neuter	Masculine	Feminine
Noun	418	141	150	114	261	224	119	659	387
Pronoun	27	41	73	56	39	24	78	70	33
Article	0	0	0	0	158	128	127	0	0
Adjective	69	42	0	0	2	3	2	118	77
Verb	122	36	0	0	0	0	0	348	126
Adverb	0	0	0	0	0	0	0	16	11
Postposition	0	0	0	0	0	0	0	196	107
Total gendered words	636	260	223	170	460	379	326	1407	741

This table shows the differences among Arabic, English, German, and Punjabi on the basis of their treatment of gender in language with the help of their frequencies in the collected corpus. The total frequencies of gendered words are comparable but the number of words having grammatical gender varies a lot. In other words, a large number of gendered words having less frequency are also present in the corpus of Arabic, German, and Punjabi, but all of them are not represented in the table. So the overall frequencies are unable to represent the prominent gender-based distinctions of these languages, which can be represented by counting all the unique words encoding gender.

4.2. Qualitative Analysis

The above quantitative findings provided the background for the qualitative discussion of the differences among Arabic, English, German, and Punjabi in terms of gender representation.

The frameworks of Distributed Morphology and Language Typology are used to describe and compare these differences.

Arabic has an entirely different system of gender markings as compared to English, German and Punjabi. Despite having only two genders, it is more complex than German (having three genders). Punjabi also has two genders but the words of English, German, and Punjabi perform only one grammatical role at a time and convey isolated ideas. But in Arabic, a single word can convey the meaning of the whole sentence. For example, أخذها means “she took it”, جلس means “he sat down” etc. The suffix ة after each masculine word indicates the feminine endings of nouns, and pronouns. The suffix ه means his or him, while ها means her. For a masculine subject, the verb begins with ي (ya) and for a feminine subject with ت (ta). Due to such difficulties the classifications of the whole word in one gender category is very difficult (Kramer & Winchester, 2018). Therefore, this expresses more complex gender system in Arabic than English, German and Punjabi, also evident by the social roles assigned to different genders (Alali, 2023).

In English short stories, there were a fewer words occurred frequently to represent the natural genders only i.e. father, mother or the names of people etc. Moreover, only the personal pronouns referring to third person singular were found to be gendered like he, she, his, him, her etc. This proves the minimal gender representation in English language. While, many nouns and pronouns are found to be common for both genders, for instance, friends, parents, students, teacher, people, they, anyone, someone etc. revealing the gender-inclusive nature of English (Papadopoulos, 2021). This is also affecting the social roles of gender i.e. promoting gender equality due to less distinctions of gender in language (Prewitt-Freilino, Caswell, & Laakso, 2011).

Contrarily, German assign gender to inanimate objects in addition to the animate objects like Arabic and Punjabi and this assignment is random without any specific rule. In German short stories, there was a robust system of gender showing masculine, feminine, and neuter gender in nouns, pronouns, articles and a few adjectives. For example, nouns like Mann (masculine), Frau (feminine), and Haus (neuter) were followed by the articles der, die, and das respectively (e.g., der Mann, die Frau, das Haus). The pronouns used for these genders are er (he), sie (she), and es (it) respectively. Adjectives also showed inflections on the basis of gender, case and number. For example, kleiner for masculine, kleine for feminine, and kleines for neuter gender. Such inflections help to understand which gender is being expressed (Mills, 2012).

In case of Punjabi, there are more gender markings than Arabic, English, and German. For instance, nouns like استاد is used for male teacher and استانی for female teacher; pronouns such as اپنا for his and اپنی for her; adjectives such as چنگا and چنگی both mean good for male and female gender respectively; verbs e.g. آیا and آئی both mean came for male and female subject respectively; adverbs like کٹھے masculine meaning together, while ٻکد feminine meaning some times; and postpositions such as دا and ٻد used for possessive postpositions with masculine and feminine nouns respectively. In both the Gurmukhi and Shahmukhi scripts, these structures are common (Arslan, Kanwal, Mehmood, & Haroon, 2023), so these findings can be generalized for the Punjabi language as a whole.

Punjabi language is similar to Arabic not only in terms of script but also in terms of gender assigning. However, the distinctions exist in the structures and word orders. For example, Arabic structure allows various morphemes within one word-like element, covering the meaning

of whole sentence sometimes within a single element, as explored by Yule (2016) in Sawahili. Punjabi, on the other hand, resembles to German and English in this respect. However, its word order is entirely different from English and German. Punjabi has postpositions rather than prepositions and these postpositions can also be gendered. While, in case of Arabic, the adjectives are placed after the noun, showing variation from the other three languages.

5. DISCUSSION

The findings are now discussed within the framework of Distributed Morphology and Language Typology. According to Distributed morphology, the minimal gender system in English and its gender-inclusive nature is the consequence of late insertion of features which removed the grammatical gender from English making it a gender neutral language (Halle & Marantz, 1993). Further developments in English are replacing the gendered nouns and pronouns with common gender words like sibling instead of brother or sister and they instead of he or she (Konnolly & Cowper, 2020). Through the lens of Language Typology, English has simplified grammatical system due to the non-binary development, but it has limited the expressiveness of language in terms of gender.

German, in contrast, has a rich gender system that influences the grammatical agreements of nouns, pronouns, determiners, and adjectives. Distributed Morphology considers this complexity as a result of development of gender as a feature during morphological derivation of German, as argued by Jassem (2013) using Lexical Root Theory. Language Typology explains this development as a positive characteristic in German for clear and precise expression of gender despite of its complex structure. For example, the English gender neutral word professor indicates the male professor in German, for the female the suffix *-in* is used i.e. professorin is the female professor in German. So the idea is clear about gender by this distinction.

The binary gender system of Punjabi is evident by almost all the grammatical categories. As the script of Punjabi language is derived from Arabic, the gender markings are also similar to those of Arabic with a few variations developed over time as described by Distributed Morphology theory. These variations are suffix-based morphological rules (Arslan, Mahmood, Shoaib, Idrees & Tariq, 2023). While, Language Typology explains linguistic gender in Punjabi language as a significant part of grammar, that maintains grammatical coherence but appears challenging for language learners.

The linguistic gender in Arabic is exhibited through the varied grammatical markings. The gender is assigned to each noun either masculine or feminine that is also reflected by the adjectives, pronouns and verbs through various inflections. Distributed Morphology finds these inflectional patterns emerged from the root-based morphological structures. The gender agreements in Arabic appear syntactic rather than morphological effect (Kramer & Winchester, 2018). Whereas Language Typology considers this obligatory nature of Arabic gender system as greater degree of grammatical coherence, that seems challenging for Arabic learners especially those who have a gender neutral language as their native language.

The comparative analysis of Arabic, English, German, and Punjabi at gendered morphological level shows distinctions among each language gender system. However, on the basis of extent of grammatical gender, Punjabi must be classified in the language family of Arabic, and German seems close to Arabic and Punjabi as compared to English despite being a member of different family than Arabic and Punjabi. The reason given by Jassem (2013) is that the Arabic language is the root for the emergence of many Indo-European languages. Later on,

late insertions occur and languages developed differently. The classification system of languages that classify English, German, and Punjabi in same language family and Arabic in different does not address gender-based similarities and differences among these languages. So, the existing classification system of languages is challenged in this research as Papadopoulos (2021) did in his study finding no accurate system to analyze gender in languages. There is a need for some more comprehensive system of classification that covers various aspects of languages including gender. In such system, Indic languages would be classified with Afro-Asiatic languages rather than European languages, and English would be placed separately with the gender-neutral languages such as Mandarin Chinese.

5.1. Limitations of the Present Study

The major limitation of this study is its focus on the morphological analysis of four languages only. The syntactic and semantic aspects, as well as the impacts of this gender-based analysis on societies, cultures and people's thought process are beyond the scope of this research. The data collection is also limited to short stories for language learners so the findings may not be generalized on informal conversations or complex language use in these languages. Besides, only the standard grammar provided in the short stories book is analyzed in this study, neglecting the colloquial and dialectal variations among languages. Finally, direct parallels between the gender systems of these languages could not be drawn due to the comparison of language from different families.

5.2. Suggestions

On the basis of this study, it is recommended to the educators to focus on the gender based differences and rules of languages along with their everyday examples that how gender influences the words construction in the language being taught. The curriculum should be designed in the way that learners from different backgrounds may internalize the rules regarding the integration of gender in that particular language though repetitive exercises and real time examples. It is also recommended to the linguists to modify the language classifications into families based on certain features other than evolutionary history i.e. on the basis of gender, ideologies, region, script etc. so that there must be a complete description of similarities and differences of languages in a family by the study of that particular family in a particular classification system. For examples, in the classification system based on gender languages with minimal gender system (i.e. English) should be classified separately, those with binary gender system (like Arabic and Punjabi) separately, and those having three genders (such as German) separately. Similarly, the languages showing gender patterns at one or two grammatical categories must be differently classified from the languages having more integration of gender at various levels.

5.3. Recommendations for Future Research

In future, researchers may explore the differences among gendered languages functionally instead of structurally i.e. at semantic or pragmatic level. The influence of such languages on the thought processes or the social practices of the speakers may also be investigated. Moreover, other languages must also be studied by utilizing this framework. The comparison can also be made on the basis of other similarities or differences rather than gender. A research on the influence of cognition on acquisition that how learners acquire such gendered grammatical patterns can also be done in future.

5.4. Contributions of this Research

Despite some limitations, the research can have various valuable contributions in the fields of linguistics and education. Through the cross-linguistic gender analysis, the study portrays the gender encoding patterns of different languages. It may also help to investigate the influence of such linguistic genders on the thought process of the speakers of that particular language. The features highlighted in this research may also help the language learners to overcome the difficulties faced in second language acquisition by enhancing the pedagogical methods. Arslan, Mehmood, and Kanwal (2024) also claimed that morphological awareness can play a crucial role in language learning, particularly when learning morphologically rich languages as second language. For example, the learners of Arabic, German, and Punjabi need to pay more attention on the gender based rules as compared to the English learners. This study also has some technology based implications in development of machine learning and translation for accuracy in gender agreements.

6. CONCLUSION

This comparative study on the basis of gendered morphology across Arabic, English, German, and Punjabi showcased the distinct gender systems of these languages. The findings revealed how gender is encoded in language and how different languages deal with gender differently. For example, English has minimal gender system, reflected only through some nouns and pronouns. On the other hand, Arabic, German and Punjabi are rich in gendered morphology exhibiting gender through various parts of speech. However, the characteristics of Arabic to communicate a whole sentence in a word or two, often using more than one gender (at subject and object level) makes it different from German and Punjabi. Still in comparison with English, these three languages being rich in grammatical gender seem closer to each other. Thus, there must be many other classifications of languages in families on the basis of various similar of different features including gender, other than their evolution. The study can have useful implications on language pedagogy, helping learners to understand the linguistic differences while learning a new language. Future researchers may explore the sociological interpretation of gendered language or may find out more similarities and differences within these languages studied under this research.

REFERENCES

- Alali, H. (2023). Developmental trajectory of grammatical gender: evidence from Arabic. <https://hdl.handle.net/1842/39731>
- Arslan, M. F., Kanwal, A., Mehmood, M. A., & Haroon, H. (2023). Morphemic Analysis Of Case Markers In Shahmukhi Punjabi Nouns. *Journal of Namibian Studies: History Politics Culture*, 33, 6501-6517. <https://doi.org/10.59670/tg8eqh42>
- Arslan, M. F., Mehmood, M. A., & Kanwal, A. (2024). Punjabi Adjectives: A Morphological Perspective. *Pakistan Languages and Humanities Review*, 8(1), 543-552. [https://doi.org/10.47205/plhr.2024\(8-1\)48](https://doi.org/10.47205/plhr.2024(8-1)48)
- Arslan, M. F., Mahmood, M. A., & Rasool, A. (2020). A corpus-based comparative study of derivational morphemes across ENL, ESL, EFL learners through ICNALE. In *Linguistic Forum-A Journal of Linguistics* (Vol. 2, No. 4, pp. 1-12). <https://doi.org/10.53057/linfo/2020.2.4.1>

- Arslan, M. F., Mahmood, M. A., Shoaib, M., Idrees, S., & Tariq, Z. (2023). Morphological Description Of Nouns In Shahmukhi Punjabi; A Corpus Based Study. *Journal of Positive School Psychology*, 1259-1269.
- Fassi Fehri, A. (2016). Semantic Gender Diversity and Its Architecture in the Grammar of Arabic. *Brill's Journal of Afroasiatic Languages and Linguistics*, 8(1), 154-199. <https://doi.org/10.1163/18776930-00801007>
- Halle, M., & Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 111–176). The MIT Press.
- Jassem, Z. A. (2013). The English, German, and French cognates of Arabic back consonants: A lexical root theory approach. *International Journal of English and Education 2* (2), 108-128.
- Kokab, S. (2020). Comparative Study of Gender Marking in Punjabi and Urdu. *EUROPEAN ACADEMIC RESEARCH*, VIII(9), 5102-5116.
- Konnolly, L. & Cowper, E., (2020) “Gender diversity and morphosyntax: An account of singular *they*”, *Glossa: a journal of general linguistics* 5(1): 40. doi: <https://doi.org/10.5334/gigl.1000>
- Kramer, R. (2020). Grammatical gender: A close look at gender assignment across languages. *Annual Review of linguistics*, 6(1), 45-66. <https://doi.org/10.1146/annurev-linguistics-011718-012450>
- Kramer, R., & Winchester, L. (2018). Number and gender agreement in Saudi Arabic: morphology vs. syntax. *The Proceedings of the 17th Texas Linguistic Society*, 39-53.
- Mills, A. E. (2012). *The Acquisition of Gender: A study of English and German (Vol. 20)*. Springer Science & Business Media.
- Papadopoulos, B. (2022). The definitional dilemma of gender in language. *Hesperia: Yearbook of Hispanic Philology*, 24 (2). <https://doi.org/10.35869/hafh.v24i2.4106>
- Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex roles*, 66(3), 268-281. <https://doi.org/10.1007/s11199-011-0083-5>
- Shafiq, Z., & Iqbal, S. (2023). A Cross-linguistic Analysis of Punjabi Versus English Gender-System. *Global Sociological Review*, VIII, 8, 1-8. [https://doi.org/10.31703/gsr.2023\(VIII-II\).01](https://doi.org/10.31703/gsr.2023(VIII-II).01)
- Steriopolo, O. (2021). Grammatical gender reversals: A morphosyntactic and sociopragmatic analysis. *Open Linguistics*, 7(1), 136-166. <https://doi.org/10.1515/opli-2021-0008>
- Veeman, H., Allasonnière-Tang, M., Berdicevskis, A., & Basirat, A. (2020). Cross-lingual embeddings reveal universal and lineage-specific patterns in grammatical gender assignment. In *CoNLL 2020: 24th SIGNLL Conference on Computational Natural Language Learning. Online November 19-20, 2020*. (pp. 265-275). DOI: [10.18653/v1/2020.conll-1.20](https://doi.org/10.18653/v1/2020.conll-1.20)
- Yule, G. (2016). *The study of language*. Cambridge University Press.