ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 14, No. 10, pp. 3057-3066, October 2024
DOI: https://doi.org/10.17507/tpls.1410.06

The Role of Culture in Abusive Language on Social Media: Examining the Use of English and Arabic Derogatory Terms

Nahla Alshalabi
Al Ain University, UAE

Hanene Lahiani
Al Ain University, UAE

Ayman Yasin
Princess Sumaya University of Technology, Jordan

***Abstract*—Although several studies have dealt with the use of derogatory terms on social media, only few compared the phenomenon across languages from a sociocultural aspect. This study used a mixed-method comparative analysis of 920 Arabic and English abusive tweets. The researchers used content analysis to annotate the tweets according to their type and severity. They also used qualitative thematic discourse analysis to interpret the linguistic themes. Furthermore, they used frequency analysis to statistically identify the most common targets and lexical items and to identify the sociolinguistic patterns behind them. The results reveal that Arabic tweets have higher frequencies of gender abusive terms, and they are more severe than the English ones. However, English showed greater reliance on vulgar terms because of cultural taboos. English communication was also dominated by implicit insults, while Arabic favored explicit offense in accordance with direct/indirect cultural values. Both languages used emojis intensively, but Arabic used more diverse registers within messages. Anonymity boosted prejudices for both languages. In conclusion, the difference in online toxicity between the languages is the result of linguistic differences and the cultural norms and the interaction between the two.**

***Index Terms*—corpus, cultural norms, derogatory terms, discourse analysis**

I. Introduction

Abusive language and derogatory terms are a clear phenomenon in human communication on social media platforms. In order to control filthy content and promote respectful discourse across languages and communities, it is very important to understand cultural variations behind the use of offensive language (Bilewicz & Soral, 2020). Upon reviewing the literature, it appeared that a few researchers have studied profanity on social media, mainly in English.

However, limited interest is shown towards cross-cultural study of the phenomenon via comparisons with other languages like Arabic (Kerkam, 2015). Sood et al. (2012) have analyzed over 300,000 tweets and concluded that the most frequent insults on twitter focused on stupidity, sexuality, race and physical appearance. Specifically, terms like "idiot", "stupid," and racist slurs were widely used to attack and degrade others.

Similarly, Cheng et al. (2017) conducted a computational analysis of millions of YouTube comments and found that the top derogatory terms consistently referred to low intelligence. These studies indicate that intelligence-based insults represent a crucial aspect of offensive English language online. Additionally, identity-focused attacks through racist, homophobic or sexist slurs also appear to be prominent in abusive English discourse on social networking platforms.

In contrast, less is known about Arabic offensive language norms and practices on social media sites. To help address this gap, this study manually classified Shammur's (2020) Excel dataset that contains 920 highly toxic and obscene Arabic tweets (see appendix). The tweets were classified as "Offensive", "Vulgar," or "Hate Speech" by the researchers. A preliminary analysis of terms by Mubarak et al. (2020) sheds light on some distinctions from English. Common offensive Arabic words like *kalb* (dog) and *ɦimar* (donkey) imply lack of intelligence or animal-like behavior. However, terms directly referring to private body parts or sexual references/acts indicating vulgar insults like *dajju:θ* (cuckold) or *kaððab* (liar) have been observed to play a relatively more salient role than in English.

Interestingly, identity-focused attacks that are seen frequently in English, such as racist insults, occur less regularly within this Arabic dataset (Mubarak et al., 2020). Rather, many Arabic tweets either condemn persons religiously or accuse them of hypocrisy or immorality.

Existing literature offers contextual evidence for these cross-linguistic differences. Mubarak and Darwish's (2019) rates (1-2% of Arabic tweets contain toxic content) are considerably lower than rates stated for English. Also, Al-Jarf (2020) argued that standard Arabic is generally used for rather serious topics like politics and religion whereas dialectal

and colloquial language is used for everyday issues and discussions. This implies that more situations justify the use of formal Arabic language in online spaces than English. Previous research shows that offensive language is a widespread problem across cultures in social platforms. However, the present comparison of existing English and newly provided Arabic datasets gathered by researchers, including Fahmy (2021), indicates offence construction and preferred targets that exhibit some divergences linked to social-cultural norms between linguistic communities. Specifically, stupidity attacks appear highly prominent in English, while religious condemnation and vulgar sexuality references assume more importance in Arabic.

Further exploring such differences through in-depth analysis of language examples within these resources can yield novel insights into how culture shapes inappropriate online discourse in an important yet distinct way, as there is a lack of comparative linguistic research exploring how culture shapes inappropriate online discourse differently between Arabic and English on social media. Consequently, the current study aims to examine the attitude and the cultural element associated with the use of derogatory terms by Arabs on social media and the differences between the use of offensive languages in their native language (Arabic) and a second language (English). The study aims to compare online derogatory language in Arabic and English to understand how cultural norms influence offensive discourse differently across languages. The study also examines individuals' experience and exposure to derogatory items, their impact on cultural aspects, and how they impact various factors and users' response to abusive comments.

II. Literature Review: Cross-Cultural Differences in Abusive Language on Social Media

According to Hall's (1976) theory of high-context and low-context cultures, abusive use of language differs across cultures according to the divergent preferences for implicit/explicit communication. For Hall, high-context cultures generally rely more on implicit meanings, body language and shared knowledge rather than overt messages. Conversely, low-context cultures select clear, direct messages (Hall, 1976).

Several studies, including Al-Ibrahim et al. (2023) dealt with offensive language use on Arabic versus English-speaking social media platforms. Shummer (2020) and Mubarak (2023) analyzed offensive tweets and collected common insults and derogatory remarks that included religious offenses like *huwwa ʔiɲna kafara willa jahu:d willa madʒu:s* (Are we disbelievers, Jews or fire worshippers?) as well as insults with sexual references that used implications rather than explicit expressions. By contrast, Byrne (2017) dealt with derogatory English tweets. He found insults tended to incorporate more explicitly crude, profane, homophobic and racist terms directly referencing body parts and excretory acts. Standard derogatory English terms included "*shit*", "*bitch*", "*cunt*", "*fag*", and racial slurs like the n-word. Several studies also found "rape jokes" and rape threats were employed more frequently in English than in Arabic on social media (Jane, 2014).

Furthermore, the implicit Arabic insults depended more on contextual understanding of cultural/religious references, while English derogatory terms tended to be directly profane or to rely on scatology, sexually explicit language and racist/homophobic slurs (Hill, 2009). Gross et al. (2022) noted some English speakers also adopted more implication-based insults influenced by code-switching with other languages (Gross et al., 2022). Moreover, keeping with their predominantly low-context communication norms, derogatory English terms displayed a marked preference for crude explicitness over nuanced implicit meanings when intending to offend others (Dayem, 2019). This, in fact, indicates how the cultural context affects abusive language conventions.

Also, Khenfar et al. (2022) reported considerable code-switching on multilingual platforms where Arabic speakers sometimes used rather direct abuse in English. Variation in norms of offensive language use may influence perceptions of abuse. Mansour (2017) argued that direct profanity and slurs are most offensive. He contended that high-context societies interpret implicit messages as abusive, while low-context cultures highlight explicit denotative meanings (Mansour, 2017). Overall, the existing literature aligns with Hall's theoretical framework in showing cross-cultural variation in abusive language associated with high-context versus low-context communication norms. Arabic social media displays a preference for implicitly versus directly offensive insults, reflecting differences from English platforms. However, globalization and online code-switching behaviours may be lessening some distinctions. More comparative studies could further explore how communication context impacts the interpretation and use of vulgarity across societies.

Social and Psychological Effects of Exposure to Derogatory Remarks on Social Media

The use of abusive language containing derogatory remarks on social media platforms can negatively impact both individuals and society. When confronted with hostile speech targeting one's social identity, it can undermine an individual's sense of belonging and self-esteem. The offensive tweets collected by Samoshyn (2020) of offensive and derogatory tweets in English and Arabic tweets and those collected by Shammur (2020) demonstrate how derogatory terms like "bitch" are commonly used, especially towards women. In contrast, in Arabic, this is more based on the conduct, race and rank as some of the most recurring words used for women were *nisa:ʔ ɦa:qida:t* (spiteful women) or comparing them to animals such as *mara miΘl l-baqara* (a woman as fat as a cow), while others considered Arab women as pure, but abused women of other origins.

Additionally, associations between exposure to online hostility and increased psychological distress have also been found by researchers (Bilewicz & Soral, 2020). One study analyzed tweets mentioning feminine hygiene brands and found that 9% contained abusive content, with terms like "bitch" being the most prevalent (Kumar et al., 2021). Those

tweets elicited stronger negative emotions of anger and sadness from female readers. Another study showed cyberbullying victims were more likely to experience low self-esteem, depression, and suicidal thoughts (Maurya et al., 2022). Exposure to such hostile speech threatens core aspects of one's identity. Social Identity Theory argues that individuals will engage in identity protection strategies like distancing themselves from the target domain or attacking the credibility of the hostile out-group, as seen in some responses to harmful tweets (Tajfel et al., 1979). However, prolonged exposure undermines well-being. Samoshyn (2020) and Shammur (2020) show that abusive language is often targeted at people based on gender or ethnicity, marginalizing them in online spaces dedicated to communication and expression.

On a societal level, abusive language can normalize hostile speech and can enable further harassment or radicalize political views (Bilewicz & Soral, 2020). Content analyses of Twitter data found discussions involving political polarization while misinformation often contain toxic rhetoric attacking opponents with terms like "bitch", seen in the English language dataset (Bovet & Makse, 2019). Such interactions negatively affect online discourse and mutual understanding. Looking specifically at the data collected by Shammur (2020) of the derogatory tweets, it was observed that these remarks pose challenges from a political perspective. Additionally, issues like cyberbullying disproportionately impact marginalized groups facing multilayered discrimination online, like women and ethnic/religious minorities, as some tweets target appearance and religious identity. Oh (2022) suggests that anti-bullying interventions should consider identity-based harassment to promote online safety.

In essence, the negative social and psychological exposure to online hostility demonstrates the importance of promoting welcoming, inclusive online communities. While technology facilitates connection, abusive language threatens well-being, marginalizes groups, and undermines civic cooperation, highlighting an ongoing need for balance and culturally sensitive digital literacy programs.

III. Research Methodology

This research employs a convergent mixed-methods approach to conduct a nuanced comparative analysis of offensive language use on social media in English and Arabic. Both top-down deductive content analysis and bottom-up inductive thematic analysis are used. The deductive process involves a priori categorization of tweets into insult types based on established constructs including references to gender, ethnicity, religion, appearance, and ability-based attributes. This allows assessing the prevalence and distribution of harmful content quantitatively. Descriptive statistics is used to determine which derogatory lexical items frequently occurred in each language.

In parallel, meaning-based themes are generated from the data that underwent an inductive thematic analysis. The wider social and cultural contexts have thus been considered to understand both implicit and explicit abuse communicated through language choice, register variation, and socio-pragmatic norms, e.g. perceptions of what is acceptable vs. what is intolerable speech standards across cultures.

A. Data Collection

This study adopts a multi linguistic comparative approach, using quantitative and qualitative analysis of derogatory terms used in social media in both Arabic and English.

1. Arabic Corpus: A dataset of 920 tweets collected by Shammur (2020) forms the Arabic corpus. Tweets focus on derogation, targeting a diverse range of social identity constructs through pejorative lexical choices. Using a manual discourse-based annotation, tweets were categorized according to the type of insult they involved like gender (e.g *ʕa:hira* ‘slut’), religion (e.g. *waƟani* ‘atheist’), ethnicity (e.g. *ʕabd* ‘slave’), and culture (e.g. *ʔmaʔa xali:ʕa* ‘a whore’). This compilation represents the breadth of derogatory language varieties observed within Muslim-majority cyber communities.

2. English Corpus: The English corpus was taken from Samoshyn's (2020) collection of 1,500 discriminatory English tweets which explored racial, cultural, social and identity-based offence (e.g. "bitch", "faggot"). Computational text mining was employed to automatically identify overtly abusive lexical items (e.g. expletives) and implicit micro-aggressions.

The multi-genre, mixed-method assemblage of vernacular social media texts in these two corpora facilitates systematic cross-linguistic comparison of discursive strategies and sociocultural dynamics that uphold or challenge norms surrounding identity-targeted derogatory language on online platforms.

B. Data Analysis

A mixed-method approach categorizes and compares identity-targeting derogation in Arabic and English tweets. Digital language processing software automatically tags tweets with insults against gender (e.g. *zawrāʾ*, "bitch"), religion (e.g. *waƟani*, "atheist"), ethnicity (e.g. *ʕabd*, “slave”), appearance (*ʕawway* “like a barking dog”), and ability (*muʕa:q* “handicapped”).

The quantitative analysis assesses insult prevalence, distribution by identity construct, and frequent lexical items (e.g. *ʕa:hira*, "whore") to understand sociolinguistic typologies. Text frequency reports gauge derogation intensity across languages. Qualitative thematic discourse analysis then interprets emergent themes around ideological functions,

context-dependent interpretations, and cultural dynamics shaping offence norms. Implicit micro-aggressions are identified through critical vocabulary analysis.

Interactional sociolinguistic and cross-cultural perspectives reflect derogation's cognitive, psychological and relational impacts. Analysis contextualizes offence within macro-level cultural values to draw nuanced comparisons between Arabic and English cyber-communities' linguistic construction and signify social boundaries.

IV. Results

The frequency analysis findings presented in Table 1 indicate that gender and race faced a significant number of derogatory attacks at 42.4% and 28.4%, respectively. This targeting of identities suggests underlying issues related to sexism and racism which prominently reflect a form of offensive discourse in Arabic on social media. By exploring why and how these social groups experience a considerable level of abuse, it seems that implicit biases and sociocultural norms target marginalized communities. Thus, insights into tackling discrimination could aid in encouraging more inclusive online spaces.

Table 1
Targeted Insults

		Freq.	%	Valid %	Cumulative %
Valid	GENDER	212	42.4	42.4	42.4
	RACE	142	28.4	28.4	70.8
	RELIGION	52	10.4	10.4	81.2
	CULTURE	25	5.0	5.0	86.2
	APPEARANCE	69	13.8	13.8	100.0
	Total	500	100.0	100.0	

The results in Table 2 show that around half of tweets were mild, while about a third were moderate. Still, a good percentage advocated severe hostility. This range of derogatory Arabic discourse on social media reflects the varying degrees of severity that social identities experience when they are subjected to offensive remarks. In fact, the levels of insult displayed suggest sociocultural feelings about what constitutes unacceptable boundaries in online platforms governed by Arabic communication customs.

Table 2
Severity Level

		Freq.	%	Valid %	Cumulative %
Valid	Mild	226	45.2	45.2	45.2
	Moderate	155	31.0	31.0	76.2
	Severe	119	23.8	23.8	100.0
	Total	500	100.0	100.0	

The findings in Table 3 reveal that more than half of the Arabic tweets were deemed offensive (54.2%), while only vulgar tweets were 25.8%, signaling that offence extends beyond insensitivity to incorporate diverse harms. Distinguishing subtleties in how offence manifests linguistically across contexts deepens our understanding of this complex phenomenon. The data set revealed vulgar derogatory terms degrade respectability, e.g. *ʕa:hira* “bitch”, and *waqiħa* “rude”. These derogatory terms divide abusers into different groups. Offence does indeed come in numerous forms, both within and across communities.

Table 3
Vulgar or Offensive

		Freq.	%	Valid %	Cumulative %
Valid	VULGAR	129	25.8	25.8	25.8
	OFFENSIVE	271	54.2	54.2	80.0
	NONE	100	20.0	20.0	100.0
	Total	500	100.0	100.0	

The Arabic tweets corpus analysis in Table 4 revealed a relatively close balance between overt and covert types of discrimination. About 44% used hints and insinuations of implicit constructs, whereas the other 56% chose a more direct and explicit language. This distribution indicates that some people opt to contempt others through subtle hints rather than express it directly and crudely.

Table 4
Nature of Tweets

		Freq.	%	Valid %	Cumulative %
Valid	IMPLICIT	219	43.8	43.8	43.8
	EXPLICIT	281	56.2	56.2	100.0
	Total	500	100.0	100.0	

Using emoji was one of the most prominent features, as shown in Table 5. Almost half of the participants chose to add iconic signs (47%) to their entries. This demonstrated how the visual signs have become more and more integrated into the vulgar lexicons used to disparage others. Analyzing the complex socio-pragmatic significance and effects of multichannel denunciation through a combination of written and visual forms improves our understanding of the constantly evolving sophisticated strategies some commentators create and use in online forums to convey their dislike and contempt of other people through a variety of multimodal channels.

TABLE 5
USE OF EMOJIS

		Freq.	%	Valid %	Cumulative %
Valid	Yes	235	47.0	47.0	47.0
	None	265	53.0	53.0	100.0
	Total	500	100.0	100.0	

The range of language registers used, presented in Table 6, reflects that half of derogatory tweets in Arabic used the official lexis of Modern Standard Arabic, while almost a third used colloquial dialects, and around a fifth used even more derogatory slang language. This spectrum implied that contempt and degradation is common in written Arabic discourse regardless of standard or regional variations in linguistic style. A comparison of the terms used across the range of language registers improves our understanding of the connections between language diversity, identity, and the expressions of hatred.

TABLE 6
USE OF LANGUAGE

		Freq.	%	Valid %	Cumulative %
Valid	MSA	250	50.0	50.0	50.0
	DIALECT	150	30.0	30.0	80.0
	SLANG	100	20.0	20.0	100.0
	Total	500	100.0	100.0	

We now move to English tweets. As Table 7 shows, feminist or masculinist mockery dominated over half of the corpus, accounting for 49.6% of the online defamations. This may reveal a pattern of objectifying rhetoric that does not respect people's humanity. This is a blunt example of how the anonymity and low-context nature on social media enables users to degrade and humiliate others with no consideration for empathy.

TABLE 7
TARGETED INSULTS _ENGLISH

		Freq.	%	Valid %	Cumulative %
Valid	GENDER	248	49.6	49.6	49.6
	RACE	52	10.4	10.4	60.0
	RELIGION	34	6.8	6.8	66.8
	CULTURE	68	13.6	13.6	80.4
	APPEARANCE	98	19.6	19.6	100.0
	Total	500	100.0	100.0	

Table 8 indicates that moderate attacks on others made up half of the English tweets (49.8%), whereas sharply critical remarks made up a third (30.2%), indicating that contempt is prevalent in varying aggressive forms. Some people may become uncontrollably willing to criticize with greater ferocity than face-to-face interaction due to the safe distance and diminished responsibility of the online environment. A deeper examination of the most indelicate criticisms reflects a set of expressions that, in their sharp vulgarity, reveal a glaring lack of empathy that transcends human nature.

TABLE 8
SEVERITY LEVEL _ENGLISH

		Freq.	%	Valid %	Cumulative &
Valid	Mild	100	20.0	20.0	20.0
	Moderate	249	49.8	49.8	69.8
	Severe	151	30.2	30.2	100.0
	Total	500	100.0	100.0	

Analyzing the use of derogatory language on social media directed towards others (Table 9 below) reveals that seven out of ten tweets have vulgar expressions. This striking percentage suggested that certain participants are very offensive when they intend to hurt others. The lack of responsibility encourages the free expression of harsh opinions that, in less private contexts, could be met with opposition.

TABLE 9
VULGAR OR OFFENSIVE _ENGLISH

		Freq.	%	Valid %	Cumulative %
Valid	VULGAR	340	68.0	68.0	68.0
	OFFENSIVE	98	19.6	19.6	87.6
	NONE	62	12.4	12.4	100.0
	Total	500	100.0	100.0	

Table 10 reflects that more than half of the English tweets combined emotionally charged pictograms with their written words, accounting for 54.2% of the derogatory tweets. This astounding frequency suggested that emojis and symbols had been appropriated to the vocabulary of tweets. The disengaged environment of social media, which removes participants from immediate repercussions, may encourage users to reinforce aggressive statements with intentionally demeaning visual cues.

TABLE 10
USE OF EMOJIS _ENGLISH

		Freq.	%	Valid %	Cumulative %
Valid	Yes	271	54.2	54.2	54.2
	None	229	45.8	45.8	100.0
	Total	500	100.0	100.0	

Table 11 shows that in the context of derogatory language and offensive statements on social media, Standard English (about 33%) and slang (around 31.2%) are used almost equally. This implies that**,** despite linguistic differences, unpleasant language and insulting phrases are pervasive in internet communication. Such statements are common in the digital sphere, whether in formal language or informal slang.

TABLE 11
USE OF LANGUAGE _ENGLISH

		Freq.	%	Valid %	Cumulative %
Valid	Standard English	165	33.0	33.0	33.0
	Dialect	179	35.8	35.8	68.8
	Slang	156	31.2	31.2	100.0
	Total	500	100.0	100.0	

The derogatory remarks were found to be implicit, as demonstrated in Table 12, rather than explicit (71.2%). This is an interesting point to address in online communication: In English insults are often more subtly communicated than bluntly stated. It shows that the English derogatory tweets used indirect derogation due to their low-context nature, hence implying their negative thoughts rather than using overtly offensive language.

TABLE 12
NATURE OF TWEETS _ENGLISH

		Freq.	%	Valid %	Cumulative %
Valid	Implicit	356	71.2	71.2	71.2
	Explicit	144	28.8	28.8	100.0
	Total	500	100.0	100.0	

A. Differences Between Arabic and English Derogatory Tweet

Table 13 presents both similarities and differences in how derogation manifests linguistically between Arabic and English, reflecting underlying cultural attitudes. While both languages use offensive tweets, targets of insult and severity of criticism were higher in English tweets, indicating looser social boundaries. Conversely, English tweets employed more vulgar/offensive terms, consistent with norms. In addition, English speakers often derogate implicitly, unlike direct Arabic speech. Emojis supported aggression in both languages, yet Arabic diversified registers more. Fundamentally, the data uncovers how cultural contexts shape- but do not determine- the language of prejudice on social media.

TABLE 13
PAIRED SAMPLES TEST

		Paired Differences					t	Df	Sig. (2-tailed)
		Mean	Std. Dev.	Std. Err.	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Targeted_Insults-Targeted_Insults_Eng	-.238	1.743	.078	-.391	-.085	-3.053	499	.002
Pair 2	Severity_Level-Severity_Level_Eng	-.316	.465	.021	-.357	-.275	-15.183	499	.000
Pair 3	Vulgar_or_Offensive-Vulgar_or_Offensive_Eng	.498	.750	.034	.432	.564	14.839	499	.000
Pair 4	Nature_of_Tweets-Nature_of_Tweets_Eng	.274	.446	.020	.235	.313	13.723	499	.000
Pair 5	Use_of_Emojis-Use_of_Emojis_Eng	.072	.455	.020	.032	.112	3.537	499	.000
Pair 6	Use_of_Language-Use_of_Lang._Eng	-.282	.450	.020	-.322	-.242	-14.000	499	.000

B. Thematic Analysis

Below is a thematic analysis of the types of insults found in both languages.

THEME 1: Gender Targeted Insults

Gender and race targeting was notably more common in the Arabic tweets, reflecting underlying issues of sexism and racism in some Middle Eastern cultures. The tweets referred to women as whores and bitches as shown in the following tweets (Translated from Arabic):

“You guys amaze me; you seem so loyal and low key but you get brought down by some bitch women”

“We can buy just as much as those dirty and disgraceful women of yours”

“As a woman, you should not complain about cleaning your house because that is what you are here to do and at night you have to please us”

Several tweets insulted or degraded women directly. None of the English tweets contained such explicit targeting. However, cursing and insulting women had been observed to be pretty common in the English tweets discourse, which can be considered as vulgar:

“Keeks is a bitch. She curves everyone " lol, I walked into a conversation like this. Smh”

“Fu*k no, that bitch doesn't even suck d**k."

“These ho*s like niggas that spend money, not talk bout it."

THEME 2: Severity Leve

The severity levels of criticism and insults appeared higher in the English tweets according to the analysis as shown below:

“Get worshipping bi**h! woof woof”

“@C**lanG15, that nigga was eating that h*e lol" Hell Yea, lol John Paul," nigga said John Paul..”

"@Cha**rParsons: How bout them Cowboys!!!!" Shutup pu**y”

The English texts included more extreme or confrontational attacks compared to the Arabic samples as they are often criticizing in political contexts as:

“The Egyptian media is known all over the world for its filth, and this media incites the killing of its own people and therefore does not incite the killing of other peoples! (Translated from Arabic)”

“You can find it in the back of Tamim or Rajab Tayouza (Translated from Arabic)”

“These Egyptians work like American dolls and they suck their teams, the effect is that women can't get enough of them (Translated from Arabic)”

This indicates that linguistic or cultural boundaries around offense may be looser for English speakers on online media compared to the Arabic groups.

THEME 03: Implicit and Vulgar Vs Explicit and Offensive

The collected data by Samoshyn (2020) and Shammur (2020) revealed that instances of vulgar or highly offensive language were significantly more frequent in English tweets. This aligns with differences in cultural and linguistic norms around vulgarity and taboo topics across English and Arabic societies as illustrated below:

Arabic Tweets:

“@tamir I just f**ted and it was so loud my dog started barking (Translated from Arabic)”

“I had a great time playing with myself last night (Translated from Arabic)”

English Tweets:

“@ja**n_aw***: Had an awesome time last night; if only my roommate weren't such a prude, I would have shared more details."

“Why did the chicken cross the road? To get away from me because I'm a total a***ole."

The above-presented discourse was more common in the English tweets. This coincides with prior research showing online effects can reduce explicitness in some contexts. English speakers may have felt less social pressure to state criticisms directly than their Arabic counterparts.

THEME 4: Diversity and Visual Appeal

According to the quantitative analysis, the use of emojis was substantial in both datasets as Emojis function prominently as a universal language of expression online. However, the analysis found emojis to have appeared with significantly higher frequency in English tweets.

"Black people are all criminals, that's why they always end up in jail 😉😁".

"Shut up, you black piece of shit, or else I will slap you back into slavery 🙂."

Along with the emojis, the Arabic tweets employed a more diverse range of language registers, moving between formal, informal, and slang varieties within individual texts, as shown below:

"I hate it when women try to act tough, they look ugly and desperate 😐🔥" (Translated from a mix of formal and informal Arabic).

“Asian girls are very ugly, their eyes are slanted and their skin is dark 😳” (Translated from informal and slang Arabic).

In General, unlike the Arabic versions, the English samples tended to be more consistent in using either standardized or slang language styles. Anonymity and lack of non-verbal signs in online platforms can help surpass cultural differences in relation to forms of unpleasant communication. The findings show how cultural contexts determine linguistic choices even for biased expressions in a significant yet incomplete way.

Generally, the datasets of abusive tweets in Arabic and English showed a few important differences and similarities mainly related to gender targeting, insult severity, register use, emoji use and communicative styles. The study also showed paramount issues related to societal prejudices that seemed to percolate easily through social media.

V. Discussion

The study at hand has found that online offensive discourse is highly influenced by cultural norms. This strongly matches the findings of Al-Qattan (2021) and Bednarek (2019) who argue that Americans tend to use more vulgar expressions than middle easterners who are constrained by societal taboos. Also, the study at hand argues that English tweets are higher in vulgarity than their Arabic counterparts. Cultural settings also influence language choices. Like Almusallam (2018) and Chen (2018), directness is preferred in Arabic, so tweets are abusive in an explicit rather than implicit way. In addition, English users felt less social pressure when online, but Arabs preserved politeness forms to save face. This covert-overt distinction echoes Almusallam’s (2018) findings on high-low context cultural diversity.

It is worth noting that this study approved that emojis are integrated into toxic messages in both Arabic and English tweets, which was also argued by Husain (2021) and Kim et al. (2022). Emojis had the power to intensify insults and transcend linguistic barriers thanks to their visual iconicity. While emojis are globally understood online, Arab users diversified registers more within texts, reflecting the language’s multifaceted communicative identity. More noticeable in Arabic is gender targeting, as well. This underscored the underlying societal sexism, which is a cross-cultural root issue. This study is also consistent with the findings of Alsafari et al. (2020) who claimed that English and Arabic are not linked to systemic oppression and that abuses could be curbed by patriarchal dynamics. English scaled higher abuse severity, indicating unlimited offence boundaries for English speakers (Jay, 2018).

Anonymity, regardless of language, reinforces bias because of disinhibition effects. However, cultural bias spread differently. Although both languages tended to objectify women, English derogation objectified more, thus splitting humanity through gendered abuse sanctioned by western masculinity (Farwaneh, 2005). Worth noting is that standard-slang variation is balanced in English, unlike in Arabic. The variance between the languages implies some linguistic inequality between English and Arabic offensive expressions (Gauthier, 2017). Comparisons reveal links between language diversity, identity politics and hate expression that merit exploration to fight bias in a systematic way.

VI. Conclusion

The study’s findings provide strong insights into how abusive language is employed differently across English and Arabic on social platform, mirroring underlying sociocultural influences. Arabic tweets proved to use higher gender targeting level while English tweets used more vulgar expressions in accordance with the societal taboos. Register diversity and other linguistic variables also differed according to cultural norms. However, themes like prejudice transcended borders online with the help of anonymity and disinhibition effects. Particularly, emojis appeared as a global aggressive device underscoring shared online behaviors. However, offline topics related to identity politics and power differentials influenced minor groups distinctively in the context of each language. Overall, the study shed light on the intersection of linguistic, cultural and technological factors controlling online abuse across Arabic and English groups.

REFERENCES

[1] Al-Ibrahim, R. M., Ali, M. Z., & Najadat, H. M. (2023). Detection of Hateful Social Media Content for Arabic Language. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(9), 1-26.

[2] Al-Jarf, R. (2019). *Effect of Social Media on Arabic Language Attrition*. Retrieved on May, 22, 2024. https://www.researchgate.net/publication/353739228_Effect_of_Social_Media_on_Arabic_Language_Attrition

[3] Almusallam, I. I. A. (2018). *A discursive approach to politeness: Negotiating offers in women's talk by Saudi Arabic and British English speakers*. PhD thesis, University of Leeds.

[4] Al-Qattan, A. (2021). Swearing on Twitter: Khaleeji Dialect. *European Scientific Journal*, *17*(33). DOI: https://doi.org/10.19044/esj.2021.v17n33p151

[5] Alsafari, S., Sadaoui, S., & Mouhoub, M. (2020). Hate and offensive speech detection on Arabic social media. *Online Social Networks and Media*, *19*, 100096.

[6] Bednarek, M. (2019). 'Don't say crap. Don't use swear words.'–Negotiating the use of swear/taboo words in the narrative mass media. *Discourse, Context & Media*, *29*, 100293.

[7] Bilewicz, M., & Soral, W. (2020). Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, *41*, 3-33.

[8] Bovet, A., & Makse, H. (2019). Influence of fake news on Twitter during the 2016 US presidential election. *Nature Communications*, *10*. https://doi.org/10.1038/s41467-018-07761-2

[9] Byrne, E. (2017). *Swearing is good for you: The amazing science of bad language*. Profile Books. House of Anansi Press Inc.

[10] Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2021). Antisocial Behavior in Online Discussion Communities. *Proceedings of the International AAAI Conference on Web and Social Media*, *9*(1), 61-70. https://doi.org/10.1609/icwsm.v9i1.14583

[11] Dayem, A. A. (2019). A Cross-linguistic and Cross-cultural Study of Explicitness in English and Arabic Discourse. *Journal of Basra Research for Human Sciences*, *44*(3), 86-91.

[12] Fahmy, A. (2021). *Detecting Offensive Language in Multi-Dialectal Arabic Social Media*. Available at SSRN 4120709.

[13] Farwaneh, S. (2005). Asymmetries of Male/Female Representation in Arabic. In: Jule, A. (eds) *Gender and the Language of Religion*. Palgrave Macmillan, London. https://doi.org/10.1057/9780230523494_4

[14] Gauthier, M. (2017). *Age, gender, fuck, and Twitter: A sociolinguistic analysis of swear words in a corpus of British tweets.* Université Lumière Lyon, 2.

[15] Gross, M. C., López González, A. C., Girardin, M. G., & Almeida, A. M. (2022). Code-Switching by Spanish- English Bilingual Children in a Code-Switching Conversation Sample: Roles of Language Proficiency, Interlocutor Behavior, and Parent-Reported Code-Switching Experience. *Languages*, *7*(4), 246. Retrieved on Jun, 2, 2024. https://www.mdpi.com/2226-471X/7/4/246

[16] Hall, E. T. (1976). Beyond Culture. Garden City, New York: *Doubleday*, pp. 45-53.

[17] Mubarak, H., Darwish, K. (2014). Using Twitter to Collect a Multi-Dialectal Corpus of Arabic. In *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing* (ANLP), pages 1–7, Doha, Qatar. Association for Computational Linguistics

[18] Hill, J. H. (2009). *The everyday language of white racism*. John Wiley & Sons.

[19] Husain, F. A. (2021). *Arabic Offensive Language Detection in Social Media*. George Mason University.

[20] Jane, E. A. (2014). "Your a ugly, whorish, slut" understanding E-bile. *Feminist Media Studies*, *14*(4), 531-546.

[21] Jay, T. (2018). Swearing, moral order, and online communication. *Journal of language aggression and conflict*, *6*(1), 107-126.

[22] Kerkam, Z. M. (2015). *A comparison of Arabic and English directness and indirectness: Cross-cultural politeness.* Sheffield Hallam University (United Kingdom).

[23] Khenfar, I., Oulmi, R., & Ayadi, K. (2022). *Algerian Arabic English code-switching* by Master 2 students at Oum El Bouaghi University.

[24] Kim, J., Wohn, D. Y., & Cha, M. (2022). Understanding and identifying the use of emotes in toxic chat on Twitch. *Online Social Networks and Media*, *27*, 100180.

[25] Kumar, P., Gruzd, A., & Mai, P. (2021). Mapping out Violence Against Women of Influence on Twitter Using the Cyber–Lifestyle Routine Activity Theory. *American Behavioral Scientist*, *65*(5), 689-711. https://doi.org/10.1177/0002764221989777

[26] Mansour, D. M. (2017). *Functions and contextual triggers of offensive language on Twitter*. [Master's Thesis, the American University in Cairo]. AUC Knowledge Fountain. Retrieved on Jun, 4, 2024 https://fount.aucegypt.edu/etds/683

[27] Maurya, C., Muhammad, T., Dhillon, P., & Maurya, P. (2022). The effects of cyberbullying victimization on depression and suicidal ideation among adolescents and young adults: a three-year cohort study from India. *BMC Psychiatry*, *22*(1), 1-14.

[28] Mubarak, H., Rashed, A., Darwish, K., Samih, Y., & Abdelali, A. (2020). Arabic offensive language on Twitter: Analysis and experiments. arXiv preprint arXiv:2004.02192. Proceedings of the *Sixth Arabic Natural Language Processing Workshop*, pages 126–135.

[30] Mubarak, H., Hassan, S., Chowdhury, S.A., (2023) Emojis as anchors to detect Arabic offensive language and hate speech. *Natural Language Engineering*, *29*(6), 1-22 DOI: 10.1017/S1351324923000402.

[31] Oh, D. (2022). *Incivility and intolerance on Twitter: A case study of political tweets about abortion in Ireland (2018) and the United States (2020)* Loughborough University.

[32] Samoshyn, Andrii (2020). *Hate Speech and Offensive Language dataset.* Kaggle. Retrieved on May 12, 2024. https://www.kaggle.com/datasets/mrmorj/hate-speech-and-offensive-language-dataset

[33] Shammur. (2020). *GitHub - shammur/Arabic-Offensive-Multi-Platform-SocialMedia-Comment-Dataset: Arabic Dialectal Offensive Language dataset from social media comments on news posts from Facebook*, Twitter and YouTube platforms. GitHub. Retrieved on May 25, 2024. https://github.com/shammur/Arabic-Offensive-Multi-Platform-SocialMedia-Comment-Dataset

[34] Sood, S.O., Churchill, E.F., & Antin, J. (2012). Automatic identification of personal insults on social news sites. *J. Assoc. Inf. Sci. Technol*., *63*, 270-285.
[35] Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational identity: A reader*, *56*(65), 9780203505984-97802035059-16.

Nahla Al-Shalabi is an Assistant Professor in the Department of Arabic Language and Literature at Al Ain University (United Arab Emirates). She has worked at Jordanian universities as well. Dr. Al-Shalabi is specialized in linguistics and grammar. She graduated from Yarmouk University in Jordan. Her research interests include syntax, morphology, and discourse analysis. She has published several research papers in peer-reviewed scientific and international journals.

Hanene Lahiani is an associate Professor in English Language and Linguistics in Al Ain University (AAU), UAE. She was born in Sfax, Tunisia in 1971. She earned her bachelor degree in English Language from Laval University, Canada. She got her Master degree and Ph.D. degree in Linguistics from the University of Manouba Tunisia in 2001 and 2017 respectively. She also taught at the University of Sfax, Tunisia. Dr. Lahiani has taught in B.A., Diploma and M.A. levels. Her teaching interest includes general Linguistics, semantics, phonology, morphology, semiotics and Business English. Her research interest encompasses visual semiotics, social semiotics, visual rhetoric and linguistics.

Ayman Yasin is currently the head of the Coordination Unit for Service Courses at Princess Sumaya University for Technology in Jordan. He studied linguistics at Purdue University in Indiana, USA (2008-2012). He is interested in social sciences in general and in linguistics in particular. He has published a number of papers in syntax, phonology, translation, and prosody. Dr. Yasin has also taught MA and PhD courses in linguistics and translation at the University of Jordan (part timer). He is supervisor and an external examiner for many MA and PhD theses in various Jordanian universities. He has experience in IELTS testing and he serves as a reviewer in several journals. Dr. Yasin has taught different language courses at university and at language centers that focus on the four skills: listening, reading, writing and speaking. He adopts the communicative approach which makes his classes very fun and interactive.